

Forecasting Chaotic Time Series with Wavelet Neural Network Improved by Particle Swarm Optimization

Hui Li^{1,2}, Dechang Pi¹ and Min Jiang¹

¹ College of Computer Science and Technology, Nanjing University of Aeronautics
and Astronautics

29 Yuda Street, Baixia District, NanJing

² Department of Information Technology, Jinling Institute of Technology
99 Hongjing Ave., NanJing

¹lihui@jit.edu.cn, ²dc.pi@nuaa.edu.cn, ³alicehh_ok@126.com

Abstract

The prediction of chaotic time series is an important research issue. To improve the prediction accuracy, a hybrid approach called WNN-PSO is proposed, which based on the self-learning ability of wavelet neural network, whose parameters are optimized by particle swarm optimization. The WNN-PSO method has higher prediction accuracy, fast convergence, and heightens the ability of jumping the local optimums. The experiment results of the prediction for chaotic time series show the feasibility and effectiveness of the proposed method. Compared with wavelet neural network and BP neural network, the proposed method are superior to them. Finally, the WNN-PSO is applied to predict the life energy consumption of china in our lives.

Keywords: Chaotic time series, Wavelet neural network, Particle swarm optimization, Forecast

1. Introduction

Chaos seems an intrinsic random feature of deterministic system, and exists everywhere in our daily lives. It is not an individual accidental phenomenon. The dynamic system of chaos possesses fixed characteristics and rules. If we can find the determination rules, it is possible to predict the performance of the non-linear dynamic system in the future [1]. Chaotic time series is an important branch of chaos theory. It has widely applied in various fields of natural science and social science, such as, hydrological forecasts, stock market, sunspots, image processing, which has important practical value and significance. Therefore, the reconstruction model and prediction of chaotic time series are an important research issue.

The classical chaotic time series forecasting methods are global prediction approach [2], local prediction method [3], adaptive prediction [4], the largest lyapunov exponents prediction [5], and so on. Global prediction approach is clear in concept, but it is difficult with larger noise. Local prediction method found the adjacent points of prediction points in phase space, and took the next point of the most nearest point on the track as the prediction output. Once the attractor overstepped the corresponding area, the model would fail with poorer accuracy. The adaptive prediction had a higher demand for tracking identification and real-time recursive capabilities. The largest lyapunov exponents were not an intelligent technology and needed to calculate LEs for predicting chaotic time series data. Scholars at

home and abroad had also proposed a variety of intelligent prediction algorithms, such as chaotic prediction based on neural networks [6, 7, 8], gray theory prediction method [9], fuzzy neural network [10], time series prediction method based on the reserve pool [11], support vector machine method [12], wavelet neural network prediction method [13, 14]. These approaches have been applied in some fields, but with poor accuracy rate or slow speed because of single algorithm.

Recently, in order to improve the accuracy and speed of time series prediction, some combined forecasting algorithms are proposed. WANG [15] *et al.*, gave an approach of BP neural network based on improved particle swarm optimization to heighten the ability of jumping the local optimums because of introducing crossover operator and mutation operator. LI Xiang *et al.*, [16] presented an improved wavelet neural network combined with PSO algorithm to reduce the curse of dimensionality and to increase the precision. Meng Yuebo *et al.*, [17] provided a WNN-IPSO algorithm to improve the modeling capability of WNN. Georgios Sermpinis *et al.*, [18] proposed hybrid neural network architecture of particle swarm optimization and Adaptive Radial Basis Function (ARBF-PSO) to forecast foreign exchange rates. These methods were used for predicting time series. Arash Miranian and Majid Abdollahzade [19] introduced a local neuro-fuzzy approach based on the LSSVMS algorithm. It is powerful in modeling and predicting for nonlinear and chaotic time series, but it was a local modeling method, which was a little difficult.

Artificial Neural Network (ANN) was put forward based on the research results of modern neuroscience, which had already applied in a wide range of fields. The wavelet neural network (WNN) is a combination of wavelet analysis theory and ANN, which takes full advantage of time-frequency characteristics of wavelet transform and self-learning ability of the neural network. Compared with ANN, WNN takes the wavelet basis function as the transfer function of the hidden layer, with the signal forward propagation, and deviation back propagation, which shows better prediction accuracy, convergence rate and fault tolerance to the complex nonlinear, uncertain, unknown system [20, 21]. In addition, WNN has the advantages of adaptive learning ability, strict theory bases and high nonlinear mapping capability. The activation function of WNN introduces translation factor, dilation factor and other parameters, which makes the network have a better predictive ability and fault tolerance than other networks. Due to fewer research on using particle swarm optimization (PSO) algorithm to establish the chaotic time series prediction model, we introduce PSO algorithm into the field of prediction modeling to train the WNN for better performance, which has an important significance in theory and practice.

In this paper, a hybrid model based on WNN and PSO is presented. The proposed method has some features: (1) it owns better prediction accuracy than single existed algorithm; (2) it leads fast convergence and global optimum in our study; (3) it can improve the learning ability of WNN. The main idea of the proposed method is as follows: firstly, PSO was used to tune the parameters of WNN for satisfying the requirements of chaotic time series, and WNN was trained to learn and extract the features of chaotic time series. Then, the hybrid model was built to predict the time series data. Finally, simulation results showed the effectiveness and efficiency of the proposed method. In addition, the model was applied to predict life energy consumption in china.

The rest of the paper is organized as follows. The hybrid prediction model for chaotic time series is presented in Section 2. Section 3 discusses the experimental results for the proposed algorithm. Conclusions are given in Section 4.

2. Hybrid Model of WNN and PSO

2.1. Particle swarm optimization

Where there is great love, there are always miracles. Love is like a butterfly. It goes where it pleases and it pleases where it goes. If I had a single flower for every time I think about you, I could walk forever in my garden. Within you I lose myself, without you I find myself wanting to be lost again. At the touch of love everyone becomes a poet.

Particle Swarm optimization (PSO) [21] is a typical swarm intelligence algorithm, which is used to seek the optimal solution of the optimization problem. It stems from the simulation of bird predation. When the birds prey, the most effective and simplest way of each bird finding food is to search around the location of the bird closest to the food. PSO algorithm first assumes that every possible solution is a particle in the search space; the features of each particle are expressed as position vector X , velocity V , and fitness value f . Among them, the particle's position means a temporary solution in the solution space; the velocity represents the direction and distance of particles' motion. In order to get the best value in the solution space, they will be adjusted accordingly by the particle itself and other particles.

Generally, in the process of search, a particle updates itself by following two extreme values in the solution. One extreme value named P_b , is the optimal solution gained by the particle itself, and is the best fitness value calculated by the fitness function. The other one named P_g , is the optimal solution found by all particles at the present, and the best fitness value of all particles in the solution space. During the updating process, if the position of a particle is moved, the corresponding fitness value will be re-calculated; P_b and P_g of the particle will be renewed according to the new fitness value. Fitness value of the particle is calculated by a pre-defined fitness function, whose value indicates the relationship of the pros and cons of the particles in the search space. After constantly moving by own experience and its neighbors experience, the particle will eventually seek to the optimal value in the solution space.

Assume a swarm of particles named $x = (x_1, x_2, \dots, x_n)$, the position and the velocity of the i -th particle in D-dimensional solution space are represented as $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})^T$ and $V_i = (V_{i1}, V_{i2}, \dots, V_{iD})^T$ respectively.

Where, X_i represents a potential solution in the search space. The best individual extreme value of the particle i is expressed as $P_{bi} = (P_{bi1}, P_{bi2}, \dots, P_{biD})^T$, and $P_g = (P_{g1}, P_{g2}, \dots, P_{gD})^T$ presents the best position among all particles in the population X . By iterative calculation again, the velocity and position of the particles are updated through P_b and P_g according to the following formulas:

$$V_{id}^{k+1} = wV_{id}^k + c_1r_1(P_{bid}^k - X_{id}^k) + c_2r_2(P_{gd}^k - X_{id}^k) \quad (1)$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \quad (2)$$

Where $d = 1, 2, \dots, D$, D is the dimension of the space; $i = 1, 2, \dots, n$, n is the swarm size; k is the times of evolutionary iteration; c_1 and c_2 are the acceleration coefficients which are nonnegative constants, Here the value is 2; r_1 and r_2 are two generated uniformly distributed random numbers in the range of $[0, 1]$; w is the internal weight coefficient. In order to prevent

the particles falling into local optimal solution, or blindly searching in the solution space, V_i and X_i of the particle are limited within a set range $[-V_{\max}, V_{\max}]$ and $[-X_{\max}, X_{\max}]$ respectively.

Particle's fitness value is calculated by the fitness function, whose value is good or bad represents the pros and cons of the particles. Fitness function is defined by equation (3).

$$y = -c_1 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{j=1}^n x_j^2}\right) - \exp\left(\frac{1}{n} \sum_{j=1}^n \cos(2\pi x_j)\right) + c_1 + e \quad (3)$$

Where $c_1 = c_2 = 20$, $e = 2.71282$, $n = 2$.

2.2. Framework of hybrid structure

Traditional wavelet neural network uses gradient descent method to search, which is a typical local search algorithm. The network is easy to fall into local minima in the training process. The initial parameters of WNN are assigned based on prior knowledge, which is subjective and cannot reach a good accuracy of prediction. In PSO algorithm, a particle represents a possible solution during the optimization iteration. In the iteration process, the formula is simple; the calculation speed of the method is much faster than the gradient descent. In this paper, the PSO algorithm is used to optimize the structural parameters of the wavelet neural network, which does not need to get the derivative and differential of the activation function. Simultaneously, the parameters are adjusted through the iterative formula easily, so the network can jump out of local extreme.

The framework of prediction model based WNN improved by PSO is shown in Figure 1. Firstly, the WNN model is used to build the initial network topology. Then, PSO algorithm is employed to search for the optimal particle in the solution space, which is the optimal solution, and the initial parameters of wavelet neural network are assigned. Thirdly, the predicted data is decomposed into two parts. The characteristics of each part of the data are analyzed, and then train the network. Finally, the data are forecasted using the proposed model. Compared the predicted results with the actual data, the performance of the network is analyzed.

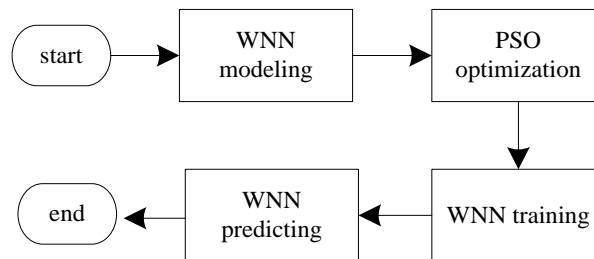


Figure 1. The prediction model of WNN improved by PSO

During the modeling process, the optimal structure of WNN depends on the number of nodes of different layers, especially the number of hidden layer nodes. If the number of hidden layer nodes is too small, the training effect is not very good. If too many, the training speed will be affected. The network topology used in the proposed model is the three-layered

structure. The structure is 4-6-1. That is, the number of the input layer nodes is 4. The number of hidden layer nodes is 6. The number of the output layer nodes is 1.

2.3. Wavelet neural network improved by PSO

In order to improve the accuracy of the prediction, PSO algorithm is used to optimize the weights and the wavelet coefficients of wavelet neural network (WNN-PSO). The network structure parameters of WNN are expressed as the position vector of the particle X , and are tuned through the iteration of equation (1) and (2) of PSO to find the optimal solution, where the fitness of the particle is calculated by equation (3). The process is repeated many times until it reached the pre-defined accuracy. The training process flowchart of PSO-WNN is shown in Figure 2.

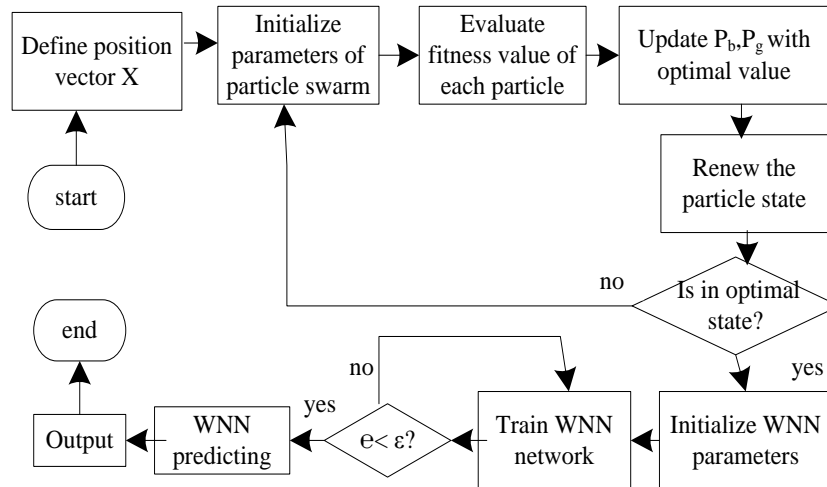


Figure 2. The flow chart for learning algorithm in WNN with PSO

The specific process of proposed model is described as follows:

Step 1: Determine the dimension of position vector X and the velocity vector V of each particle, Dimension $D = \text{number of hidden layer nodes} * \text{number of the input layer nodes} + \text{number of output layer nodes} * \text{number of hidden layer nodes} + \text{number of translation parameters} + \text{number of dilation parameters}$, namely, $D = I * J + J * K + 2 * J$.

Step2: Initialize the population, ensure the structure parameters of the network, and generate the initial particle swarm.

Step 3: Calculate the fitness value of each particle in particle swarm.

Step 4: Compare the fitness value of each particle in particle swarm with P_b , if the fitness is better, then update P_b .

Step 5: Compare the fitness value of each particle in particle swarm with P_g , if the fitness is better, then update P_g .

Step 6: Update the velocity v and position X of the particle in accordance with the formula (1) and (2).

Step 7: If the particle reaches a good position or the maximum number of iterations, the calculation process will be halted. Otherwise, go to step 3, continue to find the optimal values.

Step 8: Initialize the parameters of WNN utilizing the optimal solution, and train the network. Then, predict the data. Finally, output the forecast results.

During the entire optimization process, the velocity and position of each dimension of the particle cannot exceed the set interval of the algorithm, which are in the range $[-V_{max}, V_{max}]$ and $[-x_{max}, x_{max}]$ respectively. The range can be set according to the actual needs of different network parameters. In the iterative process, reducing V_{max} can improve the local search ability of particles, and then jump out of local optimum.

3. Empirical Results

The proposed mixed model WNN-PSO is used to predict chaotic time series data. Compared with existing models (BP, WNN), the proposed method can improve the forecast accuracy. Finally, the model is applied to our real life to predict the life energy consumption of china, and the prediction performance is analyzed.

In order to evaluate and analyze the experimental results, the mean square error (MSE) is used to measure the performance of different methods. The formula is shown in equation (4).

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - Y_i')^2 \quad (4)$$

Where Y and Y' represent the actual value and predicted value of chaotic time series respectively.

3.1. Prediction of Mackey-Glass Time Series

To validate the algorithm, the experimental data are produced by the Mackey-Glass differential delay equation. The equation is:

$$\frac{dx(t)}{dt} = \frac{ax(t-\tau)}{1+x^\gamma(t-\tau)} - \beta x(t) \quad (5)$$

Choose $\alpha = 0.2, \beta = 0.1, \gamma = 10$, the formula (5) is transformed into equation (6).

$$\frac{dx(t)}{dt} = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t) \quad (6)$$

The time series is the function of parameter τ , when $\tau > 16.8$, the system shows chaotic behavior, so let $\tau = 17$. Set the initial conditions, it can get the time series $x(t)$. Figure 3 shows the chaotic time series data when $\tau = 17$.

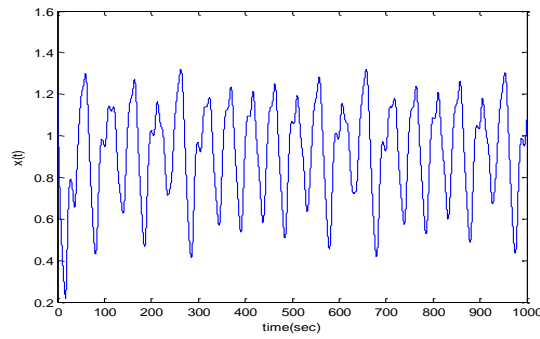
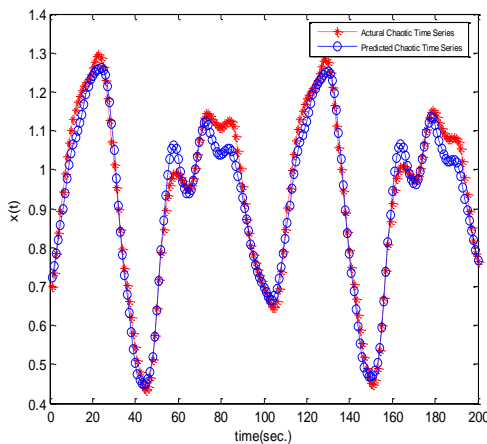


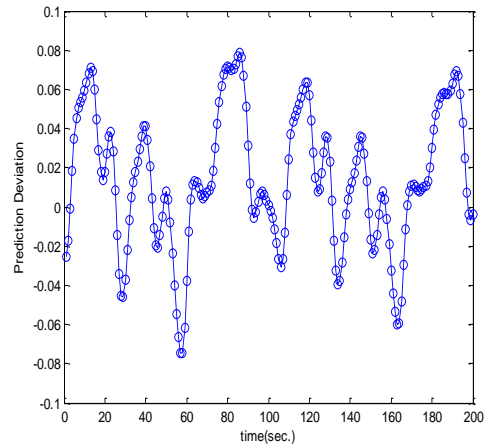
Figure 3. Mackey-Glass time series data

The prediction results and MSE are used to analyze the performance of BP, WNN, and WNN-PSO. The chaotic time series data used in the experiment contain 1000 time points and the corresponding data values, which selected as the sample data. The former 800 data points were taken as the training samples, the other 200 data points were chosen as the test data samples. Three kinds of networks trained carried out 100 times on the average. The below experiment figures give the distribution of predicted and desired results. Then, compared the forecast data with the real data, the prediction accuracy of the various algorithms were analyzed.

Firstly, BP neural network is used to predict the test data, the prediction results can predict the trend of chaotic time series, but the larger deviation is appeared. Figure 4 shows the prediction results and the prediction deviation.



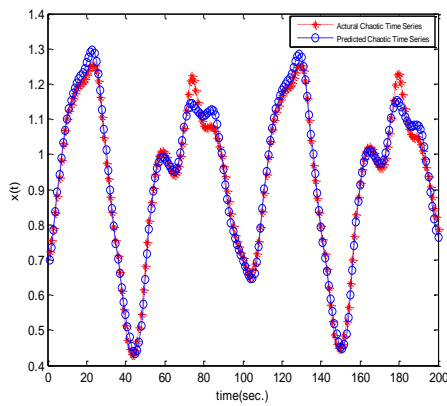
(a) The prediction results of BP network



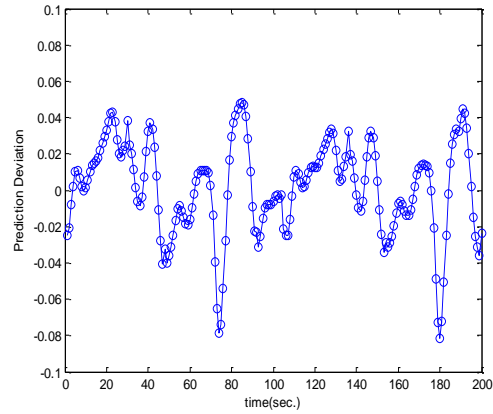
(b) The prediction deviation of BP network

Figure 4. The Experimental results by BP Network

Then, WNN method is used to experiment on the data. Compared with BP network, the results has been slightly improved. Figure 5 shows the results of WNN prediction, and the prediction error.



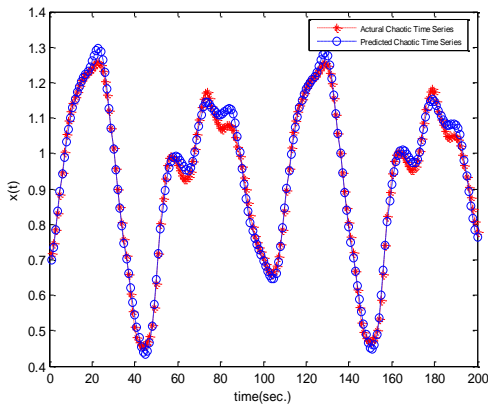
(a) The Experimental results of WNN



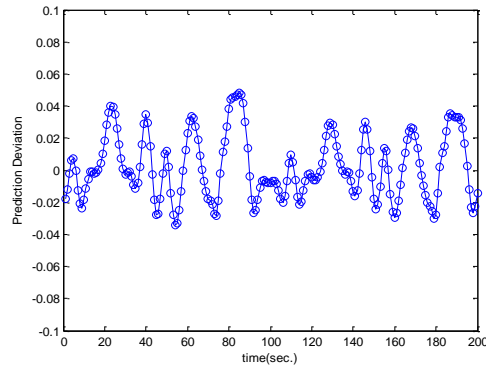
(b) The Experimental deviation of WNN

Figure 5. The Experimental results by WNN network

In order to obtain a better solution, PSO is used to optimize wavelet neural network, then WNN is trained to predict the test data. The experimental results is shown in Figure 6(a), the prediction error is shown in Figure 6(b).



(a) The prediction results of WNN-PSO



(b) The prediction deviation of WNN-PSO

Figure 6. The Experimental Results of WNN-PSO

As can be seen from the above experimental results, the proposed WNN-PSO can bring a higher degree of accuracy, which makes the predicted experiment results be closer to the real data. In order to assess the prediction errors, the MSEs of three network models above are calculated. As shown in Table 1, the WNN optimized by PSO owns the smallest prediction deviation when predicting chaotic time series.

Table 1. The comparison of MSE of three prediction model

Network Topology	BP	WNN	PSO-WNN
MSE	9.6814×10^{-6}	5.7267×10^{-6}	3.6908×10^{-6}

The prediction accuracy of WNN depends on the optimal solution of PSO algorithm, which affected by the quantity of swarm. The experiment sets the number of swarm for 20, 30, 40 and 50, then the proposed WNN-PSO method is used to predict the data respectively. As the population increasing, the MSE decreases. The results of the comparison are shown in Table 2.

Table 2. The comparison of MSE with increased population in WNN-PSO

Population	20	30	40	50
MSE	3.6908×10^{-6}	3.5626×10^{-6}	3.4773×10^{-6}	3.3964×10^{-6}

3.2. Real-world application prediction of life energy

In this section, the proposed WNN-PSO model is applied to the prediction of life energy data. The data sets used in the experiments, include the total consumption from the year 1978 to 2009 of the life energy, which comes from the National Bureau of Statistics website [23]. The data contains 32 data points.

The training data of the experiments of three methods are the life energy consumption between the years 1978 to 2004, the testing data sets are from the year 2005 to 2009. The number of training times for each network is 100 times, and the prediction errors of different models are given to measure the accuracy of the prediction.

The MSE of three models are given in Table 3. As shown in Table 3, the proposed model WNN-PSO can provide better prediction accuracy compared with other two models.

Table 3. The comparison of MSE of three prediction model

Network Topology	BP	WNN	PSO-WNN
MSE	7.2736×10^4	4.6534×10^4	2.8921×10^4

4. Conclusions

The chaotic data exists in our daily lives, the prediction of chaotic time series is a research hotspot in chaotic areas [24, 25]. Because of self-learning, fault tolerance, non-linearity of wavelet neural network, the features of optimization and simplicity of PSO, the forecast model based on these advantages is built for predicting chaotic time series. The proposed model avoids the weakness of falling into local optimization easily, slow convergence, poor prediction accuracy, which owns better global feature and convergence rate. Compared with traditional classical models, the proposed method can further improve the accuracy of the chaotic time series. Moreover, the achievement of the model is efficient enough to be applied in life energy consumption data.

The research gives a short-term prediction of chaotic time series, and cannot give an effective long-term prediction of it, which is a difficult issue in chaotic field. The long-term forecast of chaotic time series will be the focus of further research.

Acknowledgements

This work was supported by Qing Lan Project of Jiangsu province, 333 Project of Jiangsu province, Aeronautical Science Foundation of China (No. 20111052010), the National 863 Program (No. 2007AA01Z404), Universities Natural Science Research Project of Jiangsu Province (No. 13KJD520005), Modern Educational Technology in Jiangsu province (2013-R-26144).

References

- [1] H. Yoshito, O. Makito and A. Kazuyuki, "Chaos in neurons and its application: Perspective of chaos engineering", *Chaos*, vol. 22, no. 4, (2012).
- [2] S. F. Su, C. B. Lin and Y. T. Hsu, "A high Precision global Prediction approach based on Local prediction approaches", *IEEE Transactions on Systems, Man and Cybernetics, Part C*, vol. 32, no. 4, (2002).
- [3] J. D. Farmer and J. Sidorowich, "Predicting chaotic time series", *Physical Review Letters*, vol. 59, no. 8, (1987).
- [4] J. C. Gan and X. C. Xiao, "Adaptive predict-filter of chaotic time series constructed Based on the neighbourhood in the reconstructed phase space", *Acta Physica Sinica*, vol. 52, no. 5, (2003).
- [5] J. Zhang, K. F. Man and J. Y. Ke, "Time series prediction using Lyapunov exponents in embedding phase space", *Computers & Electrical Engineering*, vol. 30, no. 1, (2004).
- [6] B. Sivakumar, A. W. Jayawardena and T. M. K. G. Fernando, "River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches", *Journal of Hydrology*, vol. 265, no. 1, (2002).
- [7] M. Han, J. Xi, S. Xu and F. Yin, "Prediction of chaotic time series based on the recurrent predictor neural network", *IEEE Transactions on Signal Processing*, vol. 52, no. 11, (2004).
- [8] C. Wong and M. Versace, "CARTMAP: a neural network method for automated feature selection in financial time series forecasting", *Neural Comput. & Applic.*, vol. 21, no. 5, (2012).
- [9] E. Kayacan, B. Ulutas and O. Kaynak, "Grey system theory-based models in time series prediction", *Expert Systems with Applications*, vol. 37, no. 2, (2010).
- [10] N. K. Kasabov and Q. Song, "DENFIS: Dynamic Evolving Neural-Fuzzy Inference System and Its Application for Time-Series Prediction", *IEEE TRAN. On Fuzzy System*, vol. 10, no. 2, (2002).
- [11] M. Han, Z. W. Shi and W. Guo, "Reservoir neural state reconstruction and chaotic time series prediction", *Acta Physica Sinica*, vol. 56, (2007).
- [12] B. Zhu and Y. Wei, "Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology", *Omega*, vol. 41, (2013).
- [13] Y. Chen, B. Yang and J. Dong, "Time-series prediction using a local linear wavelet neural network", *Neurocomputing*, vol. 69, no. 4, (2006).
- [14] Z. Zainuddin and O. Pauline, "Modified wavelet neural network in function approximation and its application in prediction of time-series pollution data", *Applied Soft Computing*, vol. 11, no. 8, (2011).
- [15] Z. Wang, X. Zhao, B. Wan and J. Xie, "Research of BP Neural Network based on Improved Particle Swarm Optimization", *Journal of Networks*, vol. 8, no. 4, (2013).
- [16] X. Li, S. D. Yang, J. X. Qi and S. -x. Yang, "Improved wavelet neural network combined with particle swarm optimization algorithm and its application", *J. Cent. South. Univ. Technol.*, vol. 13, no. 3, (2006).
- [17] Y. B. Meng, J. H. Zou and X. S. Gan, "Research on WNN aerodynamic modeling from flight data based on improved PSO algorithm", *Neurocomputing*, vol. 83, no. 4, (2012).
- [18] G. Sermpinis, T. Konstantinos and K. Andreas, "Forecasting foreign exchange rates with adaptive neural networks using radial-basis functions and Particle Swarm Optimization", *European Journal of Operational Research*, vol. 225, no. 3, (2013).
- [19] A. Miranian and M. Abdollahzade, "Developing a Local Least-Squares Support Vector Machines-Based Neuro-Fuzzy Model for Nonlinear and Chaotic Time Series Prediction", *IEEE TRAN. On Neural Networks and Learning Systems*, vol. 24, no. 2, (2013).
- [20] Q. H. Zhang and A. Benveniste, "Wavelet network", *IEEE Trans. on Neural Network*, vol. 3, no. 6, (1992).
- [21] Q. H. Zhang, "Using wavelet network in nonparametric estimation", *IEEE Trans. Neural Networks*, vol. 8, no. 2, (1997).
- [22] J. Kennedy and R. C. Eberhart, "Particle swarm optimization", *Proc. IEEE Int. Conf. Neural Networks IV*, (1995), pp. 1942-1948.
- [23] <http://219.235.129.58/welcome.do>.

Authors



Hui Li

She received her M.Sc. in Computer Sciences and Technology (2006) from Northwestern Polytechnical University. Now she is a Ph.D. candidate at the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics. Since 2006 she is a lecture of Jinling Institute of Technology. Her current research interests include different aspects of Artificial Intelligence and Data Mining.



Dechang Pi.

He received his M.Sc. in Computer Sciences and Technology (1996) and PhD in Computer Sciences (2002) from Nanjing University of Aeronautics and Astronautics. Now he is full professor and Ph.D. supervisor of computer at the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics. His current research interests include different aspects of Artificial Intelligence and Data Mining.



Min Jiang

He received his M.Sc. in the College of Computer Science and Technology (2008) from Nanjing University of Aeronautics and Astronautics. Now he is a full research engineer in a research agency. His current research interests include different aspects of Artificial Intelligence and Distributed Systems.

