International Journal of Hybrid Information Technology
Vol.8, No.1 (2015), pp.1-10
http://dx.doi.org/10.14257/ijhit.2015.8.1.01

Telephone Traffic Forecasting Based on Grey Neural Network Optimized by Improved Particle Swarm Optimization Algorithm

Xiuting Yu[1], Xizhong Qin[2], Zhenhong Jia[3], Chuanling Cao[4] and Chun Chang[5]

[1,2,3] *Xinjiang University*
[4,5] *China Mobile group Xinjiang Co., Ltd.*
liulangyu_001@163.com

Abstract

To solve the problem that the parameters in grey neural network (GNN) are difficult to determine, the improved Particle Swarm Optimization (IPSO) algorithm is employed to search the optimums by the introduction of a threshold of velocity. When the particle velocity is less than the threshold, an accelerated momentum is applied on the particle to reinitialize the particle velocity and position. The proposed approach is used to predict the telephone traffic of two regions. The forecasting results are compared with those of GNN, Grey Neural Network optimized by Particle Swarm Optimization (PSO-GNN) and Back-Propagation Neural Network (BPNN). The experimental results show high prediction accuracy.

Keywords: *Particle Swarm Optimization, Grey Neural Network, speed threshold, Spearman correlation matrix, telephone traffic forecast*

1. Introduction

Grey system theory [1] can provide a new way to solve some system problems in the case of poor information. The theory studies the grey amount with the method of data generating which can organize raw data of weak distribution regularity into columns, but not with statistical laws from the perspective of a large sample of data. We call the model built in the grey system as grey model (GM) [2-6], which can predict the development trend of poor information, small samples and uncertain system eigenvalues. Neural network has the advantages of processing parallelism, nonlinear mapping, and strong fault-tolerant and adaptive ability. Therefore, we can combine the advantages of grey model and neural network model to build a grey neural network [7-12] model.

The Grey Neural Network (GNN) model can make up the inadequate of solving problems using grey model or neural network technology, but it makes the network easy to fall into local optimums which may lead to weak approximation and low prediction accuracy because of the randomness of its parameters. In this paper, the initial parameters of the grey neural network are optimized by the Improved Particle Optimization swarm algorithm (IPSO-GNN) to improve the prediction accuracy, so the predicted values are closer to the real values.

2. Grey Neural Network Model

According to the fusion methods of grey system and neural network technology, the grey neural network model can be divided into Series Grey Neural Network (SGNN), Parallel Grey Neural Network (PGNN), Inlaid Grey Neural Network (IGNN) and Blending Grey Neural Network (BGNN). In this paper, we use the BGNN model to improve the prediction accuracy which can increase the network processing capabilities.

Let set the original data sequence as $x(t)$ (t=1,2,3,…,n), so the accumulated generating data sequence can be defined as $y(t)$ (t=1,2,3,…,n) and mathematically expressed as:

ISSN: 1738-9968 IJHIT
Copyright ⓒ 2015 SERSC

$$y_q(p) = \sum_{i=1}^{p} x(i) \qquad q = 1,2,3\ldots,n \tag{1}$$

The differential equation of the grey neural network model built with n parameters is as follow:

$$\frac{dy_1}{dt} + ay_1 = b_1 y_2 + b_2 y_3 + \ldots b_{n-1} y_n \tag{2}$$

Where $y_2, y_3,\ldots,y_n$ are input data, y_1 is output data, $a, b_1, b_2,\ldots,b_{n-1}$ are the equation parameters. The time response is expressed as:

$$Z(t) = (y_1(0) - (\frac{b_1}{a} y_2(t) + \frac{b_2}{a} y_3(t) + \ldots + \frac{b_{n-1}}{a} y_n(t)))\, e^{-at} + \frac{b_1}{a} y_2(t) + \frac{b_2}{a} y_3(t) + \ldots + \frac{b_{n-1}}{a} y_n(t) \tag{3}$$

Where $z(t)$ is the predicted value. If defining $(b_1/a)y_2(t)+ (b_2/a)\, y_3(t)+\ldots+ (b_{n-1}/a)y_n(t)$ as d, then the formula (3) can be converted as follow:

$$Z(t) = (y_1(0) - d)e^{-at} + d = ((y_1(0) - d) - y_1(0)\frac{1}{1+e^{-at}} + 2d\frac{1}{1+e^{-at}})\left(1+e^{-at}\right) \tag{4}$$

So the GNN model can be obtained when the formula (4) is mapped to an extended BPNN model, and its topology is showed in Figure 1.

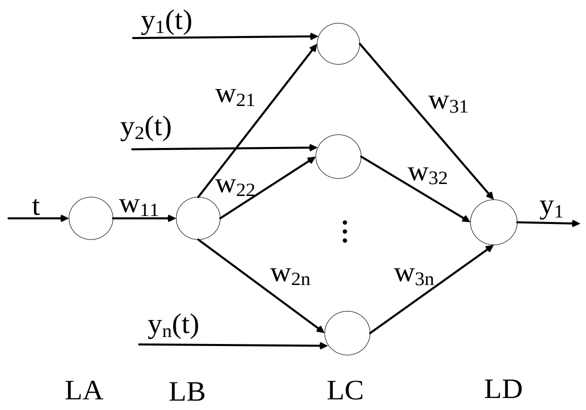

Figure 1. The Topology of GNN

As is shown in Figure 1, the GNN is divided into four layers that are recorded as LA, LB, LC and LD. Where t is the number of input parameters, $y_2(t), y_3(t), y_4(t),\ldots,y_n(t)$ are the input variables, y_1 is the output variable, w_{11} is the value of connection weight from LA layer to LB layer. $w_{21},w_{22},\ldots,w_{2n}$ are the values of connection weight from LB layer to LC layer, $w_{31}, w_{32},\ldots,w_{3n}$ are the values of connection weight from LC layer to LD layer. Setting the initial parameters of the network are that $w_{11}=a$, $w_{21}=-y_1(0)$, $w_{2i}=2b_{i-1}/a$ $(i=1,2,\ldots,n)$, $w_{3j}=1+e^{-at}$ The learning steps of GNN are as follows：

Step one: According to the characteristics of the training data, build neural network and initialize values of parameters: a, b_i.

Step two: Calculate the parameters of $w_{11}, w_{21}, w_{22} \ldots w_{2n}\, w_{31}, w_{32}, \ldots, w_{3n}$.

Step three: For each input sequence $y(t)$ (t =1,2,3,..., n), the output values of the neurons in each layer can be calculated by the following formulas.

$$l_1 = w_{11} t \tag{5}$$

 Copyright ⓒ 2015 SERSC

$$l_2 = f(w_{11}t) = \frac{1}{1 + e^{-w_{11}t}} \tag{6}$$

$$l_{31} = l_2 w_{21}, l_{32} = y_2(t) l_2 w_{22}, l_{33} = y_3(t) l_2 w_{23}, \ldots, l_{3n} = y_n(t) l_2 w_{2n} \tag{7}$$

$$l_4 = l_{31} w_{31} + l_{32} w_{32} + , \ldots, l_{3n} w_{3n} - \theta_{y_1} \tag{8}$$

Where $\theta=(1+e^{-at})(d-y_1(0))$ is the threshold value of output layer, $f(x)=1/(1+e^{-x})$ is the transfer function of neurons in LB layer, and $f(x)= x$ is the transfer function of neurons in other layers.

Step four: During the training processing, the weights and threshold of GNN are adjusted by the error between the expected and the predicted values. And the iterative formulas are as follows:

$$\sigma = l_4 - y_1(t) \tag{9}$$

Where σ is the error of LD layer.

$$\sigma_k = \sigma(1 + e^{-w_{11}t}), (k = 1,2,\ldots, n) \tag{10}$$

Where σ_k is the error of LC layer.

$$\sigma_{n+1} = \frac{1}{1 + e^{-w_{11}t}}\left(1 - \frac{1}{1 + e^{-w_{11}t}}\right)(w_{21}\sigma_1 + w_{22}\sigma_2 + \ldots + w_{2n}\sigma_n) \tag{11}$$

Where σ_{n+1} is the error of LB layer.

And the values of weights and threshold are updated by the formula (12) and (13).

$$w_{11} = w_{11} + a\sigma_2 t, w_{21} = -y_1(0), w_{22} = w_{22} - \frac{2b_1}{a}\sigma_2 l_2, \ldots, w_{2n} = w_{2n} - \frac{2b_{n-1}}{a}\sigma_n l_2 \tag{12}$$

$$\theta = (1 + e^{-w_{11}t})\left(\frac{w_{22}}{2} y_2(t) + \frac{w_{23}}{2} y_3(t) + \ldots + \frac{w_{2n}}{2} y_n(t) - y_1(0)\right) \tag{13}$$

3. IPSO-GNN Algorithm

3.1. PSO Algorithm

In PSO algorithm [13-20], all particles have fitness values which are evaluated and optimized by the fitness function, and have velocities which direct the flying of the particles. PSO is initialized with a group of random solutions and the optimized result is studied by updating generations. In iterations, each particle is updated by following two "best" values, best particle (represent with p_{best}) and the best from particles in the whole population by the PSO algorithm (represent with g_{best}). Particle velocity, position and fitness function of each generation are calculated by the following formulas:

$$v_{id}(k+1) = wv_{id}(k) + c_1 r_1 (p_{id}(k) - x_{id}(k)) + c_2 r_2 (p_{gd}(k) - x_{id}(k)) \tag{14}$$

$$x_{id}(k+1) = x_{id}(k) + v_{id}(k) \tag{15}$$

$$E = \frac{1}{N}\sum_{i=1}^{N}(y_i - y_i')^2 \tag{16}$$

Where i is the number of particles, $X_i=(x_{i1},x_{i2},\ldots,x_{id})$ $(1 \le d \le D)$ denotes the i-th particle, $P_i=(p_{i1},p_{i2}\ldots p_{id})$ $(1 \le d \le D)$ denotes the optimal position of the i-th particle and

$v_i=(v_{i1},v_{i2},\ldots,v_{id})$ ($1\leqslant d\leqslant D$) denotes velocity of the i-th particle, c_1 and c_2 is accelerating constants, rand ( ) produces random numbers, uniformly distributed in [0, 1], N is the number of particles in the population, y_i is the actual output of the i-th particles, y' is the expected output of the i-th particles, w denotes inertia weight and iterative updated by the following formula:

$$w = w_{\max} - k(w_{\max} - w_{\min})/T_{\max} \tag{17}$$

Where k is the evolutionary time, $w_{\min}$ and $w_{\max}$ is the maximum and minimum of inertia weight, $T_{\max}$ is the maximum number of iterations.

3.2. IPSO Algorithm

PSO algorithm is a random search algorithm based on swarm cooperative and prone to premature convergence in the latter part. Giving a threshold named as h which is close to 0. When the particle velocity is less than this threshold, we can add an accelerated momentum to re-initialize the speed which is between $-v_{max}$ and v_{max}, and then let the particle re-search to find the global optimum position. Above process can be expressed as:

$$v'_{id} = (1+\lambda)mv_{\max} \tag{18}$$

$$p_{id} = x'_{id} \tag{19}$$

Where v'_{id} denotes the improved speed of the particle, λ denotes optimized coefficient, m denotes a random variable parameter, x'_{id} denotes the partical position and p_{id} denotes the best partical position after velocity changed.

The optimized coefficient λ plays an important role in the algorithm and affects optimized results. If λ is too small, it may lead the particle velocity changed little and the optimized results are not obvious. If λ is too large, it may lead the particle velocity re-initialized too large and affects the convergence of particles, and it also may make the particles skip the global optimum. The value of optimized coefficient is determined by the following formula:

$$\lambda = \begin{cases} 0.5 & 0 < m \le 0.5 \\ -0.5 & 0.5 \le m < 1 \end{cases} \tag{20}$$

3.3. IPSO-GNN Algorithm

Realizing IPSO-GNN can be divided into three stages: determining GNN model structure, improving PSO algorithm and simulating network .The specific steps are as follows:

1. Based on the given training samples, GNN model is built with the initialize particle parameters of a, b_i, connection weights and threshold.
2. Normalize and pretreat the training sample.
3. Set the parameters a, b_i of GNN model as particles in particle swarm optimization algorithm. Initialize particle accelerating constant, maximum, minimum of inertia weight and maximum number of iterations; Initialize the particle velocity and position to determine the optimal value and the global optimum of the initial particle.
4. Compute the fitness value using formula (16). If the current position value is better, assign the current position to p_{best}. If the current best fitness value is better than the g_{best}, then assign the current best fitness value to g_{best}.

Copyright ⓒ 2015 SERSC

5. Compare the threshold with the current speed of the particle. If the particle velocity is less than *h*, then calculate a new speed with an accelerating momentum using formula (18) and return to step 2 to re-optimization.
6. Repeat steps 2 - 6 until a stop criterion is satisfied or the predefined number of iterations is completed. Then set the global optimum results as GNN model parameters' values of a, b_i.
7. Input training sample, then compute the error using formula (9), (10) and (11), and update the values of weights and threshold using formula (12) and (13).
8. Predict the telephone traffic using the trained network, and compare predicted values with actual values.

4. Simulation

4.1. Data Sources and Processing

As telephone traffic is influenced by many factors, we collect the traffic data and the main factors of A, B regions at 19:00 in January and February 2014, and analysis the correlation of the main factors using non-parameter statistical analysis such as Spearman correlation matrix. As it can effectively overcome the shortcoming of only suitable for describing linear correlation by the product moment correlation coefficient of Pearson, and provide the covariant trend extent of two random variables in a linear or non-linear correlation, the Spearman correlation matrix is often used to determine whether two random variables have covariant trend, and its applicability is better compared with other methods with the same parameters. Based on the literature [21], the correlation coefficient of Spearman is calculated with the software of Spss22 and the results are shown in Table 1 and 2.

The original hypothesis in correlation analyzing is that the correlation coefficient is zero, which means that there is no significant correlation between two random variables. If the concomitant probability of p is less than the significant level of a (here is a=0.01) which is previously given for hypothesis testing, we can consider that the possibility of the correlation coefficient is zero very low, meaning that there is obvious correlation between the two random variables. If the absolute value of the correlation coefficient is larger, the correlation between the two variables will be stronger. Generally, when the correlation coefficient of $| r |$ is less than 0.3, it means that there is no correlation; when $0.3 \leq | r | <0.5$, it means that there is low correlation; when $0.5 \leq| r |<0.8$, it means that there is significant correlation; when $| r | \geq 0.8$, it means that there is high correlation.

As seen in Table 1 and 2, the correlation coefficient of selected factors is greater than 0.5 and the concomitant probability of p is less than a. The total number of busy VLR users and power VLR users are significant correlation with the telephone traffic and others are high correlation with the telephone traffic. As the correlation coefficient of the number of system collect, the total number of system response, the total number of called call, the total number of called response and the total number of 2G originating response is greater, we can select the original data of them as the input data and the telephone traffic as the output data of network to train and forecast network.

4.2. Predict and Analysis Results

From the collecting data, we select the first 49 data as training sample and the others as test sample, and use the IPSO-GNN model for optimizing, training and prediction.

The structure of GNN is 1-1-6-1, and the initial parameter settings of IPSO algorithm include that accelerating constants of c_1 and c_2 are 2, inertia weights of $w_{\min}$ and $w_{\max}$ are 0.2 and 0.9, particle velocities of $v_{\min}$ and $v_{\max}$ are -1 and 1, the maximum number of iterations of T_{max} is 100. The predicted results of telephone traffic data of A, B regions using the IPSO-GNN, PSO-GNN， GNN and BPNN are shown in Figure 2 and 3, and the prediction error evaluated by formula (21) are shown in Table 3.

$$MAPE_i = \frac{1}{K}\sum_{j=1}^{k}\left|\frac{G_{ij} - G'_{ij}}{G_{ij}}\right|, (i = 1,2,..., 10) \quad (21)$$

Where G is the predicted value, G' is actual value, k is training time.

Table 1. Correlation Analysis of Telephone Traffic Factors

	Total number of busy VLR users	Total number of power VLR users	Total number of system test call	Number of system collect	Total number of system response	Total times of Network test call	Total times of System page
correlation coefficient	0.639	0.656	0.924	0.942	0.952	0.931	0.933
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 2. Correlation Analysis of Telephone Traffic Factors

	Total number of called call	Total number of Called response	Total number of switch test call	Total number of switch collect	Total number of 2G originating response	Total number of 2G terminating response
correlation coefficient	0.938	0.936	0.921	0.929	0.948	0.934
p	0.000	0.000	0.000	0.000	0.000	0.000

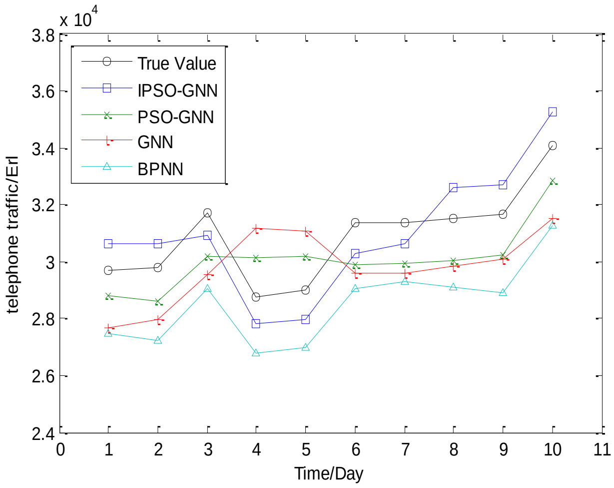

Figure 2. Comparison between Predicted Values and Real Values in A Region

 Copyright ⓒ 2015 SERSC

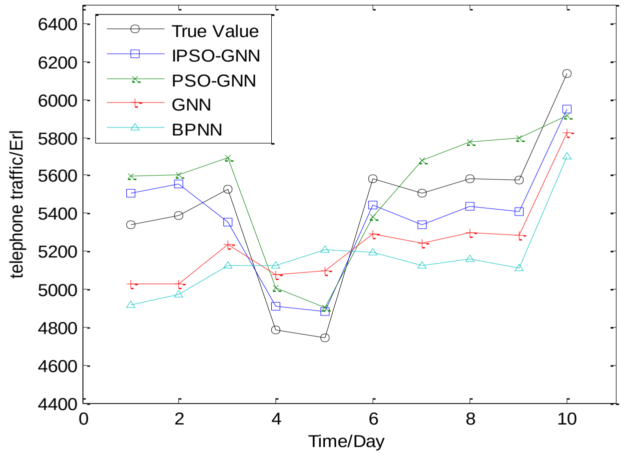

Figure 3. Comparison between Predicted Values and Real Values in B Region

Table 3. Statistics of Predicted Rrror

10 days	IPSO-GNN		PSO-GNN		GNN		BPNN	
	Error in A region (%)	Error in B region (%)	Error in A region (%)）	Error in B region (%)	Error in A region (%)	Error in B region (%)	Error in A region (%)	Error in B region (%)
1	3.11	2.97	3.05	4.50	7.35	6.26	8.02	8.71
2	2.63	2.93	4.24	3.75	6.61	7.15	9.55	8.38
3	2.56	3.21	5.16	2.92	7.33	5.57	9.24	7.76
4	3.28	2.54	4.61	4.43	7.78	5.73	7.34	6.62
5	3.86	2.91	3.90	3.35	6.68	7.00	7.51	8.95
6	3.60	2.51	4.95	3.62	5.95	5.46	8.02	7.46
7	2.46	3.16	4.74	3.02	6.10	5.08	7.15	7.50
8	3.25	2.62	4.99	3.39	5.77	5.40	8.47	8.18
9	3.22	3.09	4.72	3.81	5.12	5.44	9.48	8.99
10	3.46	3.10	3.70	3.72	7.99	5.33	8.98	7.62
mean	3.14	2.90	4.41	3.65	6.67	5.84	8.38	8.02

As can be seen from Figure 2, Figure 3 and Table 3, In terms of the predicting errors of two region in Xinjiang which are named as A and B, the errors of BPNN model are 8.38% and 8.02% respectively while those of GNN are 6.67% and 5.84%. It means that GNN outperforms BPNN when there are few observations. The errors of PSO-GNN model are 4.41% and 3.65% while the errors of IPSO-GNN are 3.14% and 2.90%. It demonstrates that the modified PSO has better global search ability to find the global optima. We can draw the conclusion that IPSO-GNN model has the most precise predicting results which fits the observations accurately compared to PSO-GNN, GNN and BPNN approaches. The proposed approach which combines the superiorities of PSO, GM and BPNN can describes the complex linear feature and nonlinear feature in the series of telephone traffic data. Hence, it significantly improves the predicting precision.

5. Conclusion

To make full use of the advantages of small sample, simple method in grey model and strong nonlinear mapping in neural network, the GNN model is built. The new model absorbs the advantages of single models and overcomes the shortcomings, so the model makes a better predict performance. As a mathematical optimization method, the PSO algorithm is emerging in recent years. To solve the problem that the PSO model may

Copyright ⓒ 2015 SERSC

fall into a local optimum and cannot obtain the global optimums when the particle's velocity below a certain threshold, an accelerated momentum applied on the particle to reinitialize the particle and position, so that the IPSO model is built to get the global optimum. Then the IPSO-GNN model is proposed and successfully applied to predict telephone traffic data. The results show that, compared to the PSO-GNN, GNN and BPNN model, IPSO-GNN model，it has higher prediction accuracy and generalization ability, and is a reliable method for telephone traffic prediction.

Due to time constraints, the IPSO-GNN model is only used to predict telephone traffic data in this paper. As well as, it can be used in other aspects。

Acknowledgements

This work was supported by the Foundation of China Mobile Communications Corporation Xinjiang Co., Ltd. Research and Development Project under contract No.XJM2013-2788.

References

[1] E. Kayacan, B. Ulutas and O. Kaynak, "Grey system theory-based models in time series prediction", Expert Systems with Applications, vol. 2, no. 37**, (2010)**.

[2] G. D. Li, S. Masuda and M. Nagai, "Predictor design using an improved grey model in control systems", International Journal of Computer Integrated Manufacturing, ahead-of-print **(2014)**.

[3] C. Hamzacebi and H. Avni Es, "Forecasting the annual electricity consumption of Turkey using an optimized grey model", Energy, vol. 70, **(2014)**.

[4] L. F. Wu, S. F. Liu and W. Cui, "Non-homogenous discrete grey model with fractional-order accumulation", Neural Computing and Applications, **(2014)**.

[5] B. Wang, X. J. Liang and H. Zhang, "Vulnerability of hydropower generation to climate change in China: Results based on Grey forecasting model", Energy Policy, vol. 65, (**2014**).

[6] D. Ma, Q. Zhang and Y. Peng, "A particle swarm optimization based grey forecast model of underground pressure for working surface", Electronic Journal of Geotechnical Engineering, vol. 6, **(2011)**.

[7] D. Zhang, Z. Ren and Y. Bi, "Power load forecasting based on grey neural network", Industrial Electronics, 2008, ISIE 2008, IEEE International Symposium on. 1885-1889, **(2008)**.

[8] J. Zhou and Y. Peng, "Variable weight combined forecast of China's energy demand based on grey model and BP neural network", Journal of Chemical and Pharmaceutical Research, vol. 4, no. 6, **(2014)**.

[9] Y. Ma and L. Peng, "Coupled Model of Artificial Neural Network and Grey Model for Tendency Prediction of Labor Turnover", Mathematical Problems in Engineering, 2014 **(2014)**.

[10] J. Yuan, L. Zhong and X. Li, "The Research and Development of Grey Neural Network", Journal of Wuhan University of Technology, vol. 3, no. 025, **(2009)**.

[11] Y. Wang, C. Yu and R. Tu, "Fire detection model in Tibet based on grey-fuzzy neural network algorithm", Expert Systems with Applications, vol. 8, no. 38, **(2011)**.

[12] Q. Bai, Y. Jing and B. Zhang, "Forecast of Gas Emission Based on Grey-Neural Network Combined Model", Coal Technology, vol. 8, no. 044, **(2009)**.

[13] A. Khare and S. Rangnekar, "A review of particle swarm optimization and its applications in Solar Photovoltaic system", J. Applied Soft Computing, vol. 5, no. 13, **(2013)**.

[14] A. I. Selvakumar and K. Thanushkodi, "A new particle swarm optimization solution to nonconvex economic dispatch problems", Power Systems, IEEE Transactions on, vol. 1,no. 22, **(2007)**.

[15] F. Valdez, P. Melin and O. Castillo, "Modular Neural Networks architecture optimization with a new nature inspired method using a fuzzy combination of Particle Swarm Optimization and Genetic Algorithms", Information Sciences, vol. 270, (**2014)**.

[16] S. Agrawal and S. Silakari, "Fletcher–Reeves based Particle Swarm Optimization for prediction of molecular structure", Journal of Molecular Graphics and Modelling, vol. 49, **(2014)**.

[17] L. Wang, B. Yang and Y. Chen, "Improving particle swarm optimization using multi-layer searching strategy", Information Sciences, vol. 274, **(2014)**.

[18] M. Pulido, P. Melin and O. Castillo, "Particle swarm optimization of ensemble neural networks with fuzzy aggregation for time series prediction of the Mexican Stock Exchange", Information Sciences, vol. 280, **(2014)**.

[19] H. J. Wang, M. Bai and Z. Jia, "Futures prices forecasting based on PSO neural network", Computer Engineering and Design, vol. 10, no. 025, **(2009)**.

[20] M.-S. Leu, M.-F. Yeh and S.-C. Wang, "Particle swarm optimization with grey evolutionary analysis. Applied Soft Computing, vol. 10, no. 13, **(2013)**.

 Copyright ⓒ 2015 SERSC

[21] S. Tao, "Applied Mathematical Statistics Method, China Environmental Science Press, Beijing, **(1994)**.

Authors

Xiuting Yu, She received her BE in Electronic and Information Engineering from Chaohu University in 2011. She is now a graduate student in Xinjiang University. She research interests include distributed computing and mobile Internet etc. Email: liulangyu_001@163.com

Xizhong Qin, He received his ME from Southeast University in1990. 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 and information processing, wireless communications.

Zhenhong Jia, He received his BE from Beijing Normal University in1986. 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. He has published over 100 papers. He was select for new century excellent talents plan in 2006. He is a member of Chinese Optical Society of Holography and Optical Information Processing Committee.

Chuanling Cao, She is an employee of China Mobile group Xinjiang Co., Ltd. Her major is business network.

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

Copyright ⓒ 2015 SERSC