



A NETWORK TRAFFIC PREDICTION MODEL BASED ON QUANTUM INSPIRED PSO AND WAVELET NEURAL NETWORK

Kun Zhang

Department of Mathematics, Chuxiong normal university, 675000, Chuxiong, Yunnan, China zhangkun@cxtc.edu.cn

Abstract- Network traffic flow prediction model is fundamental to the network performance evaluation and the design of network control scheme which is crucial for the success of high-speed networks. Aiming at shortcoming of the conventional network traffic time series prediction model and the problem that BP training algorithms easily plunge into local solution, a network traffic prediction model based on wavelet neural network and PSO-QI is presented in the paper. Firstly, the quantum principle obtained from Quantum PSO(QPSO)has been combined with standard PSO to form a new hybrid algorithm called PSO with Quantum Infusion(PSO-QI). Then, the parameters of wavelet neural network were optimized with PSO-QI and the time series of network traffic data was modeled and predicted based on wavelet neural network and PSO-QI. Experiments showed that PSOQI-wavelet neural network has better precision and adaptability compared with the traditional neural network.

Key Words- BP neural network, particle swarm optimization, PSO-QI algorithm, wavelet network traffic

1. INTRODUCTION

With the rapid development of computer and network communication technology, Internet traffic prediction plays a important role in network resource allocation and congestion control [1][2]. The network traffic prediction model is the foundation of network performance analysis and designing. Through the information provided from a network traffic prediction system, people could initiate certain network traffic control methods that help to manage route path to enhance network efficiency and know path that not overload. In the past several decades, network traffic prediction models use linear time series models, e.g. AR (Auto-regressive models), MA (moving average models), ARMA (Auto-regressive moving average model), and ARIMA (Auto Regressive Integrated Moving Average),etc [3][4]. A number of studies of network traffic time series is self-similar or long-range dependent(LRD) in nature. Since the network traffic time series models cannot explain and capture self-similarity and LRD features of the traffic [5].

Artificial neural network(ANN) are able to give better performance in dealing with the nonlinear relationships between the output and the input theoretically [6]. Neural networks have been successfully used for modeling complex nonlinear systems and predicting signals for a wide range of engineering applications. At present, whether a single model or combined model the back-propagation (BP) training algorithm, is probably the most frequently used type of neural network in practical applications [7]. Unfortunately, the BP training algorithm has some inherent defects [8], such as low learning speed, existence of local minima, and difficulty in choosing the proper size of network to suit a given problem.

Computational intelligence based algorithms have gained popularity in the training of neural networks because of their ability to find a global solution in a multidimensional search space[9].In this study, the quantum principle obtained from Quantum PSO has been combined with standard PSO to form a new hybrid algorithm called PSO with Quantum Infusion(PSO-QI)[10]. To overcome these shortcomings of BP training algorithm, the PSO-QI algorithm is applied to the wavelet neural network in the training phase, to obtain a set of weights that will minimize the error function in competitive time. Weights are progressively updated until the convergence criterion is satisfied. The objective function to be minimized by the PSO-QI algorithm is the predicted error function.

This paper is arranged as follows. In section 2, the particle swarm optimization algorithm is introduced. In section 3, the PSO-QI algorithm is presented in detail. In section 4, wavelet neural network trained by PSO-QI algorithm are proposed. Finally, experimental results and. some conclusions are given in section 5.

2. PARTICLE SWARM OPTIMIZATION ALGORITHM

The particle swarm optimization (PSO) is a new global optimization method based on Swarm Intelligence developed by Kennedy and Eberhart implements the metaphor of social behavior of the interaction between flock of birds and school of fish when searching for food in a given region [11]. In the particle swarm optimization (PSO) algorithm, the so-called particle is represented by a vector of size equal to the size of the search space, and is initialized with random velocity and position. Every particle has a velocity by which the direction and distance of the flying of the particle are determined, and a fitness that is decided by the optimized function. Particles move in the solution space; the moving direction and distance are determined by the speed vector[12]. Particles fly through the search space and adjust their velocities dynamically according to their historical behaviors. This phenomenon leads the particles to fly towards the better and better search area in the search space [13].

Assuming in a *D*-dimensional search space, the position of the ith particle is denoted as $X_i=(x_{i1},x_{i2},...,x_{iD})$, Each particle maintains a memory of its previous best position $P_i=(p_{i1},p_{i2},...,p_{iD})$. The best one among all the particles in the population is represented as P_{gd} , which is called pbest. Vector $P_g=(p_{g1},p_{g2},...,p_{gD})$. is the best position discovered by the whole population, which is called gbest. The velocity of ith particle is represented as $V_i=(v_{i1}, v_{i-2},..., v_{i-D})$. After the two extreme values are discovered, every particle updates its velocity and position based on

$$v_{id}(t+1) = wv_{id}(t) + c_1 r_1 [p_{id}(t) - x_{id}(t)] + c_2 r_2 [p_{gd}(t) - x_{id}(t)]$$
(1)

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1)$$
(2)

where c_1 and c_2 are two positive constants, known as the cognitive and social coefficients, which control the relative proportion of cognition and social interaction, respectively, and c_1 and c_2 are two random values in the range [0,1]. Parameter *w* is known as the inertia factor and establishes a commitment between diversity and search refinement. Its value can be fixed or fall linearly [14].

3. PSO-QI ALGORITHM

PSO-QI is a hybrid algorithm that uses the quantum principle from QPSO to create a new offspring in PSO[15]. In PSO-QI algorithm, the positions and velocities of the particles are updated using standard PSO equations. A randomly chosen particle from PSO's *pbest* population is utilized to carry out the quantum operation, and creates an offspring by mutating the *gbest*. The position and velocity of a particle in the quantum world cannot be determined simultaneously, the probability of finding a particle at a particular position in the quantum search space is mapped into its actual position in the solution space by a technique called "collapsing". By using Monte Carlo Simulation, the position equation in PSO-QI is given by the following equation [16]:

$$x(t) = J(t) \pm \frac{L(t)}{2} \ln\left(\frac{1}{u}\right)$$
(3)

where u is a uniform random number in the interval [0,1], The particle's local attractor point J has coordinates given by the following equation:

$$J_d(t) = \alpha_1 p_{gd}(t) + \alpha_2 p_{id}(t)$$
(4)

where p_{id} is the *i*th *pbest* particle in the *d*th dimension and p_{gd} is the *d*th dimension of the *gbest* particle obtained from PSO. *L* is the length of the potential field given by

$$L(t) = 2\beta \left| J(t) - x(t) \right| \tag{5}$$

where β is called the creativity coefficient and is responsible for the convergence speed of the particle.

The mean best position is represented as *mbest*, and is defined as:

$$mbest(t) = \frac{1}{s} \sum_{i=0}^{s} p_i(t) = \left(\frac{1}{s} \sum_{i=0}^{s} p_{i1}(t), \cdots, \frac{1}{s} \sum_{i=0}^{s} p_{iD}(t)\right)$$
(6)

where s is the size of the population, D is the number of dimensions and p_i is the *pbest* position of each particle.

Eq. (5) is replaced by *mbest* to form (7) as follow:

$$L(t) = 2\beta \left| mbest(t) - x(t) \right| \tag{7}$$

By using(3) this can also be written as follows to show the mutation on *gbest*, where the addition or subtraction is carried out with 50% probability:

$$x(t+1) = \alpha_1 p_{gd}(t) + \alpha_2 p_{id}(t) \pm \beta \left| mbest(t) - x(t) \right| \ln\left(\frac{1}{u}\right)$$
(8)

In PSO-QI ,the position update Eq.(8) has been used to mutate the best particle obtained from PSO.

4. PSOQI -WAVELET NEURAL NETWORK

4.1. Wavelet neural network

Wavelet neural network (WNN) is a new artificial neural network model based on wavelet transform, in which discrete wavelet function is used as the node activation function [17]. The topological structure of the WNN is shown in Figure 1. The wavelet neural network commonly consists of three layers: input layer, hidden layer, and output layer. All the nodes in each layer are connected to the nodes in the next layer [18].Let us assume that the structure of WNN model has r nodes in the input layer, m nodes in the output layer, n nodes in the hidden layer. The output y of the WNN can be expressed as expressed as follows:

$$y_k \approx \sum_{j=1}^n \omega_{jk} \psi \left(\frac{\sum_{i=1}^r u_{ij} x_i - b_j}{a_j} \right) , k = 1, 2, \cdots m,$$
(9)

on where, ω_{jk} is the connection weight between the *j*th node of hidden layer and the *k*th node of output layer. u_{ij} is the connection weight between the *i*th node of input layer and the *j*th node of hidden layer. b_j is the translation factor of the *j*th node in hidden layer. a_j is the expansion and contraction factor of the *j*th node in hidden layer. x_i is the *i*th node of input layer. In the hidden layer, the activation function of hidden nodes is Morlet wavelet function [19]. The Morlet wavelet function $\varphi(t)$ is often taken as:

$$\varphi(t) = \cos(1.75t)e^{-\frac{t^2}{2}}$$
(10)

Let us define error function e(t) as[20]

$$e(t) = y_e(t) - y(t) \tag{11}$$

where, $y_e(t)$ is the model actual output and y(t) is the desired output at time *t*. then the cost function *E* can be defined as

$$E = \frac{1}{2} \sum_{i=1}^{N} \left(y_e(t) - y(t) \right)^2 = \frac{1}{2} \sum_{i=1}^{N} \left(e(t) \right)^2$$
(12)

and can be minimized by all adjustable parameters using an iterative computational scheme parameters adjusted in the WNN model are as follows[21]:

$$\Delta \omega_{jk} \left(l+1 \right) = -\eta \frac{\partial E}{\partial \omega_{jk} \left(l \right)} + \alpha \Delta \omega_{jk} \left(l \right) \tag{13}$$

$$\Delta u_{ij}(l+1) = -\eta \frac{\partial E}{\partial u_{ij}(l)} + \alpha \Delta u_{ij}(l)$$
(14)

$$\Delta a_{j}(l+1) = -\eta \frac{\partial E}{\partial a_{j}(l)} + \alpha \Delta a_{j}(l)$$
(15)

$$\Delta b_{j}(l+1) = -\eta \frac{\partial E}{\partial b_{j}(l)} + \alpha \Delta b_{j}(l)$$
(16)

$$\omega_{jk}(l+1) = \omega_{jk}(l) + \Delta\omega_{jk}(l+1) \tag{17}$$

$$u_{ij}(l+1) = u_{ij}(l) + \Delta u_{ij}(l+1)$$
(18)

$$a_j(l+1) = a_j(l) + \Delta a_j(l+1) \tag{19}$$

$$b_j(l+1) = b_j(l) + \Delta b_j(l+1) \tag{20}$$

Where *l* represents the backward step number and η and α being the learning and the momentum constants, differing in the ranges 0.01 to 0.1 and 0.1 to 0.9, respectively. Fig. 1 shows a typical three layer wavelet neural network.

hidden layer



Figure 1. The schematic diagram of the WNN structure

4.2. WNN trained by PSO-QI

In WNN, the basic back-propagation (BP) algorithm adjusts the network parameters. The back propagation algorithm (BP) is a classical domain-dependent technique for supervised training [22].It works by measuring the output error, calculating the gradient of this error, and adjusting the WNN parameters in the descending gradient direction. In this algorithm, learning takes place during the propagation of input patterns from the input nodes to the output nodes. The outputs are compared with the desired target values and an error is produced. The back propagation algorithm (BP) through time suffer from local minima [23].Computational intelligence based algorithms have gained popularity in training of neural networks because of their ability to find a global solution in a multi-dimensional search space. The PSO-QI algorithm, on the contrary, is a global algorithm, which has a strong ability to find global optimistic results. Therefore, by combining the PSO-QI with the WNN, a new algorithm referred to as PSOQI–WNN algorithm is formulated in this paper.

K. Zhang

When a PSO-QI is used to train the WNN, a decision vector represents a particular group of network parameters including the connection weight, translation factor and contraction factors. It is further denoted as

$$X_{i} = \left(IH_{11}^{i}, \cdots, IH_{m}^{i}, HO_{11}^{i}, \cdots, HO_{nm}^{i}, A_{1}^{i}, \cdots, A_{n}^{i}, B_{1}^{i}, \cdots, B_{n}^{i}\right)$$
(21)

where IH_{kl}^{i} is the connection weight between input layer k^{th} node and hidden layer l^{th} node, HO_{kl}^{i} is the connection weight between hidden layer k^{th} node and output layer l^{th} node, A_{k}^{i} is the expansion and contraction factor of the *k*th node in hidden layer, B_{k}^{i} is the translation factor of the *k*th node in hidden layer. Since a component of the position corresponds to a network parameter, a WNN is structured according the particle's position vector. Training the corresponding network by inputting the training samples, we can obtain an error value computed by the following formula. *E* denotes the error function (cost function), in this paper, which can be used Eqs(12).In a word, the error function (12) is adopted as the objective function to be minimized in QPSO based wavelet neural network.

The specific procedure for the PSOQI–WNN algorithm can be summarized as follows:

step 1: Define the structure of the WNN according to the input and output sample.

step 2: Initialize the population by randomly generate the position vector X_i of each particle and set $pbest_i=X_i$.

step 3:Structure a WNN by treating the position vector of each particle as a group of network parameter.

step 4: Evaluate the fitness value of each particle by the equation (14), update the personal best position $pbest_i$ and obtain the global best position gbest across the population.

step 5: According to PSO-QI, update the inertia weights and factors. Judge the stopping criteria, if the maximal iterative times are met, stop the iteration, and the positions of particles are the optimal solution. Otherwise, the procedure is repeated from step 4.

5. NETWORK TRAFFIC PREDICTION

5.1. The background and pretreatment of the network traffic time series

The network traffic data used in the paper comes from monitor the traffic between clients in our campus network and servers, the network traffic data were given. In this paper, the minimal time interval in the network traffic data is 5 minutes. The traffic data contains one weeks' traffic records from May 10, 2013 to May 15,2013. Figure 2 shows network traffic time series.



Figure 2. Network traffic time series

The network traffic data are normalized the sample data groups as follows

$$\overline{x}_t = \frac{x_t - x_{\min}}{x_{\max} - x_{\min}}$$
(22)

$$x_t = \overline{x}_t \left(x_{\max} - x_{\min} \right) + x_{\min}$$
(23)

where x_t is defined as an initial group of data among the collected current network traffic data groups. x_{min} and x_{max} is respectively the minimum and maximum values of the network traffic time series.

5.2. Network traffic prediction model based on PSOQI–WNN

In this paper, we could predict the future network traffic level by constructing a prediction model which takes into account the past observations. To be more specific, assume that exists a smooth map :

$$y(k) = F(y(k-1), \dots, y(k-n))$$
 (24)

Then the functional relation $F(\bullet)$ is fit by adopting the QPSQI-wavelet neural network, and the predicting value can be obtained.

5.3. Experiment result

In order to test the performance of the algorithm, the network traffic data were choosed experiment data. The front 700 is the training data, and the latter 400 is the prediction data. The number of time series windows was set as 3, which meant that the forth of measurement data would be predicted from the past four of measurement data. A new three-layer wavelet neural network model was adopted, in which the number of input layer neurons is three, output layer neurons is one, and hidden layer neurons are six. The QPSQI parameters used in the study are: $c_1 = c_2 = 2$, and ω is linearly decreasing from 0.95 to 0.3. The population size of 50 particles, and *D*-dimensional

search space of particle is 36. β parameter of PSOQI is linearly increasing from 0.2 to 1.The results of PSOQI are compared with that obtained using PSO. After 100 times iterations, the cost function *E* of PSOQI–wavelet neural network was 0.1379. The connection weight between input layer and hidden layer is given by:

$$W_{1} = \begin{bmatrix} 0.5852 & -0.0901 & 1.0671 & 0.0014 & 0.9975 & 2.0061 \\ 1.9498 & 0.5766 & 1.2903 & 0.4234 & -0.5341 & 1.4959 \\ 1.6357 & -0.9155 & 1.6372 & 1.7166 & 1.4300 & 2.0339 \end{bmatrix}$$
(24)

The connection weight between hidden layer and output layer is given by:

$$W_{2} = \begin{bmatrix} -0.0392 \\ -0.3588 \\ 0.0511 \\ 0.4013 \\ 0.4654 \\ -0.009 \end{bmatrix}$$
(25)

The translation factors vector is given by:

 $B = \begin{bmatrix} 0.4730 & 1.4781 & -0.0865 & 2.1722 & 2.2688 & 2.5730 \end{bmatrix}$ (26) and the expansion and contraction factors vector is given by:

 $A = \begin{bmatrix} -0.0064 & -0.0365 & -0.1560 & 2.2169 & 2.8218 & 2.7560 \end{bmatrix}$ (27)

Figure 3(a) shows predicting results with PSOQI–wavelet neural network method. Figure 3(b) shows the error between the actual value and its predicted value with wavelet neural network method. From the prediction analysis figure, we can conclude that the prediction trend and value are all conformable. The model operated well with much higher prediction precision.



Figure 3. Prediction with PSOQI-wavelet neural network

We compare the forecast results by PSOQI-wavelet neural network and BP neural network, . Figure 4(a) shows predicting results with BP method after 1000 times iterations. Figure 4(b) shows the error between the actual value and its predicted value with BP method.





We compare the forecast results by PSO-BP algorithm and the other algorithm as shown in Table 1. It's obvious that PSOQI–wavelet neural network algorithm has more small deviation than the other algorithm. The PSOQI–wavelet neural network prediction results are very close to the target output.

K. Zhang

No	actual value	predict value				
		BP	RBF	Elman	PSOQI-	
					WNN	
701	0.7552	0.5890	0.7320	0.7430	0.7433	
702	0.7610	0.5858	0.7411	0.7421	0.7576	
703	0.7564	0.5822	0.7723	0.7814	0.7612	
704	0.7375	0.5752	0.7014	0.7283	0.7385	
705	0.7356	0.5737	0.7359	0.7167	0.7230	
706	0.7336	0.5722	0.7106	0.7256	0.7412	
707	0.7284	0.5682	0.7004	0.7144	0.7205	
708	0.7196	0.5613	0.7113	0.7109	0.7121	
709	0.7308	0.5701	0.7088	0.7208	0.7267	
710	0.7764	0.5978	0.7354	0.7559	0.7687	

Table 1. Simulation results with different artificial neural network

The error analysis was used to check the performance of the developed model. The accuracy of correlations relative to the measured values is determinated by various statistical means. The criteria exploited in this study were the Root Mean Square Error (RMSE), the Relative Error and Mean Absolute Percentage Error (MAPE) given by[24][25]:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2}$$
(28)

$$MAPE = \frac{100}{N} \sum_{i=1}^{N} |\hat{y}_i - y_i|$$
(29)

where, y stands for actual data, \hat{y} for prediction data by models, N for the number of data.. In this paper, the prediction methods, such as BP neural network, RBF neural network, and Elman neural network were applied, Tables 1 give the RMSE,MSE and MAPE of all models. Tables 2 give the relative errors of these neural networks prediction models.

PREDICTION	TIMES	MSE	RMSE	MAPE
MODEL	TRAINING			
BP	1000	0.035	0.221	0.621
RBF		0.025	0.456	0.845
Elman	1000	0.014	0.520	0.765
PSOQI–WNN	100	0.001	0.033	0.093

Table 2. Values of different statistical indicators for different algorithms

6. CONCLUSION

In this paper, a wavelet neural network based PSO-QI algorithm is applied to prediction traffic networks. The PSOOI-wavelet neural network algorithm takes full

advantages of the better performance of global optimized search of PSO-QI and the local optimized search of wavelet neural network. The result also suggests that the model presented above can provide highly accuracy in the field of prediction.

Acknowledgments- This research was partially supported by the National Natural Science Foundation of China (No. 11261001).

7. REFERENCES

1. S. Ostring, H. Sirisena and I. Hudson, Designing and Managing Networks for Self-Similarity, *Proceedings of EMACS* 2000, Sept. 2000.

2. K. Shah, S. Bohacek, E. Jonckheere. On the predictability of Data Network Traffic, *Proceedings of the 2003 American Control Conference*, Vol.2, pp.1619-1624, June. 2003.

3. A. Sang, San-qi Li, A predictability analysis of network traffic, *Computer networks* (9), 329-345, 2002.

4. A. Adas, Traffic models in broadband networks, *IEEE Communication Magazine* **35**, 82-89, 1997.

5. C. Bor-Sen, S.C. Peng, K.C. Wang. Traffic Modeling, Prediction, and Congestion Contol for High-Speed Networks: A Fuzzy AR Approach, *IEEE Trans. Fuzzy Systems* **8**(5), 491-508, 2000.

6. M. Hansegawa, G. Wu, M. Mi zuno, Applications of Nonlinear Prediction Methods to the Internet Traffic. *The 2001 IEEE International Sympsium on Circuits and Systems*, 169-172, 2001.

7. S. Guang, Network Traffic Prediction Based on the Wavelet Analysis and Hopfield Neural Network. *International Journal of Future Computer and Communication* **2**(2), 101-105, 2013.

8. G. Venayagamoorthy, special issue: Online design of an echo state network based wide area monitor for a multi-machine power system, *Neural Networks* **20**(**3**), 404-413, 2007.

9. X. Cai, D. Prokhorov, D. Wunsch, Training winner-take-all simultaneous recurrent neural networks, *IEEE Transactions on Neural Networks* **18**(**3**), 674-684, 2007.

10. J. Sun, B. Feng, W. Xu, Particle swarm optimization with particle shaving quantum behavior. *In Proceeding of congress on Evolutionary Computation* **1**, 325-331, 2004.

11. N. Singh, S.B. Singh, Personal Best Position Particle Swarm Optimization, *Journal of Applied Computer Science & Mathematics* **12** (6), 69-76, 2012.

12. J. Zhao, H. Sun, Multi-Stage Multi-Model Improved Particle Swarm Optimization Algorithm, *Computer Engineering and Applications* **46** (**33**), 32 – 35, 2010.

13. J. Kennedy, R.C. Eberhart Particle Swarm Optimization. *In Proceedings of IEEE International Conference on Neural Networks*, Perth, Australia, IEEE Service Center, Piscataway, NJ, Vol.IV, 1942-1948, 1995.

14. W. Fang, J. Sun, W. Xu, A New Mutated Quantum-Behaved Particle Swarm Optimizer for Digital IIR Filter Design, *EURASIP Journal on Advances in Signal Processing*, 2009.

15. B. Luitel, G. K. Venayagamoorthy, Quantum inspired PSO for the optimization of simultaneous recurrent neural networks as MIMO learning systems, *Neural Networks* **23**, 583-586, 2010.

16. J. Sun, B. Feng, W. Xu, Particles warm optimization with particles having quantum behavior. *In Proc.CEC 2004 evolutionary computation congress on* 1, 325-331, 2004.

17. Y.C. Pati, P.S. Krishnaparasad, Analysis and synthesis of feed forward neural networks using discrete affine wavelet transformations. *IEEE Trans Neural Networks* **4**(1), 73–85, 1993.

18. N. Chauhan, V. Ravi, D.K. Chandra. Differential evolution trained wavelet neural networks-application to bankruptcy prediction in banks. *Expert Systems with Applications* **36**, 7659-7665, 2009.

19. A. Subasi, M. Yilmaz, H. R. Ozcalik, Classification of EMG signals using wavelet neural network, *Journal of Neuroscience Methods* **156**, 360–367, 2006.

20. M. R. Ghasemi, A. Ghorbani, Application of Wavelet Neural Networks in Optimization of Skeletal Buildings under Frequency Constraints, *International Journal of Electrical and Computer Engineering* 2(8), 516-524, 2007.

21. Q. Zhao, Z. Sun, F. Sun, J. Zhu, Appearance-based Robot Visual Servo via a Wavelet Neural Network. *International Journal of Control, Automation, and Systems* **6(4)**, 607-612, 2008.

22. X. Cai, D. Prokhorov, D. Wunsch, Training Winner-Take-All Simultaneous Recurrent Neural Networks, *IEEE Transactions on Neural Networks* **18** (3), 674–684, 2007.

23. L. Panchi, L. Shiyong, Learning algorithm and application of quantum BP neural networks based on universal quantum gates. *Journal of Systems Engineering and Electronics*, 19(1), 167–174, 2008.

24. S. Chabaa, A. Zeroual, J. Antari, Identification and Prediction of Internet Traffic Using Artificial Neural Networks, *J. Intelligent Learning Systems & Applications* 2, 147-155, 2010.

25. M. B. Roman, Z. S. Ravilya, I. L. Ekaterina, Wavelet neural network (WNN) approach for calibration model building based on gasoline near infrared (NIR) spectra. *Chemometrics and Intelligent Laboratory Systems* **93**, 58–62, 2008.