Research on Service Model of Prediction in Cognitive Networks

Li Dandan[1,2,*], Xue Wanxin[1] and Pei Yilei[1]

[1]*Management School, Beijing Union University, Beijing, 100101 P.R. China;* [2]*Business School, University of International Business and Economics, Beijing, 100029, P.R. China*

Abstract: How to provide great service to users is one of the most important jobs to cognitive networks. Cognitive networks can perceive the external environment; intelligently and automatically change its behaviour to adapt the environment. This feature is more suitable to provide security for users with Quality of Service. This paper proposes a hybrid traffic prediction model, which trains BPNN with Ant Colony Algorithm based on the analysis of the present models. Furthermore, the model includes three stages, and the model predicts the network traffic with the hybrid model. The proposed model can avoid the problem of slow convergence speed and an easy trap in local optimum when coming up with a fluctuated network flow. Thus, the traffic prediction with high-precision in cognitive networks is achieved.

1. INTRODUCTION

A cognitive network (CN) has a cognitive process that can perceive current network conditions, and then plan, decide and act on those conditions. The network can learn from these adaptations and use them to make future decisions, all while taking into account end-to-end goals [1]. With the wide application of multi-media on the Internet, Quality of Service (QoS) of the network becomes more and more important. The advancement of CN can improve the user experience.

The feature of the network traffic reflects the interaction and influence in the process of data transmission. By analyzing the traffic data, people can better learn the interior operation mechanism of the network and build a mathematical model that can depict the traffic data flow more accurately. Designing a traffic prediction model with the cognitive feature can make networks give a more reasonable bandwidth assignment, traffic control, routing control, admission control and error control, *et al.* [2]. It is a good method to improve QoS.

CN has the capacities of self-learning and self-adaption. Therefore, researching a real-time prediction model of the traffic based on CN can better solve the problem of load balancing hysteresis. And a high-precision traffic prediction model, especially short-term one, can improve the cognitive feature of CN. However, current network traffic prediction models are mostly based on regular networks, few of them specially research on cognitive features of networks. Usually, their self-learning and self-adaption capacities are not so good, cannot express the cognitive features either, so it is hard to be applied in CN directly.

We analyzed the status of network traffic prediction methods, improved BP neural network, and Ant Colony Algorithm (ACA) is employed to train weight values, the design uses a combination of neural networks, we simulate the proposed model in MATLAB environment to validation and analysis model. The conclusion is that the new model is more accurately than the traditional model.

This paper is divided into six parts. Section II gives a summary about current network traffic prediction models. In section III, we give a brief introduction of the related theories. A detail description of the new model is presented in section IV. Simulation analysis of the model is shown in section V. Section VI is the conclusion

2. BRIEF LITERATURE REVIEW

The conception of CN is developed on the basis of cognitive radios, and current researches mainly focus on link layer, so research in network traffic prediction of CN is not enough. However, fundamentally the algorism of traffic prediction in CN is to show "congnition", provide necessary data to other elements of CN. Thus, from the aspect of time and data, it demands higher efficiency and precision than traditional model.

The present technologies of network traffic prediction can be divided into linear prediction and nonlinear prediction. ARIMA model [3] and Kalman Filter model are the examples of linear prediction model, such as self-adaptive linear model proposed by Lv Jun [4]. As the linear model can hardly describe the true features of real network traffic, nonlinear prediction models are proposed by scholars. Of course, nonlinear models sometimes may include linear elements, such as multi-scale combination predication model proposed by Khotanzad A [5]. The precondition to use ARIMA model is that the network traffic must have feature of linear wide stationary processes, while the network traffic always has multi-scale feature in different time frequency

scales [6] and the nature of multi-construct and self-similarity as well [7,8]. So it is hard to express the whole feature of network traffic with ARIMA singly. In these kind models, as nonlinear elements play a more important role than linear elements, we still call them nonlinear model. The typical model of nonlinear model is NN model.

Neural Network (NN) shows its great advantage in prediction because of its capacities of nonlinear approximation, self-learning and fitting. They can be divided into two kinds. One is to put some algorithm inside NN, form the scheme as Fig. (**1**) shows.
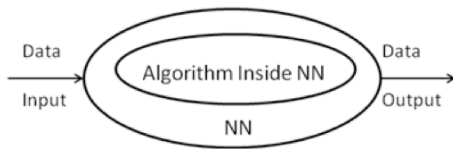

Fig. (1). Inside of NN.

WNN (Wavelet Neural Network) proposed by Wang Peng [9] just put wavelet decomposition algorithm into hidden layer of NN. We call this Prediction Model of NN with Build-in Algorithm.

The other is to separate the data processing and one fold NN's prediction as Fig. (**2**).

As for the traditional problem in prediction model proposed by Kasabov N [10], some scholars, such as Wang Peng [9] and Cheng Guang [11] have formed all kinds of improved models based on NN. Also, traditional NN models have the problem of lagging in learning and are easy to trap in local optimum, owning limited ability in coping with sudden load in the net. In order to solve this problem, the main resolution is put wavelet decomposition algorithm inside the 3rd layer of BPNN, Han Zhijie [12] properly deals with the local suboptimal, but still spends much time getting a satisfying result. References [13] and [9] employ WT and three layers BPNN; Reference [14] employs WNN in Next Generation Network (NGN); Reference [15] uses WT, RBFNN and Elman NN, then synthesizes outputs with BPNN; Reference [16] combines WT with FIRNN. These applications depict the feature of network traffic flow very well, and reach high-precision. Back Propagation (BP) algorithm is the most popular one nowadays, mainly because it is on the basis of strong theories and uses widely. But BP algorithm is based on gradient descent algorithm, so it needs a long time to train weight values, and easily falls into local optimum.
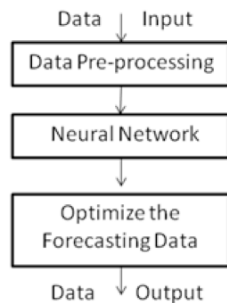

Fig. (2). Outside of NN.

ACA [17] is a new evolutionary optimization algorithm from the natural behavior of ant colony. It possesses charac-

teristics of high speed, global convergence and heuristic study. Recent years, there are some models which combine ACA with BPNN, for example [18-21], determination of appropriate BPNN structure and weight using ACA. The simulation results show that extensive mapping ability of CN and rapid global convergence of ACA can be obtained by combining ACA with NN. But because network condition is complex and uncertain, and users who access network are diverse, data like sudden traffic appears easily, thus affect the precision of traffic prediction. First, the data mean nothing to network traffic prediction. If we use them to train BP NN, they will increase the complexity of model, and reduce the accuracy of model. Second, when we deal with data using WT or normalization, it will affect the precision because of the abnormal data. So, it is necessary to deal with the original data first, to reject abnormal data. In this way [22], proposes a network traffic prediction model using double BPNN. Firstly, it judges, identifies and rejects abnormal data with BP1, then inputs them into BP2 to train. This paper will be based on it to improve the accuracy of prediction for network traffic.

3. RELATED THEORIES

In this section, we review some related theories which involved the proposed model, such as ARIMA, WT (Wavelet) and Neural network model.

3.1. ARIMA

ARIMA is a time series prediction model. It is the typical representative of linear prediction models, which can supply accurate short term prediction. Now, we introduce ARMA first. ARMA(p,q) is given by (1).

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + ... + \varphi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q} \quad (1)$$

where y_t is an observed value of period t. ε_t is an error or deviation of t period, that's to say, it is the random factors which cannot be explained by model. φ_i and θ_i are the parameters to be estimated, $i \in [1,p], j \in [1,q]$.

If a time series is non-stationary, we handle data with zero-mean and difference stationary to make it become a stationary time series. After processing, ARMA becomes ARIMA. ARIMA(p,d,q) is given by (2).

$$\phi(B)\Delta^d y_t = \delta + \theta(B)\varepsilon_t \quad (2)$$

where

$$\phi(B) = 1 - \phi_1 B - \phi_2 B - ... - \phi_p B^p \quad (3)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B - ... - \theta_q B^q \quad (4)$$

Δ is difference. Δd is d-order difference. B_k is k-steps shift operator, it means $B^k x_t = x_{t-k}$. $\varphi(B)$ is autoregressive operator. $\theta(B)$ is shift mean operator.

From the analysis we can see that, ARMA is a special example of ARIMA when d=0. It means both non-stationary and stationary time series could be expressed with ARIMA(p,d,q).

3.2. Wavelet Transform

Wavelet Transform is one of the most effective methods in dealing with non-stationary time series. Generally speaking, for an arbitrary continuous variable function or signal *f(t)*, continuous WT is given by (5), and we call it WT for short.

$$W_\psi[f](a,b) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t)\overline{\psi(\frac{t-b}{a})}dt \tag{5}$$

While discrete WT is given by (6).

$$W_\psi[f](m,n) = |a_0|^{-m/2} \int_{-\infty}^{\infty} f(t)\overline{\psi(a_0^{-m}t - nb_0)}dt \tag{6}$$

where *a* is scale parameter. *b* is step parameter. $\psi(t)$ is a basic wavelet, which is positive and negative shocks, its mean is zero. *f(t)* is a square enterable function, $f(t) \in L^2(R)$. $a = a_0^m$, $b = nb_0a_0^m$ $(m, n \in Z)$.

In this paper, we decompose network traffic with Mallat algorithm. It is an easy recursion formula for wavelet and scale parameter, based on (7).

$$\begin{cases} A_0(k) = f(k) \\ A_j(k) = 2^{-1/2}[A_{j+1}(2k) + A_{j+1}(2k+1)] \\ D_j(k) = 2^{-1/2}[D_{j+1}(2k) + D_{j+1}(2k+1)] \end{cases} \tag{7}$$

where $j \in [1,L]$, L is steps of layer. $A_j(k)$ and $D_j(k)$ are approximate and detailed signal.

3.3. Neural Network Model

Nowadays, BPNN, ENN, HNN, and KNN, etc. are the popular NN models in use. We just introduce KNN and BPNN that we used in the novel model.

KNN's main idea is to self organize the information outside into a conception in brain. As for a system, it is just to organize a corresponding presentation format in system automatically when affected by information outside the system. This includes adjustment of the NN's weight coefficient.

KNN is a typical self organizing neural network which is also called SOM. Its input-layer is monolayer and single-dimensional neurons, while the output layer are two-dimensional neurons. The format of lateral interaction between neurons in output layer is Mexican Cap. So in the output layer, KNN has the feedback character that the neurons are closer, the effect is stronger. Thus, KNN can be the detector of mode characteristics and an effective method to enhance the ability of self adaptively. In this view, KNN can be a prompt model in the use of network traffic prediction.

Based on the original NN, through self-organized learning to simulate the biological nerve reflex is the typical character of KNN. In a complex nonlinear system that changes greatly, KNN can improve the forecasting precision by the change of the content and amount of study.

The learning process of KNN is as follows:

(1) Initialize the link weights, each weight can be initialized from the training data arbitrarily;

(2) $X_k = (x_1, x_2, ..., x_n)$ is input vector, for each input vector, calculate the Euclidean distance W_{ij} which between X_{ik} and all of the output node N_j;

(3) N_j*, which has the minimum distance in output node, is the winner in competition $d_{j*} = \min\limits_{j \in \{1,2,...,m\}}\{d_j\}$;

(4) Adjust the link weights between output node Nj* and every input node Xik in geometric neighborhood: $w_{ij} = w_{ij} + \eta(t)(x_i^k - w_{ij})$, $i \in (1,2...N)$, where $\eta(t)$ is learning rate: $0 < \eta(t) < 1$;

(5) As for different t: t=1,2,…, come back to step (3).

BPNN is a multilayer feed forward network based on error back-propagation algorithm, which is one of the broadly used ANN models. BPNN can learn and restore many input-output mapping relations without mathematical equations that reveal the relation between them. The learning rule of BPNN is to make use of the steepest descent method, continuously adjust the net weights value and net threshold value by back propagation in order to get a minimum square sum of error. The learning course of BPNN is as Fig. (**3**).
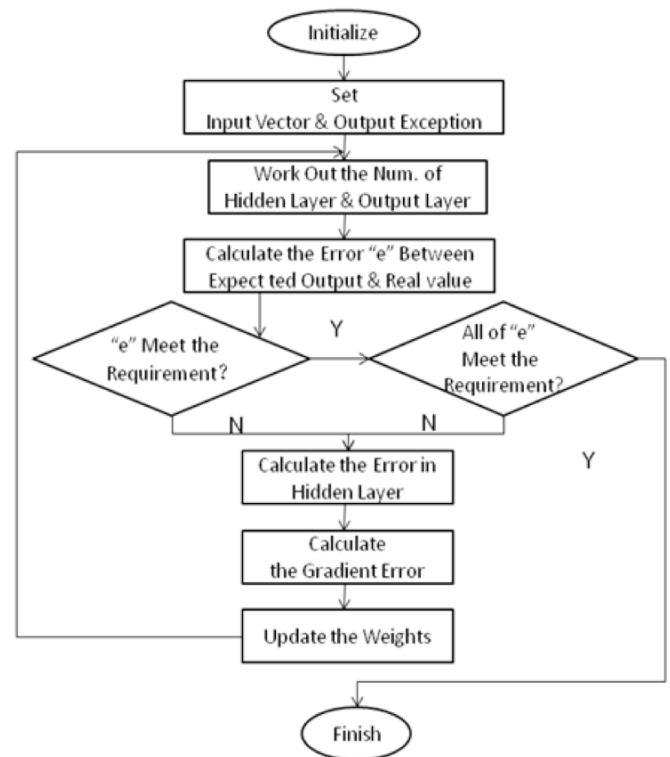

Fig. (3). The Learning Course of BPNN.

4. THE PROPOSED MODEL

4.1. Basic Thinking of New Model

As the one of the basic ANN model, BPNN works on the principle of continuous error feedback, the learning course of BPNN is as Fig. (**3**). The error is transmitted through output layer, then to the hidden layer and input layer, and weight in every layer can be corrected in the method of gradient descent algorithm. In cycles of information forward propaga-

tion and error back propagation, weights in every layer can be modified, and this is the process of NN's learning. This process will last until the output error can be accepted or the set learning times are met.

AR model has a good performance in the prediction of relative stationary traffic, and it can reduce the prediction time and improve the efficiency of the whole forecasting model.

KNN model can keep the character of wavelet's multi-scale analysis, avoid poor adaptive ability of wavelet coefficient in processing, achieve the aim of dynamic learning while reduce the NN's learning period, avoid being trapped in local optimum, thus improve the performance of prediction model.

According to Kolmogorov Theory, if the number of input unit is N, then the unit number of hidden layer is always 2N+1. In this model, we use the superposition of traffics which predicted by AR and KNN model as the input of BPNN. For BPNN's excellent ability of function approximation, we can get a more accurate forecasting result with small error.

In order to improve the accuracy of BPNN, we take [22] for reference, which trains BPNN with Ant Colony Algorithm. That is, we determine BP structure and weight value using ACA. It can avoid the problem of local optimum and lower the speed of convergence while the weight value has no business of training samples.

4.2. Procedure of New Model

Based on the analysis above, we form the structure of the new model as Fig. (**4**).

In the first stage of the model, decompose the nonlinear traffic signal with Mallat algorithm according to formula (7).

In the second stage, in line with formula (8), forecast the relatively stationary signal by AR model; as for the nonlinear and non-stationary high frequency signal, we use KNN model to make the forecasting. In the last stage, BPNN fit the superposition of former forecasting results, output a high precision final result.

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + L + \phi_p X_{t-p} + \varepsilon_t \tag{8}$$

where f_1, f_2, L, f_p are autoregressive coefficients.

Compared with the traditional single NN model, the new model avoids the problem of trapping in local optimum, and resolves the problem of long learning period and poor self-adaptively under a sudden load.

The speed of wavelet decomposition and linear forecasting can improve the efficiency of the whole model. In reality

application, regular learning cycle can be written into protocol and save the net resource. As a NN with good output performance, BPNN makes the forecasting result more accurate. Also, its ability in learning can make up the defect of long time learning and poor self-regulation mentioned above. Meanwhile, as a non-stationary model, BPNN is suitable to depict the non-stationary traffic.

Because this novel model can reduce the forecasting time, improve the forecasting precision, and receive a result with high precision under wide traffic fluctuation. So in the aspect of time and precision, this new model shows the meaning of cognition. It can work in the net under the condition of multi-background traffic or multi-scale time variation; provide a timely, accurate, steady cognitive platform to the high level algorithms.

4.3. Determination of Appropriate BPNN Weight

An ant always can find the shortest path between food source and formicary. That is because when ants cross the road, they leave something volatile, we call it pheromone. So the more ants walk across the road, the more pheromone will be left; while the more pheromone, the more ants follow it on this path. If there is no ant on the path, the path pheromone volatilizes itself, which is called positive feedback effect.

In this section we take [22] for reference. The parameter table is a set of parameters on each route, represent by *PTBn*. Where n is a sequence of route in network. Which is consist group of the parameters for an ant choose from *PTBn*, each parameter in *PTBn* represent by pa_{ni}, where i = 1,2,…, P (P is the number of parameter in *PTBn*). The pa_{ni} consist pheromone (τ_{ni}), connection (c_{ni}) and weight (w_{ni}), shows in Table **1**.

Detailed steps are as following.

Step 1: Initialization. Give an initial value for pa_{ni} for all *PTBn* on the routes. Set τ_{ni}=τ0, where τ_0 is the initial value of the pheromone. The rest parameters are random. Take 2×2×1 BP NN for example, Fig. (**5**) is the framework of ant k=1.

Step 2: Route traversal. An ant k starts from arbitrary route, selects pa_{ni} from *PTBn* is given by (9) and (10). It depends on a random variable q uniformly distributed over [0,1], and a parameter q_0 ($q_0 \in [0,1]$). M mean square error (MSE) is calculated. An ant k which travels all routes decodes the parameters shows in Fig. (**6**). Then BP structure is determined shows in Fig. (**7**).

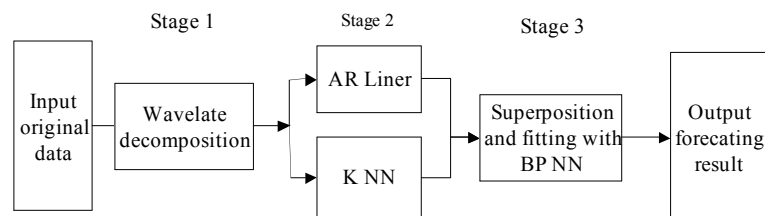$$i = \begin{cases} \arg\max_{1<i<p}\{\tau_{ni}\}, q \le q_0 \\ p_{ni}, else \end{cases} \tag{9}$$

Stage 1   Stage 2   Stage 3

Input original data → Wavelate decomposition → AR Liner / K NN → Superposition and fitting with BP NN → Output forecasting result

Fig. (4). Three-stage combined NN model.

Table 1.   The parameter table (*PTBn*) on route n.

Number	Pheromone	Connection	Weight
Pa_{n1}	τ_{n1}	c_{n1}	w_{n1}
Pa_{n2}	τ_{n2}	c_{n2}	w_{n2}
Pa_{n3}	τ_{n3}	c_{n3}	w_{n3}
……	……	……	……
Pa_{np}	τ_{np}	c_{np}	w_{np}
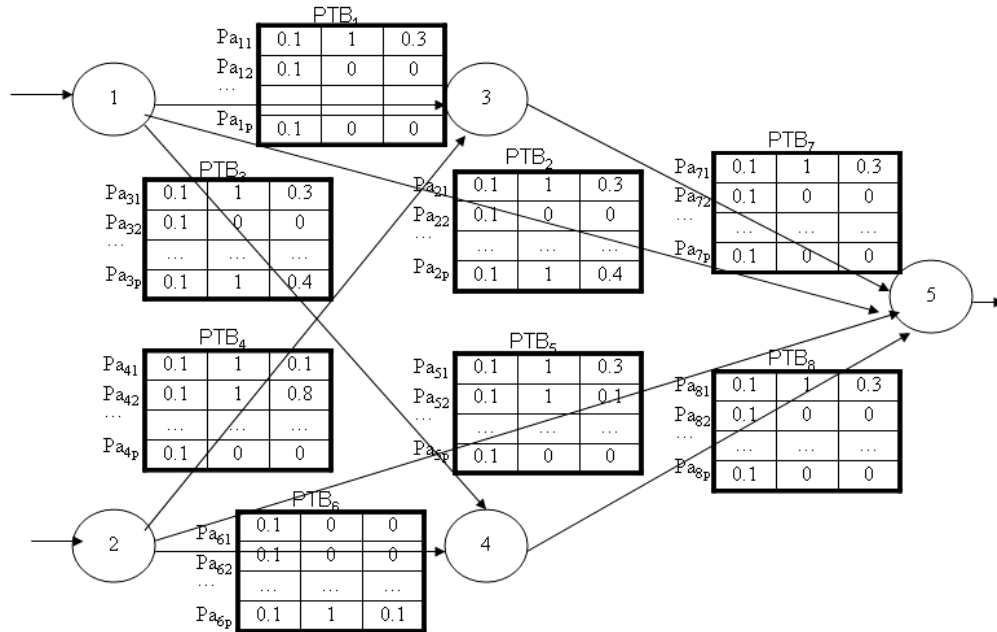
Fig. (5). The parameter table (PTBn) on every route.
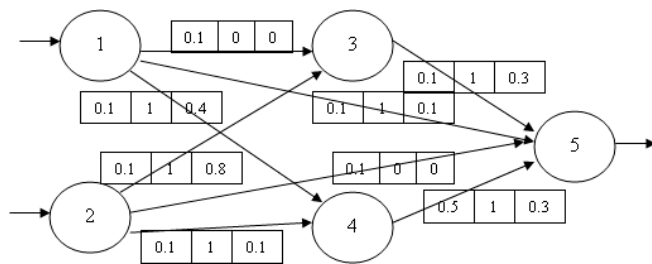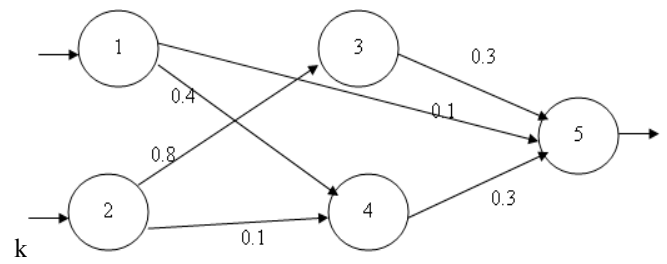

Fig. (6). The parameter of ant.

Fig. (7). The structure of ant k.
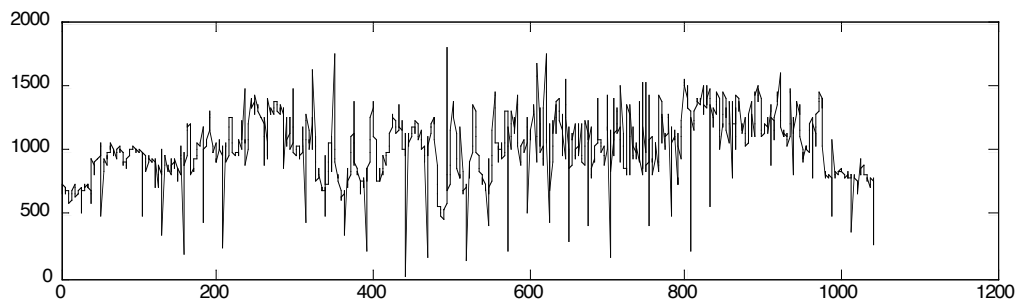
Fig. (8). The curve of original data.

$$p_{ni} = \tau_{ni} / \sum_{j=1}^{p} \tau_{nj} \tag{10}$$

Step 3: Updating local pheromone. An ant k applies this step only to the pani choosed form all PTBn in the routes, then it updates the pheromone (τ_{ni}) level is performed by applying (11).

$$\tau_{ni} = (1-\varphi)\tau_{ni} + \varphi\tau_0 \tag{11}$$

where $\varphi \in (0,1]$, and φ is the pheromone decay coefficient.

Step 4: Repeat step2-step4, for every ant.

Step 5: Updating global pheromone. A global update is applied to the pheromone on the pani belonging to the best solution or least MSE using (12) and (13).

$$\tau_{ni} = \begin{cases} (1-\rho)\tau_{ni} + \rho\Delta\tau_{ni}, i \in MSE_{best} \\ \tau_{ni}, otherwise \end{cases} \tag{12}$$

And

$$\Delta\tau_{ni} = 1 / MSE_{best} \tag{13}$$

where $\rho \in (0,1)$, and ρ is the evaporation rate, $\Delta\tau_{ni}$ is additional pheromone and MSE best is the least MSE of the tour ant.

Step 6: Judging end condition. If the maximum number of iteration is run or it reaches the error precision that is given before, go step 7; else go step 2.

Step 7: Determination of BP structure and weight value. Use c_{ni} to determine BP structure, it means deleted connections which $c_{ni}=0$, and left connections which $c_{ni}=1$. The weight value is the result that is convergence by ACA.

Step 8: The end.

From this algorithm we can see that it does not completely depend on original data using ACA, and makes the whole model posses better self-learning and self-organization ability. It embodies the feature of CN well, and more suitable in CN.
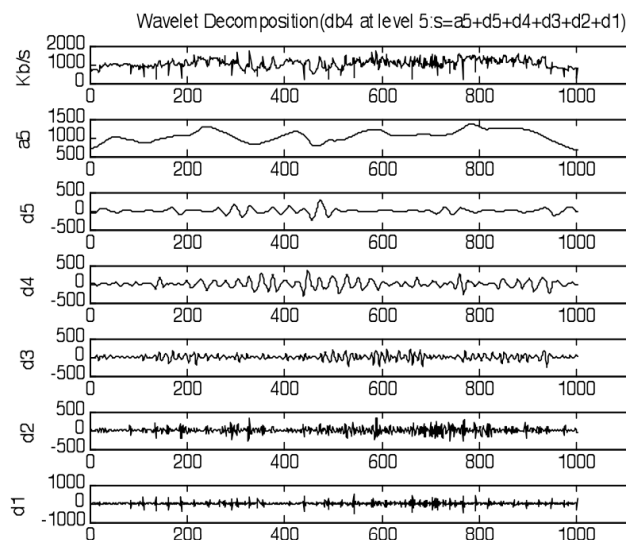

Fig. (9). The result of data Wavelate.

5. SIMULATION AND ANALYSIS

In this section, MATLAB [23] is employed to simulate and analyze the new model. The simulated data are history data from a core router of backbone of 1 hour intervals. We gathered 1040 samples, each sample is the average traffic value in an hour, and the last 140 samples have been used to test the forecasting result. The curve of the network traffic flow shows in Fig. (**8**).

For traffic data in Fig. (**8**), transmit them by (1). L=5, wavelet coefficient series after transmitting is {$D_1(k)$, $D_2(k)$, $D_3(k)$, $D_4(k)$, $D_5(k)$, $A_5(k)$}, and $t=d_1+d_2+d_3+d_4+d_5+a_5$, the result shows in Fig. (**9**).

Then we input the low frequency a5 into AR model to forecast. At the same time, normalize the high frequency d_1, d_2, d_3, d_4, d_5, and input them into KNN to forecast. The weight values and threshold values of KNN are determined by self-learning. We set the start learning rate on 0.95, the minimum learning rate on 0.001, the maximum step of training on 5000. At last, composite the AR model's forecasting result and KNN model's forecasting result, make it as the input of BPNN. According the Kolmogorov algorithm mentioned above, there are 3 hidden layers in BPNN.

The fitting result by last 140 signals shows in Fig. (**10**), it is a short time prediction for one hour.

Fig. (**11**) is the fitting result by last 140 signals which is a short time prediction for two-hour. And Fig. (**12**) is 24-hour.

From Figs. (**10**) and (**11**), we can see that the new model reaches high precision and fitting degree in short term prediction. Although in Fig. (**12**) 24-hour prediction is not good as Figs. (**10**) and (**11**), its error is so small, and it does not influence the effective of model.

6. CONCLUSION

This paper is based on NN, through analyzing the advantage and disadvantage of current network traffic prediction models. Taking into account the multi-scale and non-linear characters of the network traffic, combined with the wavelet decomposition, AR model and NN models, then a new model is proposed which has great self-learning and self-adaptation ability. BP weight does not depend on training samples using ACA at all.. Meanwhile, using hybrid model obtains high fitting and prediction prevision. Comparing with other models, the new model solves the problem of low speed convergence and local optimum, improves the prevision. In stage two, two kinds of the signals are predicted with AR model and KNN model respectively. To enhance the prediction accuracy and merge the traffic characters captured by individual models, the output of the previous models are combined using BPNN. The Simulation in MATLAB shows that performance of the novel model is better.

But the proposed model is very complex, so how to improve the efficiency with high-precision character is the researching trend of this paper.

CONFLICT OF INTEREST

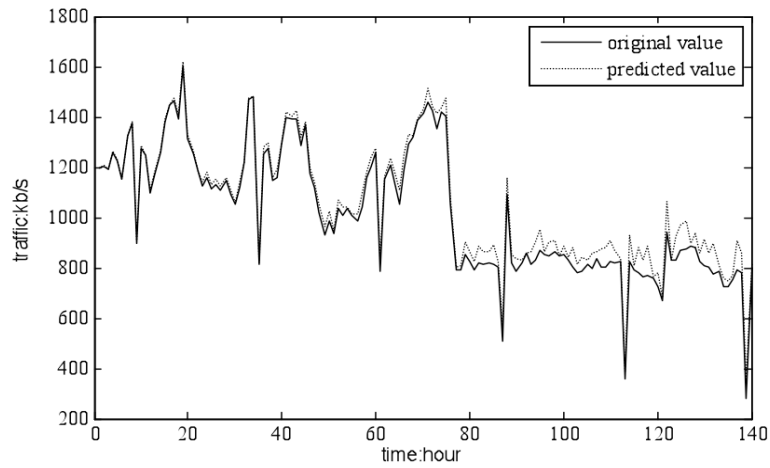The authors confirm that this article content has no conflict of interest.

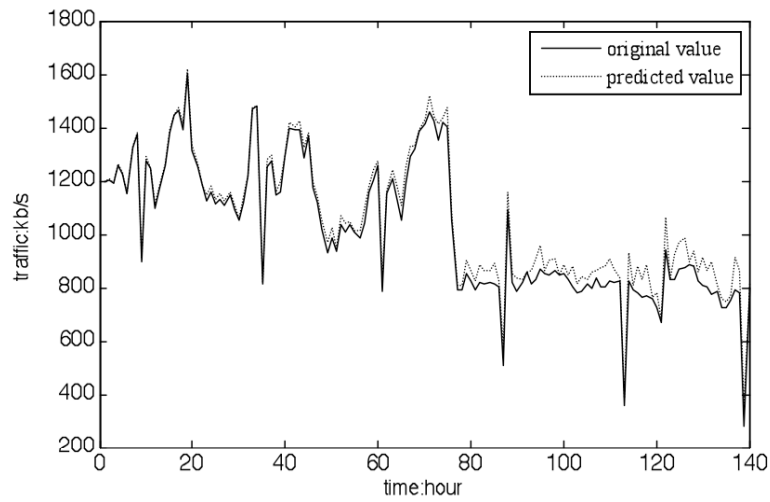**Fig. (10).** Proposed model one hour prediction.

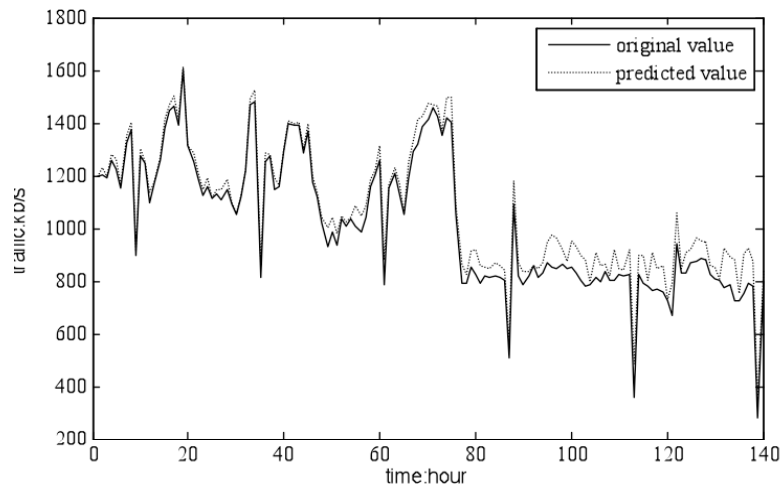**Fig. (11).** Proposed model two-hour prediction.

Fig. (12). Proposed model 24-hour prediction.

ACKNOWLEDGEMENTS

REFERENCES

[1]   R. W. Thomas, L. A. Da Silva, and A. B. Mackenzie, "Cognitive Networks", In: *Proceedings of IEEE DySPAN 200*5, November 11-15, 2005, pp. 352-360.

[2]   Z. Wang, Y. Sun, Z. Chen, and Z. Yuan, "Study of Predicting Network Traffic Using Fuzzy Neural Networks", *Journal on Communication*, vol. 26, pp. 136-140, 2005.

[3] F. Huifang, and S. Yantai, "Study on network traffic prediction techniques. wireless communications", *Networking and Mobile Computing*, vol. 2, pp. 1041-1044, 2005.

[4] L. Jun, L. Xing, C. Ran, and H. Tao, "Network Traffic Prediction and Fault Detection based on Adaptive Linear Model", *IEEE ICIT International Conference*, 2004, vol. 2, pp. 880-885.

[5] A. Khotanzad, and N. Sadek, "Multi-scale high-speed network traffic prediction using combination of neural networks", *Neural Networks,* vol. 2, pp. 1071-1075, 2003.

[6] X. Ren, Y. Yu, J. Zhang, L. Ma, and X. Ma, "Parameter estimation and application of time-varying farima model", *IJACT*, vol. 3, pp. 89-94, 2011.

[7] C. Gao, L. Han, Z. Cen, and C. Chu, "A New Multi Fractal Traffic Model Based On The Wavelet Transform", In: *Proceedings of the ISCA 14th International Conference: Parallel and Distributed Computing Systems*, Richardson, Texas, USA, 2001, pp. 157-162.

[8] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the Self-similar Nature of Ethernet traffic" (extended version), *IEEE/ACM Transactions on Networking,* vol. 2, pp. 1-15, 1994.

[9] W. Peng, and L. Yuan, "Network Traffic Prediction based on Improved BP Wavelet Neural Network", In: *4th International Conference on Wireless Communications, Networking and Mobile Computing*, Dalian, 2008, pp. 1-5.

[10] N. Kasabov, "DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction", *IEEE Transactions on Fuzzy System,* vol. 10, no. 2, pp. 144-154, 2002.

[11] C. Guang, G. Jian, and D. Wei, "Nonlinear-periodical network traffic behavioral forecast based on seasonal neural network model", *Communications, Circuits and Systems,* vol. 1, pp. 683-687, 2004.

[12] H. Zhijie, and W. Ruchuan, "Novel peer to peer network traffic prediction algorithm", *Computer Science,* vol. 9, pp. 40-41, 2008.

[13] T. Lei, and Z. Yu, "A wavelet neural network model of network traffic forecast", *Computer Application,* vol. 26, pp. 526-528, 2006.

[14] Q. Zhao, X. Fang, and Q. Li, "WNN—based NGN Traffic Prediction", In: *Proceedings of ISADS 2005*, Chengdu, China, 2005, pp. 230-234.

[15] M. Yao, Y. Liu, and G. Zhou, "Network traffic prediction model of wavelet combined neural network", *Computer Engineering and Design,* vol. 28, pp. 5135-5136, 2007.

[16] N. Tian, and L. Yu, "A WAN network traffic prediction model based on wavelet transform and FIR neural networks", *Journal of Electronics & Information Technology*, vol. 30, pp. 2499-2502, 2008.

[17] R. Schoonderwoerd, O. Holland, and J. Brute, "Ant-based load balancing in Tele-communications networks", *Adaptive Behavior,* vol. 5, pp. 169-207, 1996.

[18] B. Hong, F. Jin, and Q. Gao, "Multi-layer feed forward neural network based on ant colony system", *Journal of Harbin Institute of Technology,* vol. 35, pp. 823-825, 2003.

[19] N. Yang, Q. Fu, and S. Wang, "Improvement of wavelet neural networks model and application", *Systems Engineering-Theory & Practice,* vol. 29, pp. 168-173, 2009.

[20] W. Gao, "Evolutionary neural network based on new ant colony algorithm", *Proceedings of IEEE/2008 International Symposium on Computational Intelligence and Design*, vol. 1, pp. 318-332, 2008.

[21] N. Pokudom, "Determine of Appropriate Neural Networks Structure Using Ant Colony System", In: *Proceeding of ICROS-SICE International Joint Conference*, Fukuoka International Congress Center, Japan, 2009, pp. 4522-4525.

[22] D. Li, R. Zhang, and X. Shang, "A new application in neural network", *Advances in Information Sciences and Service Sciences,* vol. 3, no. 6, pp. 151-159, 2011.

[23] L. Liu, J. Chen, and L. Xu, "Realization and Application Research of BP Neural Network based on MATLAB", *Future Bio-Medical Information Engineering International Seminar,* 2008, pp. 130-1338.