

A Novel Network Flow Prediction Method based on Cuckoo Search Algorithm Optimizing BP Neural Network

Liqiang Fan

Abstract—Network traffic modeling and forecasting is the basis of network management and security warning. According to the characteristics of the nonlinear network flows, chaos, polygon, etc., in order to improve the prediction accuracy of network traffic, and puts forward the a cuckoo search cable calculation method and BP neural network by network traffic prediction model, BP neural network is used by the network of the learning sample book training, die quasi cloth Valley bird found nest eggs to find the optimal model parameters and the mining network flow number in simulation experiment according to measure the trial model of can. Simulation results show that compared with the reference model, CS-BPNN improves the prediction accuracy of network traffic, network traffic trends are described more accurately, provides a new research tool with network traffic prediction.

Keywords—BP neural network, Cuckoo search algorithm, Parameters optimizing, Network traffic prediction.

I. INTRODUCTION

WITH the increase in the size of the network and the growing complexity of network management, modeling and predicting the network traffic is essential for network management and early warning on network safety. Meanwhile, the network traffic is influenced by many factors, and exhibits non-stationary and chaotic properties. Accurate prediction of network traffic can not only optimize the allocation of network bandwidth but also effectively prevent network congestion. Hence, predicting the network traffic is a major issue in the network management field.

The traditional models for network traffic prediction include the linear regression analysis and the time series models. Klivansky et al. propose the multi-fractal wavelet model based on the self-similarity of the network traffic. The authors in [1] exploited the reliability model of the neural network to propose the software reliability prediction method. The authors in [2] used the multi-fractal prediction model to convert the long range dependent traffic series which is difficult to be predicted and analyzed into the series that can be predicted using the short range dependent linear model.

Liqiang Fan is with the Department of Mathematics, Langfang Normal University, Langfang065000, Hebei, China (corresponding author; e-mail: fxj510403@126.com).

Recently, network traffic prediction problem has attracted many attentions of the scholars at home and abroad, and an extensive and in-depth research, obtain good effect. Network traffic prediction method can generally be divided into traditional Markov method, time series analysis method and neural network method, etc. The principle of the autoregressive or autoregressive moving average forecasting model of the algorithm is simple, and the prediction accuracy for short-term forecasting is high, but the result of long-term forecast is bad. Literatures [15-17] proposed forecasting model based on autoregressive or autoregressive moving average (AR), paper [18] proposed autoregressive integrated moving average (ARMA). In paper [19], the researchers assumed that the network traffic appeared a smooth change of data, and proposed AR model to deal with the network traffic prediction. However, network traffic is the result of comprehensive shaping of many factors, randomness, time-varying non-stationary characteristics, therefore, the traditional model is difficult to do an accurately long-term projections for network flow.

Some scholars started to import the theory of nonlinear modeling and forecasting network traffic, such as support vector machine (SVM), grey theory and neural network prediction model of network traffic and so on. Papers [2-3] established the network traffic reliability prediction method based on the neural network model. In papers [4-7], neural network was used to study the nonparametric, nonlinear classification and prediction problems of network traffic. Paper [8] presents a time series forecasting algorithm, which adopts neural network puts forward a nonlinear mapping relationship based on the input set and expectation model.

Due to network traffic is an affected by multiple factors combination of complex system, a single prediction algorithm can only describe the part or section information, it is difficult to comprehensively and accurately predict its change rule. Some scholars have established the combination of network traffic prediction model based on the famous M-competition theory, the results showed that combination model can forecast larger limit use a variety of sample information, system for more than a single prediction model considering problems, more comprehensive [9-10]. Due to the network traffic has multi-scale features at the same time, some scholars put forward a kind of wavelet analysis and integration of ARIMA network traffic prediction model, has obtained the good

prediction effect [11]. However, this model adopts the single ARIMA model for high frequency and low frequency part modeling, and then it is difficult to accurately describe the network traffic complex changes, which means that, the forecast needs further improve [12-14].

However, many studies show that the variation of network traffic can be regarded as a non-linear non-varying, chaotic and dynamic system. The traditional methods always rely on linear modeling and is ineffective in accurately capturing dynamics of the network traffic. The advancements in the chaos theory and the artificial intelligence achieved in recent years yield the chaos theory-based non-linear prediction algorithms, such as the neural network method and the support vector machine (SVM) method. The neural network was used in [3] to study the non-parametric no-linear classification and prediction problem. The authors in [4] employed the neural network to propose a self-learning process for time series prediction and presented a non-linear mapping relationship based on the input set and the expectation pattern. These algorithms, however, have limitations. For example, the neural network algorithm based on minimization of empirical risk is prone to over learning and its generalization ability is poor.

This paper proposes a novel network traffic prediction model CS-BPNN that can effectively optimize BP neural network parameters and improve network traffic prediction accuracy. Simulation results demonstrate the ability of the proposed method to accurately predict network traffic.

II. OPTIMIZATION BP NEURAL NETWORK WITH CUCKOO SEARCH

A. BP Neural Network

BP neural network model can be formed by three layer network structure. The input signal from the input nodes through the hidden layer nodes, then the nodes in the hidden layer to the output layer, a layer of nodes on a layer of influence by input node output. Figure 1 is a diagram of a structure with three layers structure of typical BP neural network.

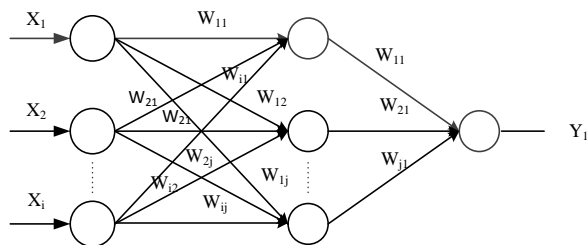


Fig. 1 The structure of BP neural network

Consider that the input to the network traffic dynamic system is shown in Equation (1).

$$X(i) = \{x(i - (m-1)\tau), \dots, x(i - \tau), x(i)\} \quad (1)$$

Let τ denote the time delay, m denote the embedding dimensionality, $X(i)$ denote the phase point after reconstruction. Then, the output function is $y(i)=x(i+1)$. The number of input nodes in the BP neural network is a function of the embedding dimensionality of the network traffic, m , and the number of nodes in the hidden layer, p . There is one output node. Mapping is constructed through $f: R^n \rightarrow R$.

The input to each node in the hidden layer is:

$$S_j = \sum_{i=1}^n \omega_{ij} x(i) - \theta_j \quad (2)$$

where ω_{ij} is the weight of link between the input layer and the hidden layer, θ_j is the threshold of the node in the hidden layer.

The output of the node in the hidden layer is:

$$b_j = \frac{1}{1 + \exp\left(\sum_{i=1}^n \omega_{ij} x_i - \theta_j\right)} \quad (3)$$

The input to the node in the output layer is:

$$L = \sum_{j=1}^p \omega_{jk} b_j - \theta_k \quad (4)$$

The output of the node in the output layer is:

$$x_{i+1} = \frac{1}{1 + \exp\left(\sum_{j=1}^p v_j b_j - \gamma\right)} \quad (5)$$

where v_j is the weight of link between the hidden layer and the output layer, γ is the threshold of the hidden layer.

B. Cuckoo Search algorithm

Cuckoo search (CS) is a novel heuristic algorithm motivated by the replication strategy used by the cuckoo for breeding their chicks. The usual Levi flight mechanism is introduced to simulate the search of the cuckoo for nest. This algorithm is based on the three ideal rules below.

- (1) Each cuckoo can produce only one egg each time and choose a random nest to store it.
- (2) The nest that hosts the best egg will be retained for the next generation.
- (3) The total number of nests, n , is fixed. The host of the nest

recognizes the foreign egg at a probability of pa . If the host recognizes foreign eggs, it will remove these eggs or directly abandon the nest and search for a new one at a different place.

The detection probability pa represents a selection strategy in CS. Based on the three rules above, the formulae for updating the search path and nest location of the cuckoo are as follows.

$$x_i^{t+1} = x_i^t + \alpha \oplus L(\lambda) \quad (6)$$

In Equation (6), x_i^{t+1} represents the nest location for the next generation, $x_i^t = (x_1^t, x_2^t, \dots, x_n^t)$ represents the location of the i th nest in the t th generation in the n -dimensional optimization problem, $\alpha > 0$ denotes the step size which is used to determine the range of random search and varies with the actual optimization problems. Equation (6) is a random walk equation in essence. Generally, a random walk is a Markov chain, where the future location depends on the current location (the first term in the equation) and the transition probability (the second term). \oplus denotes the point-to-point multiplication, $L(\gamma)$ denotes the Levy random search path, which, together with the time t , follow the Levy distribution with parameter γ . Details are given below.

$$levy \sim u = t^{-\lambda}, (1 < \lambda < 3) \quad (7)$$

In the Levy distribution-based search path, short-distance exploration is equivalent to occasional long-distance walk. Applying this search strategy to the intelligence optimization algorithm can increase the search scope, improve the population diversity, and jump out of local optimum more easily.

C. Steps to optimize BP neural network parameters using CS

CSBP algorithm, each egg represents the possibility of BP neural network optimal initial weights and thresholds, in every iteration, cuckoo search algorithm is used to optimize the initial weight threshold, optimal solution is assigned to the BP neural network, BP neural network through training function of the sample training, the cuckoo search algorithm will continue to optimize the initial weights and thresholds, until reaching the convergence condition of the number of iterations is reached or meet the (MSE) requirements. Figure 2 shows the flow chart of the CSBP model.

Step 1. Initialize the number of nests, n , P_a , and the maximum number of iterations, N_{max} .

Step 2. Randomly generate the initial locations of the n nests: $q_i^{(0)} = [x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}]^T$.

They correspond to the initial thresholds of the BP neural network and the weights of links. The BP neural network trains the training set based on parameter settings and compute the predicted result.

Step 3. Use the predicted result to find the optimal nest location $x_d^{(0)}$ of the current generation, update the nest location based on Equation (6) to obtain the new nest location.

Step 4. Compute the new nest locations, and substitute good nest locations for bad ones in the previous generation to obtain a group of better nest locations, $e_k = [x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}]^T$.

Step 5. Compare r with P_a , keep nests in e_k whose P_a is small, and update locations of nests whose P_a is larger to obtain a new group of nest locations. Substitute good nest locations for bad ones in e_k to obtain a group of better nest locations, $q_k = [x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}]^T$.

Step 6. Find the optimal nest location $x_d^{(k)}$ from q_k . If the maximum number of iterations is reached, then stop searching and output the optimal location $x_d^{(k)}$. Otherwise, return to Step (3) to perform further optimization.

Step 7. Define the parameters corresponding to the optimal nest location $x_d^{(k)}$ as the initial thresholds and weights of the BP neural network to train the training sets and construct the network traffic prediction model.

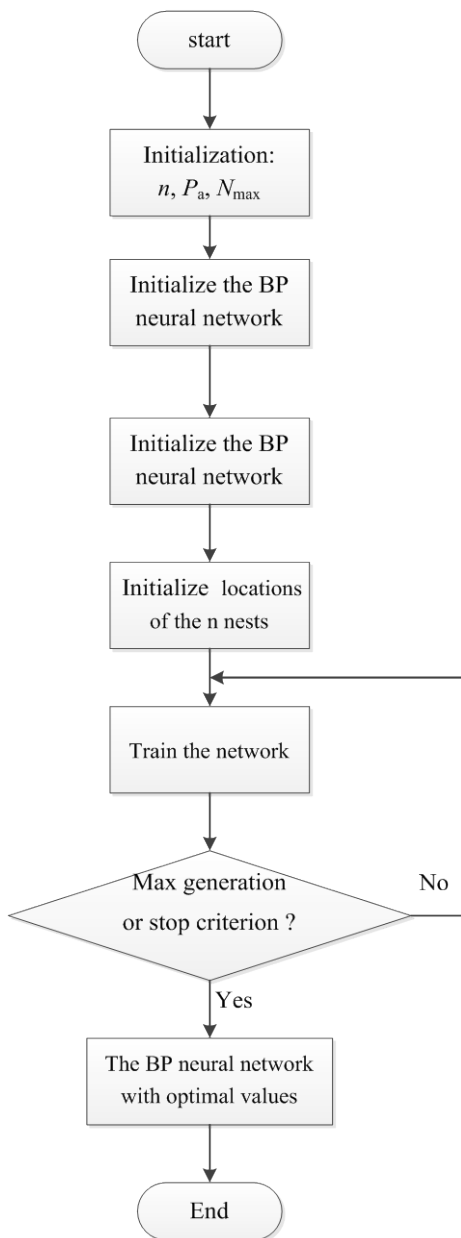


Fig.2 Flow chart of CSBP model

III. NETWORK TRAFFIC PREDICTION BY CSBP

A. Network traffic data

The network traffic data used in the experiment came from the traffic library. The traffic library collected the network traffic per hour within 50 days from July 1 to August, 19, 2014, yielding $50 \times 24 = 1200$ time series of the network traffic, as shown in Fig. 1. The time series of the network traffic within the first 45 days were used as the training set to construct the network traffic prediction model. The network traffics within the last 5 days were used as the validation set to evaluate the prediction performance of the model.

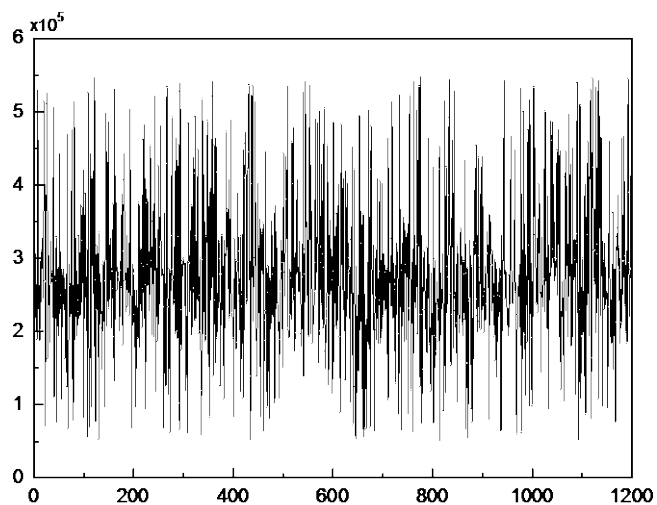


Fig.3 Network traffic time series

B. Construction of the training set

First, the training set of the network traffic time is processed using the mutual information method, yielding the time delay $\tau = 1$. Later, the Cao method is used to compute the embedding dimensionality, yielding the minimal embedding dimensionality for the reconstruction of the network traffic time series phase space, $m=12$. Finally, small data sets are employed to compute the maximum Lyapunov index as 0.01952, implying this network traffic time series is chaotic. Jointly considering Equation (1) and the parameter settings of $\tau = 1$ and $m=12$, we perform phase space reconstruction on the training set of the network traffic, obtaining 948 training samples.

BP is most sensitive to the data in $[0, 1]$. In order to improve BP training efficiency, the reconstructed training set is normalized as follows:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (8)$$

where x denotes the original data of traffic, x_{\min} and x_{\max} denote the minimum and maximum values, respectively.

C. Compared models and evaluation metrics

To further demonstrate the prediction performance of CSBP, it is compared with the combination of the BP, and the combination of the particle swarm optimization (PSO) method with PSO-BP. All methods adopt the one-step prediction strategy. That is, the first k network traffic data items are used as the original training samples to predict the $k+1$ th network traffic data item. Then, the $k+1$ th actual network traffic data item is incorporated into the original training samples for further prediction. Similar steps can be followed to obtain all

predicted values of validation samples. RMSE and MAPE are used as the evaluation metrics and they are defined below.

$$\begin{cases} RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^k (x(i) - \hat{x}(i))^2} \\ MAPE = \frac{1}{n} \sum_{i=1}^k \left| \frac{x(i) - \hat{x}(i)}{x(i)} \right| \times 100 \end{cases} \quad (9)$$

where x_i and \hat{x}_i denote the actual and predicted value of the network traffic, n is the number of test samples.

D. Result and analysis

The optimal parameters of BP neural net determined using PSO and CS are used to establish the network traffic prediction models PSO-BP and CSBP. Then, the test set is employed to make predictions. The prediction results of each model are shown in Fig. 4. Fig.4 (a) presents prediction accuracy result by BP neural net algorithm, Fig.4 (b) shows prediction accuracy result by PSO-BP neural net algorithm, Fig.4 (c) presents prediction accuracy result by CSBP neural net algorithm, and Fig.4 (d) presents comparison result of three algorithms. This figure shows that CSBP outperforms BP and PSO-BP in terms of prediction accuracy. CSBP embodies the global search ability of CS and the non-linear prediction ability of BPNN. Hence, CSBP can predict the network traffic variations more accurately, robustly and reliably.

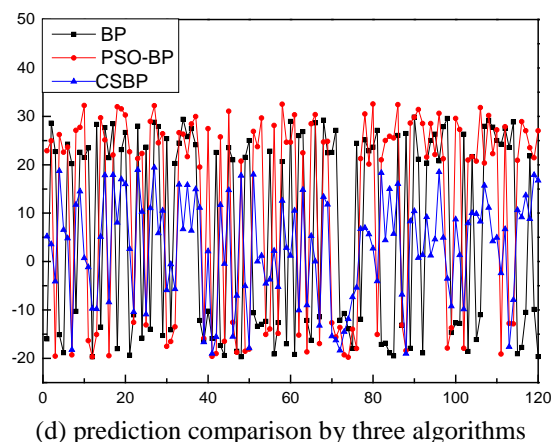
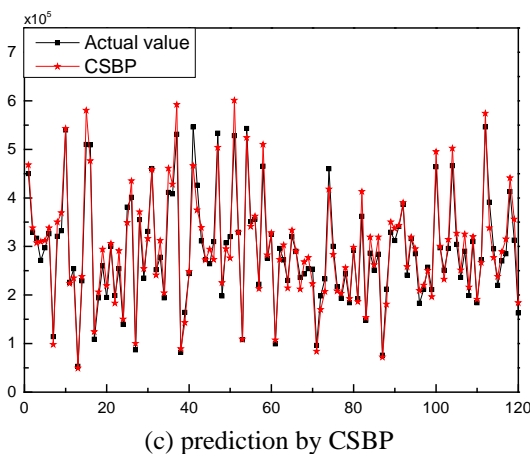
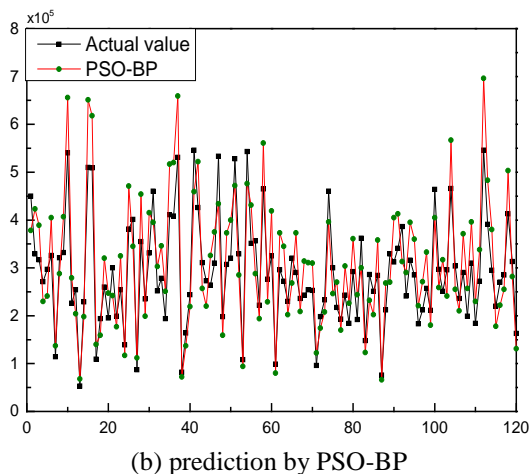
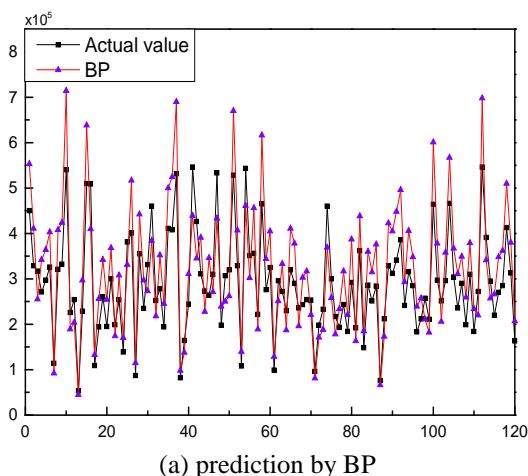


Fig.4 Comparison of the prediction effect of each model to the validation set

Furthermore, in Table 1 the model prediction result error analysis shows that, the error of proposed CSBP measuring results smaller than the other two models (BP and PSO-BP), the prediction accuracy of network traffic is improved. Results prove that CS-BP provides a more accurate network traffic prediction model.

Table 1. Error comparison of the prediction results of each model

Model	MAPE (%)	RMSE
BP	13.48	11.84
PSO-BP	8.58	7.92
CSBP	4.51	3.98

IV. CONCLUSION

Due to the large number of highly variable influencing factors, network traffic is non-linear and chaotic. The traditional method can hardly establish accurate prediction models. The performance of BP neural network is affected by parameter settings. To improve prediction performance, this paper proposes a novel network traffic prediction model CSBP and evaluates its performance via simulations. The prediction model does not depend on the accurate setting of the related parameters. Results show that CSBP solves the parameter optimization problem of the BP neural network, is able to describe complicated variations of network traffic more accurately, and improves network traffic prediction accuracy. Hence, the proposed model is very promising and provides insights into the establishment of prediction model. Further, how to apply the results of network traffic prediction to the actual traffic safety monitoring, network resource optimization and network protocol performance optimization is the main research work.

REFERENCES

- [1] Nguyen T T, Armitage G. "A survey of techniques for internet traffic classification using machine learning", *IEEE Communications Surveys and Tutorials*, vol. 10, no. 4, pp. 56-76, 2008.
- [2] Li S, Yin Q, Guo P. "A hierarchical mixture of software reliability model for prediction", *Applied Mathematics and Computation*, vol. 185, no. 2, pp. 1120-1130, 2007.
- [3] Hu Q, Xie M, NG S. "Robust recurrent neural network modeling for software fault detection and correct prediction", *Reliability Engineering and System Safety*, vol. 92, no. 3, pp. 332-340, 2007.
- [4] Ozkaya B, Demir A, Bilgili M S. "Neural network prediction model for the methane fraction in biogas from field-scale landfill bioreactors", *Environmental Modeling & Software*, vol. 22, no. 6, pp. 815-822, 2007.
- [5] Farsi H, Gobal F. "Artificial neural network simulator for supercapacitor performance prediction", *Computational Materials Science*, vol. 39, no. 3, pp. 678-683, 2007.
- [6] Byungwhan K, Jungki B. "Prediction of plasma processes using neural network and genetic algorithm", *Solid-State Electronics*, vol. 49, pp. 1576-1580, 2005.
- [7] Yang L, Dawson C W, Brown M R, et al. "Neural network and GA approaches for dwelling fire occurrence prediction", *Knowledge-Based Systems*, vol. 19, no. 4, pp. 213-219, 2006.
- [8] Liao G C. "Using immune genetic algorithm based hybrid techniques for short-term unit commitment problem", *Electrical Engineering*, vol. 87, pp. 267-279, 2005.

- [9] Bao Rong Chang, Hsiu Fen Tsai. "Improving network traffic analysis by foreseeing data-packet-flow with hybrid fuzzy based model prediction", *Expert Systems with Applications*, vol. 36, no. 3 pp. 6960-6965, 2009.
- [10] Este A, Gringoli F, Salgarelli L. "Support vector machines for TCP traffic classification" *Computer Networks*, vol. 53, no. 14, pp. 2476-2490, 2009.
- [11] Chen Y, Yang B. "Small-time scale network traffic prediction based on flexible neural tree", *Applied Soft Computing Journal*, vol. 12, no. 1, pp. 274-279, 2012.
- [12] Dong-Chul Park. "Prediction of network traffic using dynamic bilinear recurrent neural network", in *Proceedings of the 5th International Conference on Natural Computation*, Tianjin, 2009, pp. 419-423.
- [13] Shannon C E. "A Mathematical Theory of Communication", *The Bell System Technical Journal*, vol. 27, pp. 379-423, 623-656, 1948.
- [14] Jaynes, E. T. "Information theory and statistical mechanics", *Physical Review*, vol. 106, no. 4, pp. 620-630, 1957.
- [15] Jaynes, E. T. "Information theory and statistical mechanics. II", *Physical Review*, vol. 108, no. 2, pp. 171-190, 1957.
- [16] Jun Jiallg, Papavassiliou S. "Enhancing network traffic prediction and anomaly detection via statistical network traffic separation and combination strategies", *Computer Communications*, vol. 29, no. 10, pp. 1627-1638, 2006.
- [17] ZOU B X, LIU Q. "Network traffic prediction based on ARMA model", *Journal of Computer Research and Development*, vol.39, no. 12, pp. 1645-1652, 2002.
- [18] DENG Z L. "Optimal Estimation Theory with Applications—Modeling, Filtering and Information Fusion Estimation", HARBIN Institute of Technology Press, 2005.
- [19] Corradi M, Garroppo G, Giordano S, et al. "Analysis of F-ARIMA processes in the modeling of broadband traffic" in *Proceedings of the International Conference on Communications*, Helsinki, 2001, pp.964-968,.
- [20] LinShih-Wei, YingKuo-Ching, ChenShih-Chieh, LeeZne-Jung. "Particle swarm optimization for parameter determination and feature selection of support vector machines" *Expert Systems with Applications*, vol. 35, no. 4, pp. 1817-1824, 2008.

Liqiang Fan received M.A. in Computational Mathematics from Shanghai Normal University, Shanghai, China, in 2005. Now he works at Langfang Normal University and has been employed as an associate professor. His current research interests include mathematical modeling and algorithm analysis. He has published many papers in related journals.