2017 International Conference on Computer Network, Electronic and Automation (ICCNEA 2017)

Forecasting CSI 300 index using a Hybrid Functional Link Artificial Neural Network and Particle Swarm Optimization with Improved Wavelet Mutation

Tian Lu
Dept. of Math. and Physics
North China Electric power university
Beijing, China
E-mail: lutian515@163.com

Zhongyan Li
Dept. of Math. and Physics
North China Electric power university
Beijing, China
E-mail: lzhongy@ncepu.edu.cn

***Abstract*—Financial market dynamics forecasting has long been a focus of economic research. A hybridizing functional link artificial neural network (FLANN) and improved particle warm optimization (PSO) based on wavelet mutation (WM), named as IWM-PSO-FLANN, for forecasting the CSI 300 index is proposed in this paper. In the training model, it expands a wider mutation range while apply wavelet theory to the PSO, in order to exploring the solution space more effectively for better parameter solution. In the stimulating experiment, we use five benchmark functions to test the proposed method, and the results shows that IWM-PSO has greater convergence accuracy than WM-PSO and PSO. The empirical research is performed in testing the predictive effects of CSI 300 index in the proposed model compared with the back propagation functional link neural network (BP-FLANN), PSO-FLANN and WM-PSO-FLANN. The experiment utilizes two expansion functions, Chebyshev functions and trigonometric functions, to map the input data to higher dimension. The results show that the prediction performance of the proposed model displays a better performance in financial time series forecasting than other three models. Moreover, the accuracy of the input with trigonometric functions is higher, and it suggests that trigonometric function is more suitable for this kind of data type.**

Keywords-Stock index; Forecasting; Functional link artificial neural network (FLANN); Improved wavelet mutation (IWM); Particle warm optimization (PSO)

I. INTRODUCTION

The CSI 300 index is composed of the 300 largest and most liquid A-shares listed on the two stock exchanges, launched on April 8th, 2005, and it aims to reflect the price fluctuation and performance. The CSI index has been used as the basis for many financial products around the world and is also used by investors to develop and benchmark their portfolios [1].

Predicting the stock market index is of prime importance in private and institution investors when making investment decision because successful prediction of stock prices may be guaranteed benefits. Their mainly quest is to forecast, to determine the appropriate time to buy, hold or sell. The investors assume that the future trend of the stock market is based at least in part on present and past events and data. However, the price variation of stock market is a dynamic system and the disordered behavior of the stock price movement duplicates complication of the price prediction and the highly non-linear, dynamic complicated domain knowledge inherent in the stock market makes it very difficult for investors to make the right investment decisions. Many researchers in the past have applied various computing techniques to predict the movement of the stock markets.

Artificial neural network have gained its popularity due to their inherent capability to approximate any none-linear function, less sensitive to error, tolerate noise, and chaotic components. To improve predicting precision, various network architectures and learning algorithms have been developed in the papers [2-4].

Further it is well known that the artificial neural network (ANN) suffers from local minima trapping, saturation, weight interference, initial weight independence and over fitting, make ANN training difficult. Additionally, it is also very difficult to fix the parameters like number of neurons in a layer, and number of hidden layers in a network, thereby deciding a proper architecture is not that easy. Thus to overcome from this functional link artificial neural network (FLANN) based simple network may be used for the prediction the stock index data with less computational overhead than the ANN [5]. FLANN is proposed by Y.H.PAO and Y. Takefji [6] contains a single layer neural network in which nonlinearity is introduced as a functional block, thus giving rise to higher dimension input space. In FLANN, higher-order correlations among input components can be used to construct a higher –order network to perform non- linear mapping using only a single layer of units.

The researchers proposed different FLANN models to predict the stock market indices, such as DJIA, S&P 500 stock indices[7,8] Indian stocks[9,10], and there is no research about the CSI 300 index.

Particle swarm optimization (PSO) is a population-based stochastic optimization algorithm which inspired by the social behaviors of animals like fish schooling and bird flocking proposed by Kennedy J and Eberhart R [11]. The PSO has comparable or even superior search performance for many hard optimization problems with faster and more stable convergence rate. The PSO has been widely used in parameter learning of neural network [12-14]. We can harness the power of PSO to train our model to reduce the possibility to be trapped in local optimal and speed up the convergence.

The rest of the paper is organize as follows: in section 1, we will give a brief introduction of the study. Section 2 deals with model development using FLANN structure using the improved wavelet mutation based PSO algorithm. The details experiment setup including process are discuss in Section 3. The simulating study and the results obtained from experiment are provided in section 4. In the final section the conclusion has been made.

II. Model development

A. The FLANN Model

Depending the complexities of the problems, number of layers and number of neurons in the hidden layer need to be changed. As the number of layers and the number of neurons in the hidden layer associated with multi-layer neural network increases, training the model becomes more complex. In FLANN, each element of the input data undergoes functional expansion through a set of basis functions to enhance the input pattern with nonlinear functional expansion. This enables the FLANN to solve complex problems by generating non-linear decision boundaries.

The architecture consist of two parts, namely, transformation part and learning part. The transformation deals with the input feature vector to hidden layer by approximate transformable method. A simple FLANN model with a pattern of two features is shown in Fig .1.

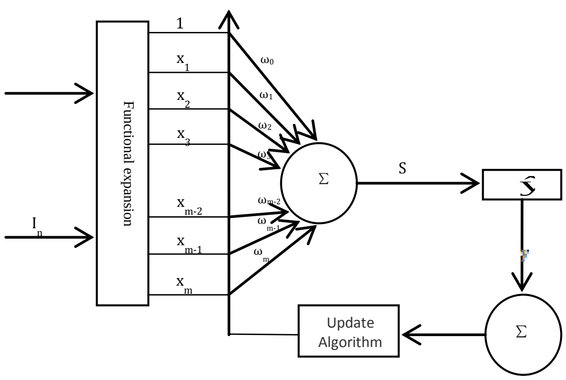

Figure 1. Structure of FLANN model with a pattern of two features

Let us consider a two dimension input sample $I = [I_1, I_2]$. This sample is mapped to a higher dimensional space by functional expanding using trigonometric functions (1):

$$X = [(I_1, \cos\pi I_1,, \sin\pi I_1 \cdots \cos N\pi I_1,, \sin N\pi I_1), (I_2, \cos\pi I_2,, \sin\pi I_2 \cdots \cos N\pi I_2,, \sin N\pi I_2)]^T \quad (1)$$

Each input sample is expanded to N sine Terms, N cosine terms plus the sample itself. If Chebyshev polynomial basis function (2) is employed

$$X = [(Ch_0(I_1), Ch_1(I_1) \cdots Ch_{n+1}(I_1)), (Ch_0(I_2), Ch_1(I_2) \cdots Ch_{n+1}(I_2))]^T \quad (2)$$

where $Ch_0(x) = 1$, $Ch_1(x) = x$, $Ch_{n+1}(x) = 2xCh_n(x) - Ch_{n-1}(x)$, and n is the order of the polynomial chosen.

The output of hidden neuron is given:

$$\hat{y} = \tanh\left(\sum_{i=1}^{m} x_i w_i - \theta\right) \quad (3)$$

Where θ is the threshold of neuron in the output layer.

B. Improved Wavelet Mutation Based Particle Swarm Optimization

In [15], it proposed an improved version of the PSO, where the constriction inertia weight factors are introduced, the velocity v_{ij}^t and the position x_{ij}^t of the element the *j*th of the particle *i*th at the *t*th iteration can be calculated using (4)(5):

$$v_{ij}^{t+1} = k\left((wv_{ij}^t + \phi_1 r_1)\left(pbest_{ij}^t - x_{ij}^t\right) + \phi_2 r_2\left(gbest_j^t - x_{ij}^t\right)\right) \quad (4)$$

$$x_{ij}^{t+1} = x_{ij}^t + v_{ij}^{t+1} \quad (5)$$

where $pbest_I^t$ is the best position in the history of particle *i*th at the *t*th iteration; $gbest^t$ is the global best position in the history at the *t*th iteration; r_1 and r_2 are uniform random number in the range of [0,1]; φ_1 and φ_2 are acceleration constants is a constriction factor derived from the stability analysis of to ensure the system to be converged but not prematurely. Mathematically, k is a function of φ_1 and φ_2 as reflected in (6):

$$k = \frac{2}{\left|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}\right|} \quad (6)$$

Where $\varphi = \varphi_1 + \varphi_2$, and $\varphi > 0$

w is inertia weight. Generally can be dynamically set with (7):

$$w = V\max\text{-}\left(V\max - V\min\right)\frac{t}{T\max} \quad (7)$$

Where t is the current iteration number, *Tmax* is the total number of iteration, and *Vmax* and *Vmin* are the upper and lower limits of the inertia weight.

However, observation reveals that the PSO sharply converges in the early stages of the searching process, but saturated or even terminates in the later stages. In [11], it proposed the operation of the hybrid PSO with a wavelet mutation. Every particle element of the swarm will have a chance to mutate that is governed by a probability of mutation p_m, which is defined by the user. For each particle element, a random number between 0 and 1 will be generated such that if it is less than or equal to p_m, a mutation will take place on that element.

The element of particle is randomly selected for the mutation (the value of x_{ij}^t is inside the particle element's boundaries [$para_{min}^j$, $para_{max}^j$]), the resulting particle is given by (8):

$$\text{mut}(x_{ij}^t) = \begin{cases} x_{ij}^t + \sigma * (para_{\max}^j - x_{ij}^t), \sigma > 0 \\ x_{ij}^t + \sigma * (x_{ij}^t - para_{\min}^j), \sigma \le 0 \end{cases} \tag{8}$$

The Morlet wavelet integrates to zero. Over 99% of the total energy of the function is contained in the interval of [-2.5, 2.5]. By using Morlet wavelet in as the mother wavelet.

$$\sigma = \frac{1}{a} e^{-\left(\frac{b}{a}\right)^2/2} * \cos\left(5 * \frac{b}{a}\right) \tag{9}$$

Hence, the overall positive mutation and the overall negative mutation throughout the evolution are nearly the same. This property gives better solution stability. $\sigma = 1$, $mut(x_{ij}^t) = para_{max}^j$;$\sigma = -1$, $mut(x_{ij}^t) = para_{min}^j$. A larger value of $|\sigma|$ gives a larger searching space for fine-tuning. b can be randomly generated form [-0.25a, 0.25a]. A monotonic increasing function governing a and $t/Tmax$ is proposed as follows:

$$a = e^{-\lg(g)*\left(1-\frac{t}{T\max}\right)^{\xi_{wm}} + \lg(g)} \tag{10}$$

where ξ_{wm} is the shape parameter of the monotonic increasing function, and g is the upper limit of the parameter a.

However, after mutation, the value of the particle element is between [$para_{min}^j$, $para_{max}^j$]. If the particle did not go by the best solution, it still has the possibility of being trapped in the local optima. So we propose the particle swarm optimization based on an improved wavelet mutation by expand the mutation range during each generation, so that the particle would be more likely to fly near the global best optima. The equation of the mutation can be changed into (11):

$$\text{mut}(x_{ij}^t) = \begin{cases} x_{ij}^t + \sigma * (\eta_{\max} * para_{\max}^j - x_{ij}^t), \sigma > 0 \\ x_{ij}^t + \sigma * (x_{ij}^t - \eta_{\min} * para_{\min}^j), \sigma \le 0 \end{cases} \tag{11}$$

Where η_{max} and η_{min} is defined by users to expand the mutation range.

III. Analysis of Datasets

A. Stimulation Study

TABLE I. BENCHMARK FUNCTIONS

Function expression	ξ_{wm}
$f_1 = \sum_{i=1}^{n} x_i^2$	2
$f_2 = \sum_{i=1}^{D}\left(\sum_{j=1}^{i} x_i^2\right)$+random[0,1)	4
$f_3 = \frac{1}{4000}\sum_{i=0}^{n-1} x_i^2 + \sum_{i=0}^{n-1}\cos(\frac{x_i}{\sqrt{i+1}}) + 1$	4
$f_4 = \sum_{i=1}^{n}(x_i^2\text{-}10\cos(2\pi x_i)+10)$	6
$f_5 = 0.5 + \frac{\sin(\sqrt{x_1^2+x_2^2})^2 - 0.5}{\left(0.1+0.001\left(x_1^2+x_2^2\right)\right)^2}$	6

A suite of five benchmark test functions is used to test the performance of the proposed model. Many different kinds of optimization problems are covered by these benchmark test functions. These five benchmark functions can be divided into three categories. The first one is the category of the unimodal functions, which is a symmetric model with one single minimum, f_1 and f_2. The second one is the category of multimodal functions with some local minima, f_3 and f_4. The third category is low dimension function, f_5. The benchmark function expressions and their parameter ξ_{wm} setting for different functions are shown in Table I.

B. Empirical Dataset

Empirical study is carried out using the date set CSI 300 from 2008/1/2 to 2017/3/15 up to 2236 trading days. Following 7:3 ratios, we use 1566 days for training and remaining 670 days for validating the model. Choose five kinds of stock prices as the input values in the input layer: daily highest price, daily lowest price, daily closing price, change rate and turnover, and the output layer is the closing price of the next trading day. The entire data is normalized to values between 0 and 1. The normalization formula in (12) to express the data in terms of the minimum and maximum value of the dataset.

$$\text{input} = \frac{\text{x} - Min.}{Max. - Min} \tag{12}$$

Where input and x represents normalized and actual value respectively.

We use both Chebyshev functions and trigonometric functions to map the data into high dimension. In trigonometric functions, we set the N is 2, and the swarm dimension would be 26, the size of the swarm we set is 60. In trigonometric functions, we set the n is 3, and the swarm dimension would be 16, the size of the swarm we set is 40. And the p_m is 0.2 and ξ_{wm} we set is 0.2.

To compare the forecasting performance of four considered forecasting models, IWM-PSO-FLANN, WM-PSO-FLANN, PSO-FLANN and BP-FLANN, we use the following criteria: the mean absolute error (MAE)(13), the root mean absolute error (RMSE)(14) and the mean absolute percentage error (MAPE)(15), the corresponding definition are given as follows:

$$\mathrm{MAE} = \frac{1}{\mathrm{N}}\sum_{t=1}^{N}\left|y_t - \hat{y}_t\right| \tag{13}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}}\sum_{t=1}^{N}\left(\left|y_t - \hat{y}_t\right|\right)^2} \tag{14}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}}\sum_{t=1}^{N}\left(\left|y_t - \hat{y}_t\right|\right)^2} \tag{15}$$

Where y and $\hat{y}$ represents the actual and forecast values; N is the total number of the data. Noting that MAE, RMSE, and MAPE are measures of the deviation between the prediction values and the actual values, the prediction performance is better when the values of these evaluation criteria are smaller. However, if the results are not consistent among these criteria, we choose the MAPE as the benchmark since MAPE is relatively more stable than other criteria.

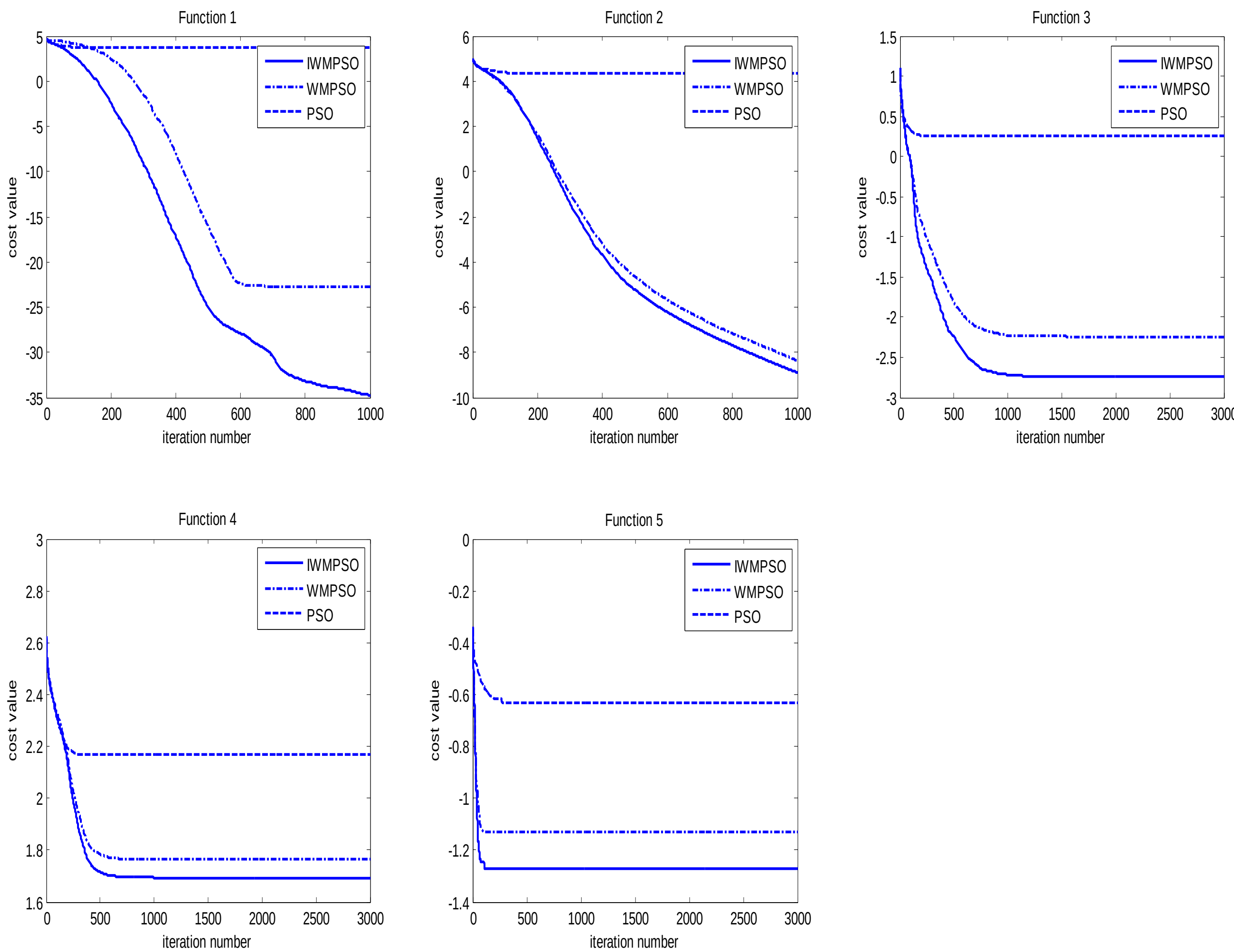

Figure 2. Convergence curves of benchmark function for different PSO method

IV. RESULTS AND DICUSSION

A. *Benchmark Test Functions*

The result for the five benchmark test functions are in Table II and Fig .2. From Fig .2, The IWM-PSO will get the same accuracy as other two models in fewer iteration times. And from Table II, results of stimulating experiment in terms of the mean cost values and the best cost value of the IWM-PSO are much better than those of the other methods. Also, the standard deviation is much better, which means that the searched solutions are more stable. From what mentioned below, we can conclude that the IWM-PSO will have better and more stable solution compared to other models.

TABLE II. COMPARISON BETWEEN DIFFERENT PSO METHOD FOR BENCHMARK TEST FUNCTIONS.

		IWM-PSO	WM-PSO	PSO
f_1	Mean	-34.7321	-22.7162	3.6758
	Best	-47.1855	-38.8609	2.6199
	Std Dev	-68.0077	-44.0165	1.2541
f_2	Mean	-8.8846	-8.3512	4.3211
	Best	-28.8609	-10.5959	3.4875
	Std Dev	-18.9078	-12.5434	0.16147
f_3	Mean	-2.7140	1.7642	2.1664
	Best	1.3407	1.2022	1.9377
	Std Dev	2.3953	2.8338	2.9743
f_4	Mean	1.6944	1.7642	2.1664
	Best	1.3407	1.2022	1.9377
	Std Dev	2.4953	2.8338	2.9743
f_5	Mean	-1.2728	-1.1299	-0.6316
	Best	-3.1565	-2.0125	-1.4325
	Std Dev	-2.0186	-1.7167	-1.3677

B. *CSI 300 Index Prediction*

Fig .3. Shows the predictive values on the test set for CSI 300 ((a) uses trigonometric functions and (b) uses Chebyshev functions). The empirical research shows that the proposed the IWM-PSO-FLANN model has the best performance. When the stock market is relatively stable, the forecasting result is near to the acturacl value. At the same time, we can see that the large fluctuation period forcasting is relatively not that accurate from these four models. But the IWM-PSO-FLANN has the best performance than other three models in this period. In Table III, the results show that the forecasting results of the proposed model are almost smaller than those by other models and these can conclude that the proposed IWM-PSO-FLANN model is better than the three other models. Morever, the model expaned with trigonometric functions has smaller values of MAE, RMSE and MAPE, which mean that trigonometric functions is more suitable for this kind of data type.

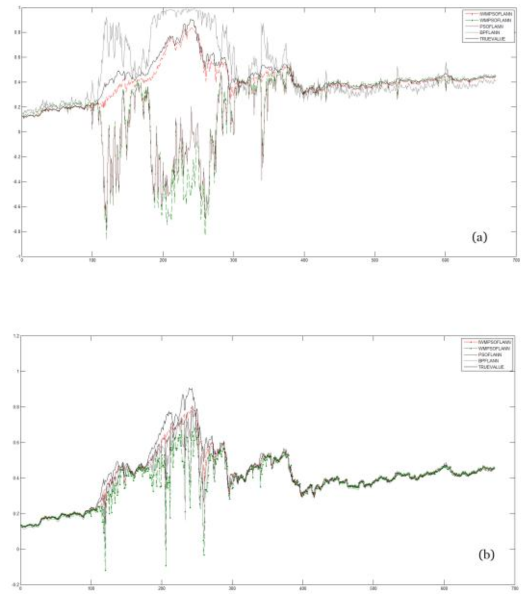

Figure 3. Forecasting result of CSI 300 stock insex for four forecasting models

TABLE III. COMPARISION OF INDEX'S PREDICTION FOR DIFFERENT FORECASTING MODEL

	trigonometric functions			
	IWM-PSO-FLANN	WM-PSO-FLANN	PSO-FLANN	BP-FLANN
MAE	0.018266	0.050491	0.025934	0.039624
RMSE	0.032258	0.111879	0.050256	0.088087
MAPE	3.766379	9.09154	5.126772	7.784697
	Chebyshev functions			
	IWM-PSO-FLANN	WM-PSO-FLANN	PSO-FLANN	BP-FLANN
MAE	0.028131	0.229882	0.213812	0.098014
RMSE	0.043821	0.456479	0.414253	0.143458
MAPE	6.255943	41.097287	38.053276	39.453523

ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression, "One of us (R.B.G.) thanks . . ." Instead, try "R.B.G. thanks". Put applicable sponsor acknowledgments here; DO NOT place them on the first page of your paper or as a footnote.

REFERENCES

[1] Biktimirov E N, Li B. Asymmetric stock price and liquidity responses to changes in the FTSE Small Cap index[J]. Review of Quantitative Finance and Accounting, 2014, 42(1):95-122.

[2] Lahmiri S. Wavelet low- and high-frequency components as features for predicting stock prices with backpropagation neural networks[J]. Journal of King Saud University - Computer and Information Sciences, 2014, 26(2):218-227.

[3] Rather A M, Agarwal A, Sastry V N. Recurrent neural network and a hybrid model for prediction of stock returns[J]. Expert Systems with Applications, 2015, 42(6):3234-3241.

[4] Oliveira F A D, Nobre C N, Zárate L E. Applying Artificial Neural Networks to prediction of stock price and improvement of the directional prediction index – Case study of PETR4, Petrobras, Brazil[J]. Expert Systems with Applications, 2013, 40(18):7596-7606.

[5] Das S, Patra A, Mishra S, et al. A self-adaptive fuzzy-based optimised functional link artificial neural network model for financial time series prediction[J]. International Journal of Business Forecasting & Marketing Intelligence, 2015, 2(1):55.

[6] Pao Y H. Adaptive pattern recognition and neural networks [M]. Addison-Wesley, 1989.

[7] Majhi R, Panda G, Sahoo G. Development and performance evaluation of FLANN based model for forecasting of stock markets[J]. Expert Systems with Applications An International Journal, 2009, 36(3):6800-6808. A

[8] Rout M, Majhi B, Mohapatra U M, et al. Stock indices prediction using radial basis function neural network[C]// International Conference on Swarm, Evolutionary, and Memetic Computing. 2012:285-293.

[9] Bebarta D K, Biswal B, Dash P K. Comparative study of stock market forecasting using different functional link artificial neural networks[J]. International Journal of Data Analysis Techniques & Strategies, 2012, 4(4):398-427.

[10] Rout M, Majhi B, Mohapatra U M, et al. Stock indices prediction using radial basis function neural network[C]// International Conference on Swarm, Evolutionary, and Memetic Computing. 2012:285-293.

[11] Kennedy J, Eberhart R. Particle swarm optimization [C]//Proceedings of the IEEE International Conference on Neural Networks, Perth, 1995: 1942–1948.

[12] Pulido M, Melin P, Castillo O. Particle swarm optimization of ensemble neural networks with fuzzy aggregation for time series prediction of the Mexican Stock Exchange[J]. Information Sciences, 2014, 280:188-204.

[13] Wang S, Zhang Y, Dong Z, et al. Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection[J]. International Journal of Imaging Systems & Technology, 2015, 25(2):153-164.

[14] Bouacha K, Terrab A. Hard turning behavior improvement using NSGA-II and PSO-NN hybrid model [J]. International Journal of Advanced Manufacturing Technology, 2016:1-20.

[15] Zhao B, Guo C X, Cao Y J. A multiagent-based particle swarm optimization approach for optimal reactive power dispatch"[J]. IEEE Transactions on Power Systems, 2005, 20(2):1070-1078.

[16] Ling S H, Yeung C W, Chan K Y, et al. A new hybrid Particle Swarm Optimization with wavelet theory based mutation operation[C]// Evolutionary Computation, 2007. CEC 2007. IEEE Congress on. 2007:1977-1984.