Prediction of Stock Price Trend Based on Wavelet Neural Network and RS Attributes Reduction

Wei Yanming

XiJing College
Xi'an, Shaanxi
710123, China
fyw2121@sohu.com

Lou Yuanwei

Air Traffic Control and Navigation College
Air Force Engineering University
Xi'an, Shaanxi, 710051, China
911yhk@163.com

Lei Lei

School of Business Administration
Henan University of Economics and Law
Zhengzhou, Henan, 050046, China
lyx521@yahoo.com.cn

Abstract—**For the prediction problem of stock price, a Wavelet Neural Network (WNN) method based on Rough Set (RS) attribute reduction is proposed. First RS attribute reduction is applied to reduce the dimensions of feature index for stock price trend, then based on RS attribute reduction the structure of WNN is optimized to establish the prediction model of stock price trend on the basis of feature index reduction, finally the built model is applied to predict the stock price trend. The simulation results indicate that, by introducing the RS attributes reduction, the structure of WNN model can be simplified to a great extent for stock price trend with improvement of the performance. The direction symmetry of prediction corresponding to SSE Composite Index is 65.75% with 1.7s training time. The prediction result is better than that of other neural network and WNN models. This verifies the feasibility and effectiveness of the method in the prediction of stock price trend.**

Keywords—Wavelet Neural Network; Rough Set; Stock Price; Prediction

I. INTRODUCTION

The purpose of stock price prediction is to explore the development law of stock market so as to provide a scientific basis for stock investment [1][2]. Due to a lot of factors affecting the fluctuation of stock price, the impact of the uncertainty change of these factors on stock market cannot be accurately grasped. Therefore, it is very difficult to the accurately predict the stock price.

The traditional methods of stock price prediction mainly have the time series method [3], grey method [4] and prosperity method [5] etc., and have been widely used in some large stock market. However, the stock market has a great many influence factors and the uncertain interaction between these influence factors is very complicated, which can result in large prediction deviations. The stock rice prediction technique based on neural network have been made great progress [6~10]. However, many problems in accuracy, convergence or local minimum etc. are also exposed to a varying extent, and are not solved by the root despite having some improvements,

which affect its application and populari- zation. Wavelet Neural Network (WNN) is a new neural network developed on the basis of wavelet analysis. Compared with the traditional neural network, WNN has a great performance improvement with strict theory bases and good mapping capability, suitable for all areas of neural network application [11]. However, to really use WNN to predict the stock price, many problems need to be solved such as feature extraction, parameters selection, and structure optimization, etc.. The research will focus on feature extraction and structure optimization for a good WNN prediction model.

Based on the analysis, a WNN prediction method based on Rough Set (RS) [12] is proposed for stock price trend. In the method, RS plays 2 roles in the method: 1. Downsize WNN to release the training burden by reducing attribute, removing noise and eliminating redundant object to the sample data; 2. Reduce the redundant nodes in hidden layer to determine the relatively optimal WNN structure. The feasibility of this method is proved by experiment simulation.

II. WAVELET NEURAL NETWORK

Basic structure of WNN is shown in Fig. 1 where $x_i(i=1,\cdots,I)$ is the input; $y_k(k=1,\cdots,K)$ is the output; $w_{ji}(j=1,\cdots,J;\ i=1,\cdots,I)$ is the connection weight between hidden and input nodes; $w_{kj}(k=1,\cdots,K;\ j=1,\cdots,J)$ is the connection weight between output and hidden nodes. I is the number of input nodes; J is the number of hidden nodes; K is the number of output nodes. $a_j,b_j(j=1,\cdots,J)$ is the dilation factor and translation factor respectively; $\varphi(\cdot)$ is wavelet base function that satisfied the framework condition, and in the paper is selected as Morlet wavelet function $\varphi(x)=\cos(1.75x)\exp(-x^2/2)$. As a matter of fact, the

training of WNN can be considered as the process of the optimal estimation to the parameter: w_{ji}, w_{kj}, a_j, b_j.

The output of the network in Fig. 1 can be given as

$$y_k(x) = \sum_{j=1}^{J} w_{kj}\varphi(\frac{X_j - b_j}{a_j}) \qquad (1)$$

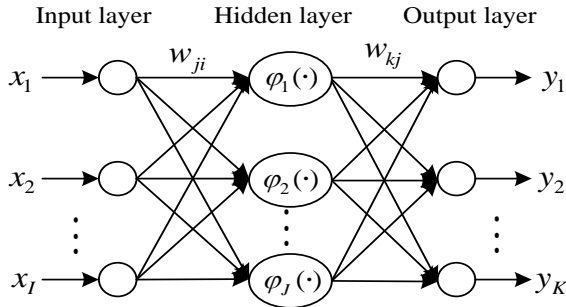where $X_J = \sum_{i=1}^{I} w_{ji}x_i$, $(j = 1, \cdots, J)$.

Fig. 1. Basic structure of WNN

III. REALIZATION BASED ON RS AND WNN

For the WNN modeling, when the input dimension is large, the network structure is very complex, resulting in long training time and low accuracy. Especially, as for the more complex system, the sample data collected from it not only have large dimension, but also contain some redundant variables, which may result in an unsatisfactory result for WNN modeling. It can be seen from the studies that, RS can not only remove the redundant information, but also simplify the dimensions of the input space by analyzing the internal relation between the data, providing a good method for solving the above problem. Therefore, by learning from each other, WNN is integrated with RS to build a new WNN, which can reduce the complexity of WNN structure with the performance improvement in theory.

Network structure design is a difficult problem in WNN modeling. It can generally be solved by top-down method, i.e. first construct a network that contains enough hidden layer nodes, and then it can be reduced by experience. The network obtained by the way is redundant, time-consuming, less stable with depending on personal experience, having the improvement room. Research shows that, in WNN, the contribution of some hidden layer nodes to the error is smaller, which has the less impact on the network; conversely, some hidden layer nodes have larger contribution to the error, which can cause significant fluctuation to the network. To obtain the relatively optimal WNN structure with performance improvement, it is required to retain the hidden layer nodes with larger contribution to network error, and to remove the hidden layer nodes with less contribution. Considering the excellent knowledge reduction ability of RS, the study using it to remove the redundant structure of WNN for reasonable network.

Based on the above analysis, the main idea of integration method based on RS and WNN is to use the attribute reduction

capability of RS to reduce the complexity of WNN modeling and structure. The integration principle is shown in Fig. 2. RS mainly plays the role in 2 aspects: first reduce the input dimension of WNN, namely, the input variables are taken as the condition attributes, the output variables as the decision attributes, constructing the decision table, and conduct the attribute reduction to eliminate the redundant variables for decreasing the input dimension of WNN; second optimize the network structure of WNN, namely, the outputs of each hidden layer node are taken as the condition attributes, the output error squares of each output layer node as the decision attributes, constructing the decision table, and conduct the attribute reduction to prune the corresponding hidden nodes and their connection weights for getting the optimal WNN topology structure.

According to the above elaboration, the realization steps of integration method based on RS and WNN are as follows:

1. The input variable $x_1, x_2, \cdots, x_I$ are taken as the condition attributes, the output variable $y_1, y_2, \cdots, y_K$ as the decision attributes, constructing the initial decision table of dimension reduction for input variables.

2. The condition attributes and decision attributes in decision table are discretized and the attribute reduction is conducted.

3. The number of input layer nodes of WNN can be determined by the attribute reduction result obtained in Step 2; the number of output layer nodes can be determined by the practical problems to be solved; the number of hidden layer nodes can be preliminarily pre-determined by empirical formula $J \geq \sqrt{I + K} + a$, a is a positive integer in [1, 10].

4. WNN model is trained and the output $h_1, h_2, \cdots, h_J$ of each hidden layer node and the squared output error $s_1, s_2, \cdots, s_K$ of each output layer node are recorded for each sample, constructing the initial decision table of structure optimization for WNN.

5. The decision table gotten by Step 4 is discretized and the attribute reduction is conducted.

6. In accordance with the decision table reduced by Step 5, the topology structure of WNN is pruned to delete the hidden layer nodes and their connection weights corresponding to the redundant attributes, simplifying the structure of WNN model.

RS analysis is mostly used to process the discrete data. However, for many practical problems, the value range of some condition attributes and decision attributes is usually continuous value, needing to be discretized before RS analysis. This is an important problem in the application of RS. Existing discretization methods mostly rely on subjective experience. Therefore, based on information entropy of decision table, a heuristic discretization algorithm [13] is introduced for the continuous attribute.

Information entropy can be defined as

$$H(X) = -\sum_{j=1}^{r(d)} p_j \log_2 p_j, \quad p_j = k_j / |X| \quad (2)$$

where $|X|$ is the examples number, the examples number of the decision attribute $j(j = 1, 2, \cdots, r(d))$ is k_j.

Input: data X

Output: discretization value and division points of continuous attributes in X

For (each continuous attribute) {
   For k non-repeated attribute value, calculate the probability p_k and entropy $H(p_k)$ by formula (2), and save the division points;
   Suppose that $Symbol = 1$, $k_0 = k$, $T_k = 0$;

While ($Symbol$) {

   Select 2 adjacent intervals and combine to minimize the difference of the entropy before and after combination, and reset the division points, save the entropy after combination;

   Calculate: $T_{k-1} = k_0 H(p_{k-1}) - H(p_k)(k-1)$

   If ($H_{k-1} > H_k$)

   { Recalculate the interval probability }
   else
   { $Symbol = 0$, $k = k - 1$, save the division points }

   }}

Output the continuous value and division points of continuous attribute.

IV. EXPERIMENT SIMULATION

A. Selection of Experimental Data

To verify the performance of proposed method, all experimental data used is from a famous stock indices: SSE Composite Index (China). The original data is in the range of 4/10/2009~6/24/2014 from SSE Composite Index. The original data include: opening price, top price, bottom price, closing price and trading volume. To facilitate the modeling and testing, the original data need to be transformed into 15 index feature quantities as shown in Table 15 in advance by calculation formula, and these index feature quantities are used as the inputs of the prediction model to be established. In addition, the prediction output of the model is the change trend of stock price. According to the percent yield, its change trends can be defined as: Falling: $R \in (-\infty, -0.5\%]$, Invariability: $R \in (-0.5\%, 0.5\%)$ and Climbing: $R \in [0.5\%, \infty)$, R denotes the percent yield that can be calculated by

$$R_t = (P_t - P_{t-1}) / P_{t-1} \times 100\% \quad (3)$$

where P_t is the series of stock price.

In addition, because the different index features usually have the different value ranges, the normalization method needs to be used to standardize the sample data before modeling. The value a of the attribute A can be mapped into the normalization interval (L, H) by the following formula

$$a' = L + (H - L) \cdot \frac{a - a_{\min}}{a_{\max} - a_{\min}} \quad (4)$$

where $a_{\min}$, $a_{\max}$ is the minimum and maximum of the attribute A respectively. The normalization interval $(-0.5, 0.5)$ is set to be uniform in the paper.

B. Evaluation Index of Model Performance

To accurately evaluate the prediction model of stock price trend, the reasonable evaluation indices needs to be selected in advance. Through the analysis and study on relevant literature, the four evaluation indices are determined to evaluate the performance of built models including: Root Mean Square Error (RMSE), Mean Absolute Difference (MAD), Mean Absolute Percent- age Error (MAPE); Directional Symmetry (DS). RMSE, MAD and MAPE are used to measure the deviation between actual value and predicted value. DS is applied to measure the accuracy of the prediction. DS is used to describe the accuracy of stock index direction prediction in percentage form. The calculation formulas of the six performance evaluation indexes are as follows:

$$RSME = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - T_i)^2} \quad (5)$$

$$MAD = \frac{1}{n} \sum_{i=1}^{n} |P_i - T_i| \quad (6)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} |\frac{P_i - T_i}{P_i}| \times 100\% \quad (7)$$

$$DS = \frac{100}{n} \sum_{i=1}^{n} d_i, \quad (8)$$

$$d_i = \begin{cases} 1 & (P_i - P_{i-1})(T_i - T_{i-1}) \geq 0 \\ 0 & Other \end{cases}$$

where T is the actual value, P is the predicted value, n is the total number of samples.

C. Model Demonstration and Result Analysis

Experimental scheme: Analyze the effect of training set size on the performance of 2RS-WNN prediction model for stock price trend. First 100, 300, 500, 600, 800, 1000 samples are randomly selected from the data of SSE Composite Index to constitute the training set respectively. At the same time, the equivalent of 20% training set size are randomly selected from the remaining data to construct the corresponding test set. Then 2RS-WNN prediction model of stock price trend is established for each training set and be tested for analyzing the effect of training set size.

Parameter settings: w_{ji}, w_{kj}, a_j, b_j can be initialized as the random number in $[-0.5, 0.5]$, the maximum training number can be taken as 500. In the experiment, except for 2RS-WNN.

Fig. 2 shows the structure of 2RS-WNN prediction model trained by 500 training samples for stock price trend of SSE Composite Index, where 1, 2,5,7,10,11,13,15 are the index feature quantities obtained by RS attribute reduction. ①,②,③,⑤,⑥,⑨ denotes the hidden layer nodes obtained by RS attribute reduction. The prediction results and performance comparison of 2RS-WNN model obtained by different size training set are shown in Table 3 for stock price trend of SSE Composite Index.
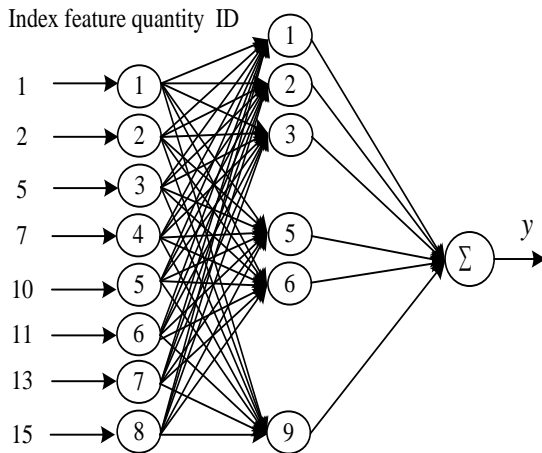

Fig. 2. Structure of 2RS-WNN prediction model for stock price trend

TABLE I.   PREDICTION RESULTS AND PERFORMANCE COMPARISON 2RS-WNN MODEL UNDER DIFFERENT TRAINING SET SIZE

Index	Samples numbers					
	100	300	500	600	800	1000
Feature obtained by RS	1, 2, 5, 11, 13	1, 2, 5, 7, 11, 13, 15	2, 5, 7, 10, 11, 13, 15	1, 2, 5, 7, 11, 12, 15	1, 2, 5, 7, 10, 11, 12, 13, 15	1, 2, 4, 5, 10, 11, 12, 13, 15
Model structure	5-3-1	7-6-1	8-6-1	7-9-1	9-8-1	9-9-1
Training time (s)	0.5	1.1	1.7	3.8	4.9	6.2
RMSE	97.05	83.62	68.50	55.67	50.01	48.24
MAD	84.11	46.35	27.18	24.52	24.19	23.05
MAPE	2.54	1.97	1.35	1.32	1.28	1.17
DS%	60.36	63.83	65.74	65.96	66.15	66.32

It can be seen from Table 1 that, for stock price trend of SSE Composite Index, with the increase of training set size, the performance of 2RS-WNN prediction model is improved in different degree, however, the improvement of DS index is not significant when the number of training samples is more than 500. DS value corresponding to the training set with 500, 600, 800, 1000 samples is 65.74%, 65.96%, 66.15%, 66.32%, respectively. Considering the balance between prediction accuracy and training time, it is more appropriate to select the training set with 500 samples for testing and comparing the model.

V.   CONCLUSION

RS has the excellent knowledge reduction ability. WNN is good at nonlinear system modeling. Combining with the both, it can reduce the input dimension and optimize the network topology structure for the modeling and prediction of WNN, decreasing the calculation complexity with the improvement of stock price trend prediction. Simulations show that the proposed prediction method reduces the size of the network, simplifies the network structure, and has a great improvement in contrast with simple WNN in training time and prediction accuracy. Thereby, it provides an effective solution for the prediction of stock price trend.

REFERENCES

[1]   Y. X. Chen, "Stock Investment," Chengdu: Sichuan University Press, 2005.

[2]   R. D. Edwards, J. Magee, W. H. C. Bassetti, "Technical Analysis of Stock Trends," Boca Raton: CRC Press, 2007.

[3]   Y. Y. Win, "Time series data mining method based on ARIMA model improvements," Journal of Harbin University of Commerce, vol. 30, No. 6, 2014, pp. 675-676, 681.

[4]   M. A. Jin, L. C. Cheng, "Grey forecasting model for the price of the stock," Journal of Yanbian University, vol. 40, No. 1, 2014, pp. 54-57.

[5]   J. Xiao, Y. Fan, B. Zhu, "Applied research on business forecast model," Soft Science, vol. 20, No. 4, 2014, pp. 12-15.

[6]   T. Kimoto, K. Asakawa, M. Yoda, et al, "Stock market prediction system with modular neural networks," Proceedings of 1990 International Joint Conference on Neural Networks, vol. 1, 1990, pp. 1－6.

[7]   F. Oliveira, C. N. Nobre, L. E. Zarate, "Applying artificial neural networks to prediction of stock price and improvement of the directional prediction index-case study of PETR4," Expert System with Application, vol. 40, No. 18, 2013, pp. 7596-7606.

[8]   E. Hadavandi, H. Shavandi, A. Ghanbari, "Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting," Knowledge-Based Systems, vol. 23, No. 8, 2013, pp. 800-808.

[9]   Z. J. Yu, S. L. Yang, Z. Zhang, J. Jiao, "Stock returns prediction based on error-correction grey neural network," Chinese Journal of Management Science, vol. 23, No. 12, 2015, pp. 20-26.

[10]   A. L. Leonel, A.S. F. Ricardo, G. L. Guilherme, "Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks," Applied Soft Computing, vol. 35, 2015, pp. 66-74.

[11] Q. H. Zhang, A. Benveniste, "Wavelet network," IEEE Transactions on Neural Networks, vol. 3, No. 6, 1992, pp. 889-898.

[12] L. Zhu, J. M. He, S. Chang, "Rough set integrated neural network model for forecasting stock price," Chinese Journal of Management Science, vol. 10, No. 4, 2002, pp. 7-12.

[13]   H. Xie, H. Z. Chang, D. X. Niu, "Discretization of continuous attributes in rough set theory based on information entropy," Chinese Journal of Computers, vol. 28, No. 9, 2005, pp. 1570-1574.