

Forecasting of Busy Telephone Traffic Based on Wavelet Transform and ARIMA-LSSVM

Weishan He¹, Xizhong Qin¹, Zhenhong Jia¹, Chun Chang² and Chuanling Cao²

¹Xinjiang University

²China Mobile group Xinjiang Co., Ltd.
xuqiangren@hotmail.com

Abstract

In order to improve the prediction accuracy of busy telephone traffic which is influenced by multiple factors, this paper proposes a combined forecasting model which takes the influence of multiple factors into consideration and combines three models—wavelet transform, autoregressive integrated moving average (ARIMA) model and least squares support vector machines (LSSVM) model, LSSVM is optimized by particle swarm optimization (PSO). Correlation analysis is firstly applied to the busy telephone traffic data to obtain the key factors which influence the busy telephone traffic. Then wavelet transform is used to decompose and reconstruct the telephone traffic data to get the low-frequency and high-frequency components. The low-frequency component is loaded into ARIMA model to predict, while the high-frequency component and the obtained key factors are loaded into PSO-LSSVM model to predict. Finally the forecasting result is achieved by the superposition of predictive values. The simulation results show that the proposed model has higher prediction accuracy and strong generalization ability.

Keywords: busy telephone traffic, wavelet transform, autoregressive integrated moving average, least squares support vector machines, combined forecasting model

1. Introduction

With the rapid development of the economy and communication technology, the number of mobile users is increasing and telephone traffic is also growing. Accurate telephone traffic forecast can provide important assist for the network management, planning and designing, as well as decision support for network congestion, coverage and obstruction. It is increasingly a hot research topic to improve the accuracy of telephone traffic prediction nowadays.

Currently, the forecasting models mainly include autoregressive moving average (ARMA) model [1, 2], Markov model [3, 4], gray model [5, 6], neural network model [7, 8] and so on. As the telephone traffic is influenced by multiple factors, most of these models are based on historical traffic data without taking the factors that influence the telephone traffic into account. For nonlinear, non-stationary, self-similarity and multi-scale characteristics of telephone traffic, single forecasting model is easy to achieve, but it is difficult to predict the complicated and variable telephone traffic accurately. The combined forecasting model works through making the best use of the advantages of the single models, which can improve the prediction accuracy. It is a new trend in the study of telephone traffic prediction.

Most of the combined forecasting models are based on wavelet transform which can make multi-scale detailed analysis on signal, and make the signal be a simple single frequency component, so as to be much easier to predict. ARIMA model is a linear prediction model which can predict a smooth sequence more accurately [9, 10]. PSO-LSSVM model with powerful generalization ability can find the optimal solution

and predict the non-stationary sequences more accurately [11, 12]. According to the characteristics of telephone traffic data and the advantages of wavelet transform, ARIMA and PSO-LSSVM, the paper proposes a combined forecasting model considering the influence of multiple factors based on wavelet transform and ARIMA-LSSVM. From the simulation experiment results based on the actual telephone traffic data, it concludes that the combined model has higher prediction accuracy and strong generalization ability.

2. Fundamental Theories

2.1. Gray Correlation Analysis Method

Grey correlation analysis is based on the geometry of the sequence curve to determine whether they contact closely according to their similarity degree, so the closer the geometry curve [13]. The greater the correlation, the steps are as followed:

Step 1: If the system characteristic sequence for the X_1 and multi-factor sequences for X_k

$$X_1 = (x_1(1), x_1(2), \dots, x_1(n)) \quad (1)$$

$$X_k = (x_k(1), x_k(2), \dots, x_k(n)) \quad (2)$$

$$k = (2, 3, \dots, n)$$

Step 2: To deal with each sequence

$$Y_k = X_k / x_k(1) = (y_k(1), y_k(2), \dots, y_k(n)) \quad k = (1, 2, \dots, n) \quad (3)$$

Step 3: Correlation coefficient $\xi_k(j)$ and correlation degree γ_k

$$\xi_k(j) = \frac{\min_k \min_j |y_1(j) - y_k(j)| + \varepsilon \max_k \max_j |y_1(j) - y_k(j)|}{|y_1(j) - y_k(j)| + \varepsilon \max_k \max_j |y_1(j) - y_k(j)|} \quad (4)$$

$$\gamma_k = \frac{1}{n} \sum_{j=1}^n \xi_k(j) \quad (5)$$

2.2. Wavelet Decomposition and Reconstruction

Decomposition and reconstruction of the raw data using the Mallat algorithm, then we can get the formulas as follows [14]:

$$\begin{cases} a_{j+1,k} = \sqrt{2} \sum_{p \in \mathbb{Z}} \bar{h}_{p-2k} a_{j,p} \\ d_{j+1,k} = \sqrt{2} \sum_{p \in \mathbb{Z}} \bar{g}_{p-2k} a_{j,p} \end{cases} \quad (6)$$

$$a_{j,k} = \sqrt{2} \sum_{p \in \mathbb{Z}} h_{k-2p} a_{j+1,p} + \sqrt{2} \sum_{p \in \mathbb{Z}} g_{k-2p} d_{j+1,p} \quad (7)$$

Where $H = \{h_j\}$ denotes a low-pass filter, $G = \{g_j\}$ denotes a high-pass filter. $A_j = \{a_{j,1}, a_{j,2}, \dots, a_{j,k}\}$ denotes low frequency component of layer j . $D_j = \{d_{j,1}, d_{j,2}, \dots, d_{j,k}\}$ denotes high

frequency component of layer j .

Mallat algorithm can make the time sequence into low frequency components and high frequency components [15], [16]. The low frequency components can reflect the trend of a sequence substantially, and the high frequency components can reflect the differences of the sequence in the details. Therefore, we can decompose the low frequency components further to obtain new low frequency components and high frequency components. At different time periods, we can get what we so called “single reconstruction” through reconstructing low frequency components and high frequency components.

2.3. ARIMA Model

ARIMA (p, d, q) model can get a new stationary data series by making non-stationary original data differential treatment for d times, the order determination methods of ARIMA (p, d, q) are generally decided by the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). Combining with the methods of the minimum information criterion (AIC) and consistency criterion (SBC), we can get the most appropriate model. The functions of AIC and SBC are defined as followed:

$$AIC = n \ln(\sigma_\varepsilon^2) + 2(p + q + 1) \quad (8)$$

$$SBC = n \ln(\sigma_\varepsilon^2) + 2(p + q + 1) \quad (9)$$

2.4. Least Squares Support Vector Machine and Particle Swarm Optimization

Suppose the training sample set as (x_i, y_i) , $i = 1, 2, \dots, l$, $x_i \in \mathbb{R}^n$ is the input data, $y_i \in \mathbb{R}$ is the output data and the support vector machine model in the feature space as following [17], [18]:

$$f(x) = w^T \cdot \varphi(x) + b \quad (10)$$

Where w is the weight vector, $\varphi(x)$ is the mapping of the input data x in the feature space, b is the bias parameter.

So we can transform the non-linear estimation function into linear estimation function in high dimensional feature space if it meets the conditions of the structural risk minimization principle:

$$\min J(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^l \xi_i^2 \quad (11)$$

The solved target equations can be converted to the equation:

$$y_i = \varphi(x_i) \omega + b + \xi_i \quad (12)$$

Where, $\|\omega\|^2$ is used to control the complexity of the model, c is the regularization parameter, and ξ_i is the relaxation factor.

The optimization problems of the least squares support vector machine is transformed into Lagrangian function according to the KKT optimality conditions and Mercer conditions. Finally we get a nonlinear regression model:

$$f(x) = \sum_{i=0}^l a_i K(x, x_i) + b \quad (13)$$

Where a_i ($i = 1, 2, \dots, l$) is the Lagrange multiplier, $K(x, x_i)$ is kernel function.

It makes a great influence on the accuracy that the selection of the penalty parameter and kernel function of least squares support vector machines, and it is difficult to determine the appropriate values in advance. The particle swarm optimization is a population-based optimization algorithm [19], [20], making the values of the penalty

parameter and kernel function initialize to a group of random particles, finding the optimal solution through iteration. In this paper, we apply the particle swarm optimization algorithm to obtain the optimal parameter of the least squares support vector machines, and the simulation results show that the parameter selection of the least squares support vector machine based on the particle swarm optimization is reliable and accurate.

3. Forecasting Model and its Process

The forecasting model is a combined model of wavelet analysis and ARIMA-LSSVM. It is shown in Figure 1.

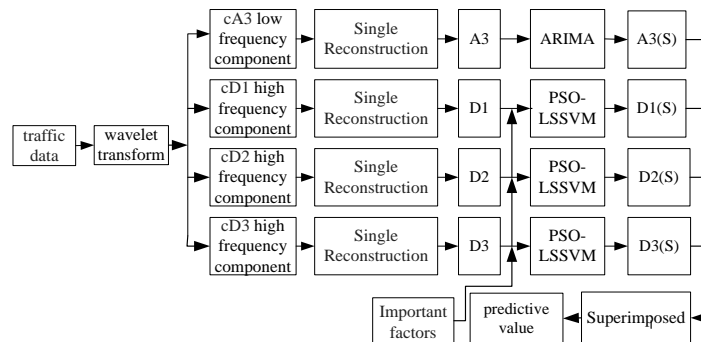


Figure 1. Model Figure

The concrete steps of the telephone traffic forecasting are as following:

- 1) We carry out the correlation analysis on the busy telephone traffic, getting the key factors affecting the telephone traffic;
- 2) In this paper, we apply bi-orthogonal wavelet to decompose the busy telephone traffic data into 3 layers, getting the low-frequency cA3 and high-frequency cD1, cD2, cD3;
- 3) We apply the Mallat algorithm to achieve the single reconstruction of low-frequency cA3 and high-frequency cD1, cD2, cD3 respectively, getting A3 and D1, D2, D3;
- 4) The low-frequency component A3 is used as input variables into the ARIMA model to predict, getting the predictive values A3(S). The high-frequency components D1, D2, D3, respectively, and important influence factors are used as input variables into the PSO-LSSVM model to predict, getting the predictive values D1(S), D2(S), D3(S);
- 5) The final results would be $Pre = A3(S) + D1(S) + D2(S) + D3(S)$.

4. Experimental Results and Analysis

4.1. Forecasting Results and Analysis in a Region

The experimental data is provided by a telecom, the data of telephone traffic and key factors are collected from April 7, 2011 to May 3, 2011, April 7, 2012 to May 3, 2012 and April 7, 2013 to May 3, 2013 in some region, and we select the maximum of daily telephone traffic data as busy data, the 81 busy telephone traffic data is shown in Figure 2.

The trend of telephone traffic data is not only affected by historical telephone traffic data, but also many other factors. In order to predict the busy telephone traffic more accurately, Firstly, we make correlation analysis on busy telephone traffic with gray correlation analysis method, choosing 4 affecting factors of bigger correlations from the 17 factors of busy telephone traffic, and the statistical results are shown in tab.1. By analyzing the results, the total number of users busy VLR, the number of the system

connectivity, the total number of switch test call and the total number of system paging were closely related to the telephone traffic. The low-frequency component is input into the ARIMA model, and the high-frequency component and the key factors of busy telephone traffic are input into the PSO-LSSVM model, and we use the previous 80 data as training set to predict the busy telephone traffic of May 3.

In order to prove the accuracy and validity of this method proposed in this paper, we compare the forecasting results with those of PSO-LSSVM model which considers the influence of multiple factors (Model 1), the Elman neural network without considering the influence of multiple factors (Model 2). Prediction error indicator applies the relative error, namely: $\xi_i = |(x'_i - x_i)/x_i|$, x'_i is predictive value, x_i is the actual value.

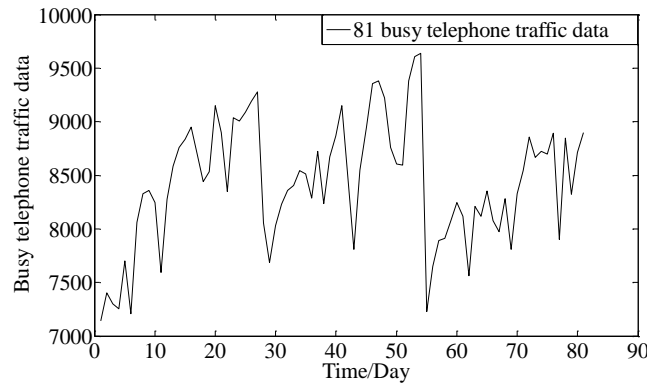


Figure 2. Busy Telephone Traffic Data

Table 1. Correlation Analysis

	Total number of users busy VLR	The number of the system connectivity	Total number of switch test call	Total number of system paging
The correlation coefficient	0.6903	0.7263	0.7107	0.7007

Figure 3 to Figure 6 show the forecasting results of each frequency component with the traffic forecasting model considering the influences of multiple factors based on wavelet transform and ARIMA-LSSVM; Figure 7 is the comparison chart of the actual value and the predictive value using the model in this paper; Figure 8 shows the forecasting results of telephone traffic based on the three models. The comparison of error is shown in Table 2.

As shown in Figure 3 to Figure 6, we get that the signals obtained from wavelet transform are more single in the frequency component, combining the advantages of the ARIMA in forecasting stationary series and that of the PSO-LSSVM in forecasting non-stationary series, so the prediction accuracy of the model for each subsequence is higher, and the final results superimposed by each component have high accuracy. As shown in Table 2, the prediction error of the combined forecasting model under the influence of multiple factors based on wavelet transform and ARIMA-LSSVM is small, mainly due to the errors of forecasting result from the single component based on the decomposition and reconstruction being positive and negative, and the components are superimposed, the prediction error can be canceled each other, taking into account the effect of the key factors to the telephone traffic, so the prediction error is reduced.

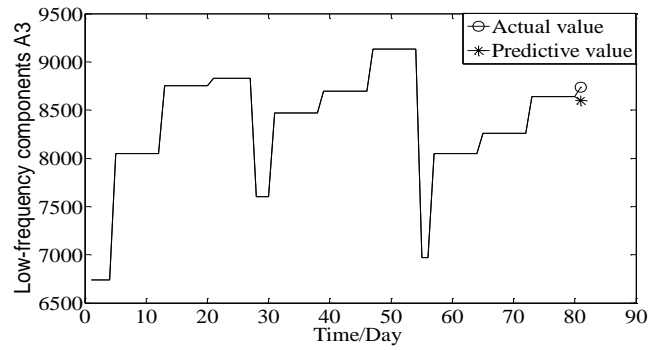


Figure 3. Actual Value and Predictive Value of Low-frequency Component A3

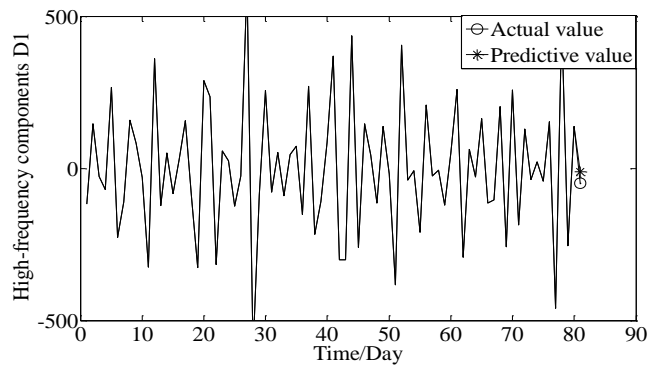


Figure 4. Actual Value and Predictive Value of High-frequency Component D1

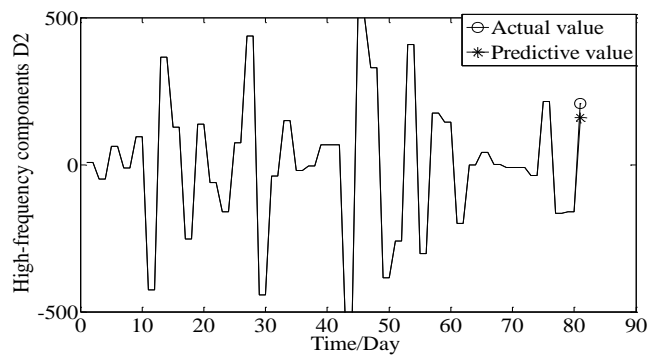


Figure 5. Actual Value and Predictive Value of High-frequency Component D2

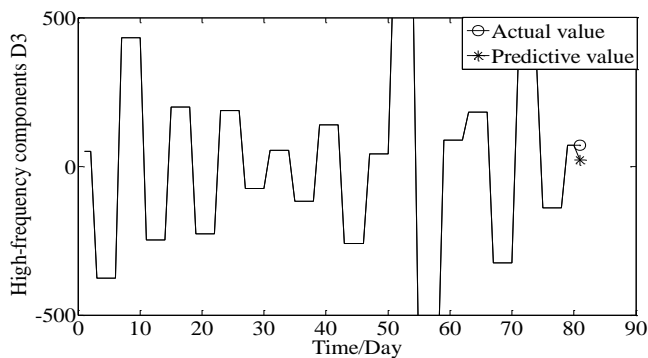


Figure 6. Actual Value and Predictive Value of High-frequency Component D3

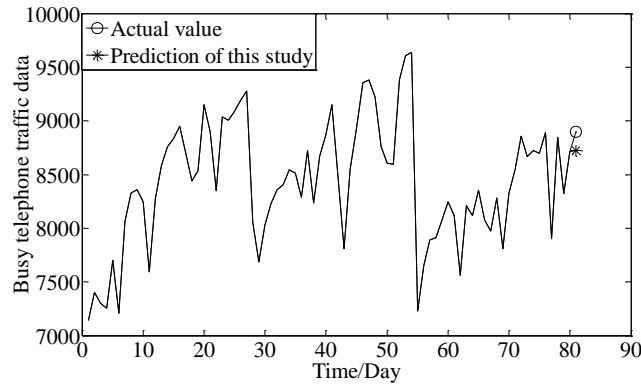


Figure 7. Actual Value and Predictive Value of this Study

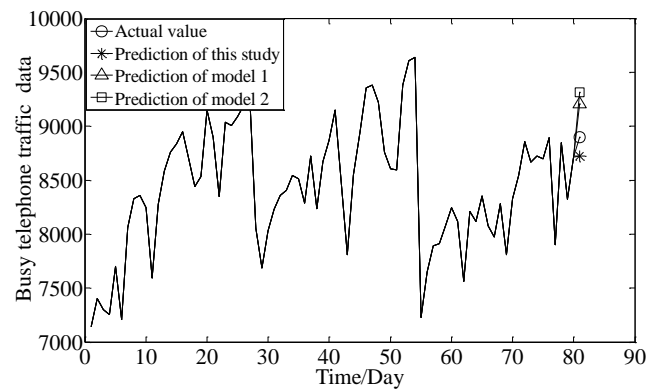


Figure 8. Predictions Comparison of 3 Models

Table 2. The Comparison of Error in 3 Forecasting Methods

Algorithm	error (%)
Proposed algorithm	1.78
Model 1	3.93
Model 2	6.54

4.2. Forecasting Results and Analysis in the other Four Regions

In order to test the generalization ability of this model, busy telephone traffic is predicted in the other four regions by the proposed method in this paper. We compare the forecasting results with those of Model 1 and Model 2. The forecasting results are shown in Table 3.

Table 3. The comparison of Results in 4 Regions

Algorithm	A error (%)	B error (%)	C error (%)	D error (%)
Proposed algorithm	1.14	2.37	3.65	1.68
Model 1	3.81	4.72	5.34	4.13
Model 2	5.92	6.76	7.81	6.09

As shown in Table 3, the proposed approach performs well on predicting busy telephone traffic in the four regions. Experimental results show that the model has strong generalization ability.

5. Conclusion

We propose a combined traffic forecasting model considering the influence of multiple factors based on wavelet transform and ARIMA-LSSVM. The single frequency components of signal are got through the wavelet decomposition and reconstruction, the better stability of each component, and easier to predict. Given the characteristics of each component, and the advantages of the ARIMA forecasting in stationary series and the PSO-LSSVM forecasting in non-stationary series, we put the low-frequency component into the ARIMA model to predict, while the high-frequency component as well as the key factors into the PSO-LSSVM model to predict. The final results superimposed by each component show the effect of the combined forecasting model is much better than those of PSO-LSSVM model which considers the influence of multiple factors, the Elman neural network without considering the influence of multiple factors. In order to test the generalization ability of the model, busy telephone traffic is predicted in the other four regions and forecasting results are highly accurate. Experimental results show that the model has higher prediction accuracy and strong generalization ability.

Acknowledgments

Thanks to my teacher for many valuable suggestions. The experimental data is provided by China Mobile group Xinjiang Co., Ltd (XJM2013-2788).

References

- [1] O. Valenzuela Rojas, F. Rojas, A. Guillen, L. J. Herrera, H. Pomares, L. Marquez and M. Pasadas, "Soft-computing techniques and ARMA model for time series prediction", *Neurocomputing*. 4, 71 (2008).
- [2] Gao Bo, Zhang Qinyu, Liang Yongsheng, Liu Ningning, Huang Chengbo, Zhang Naitong. Predicting self-similar networking traffic based on EMD and ARMA. *Journal of China Institute of Communications*. 4, 32 (2011).
- [3] Mamoun A. Awad and Issa Khalil. Prediction of user's web-browsing behavior: Application of markov model. *Systems, Man, and Cybernetics, Part B: Cybernetics*. 4, 42 (2012).
- [4] Liat Antwarg, Lior Rokach, and Bracha Shapira. Attribute-Driven Hidden Markov Model Trees for Intention Prediction. *Systems, Man, and Cybernetics, Part C: Applications and Reviews*. 6, 42 (2012).
- [5] Zhi He, Qiang Wang, Yi Shen, Yan Wang. Discrete multivariate gray model based boundary extension for bi-dimensional empirical mode decomposition. *Signal Processing*. 1, 93 (2013).
- [6] He Z, Shen Y, Wang Q. Boundary extension for Hilbert–Huang transform inspired by gray prediction model. *Signal Processing*. 3, 92 (2012).
- [7] Wei Wan , Hua Xu , Wenhao Zhang , Xincheng Hu , Gang Deng. Questionnaires-based skin attribute prediction using Elman neural network . *Neurocomputing*. 17, 74 (2011).
- [8] Weizhong Yan, Senior Member, Toward automatic time-series forecasting using neural networks. *Neural Networks and Learning Systems*. 7, 23 (2012).
- [9] Valenzuela O, Rojas I, Rojas F, et al. Hybridization of intelligent techniques and ARIMA models for time series prediction. *Fuzzy Sets and Systems* . 7, 159 (2008).
- [10] Khashei M, Bijari M, Raissi Ardali G A. Hybridization of autoregressive integrated moving average (ARIMA) with probabilistic neural networks (PNNs). *Computers & Industrial Engineering*. 1, 63 (2012).
- [11] Ci Wang, Limin Jia, Xiaofeng Li, A Study of Fault Diagnosis Method for the Train Axle Box Based on EMD and PSO-LSSVM. 2013 Third International Conference on Instrumentation, Measurement, Computer, Communication and Control. (2013) August 28-30; Nusa Dua, Bali.
- [12] Qianwen Xiang, Yukun Sun, Xiaofu Ji, Modeling inductance for bearingless switched reluctance motor based on PSO-LSSVM. 2011 Chinese Control and Decision Conference. (2011) May 23 -25; Mianyang,China.
- [13] sai, Men-Shen, and Fu-Yuan Hsu. Application of grey correlation analysis in evolutionary programming for distribution system feeder reconfiguration. *Power Systems*. 2, 25 (2010).
- [14] Bai XiangYu, YE Xinming, Jiang Hai. Network Traffic Predicting Based on Wavelet Transform and Autoregressive Model. *Computer Science*.7,34 (2007).
- [15] Zheng Sun, Li Huashi, and Ying Huizhou. Relaxing CFL limit of FDTD by DWT. *Electronics Letters*.7,50 (2014).
- [16] Lv Xuefang, Ping Tao. Mallat algorithm of wavelet for time-varying system parametric identification. 2013 25th Chinese Control and Decision Conference.(2013) May 25- 27; Guiyang, China.

- [17] Cui Qing, Ma Xiaoyi, Li Xianbo, Zhu Hui, Li Zhongjuan. Application of LSSVM in chaotic time series based on ADE algorithm. *Computer Applications and Software*.1, 31 (2014).
- [18] Ismail S, Shabri A, Samsudin R. A hybrid model of self-organizing maps (SOM) and least square support vector machine (LSSVM) for time-series forecasting. *Expert Systems with Applications*.8, 38 (2011).
- [19] LI Wenyuan, Yan Haihua, YAO Hongjie. Application of LSSVM Based on PSO to Fault Prediction of Communication Equipmen. *Microelectronics&Computer*.2, 30 (2013).
- [20] Fernandez Martinez, Juan Luis, and Esperanza Garcia Gonzalo. Stochastic stability analysis of the linear continuous and discrete PSO models. *Evolutionary Computation*. 3, 15 (2011).

Authors

Weishan He, she received her BE in Communication Engineering from Xinjiang University in 2011. She is now a graduate student in Xinjiang University. Her research interests include machine learning and mobile Internet. Email: xuqiangren@hotmail.com

Xizhong Qin, he received his ME from Southeast University in 1990. Currently, he is with the Xinjiang University where he is an associate Professor and a Master tutor of information science. His research interests include signal processing, wireless communications.

Zhenhong Jia, he received his BE from Beijing Normal University in 1986. He received his ME and PhD from Shanghai Jiaotong University in 1989 and 1993. Currently, he is with the Xinjiang University where he is a Professor and a PhD tutor of information science. His research interests include optical communications, wireless communications.

Chun Chang, he is an employee of China Mobile group Xinjiang Co., Ltd. His specialty is network planning.

Chuanling Cao, she is an employee of China Mobile group Xinjiang Co., Ltd. Her major is service planning.

