Expert Systems with Applications 36 (2009) 8849–8854

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network

Zhang Yudong *, Wu Lenan

School of Information Science and Engineering, Southeast University, Room 204, Jianxiong Yuan, Nanjing 210096, China

ARTICLE INFO

Keywords:
Bacterial chemotaxis optimization (BCO)
Stock index prediction
Back propagation neural network (BPNN)

ABSTRACT

The paper proposed an improved bacterial chemotaxis optimization (IBCO), which is then integrated into the back propagation (BP) artificial neural network to develop an efficient forecasting model for prediction of various stock indices. Experiments show its better performance than other methods in learning ability and generalization.

1. Introduction

The financial forecasting or stock market prediction is one of the hottest fields of research lately due to its commercial applications owing to the high stakes and the kinds of attractive benefits that it has to offer (Ritanjali Majhi & Panda, 2007). Unfortunately, stock market is essentially dynamic, non-linear, complicated, nonparametric, and chaotic in nature (Tan, Quek, & Ng, 2005). The time series are multi-stationary, noisy, random, and has frequent structural breaks (Oh & Kim, 2002; Wang, 2003). In addition, stock market's movements are affected by many macro-economical factors (Miao, Chen, & Zhao, 2007; Wang, 2002) such as political events, firms' policies, general economic conditions, commodity price index, bank rate, bank exchange rate, investors' expectations, institutional investors' choices, movements of other stock market, psychology of investors, etc.

Generally there are three schools of thoughts regarding such prediction: (1) the first school believes that no investor can achieve above average trading advantages based on historical and present information, the major theories of which contain Random Walk Hypothesis and Efficient Market Hypothesis (Peters, 1996), which was also vanquished by the persuasive evidence in document (Taylor, 1986); (2) the second view is fundamental analysis. It encourages analysts to undertake in-depth studies into the various macro-economic factors and look into the financial conditions and results of the industry concerned to discover the extent of correlation that may exist with the changes in the stock prices (Ritanjali Majhi & Panda, 2007); (3) the tertiary aspect is recently sprung up to develop models for foreseeing financial data. Most artificial neural network (ANN) based models use historical and present stock index data to predict future prices, and ANN have increasing gained their popularity due to their inherent capabilities to approximate any non-linear function to a high degree of accuracy (Han, 2006).

The progresses of the third school are presented as follows. In the earlier stage most studies were mainly focused on the application of only ANNs for stock market prediction. Kamijo and Tanigawa (1990) used recurrent neural networks and Ahmadi (1990) employed back propagation neural networks with the generalized delta rule to predict the stock market. Yoon and Swales (1991) utilized qualitative and quantitative data to predict. Trippi and Desieno and Choi et al. (1995) predicted daily direction of change in the S&P 500 index futures using ANNs.

However, the ANN especially back propagation (BP) ignored the tremendous noise and non-stationary characteristics in stock market data. Lawrence, Tsoi, and Giles (1996) pointed out that, when the training of a BP tends to be difficult due to the noise of data, then the networks fall into a naive solution such as always predicting the most common output. Miao et al. (2007) indicates that the solutions of the BP usually are forced to the local minimum due to the gradient descent algorithm used to get weights of connection. Engoziner and Tomesn (1995) presented that BP use some variation of the gradient technique, however, it is essentially a local optimizing method and thus has some inevitable drawbacks, such as easily trapping into the local optimal and dissatisfying generalization capability. Sexton, Alidaee, Dorsey, and Johnson (1998) proposed the fact that the gradient descent algorithm may perform poorly even on simple problems when predicting the holdout data.

Thus, Sexton, Dorsey, and Johnson (1998) suggest that, in interest of mitigating the above limitation, weighted values and thresholds of neurons in BP are optimized by global search algorithms. Newly research tends to hybridize several artificial intelligence (AI) techniques to improve the prediction performance. Tsaih, Hsu, andLai,Lai, integrated the rule-based technique and ANNs to predict the direction of change of the S&P 500 stock index futures on a daily basis. Kohara, Ishikawa, Fukuhara, and Nakamura

* Corresponding author.
E-mail address: zhangyudongnuaa@gmail.com (Z. Yudong).

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.11.028

(1997) incorporated prior knowledge, and Gao (2007) incorporated Niche Genetic Algorithm (NGA), and Chen, Chen, and Zhang (2007) united immune programming (IP) and gene expression programming (GEP) to improve the learning process of the conventional ANN. Miao et al. (2007) adopted bacterial colony radial basis function neural network (RBFNN) and Ritanjali Majhi and Panda (2007) utilized bacterial foraging optimization (BFQ) to forecast the stock price. Blanco, Delgado, and Pegalajar (2001) used GA and He, Qiu, and Liu (2005) used tabu search (TS) for optimizing ANN and achieved a certain extent of success.

In this paper the BP neural network incorporated with an improved bacterial chemotaxis optimization (IBCO) is suggested, and its ability is also demonstrated to predict stock index for both short term (next day) and long term (15 days). Simulations exhibits that the proposed approach offers superior performance.

The paper is organized as follows. Next Section 2 presents the basic of BP neural network. Section 3 introduces the fundamental principle of and Section 4 proposes an improvement on the BCO. Section 5 gives the outline of adjusting neuron weights of BP via IBCO, and demonstrates the IBCO based BP (IBCO–BP for brevity) performs better than BP in convergence rate and precision. Section 6 contains experimental data and the formula of technical indicators. Section 7 includes and discusses the simulations results. Finally conclusion is given in Section 8.

2. The error back propagation neural network

The BP, one of the most popular techniques in the filed of NN, is a kind of supervised learning neural network, the principle behind which involves using the steepest gradient descent method to reach any small approximation. A general model of the BP has a structure as depicted in Fig. 1.

Here we can see there are three layers contained in BP: input layer, hidden layer, and output layer. Two nodes of each adjacent layer are directly connected, which is called a link. Each link has a weighted value presenting the relational degree between two nodes (Basma & Kallas, 2004). Assume that there are n input neurons, m hidden neurons, and one output neuron, we can infer a training process described by the following equations to update these weighted values, which can be divided into two steps:

(I) Hidden layer stage: The outputs of all neurons in the hidden layer are calculated by following steps:

$$\mathrm{net}_j = \sum_{i=0}^{n} v_{ij} x_i \quad j = 1, 2, \ldots, m \tag{1}$$

$$y_j = f_H(\mathrm{net}_j) \quad j = 1, 2, \ldots, m \tag{2}$$

Here net_j is the activation value of the jth node, y_j is the output of the hidden layer, and f_H is called the activation function of a node, usually a sigmoid function as follow:

$$f_H(x) = \frac{1}{1 + \exp(-x)} \tag{3}$$

(II) Output stage: The outputs of all neurons in the output layer are given as follows:

$$O = f_O\left(\sum_{j=0}^{m} \omega_{jk} y_j\right) \tag{4}$$

Here f_O is the activation function, usually a line function. All weights are assigned with random values initially, and are modified by the delta rule according to the learning samples traditionally.

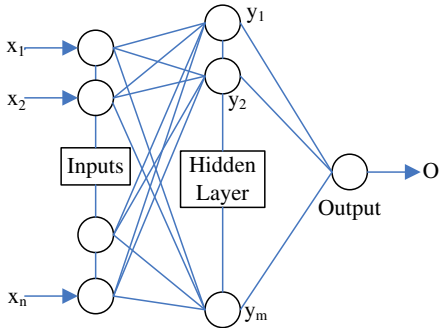


Fig. 1. The architecture of Back Propagation Neural Network.

3. Bacterial chemotaxis optimization

The optimization based on bacterial chemotaxis (BC) (Bremermann, 1974; Sibylle, Jarno, & Stefane, 2002) was inspired from bacterial foraging behavior, pioneered by Bremermann (1974) and his coworkers, and proposed by analogy to the way bacteria react to chemo-attractants in concentration gradients. Miiller, Airaghi, and Marchelo (2000), Sibylle et al. (2002) extracted the BC algorithm from the newest production found in biology field, which was testified to excel other optimization algorithms.

3.1. Description of the 2-D model

Dahlquist, Elwell, and Lovely (1976), Sibylle et al. (2002) modeled the motion of a single bacterium in two dimensions by making the following assumptions.

(1) The path of a bacterium is a sequence of straight-line trajectories.
(2) All trajectories have the same constant speed.
(3) When a bacterium turns, its choice of the new direction, the angle between two successive trajectories, and the duration of a trajectory are all regulated by a probability distribution.
(4) The probability distributions for both the angle and the duration are independent of parameters of the previous trajectory.

3.2. Algorithm steps of BCO

The processing of BCO is presented as follows:

STEP 1: Compute the velocity of a bacterium v, which is assumed to be a scalar constant value 1.
STEP 2: Compute the duration of the trajectory τ, the distribution of which satisfies the exponential probability density function (PDF) as

$$P(X = \tau) = \frac{1}{T} \exp\left(-\frac{\tau}{T}\right) \tag{5}$$

where the expectation value $E(X) = T$ and the variance $Var(X) = T^2$. The time T is given by

$$T = \begin{cases} T_0, & \text{for} \quad \frac{f_{\mathrm{pr}}}{l_{\mathrm{pr}}} \geqslant 0 \\ T_0\left(1 + b\left|\frac{f_{\mathrm{pr}}}{l_{\mathrm{pr}}}\right|\right), & \text{for} \quad \frac{f_{\mathrm{pr}}}{l_{\mathrm{pr}}} < 0 \end{cases} \tag{6}$$

where T_0 presents the minimal mean time; f_{pr} presents the difference between the actual and the previous function value; l_{pr} presents the vector connecting the previous and the actual position in the parameter space; and b is assumed as the dimensionless parameter.

STEP 3: Compute the new direction. The PDF of the angle α between the previous and the new direction is Gaussian and read, for turning right or left, respectively as follows:

$$P(X=\alpha, v=\pm\mu)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(\alpha-v)^2}{2\sigma^2}\right] \tag{7}$$

where, the expectation value $\mu = E(X)$ and variance $\sigma=\sqrt{Var(X)}$ are given by
If $\frac{f_{pr}}{l_{pr}}<0$, then

$$\mu=62^\circ(1-\cos\theta);\sigma=26^\circ(1-\cos\theta);\cos\theta=\exp(-\tau_c\tau_{pr}) \tag{8}$$

Else $\frac{f_{pr}}{l_{pr}}\geqslant 0$, and $\mu = 62^\circ$, $\sigma = 26^\circ$
where τ_c presents the correlation time, and τ_{pr} the duration of the previous step. The choice of a right or left direction as referring to the previous trajectory is determined using a uniform PDF, thereby yielding a PDF for the angle α as

$$P(X=\alpha)=\frac{1}{2}[P(X=\alpha,v=\mu)+P(X=\alpha,v=-\mu)] \tag{9}$$

STEP 4: Compute the new position,

$$x_{new}=x_{old}+n_u l \tag{10}$$

Here, x_{new} presents the new position of the bacteria; x_{old} its previous position; n_u the normalized new direction vector; and l the length of the new trajectory.

In summary, the algorithm contains the following parameters to be computed in advance: T_0, τ_c and b. Document (Sibylle et al., 2002) gives their detailed formula

$$T_0=\varepsilon^{0.30}10^{-1.73};b=T_0\left(T_0^{-1.54}10^{0.60}\right);\tau_c=\left(\frac{b}{T_0}\right)^{0.31}10^{1.16} \tag{11}$$

4. An improved BCO strategy

4.1. Mechanism of the improvement

In the optimization of BCO, every bacterium imparts information each other for ameliorating the foraging environment, nevertheless it will beget the whole colony to fall into local nutrients (Dong, Ajith, & Jae, 2007) and debilitated to cross over noxious substance. As to the algorithm, it means the searching will be trapped into local extrama.

No life-form is living single in realism. Although bacteria are microscopic and primeval, there are by all means correlations between different individuals and colonies. Investigation showed that Escherichia coli aggregate at the foraging process, and disparate colonies exchange the food information while keep certain distance from each other (Passino, 2002). They are more capable of surviving because of the enhancement of comprehension on the settings around.

4.2. Detailed improvements measures

As stated above, improvements are present here, encompassing the following features which can be separated as two stages, as shown in Fig. 2.

(I) Infra-colony Phase (Li, Wang, Zou, & Qian, 2005)
Since every bacterium seizes limited intelligence, and has the capability to regulate its locomotion by information perceived from approximate bacteria, aggregate several bacteria into one colony, which abide following pattern due to the description on the biome community distribute behavior:

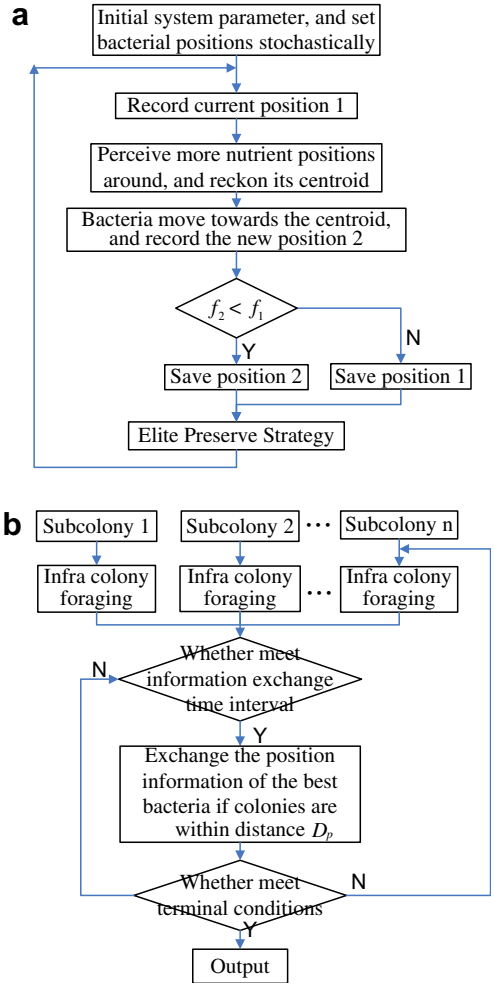


Fig. 2. Two stages of IBCO flow chart. (a) intra-colony foraging (b)inter-colony foraging.

(I–I) Before every new motion, bacteria should perceive surroundings to validate the existence of more nutrient areas. If there are, bacteria are more likely to transfer to the centroid of these areas.

$$\begin{aligned}&\text{Center}(x_{i,k})=\text{Aver}(x_{j,k}|f(x_{j,k})<f(x_{i,k}) \quad \text{and}\\ &\quad \text{dis}(x_{j,k},x_{i,k})<\text{SenseLimit})\\ &\text{Aver}(x_1,x_2,\ldots x_n)=\left(\sum_{i=1}^{n}x_i\right)\Big/n\end{aligned} \tag{12}$$

where k presents motion step, i the bacteria index, and dis($x_{j,k}$,$x_{i,k}$) the distance between bacteria i and j.

(I–II) If a bacterium trends to shift to the center of its accompanies around, then the length of the new trajectory is rand()•dis($x_{i,k}$,center($x_{i,k}$)), where rand() uniformly distribute within the range of (0, 2).While bacteria perceive the chemo-attractants vary little for a session, they will migrate by diverse forms in pursuit of more nutrient nourishment (Passino, 2002). So after continuous n_e steps with absolute differences of functional values are less than given threshold ε_e, the bacteria colony will migrate to a new place. Through the migration, the multiformity of the bacteria colony can be preserved, and the ability of bouncing out of local minima can be fortified.

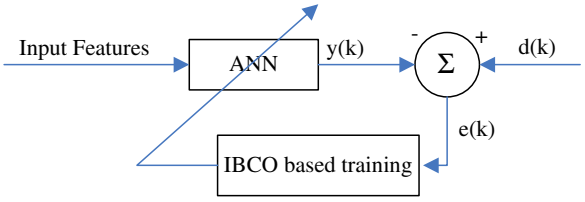
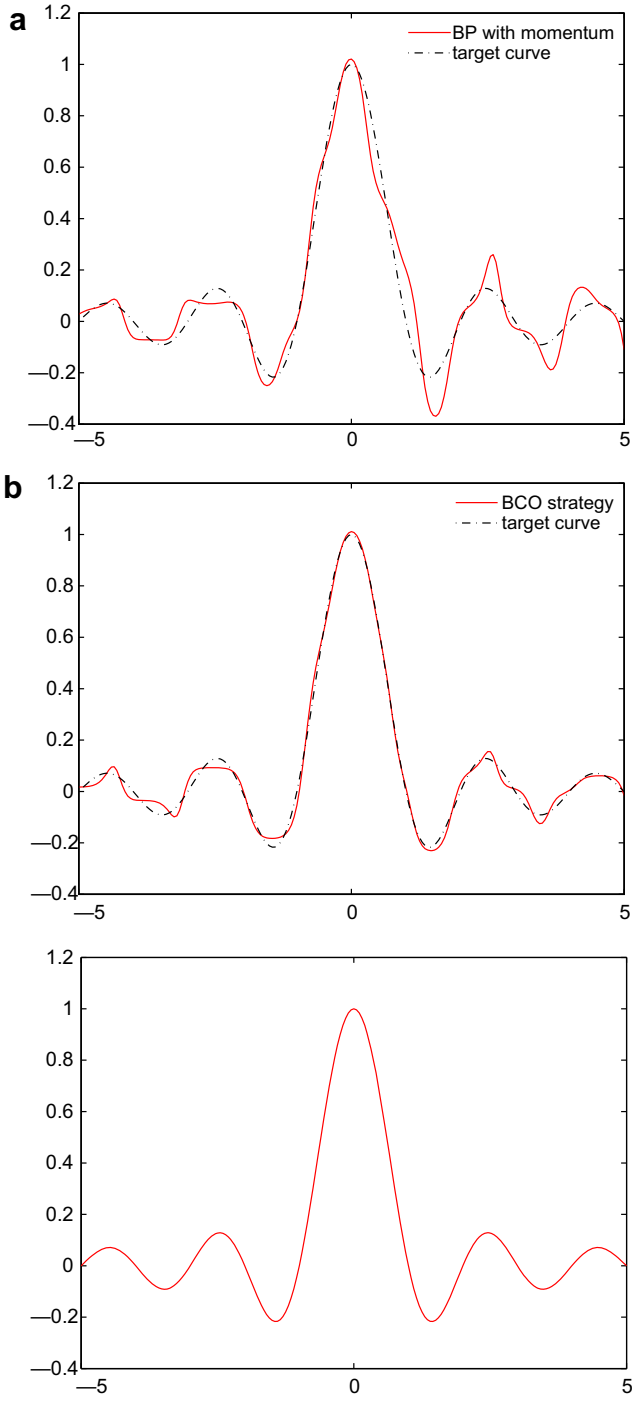


Fig. 3. The structure of IBCO for optimization of ANN weights.



(II) Inter-colony Phase

Divide the whole bacteria colony into several sub-colony, and each casts about for food independently. If two sub-colonies encounter each other with mutual distance less than a given threshold L ($L \gg$ SenseLimit), the sub-colony with poor performance should be punished to migrate. In this way it can be assured that only one sub-colony exists at the round with diameter L, which prevents the whole colony from congregating at one point, and maintains its diversity of augmenting the global search ability of origin algorithm.

5. The IBCO for training of ANNs

5.1. Outline

MSE (mean squared error) is selected as our search objective and is detailed as follow:

$$\mathrm{MSE} = \sum_{i=1}^{N} (\text{true-output})^2 \tag{13}$$

Here true presents the authentic values known to users already, output is the output values of the BP after BCO training, and N presents the number of samples. Our goal is to minimize the MSE through BCO. (See Fig. 3).

5.2. Function generalization of IBCP based BO

In interest of examining the generalization capability of our IBCO technique, a comparison to BP algorithm with momentum was implemented for sinc function approximation ($f(x)$=$sin(x)/x$). In our experiments, network architecture were set to 1-20-1; the transfer function for the hidden layer and the output layer was *tansig* and *purelin*, respectively; error function was set to MSE; the maximum iterative steps were set to 2000; the error goal precision was set to 10^{-6}; and the initial weights are all picked out randomly. The approximation comparison is shown in Fig. 4. It is evident that the approximation effect of IBCO is better than that of other algorithms.

From Fig. 4, it is obvious that our IBCO–BP is sufficient on fitting unknown functions. Considering there is also an unknown relation between stock indices and stock values, so data are gathered to train the IBCO–BP by modifying weights of neurons in order to obtain the function which is accurate enough to express the relation.

6. Simulation experiments

6.1. Experimental data

The data for the stock market prediction experiments has been collected for Standard's & Poor's 500 (S&P 500), USA. The total number of samples for the stock indices is 2350 trading days, from

23rd October 1998 to 27th February 2008. Each sample consists of the closing price, opening price, lowest price, highest price, total volume of stocks traded, and the adjacent close price.

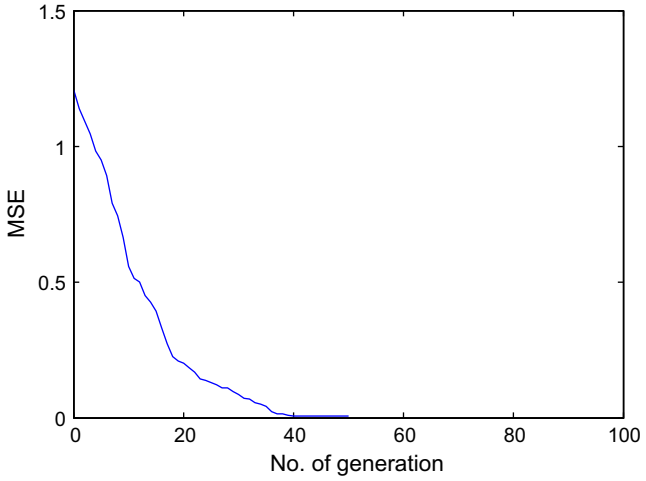


Fig. 5. Learning characteristics of S&P 500 for one day advance using our model.

Among the data, about 20% are used for holdout and 80% for training. Models are used for predicting the close price of the index one day, and fifteen days in advance, which corresponds to short time prediction and long time prediction, respectively.

6.2. Formula of technical indicators

Technique indicators are selected as feature subsets by the review of domain experts and prior research. They are computed from the raw data indicated in Table 1, where *P* presents current price, A smoothing factor, *N* time period, C.P closing price, H.P highest price, and L.P, lowest price.

7. Results and discussion

Fig. 5 shows the learning characteristics of the model obtained through simulation for one day advance for S&P, which indicates that the MSE falls substantially during training and then settles at a minimum value indicating the convergence of weights.

To compare the performance of the proposed model, a best BP model with {9–3–1} neuron is also simulated using the same technical indicators.

Figs. 6 and 7 display the actual versus predicted graph for S&P500 index for one day and fifteen days ahead prediction, respectively.

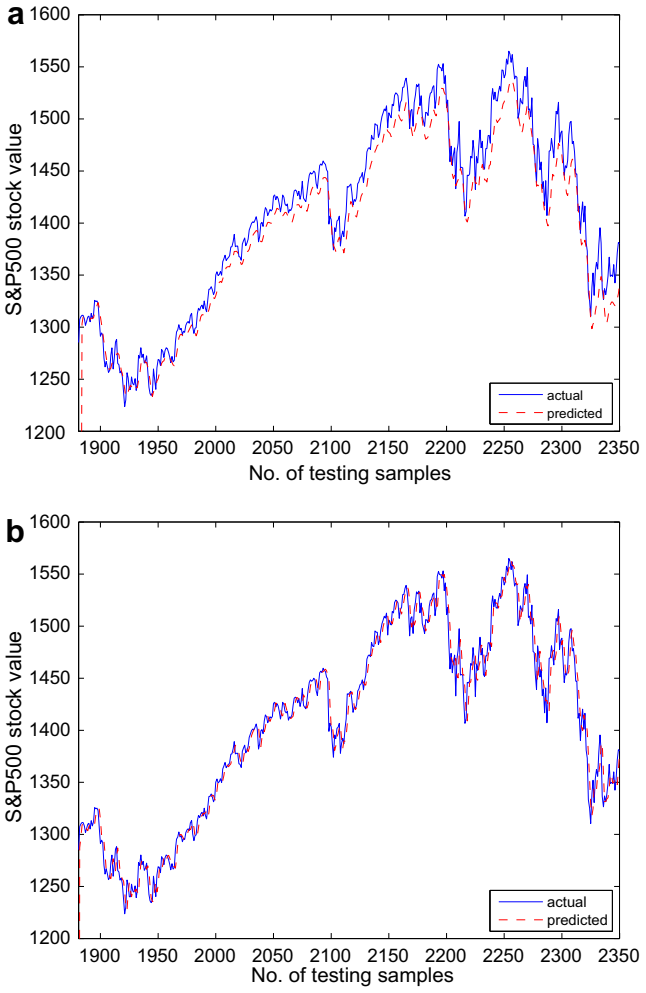


Fig. 6. Comparison of our proposed model and traditional BP model for prediction one day ahead. (a) BP model; (b) IBCO–BP model.

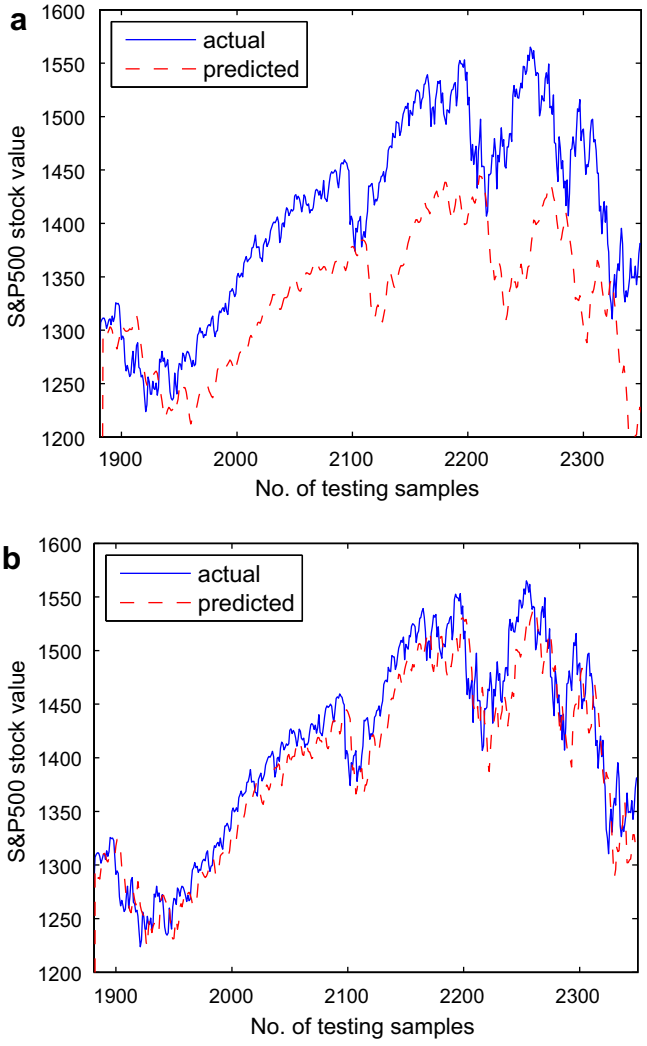


Fig. 7. Comparison of our proposed model and traditional BP model for prediction fifteen days ahead. (a) BP model; (b) IBCO–BP model.

Table 2
Comparison of computation time between IBCO model and BP.

Days ahead	1	3	5	7	9	15
Time of IBCO/min	15.4983	15.6306	15.7734	15.9040	16.1438	16.2503
Time of BP/min	24.3279	24.4821	24.6203	24.7955	24.9186	25.2044

Table 2 shows the computation time of the two models for 1 day, 3, 5, 7, 9, and 15 days in advance prediction for S&P500.

Figs. 6 and 7 indicates that the IBCO–BP model is superior to the traditional BP model. Meanwhile, observations in Table 2 indicate that the novel model takes less time for training compared to its BP counterpart.

8. Conclusion

The IBCO–BP model for prediction of stock indices is developed in this paper. Its structure of the model is basically an adaptive BP ANN whose weights are updated using IBCO tool. Performance comparison with the BP model simulated indicates that the developed model offers less computational complexity, better prediction accuracy, and less training time.

References

Ahmadi, H., (1990). Testability of the arbitrage pricing theory by neural networks. In *Proceedings of the international joint conference on neural networks (IJCNN)*, June (pp. 385–393) Washington DC.

Basma, A. A., & Kallas, N. (2004). Modeling soil collapse by artificial neural networks. *Geotechnical and Geological Engineering, 22*, 427–438.

Blanco, A., Delgado, M., & Pegalajar, M. C. (2001). A real-coded genetic algorithm for training recurrent neural networks[J]. *Neural Networks, 14*, 93–105.

Bremermann, H. J. (1974). Chemotaxis and optimization[J]. *Journal of The Franklin Institute-Engineering and Applied Mathematics*(297), 397–404.

Chen, F., Chen, Y. H., & Zhang, J. Z. (2007). Stock index modeling using IP based Gene Expression Programming. *Computer Engineering and Applications, 43*(26), 227–229 (in Chinese).

Choi, J. H., Lee, M. K., & Rhee, M. W., (1995). Trading S&P 500 stock index futures using a neural network. In R. S. Freedman, editor. *Proceedings of the 3rd annul international conference on artificial intelligence applications on Wall Street*, June (pp. 63–72), New York.

Dahlquist, F. W., Elwell, R. A., & Lovely, P. S. (1976). Studies of bacterial chemotaxis in defined concentration gradients – A model for chemotaxis toward l-serine. *Journal of Supramolecular Structure, 4*, 329(289)–342(302).

Dong, H. K., Ajith, A., & Jae, H. C. (2007). A hybrid genetic algorithm and bacterial foraging approach for global optimization[J]. *Information Sciences*(177), 3918–3937.

Engoziner, S., & Tomesn, E. (1995). An accelerated learning algorithm for multiplayer perception: Optimization layer by layer. *IEEE Transactions on Neural Networks, 6*, 31–42.

Gao, X. J. (2007). Artificial neural network based on Niche Genetic Algorithm applies to stock price prediction. *Market Modernization, 500*, 35 (in Chinese).

Han, L. Q. (2006). *Tutorial of artificial neural network*. Beijing Posts & Telecommunication Press (in Chinese).

He, Y., Qiu, Y., Liu, G., et al. (2005). Optimizing weights of neural network using an adaptive tabu search approach. In *The second international symposium on neural networks 2005 (ISNN-2005)* (pp. 672–676), Chongqing, China.

Kamijo, K., & Tanigawa, T. (1990). Stock price pattern recognition: A recurrent neural network approach. In *Proceedings of the international joint conference on neural networks (IJCNN)*, June (pp. 215–221), Washington DC.

Kohara, K., Ishikawa, T., Fukuhara, Y., & Nakamura, Y. (1997). Stock price prediction using prior knowledge and neural networks. *Int J Intelligence System Accounting Finance Manage, 6*(1), 11–12.

Lawrence, S., Tsoi, A. C., & Giles, C. L. (1996). Noisy time series prediction using symbolic representation and recurrent neural network grammatical inference. Technical report UMIACS-TR-96-27 and CS-TR-3625. Institute for Advanced Computer Studies, University of Maryland.

Li, W. W., Wang, H., Zou, Z. J., & Qian, J. X. (2005). Function optimization method based on bacterial colony chemotaxis[J]. *Journal of Circuits and Systems, 10*(1), 58–63 (in Chinese).

Miao, K., Chen, F., & Zhao, Z. G. (2007). Stock price forecast based on bacterial colony RBF neural network. *Journal of QingDao University, 20*(2), 50–54 (in Chinese).

Miiller, S., Airaghi, S., & Marchelo, J., et al. (2000). Optimization algorithms based on a model of bacterial Chemotaxis[C]. In Animals to animats 6: *Proceedings of the 6th international conference on simulation of adaptive behavior* (pp. 375–384). MIT Press.

Oh, K. J., & Kim, K.-J. (2002). Analyzing stock market tick data using piecewise non linear model. *Expert System with Applications, 22*(3), 249–255.

Passino, K. M. (2002). Biomimicry of bacterial foraging for distributed optimization and control[J]. *IEEE Control System Magazine, 22*(3), 52–67.

Peters, E. (1996). *Chaos & order in capital markets: A new view of cycles, prices, market volatility*. John Wiley & Sons Inc.

Ritanjali Majhi, & Panda, G., (2007). Stock market prediction of S&P 500 and DJIA using bacterial foraging optimization technique. In *2007 IEEE congress on evolutionary computation* (CEC 2007) (pp. 2569–2579).

Sexton, R. S., Alidaee, B., Dorsey, R. E., & Johnson, J. D. (1998). Global optimization for artificial neural networks: A taboo search application. *European Journal of Operational Research, 106*(2–3), 570–584.

Sexton, R. S., Dorsey, R. E., & Johnson, J. D. (1998). Toward global optimization of neural networks: a comparison of the genetic algorithm and back-propagation. *Decision Support Systems, 22*(2), 171–185.

Sibylle, D. M., Jarno, M., Stefane, A., et al. (2002). Optimization based on bacterial chemotaxis[J]. *IEEE Transactions on Evolutionary Computation, 6*(1), 17–19.

Tan, T. Z., Quek, C., & Ng, G. S. (2005). Brain inspired genetic complimentary learning for stock market prediction. In *IEEE congress on evolutionary computation*, 2–5th September (Vol. 3, pp. 2653–2660).

Taylor, S. (1986). *Modeling financial time series*. John Wiley & Sons, Inc.

Trippi, R. R., & Desieno, D. (1992). Trading equity index futures with a neural network. *Journal of Portfolio Management, 19*, 27–33.

Tsaih, R., Hsu, Y., & Lai, C. C. (1998). Forecasting S&P 500 stock index futures with a hybrid AI system. *Decis Support System, 23*(2), 161–174.

Wang, Y. (2002). Predicting stock price using fuzzy grey prediction system. *Expert System with Applications, 22*, 33–39.

Wang, Y. (2003). Mining stock prices using fuzzy rough set system. *Expert System with Applications, 24*(1), 13–23.

Yoon, Y., & Swales, G., (1991). Predicting stock price performance: a neural network approach. In *Proceedings of the 24th annual Hawaii international conference on system sciences (HICSS-24)*, January (pp. 156–162), Kauai, Hawaii.