

# *WNN Prediction Model of Stock Price with Input Dimensions Reduced by Rough Set*

Huang Haiqing

XiJing College  
Xi'an, Shaanxi  
710123, China  
ganxusheng123@163.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

Li Huaping

Air Force Xi'an Flight Academy  
Xi'an, Shaanxi  
710306, China  
gxsh521@163.com

**Abstract**—To improve the prediction ability of stock price, an integration prediction method based on Rough Set (RS) and Wavelet Neural Network (WNN) is proposed. First RS is used to reduce the dimensions of feature of stock price, then the WNN prediction model is established for stock price movement on the basis of feature dimension reduction; finally, the built model is applied to predict the stock price movement. The simulations on daily closing price index of SSE Composite Index indicate that, the proposed method has advantages of simple structure, strong implementation and good prediction accuracy with average correct rate 64%, and gets better stock price prediction in contrast with single neural network, genetic neural network and WNN.

**Keywords**—Wavelet Neural Network; Rough Set; Attribute Reduction; Stock Price; Prediction

## I. INTRODUCTION

The stock prediction is that, on the basis of accurate statistical investigation data and stock market information, from the history, current situation and regularity of stock market, the science methods are used to predict the future development prospects of stock market. With the development of national economy, the people's understanding of stock market has gradually deepen, and many prediction methods of stock price appeared one after another such as time series method, grey method, prosperity method and trend extrapolation method and so on. However, as a complex giant system with many influence factors and various uncertainty interaction, the price fluctuation of stock market often show the strong nonlinearity. To get an accurate prediction, we need to ensure that the selected prediction method can deal with the huge amount of information with a certain induction reasoning ability. This is the main reason that the traditional methods have poor prediction effects on stock market.

In 1988, neural network was first used to predict the daily return rate of IBM ordinary shares by White et al, but the results are unsatisfactory [1]. In 1990, based on neural network,

TPOIX prediction system was developed by Kimoo et al to predict the price-weighted average index of Tokyo Stock Exchange, it indicate that the return rate prediction of neural network is better than that of weighted average index [2]. In 1992, neural network was applied to predict the climbing and falling movement of Japanese stock prices by Baba et al, proposing a new prediction direction for stock price, however, the trend judgment of stock price has great influence on the prediction ability of built model [3]. In 1999, neural network was utilized to carry out the prediction of UK London stock index for the past 25 years by Pesaran et al, reaching close to 60% accuracy rate [4]. In 2001, Kwong used neural network to predict the stock exchange data of 7 listed companies in Australia, achieving an average accuracy 48.2% [5]. In addition, Dennis [6], Philp [7], Leonel [8], et al. has made certain achievement in the research. Altogether, the neural network has excellent characteristics such as self-organizing, self-adapting, self-learning, distributed processing, strong fault tolerance and so on, which provides a new mode for stock price prediction and changes the existing mode completely with many achievements. Nevertheless, in the application of neural network for stock price prediction, exposing some shortcomings, such as low accuracy, over-fitting, falling into local minimum easily and so on, is the fact that cannot be ignored, so that it is still difficult to obtain generally accepted prediction for stock price. Compared with the traditional neural network, Wavelet Neural Network (WNN) has a great performance improvement, suitable for all areas of neural network application [9][10]. However, to really use WNN to predict the stock price, some problems such as feature reduction need to be solved.

Based on this, a WNN prediction method of stock price based on Rough Set (RS) is proposed. RS plays the roles in the method: Conduct the attribute reduction for stock price feature to reduce the input dimensions of WNN, which can meet the accuracy and practicability of WNN prediction for stock price.

## II. RS ATTRIBUTE REDUCTION

In RS theory, the knowledge can be regarded as the division with respect with the universal, and the roughness of knowledge can be defined from the angle of indiscernibility relation, realizing the rigorous analysis and processing of knowledge based on the classification mechanism. As the basis of knowledge discovery, attribute reduction can reduce the overall attribute number under the premise of keeping the same classification ability, and has become the core research contents of RS theory [11].

Minimum attribute reduction is to obtain the reduction containing the least attribute number in all reduction under the same classification ability. It has already been proven that searching all reduction and minimum reduction is a NP hard problem, so some heuristic attribute reduction algorithms are proposed, and most of them realize the reduction by constructing the relationship between knowledge and information entropy. In these algorithms, the compatible class obtained by compatibility relation division does not generally constitute the partition to the universal, but constitutes the cover to the universal. That is, the knowledge is regarded as the cover with respect to the universal, which is a lack of rationality obviously. In the paper, on the basis of hierarchy structure and approximation accuracy concept, a heuristic attribute reduction recursive algorithm is introduced. By introducing the heuristic information different from information entropy, the search space of the problem can be narrowed to get the optimal and near-optimal solution.

For the decision information system  $S=(U,R,V,f)$ ,  $U = \{x_1, x_2, \dots, x_n\}$  denotes the object set (universal);  $R = C \cup D = \{r_1, r_2, \dots, r_m\}$  denotes the attribute set,  $C$  and  $D$  is the condition attribute subset and the decision attribute subset respectively;  $V = \bigcup_{r \in R} V_r$  ( $V_r$  is the value range of the attribute  $r \in R$ ),  $f$  can be thought of as the information function between  $U$  and  $R$  ( $f:U \times R \rightarrow V$ ,  $f(x,r) \in V_r$ ). Suppose that  $X$  is a subset of  $U$  ( $X \subset U$ ). For each  $X$  and the indiscernibility relation  $B(B \subseteq A)$ , and the lower approximation set  $B_*(X)$ : according to the knowledge  $B$  determine the set composed of the objects in  $U$  that must belong to  $X$  ( $B_*(X) = \{X \in U : B(X) \subseteq X\}$ ). The upper approximation set  $B^*(X)$ : according to the knowledge  $B$  determine the set composed of the objects in  $U$  that may belong to  $X$  ( $B^*(X) = \{X \in U : B(X) \cap X \neq \emptyset\}$ ).

For the condition attribute, if the decision value of the objects in indiscernibility relation set is the same, then the decision system is compatible; otherwise, it is incompatible. For this attribute reduction algorithm, first the equivalence relation of all condition attribute need to be determined in compatible decision information system, then the upper and

lower approximations set with respect to the decision value are calculated for each equivalence relation. The heuristic parameter can be gotten by the following formula

$$\lambda = (|B^*(X)| - |B_*(X)|) / |U| \quad (1)$$

Retain the attribute with the smallest  $\lambda$ . For the given universal, Lower approximate value can be used as the basis for pruning

$$U = U - POS_{red} \{d\} \quad (2)$$

where  $POS_{red} \{d\}$  is the object set of the decision attribute  $d$  that the conditional attribute of the reduction set  $red$  can be subsumed into the universal  $U$ . Repeat the process, and select the attributes that need to be retained, the remaining attributes are combined with the attributes in the reduction set  $red$  to generate a new equivalence relation. The algorithm is recursive, and the calculation is terminated until the given universal is empty [12].

For the system  $DS=(U,C \cup D, V, f)$ , the attribute reduction algorithm is as follows:

1. Initialize the reduction set  $red$  as  $\emptyset$  (empty set),  $C$  as all condition attributes.
2. Calculate  $\lambda$  of each condition attribute in  $C$  using the formula (1).
3. Sort  $\lambda$  in descending order, and select the attribute with the smallest  $\lambda$ . If two attributes have the same  $\lambda$ , the attribute with the least attribute value can be selected. If  $\lambda$  is the same and the attribute value is the same, an attribute can be selected at random.
4. Put the selected attribute  $a_i$  into the reduction set  $red$ , and reduce the attribute  $a_i$  from  $C$ , conducting the formula (2).
5. If  $U$  is  $\emptyset$ , go to step 7, otherwise, go to step 6.
6. Combine the attributes in the reduction set  $red$  and remaining attributes in  $C$ , repeat steps 2~5.
7. Output the reduction set  $red$ , and terminate the reduction calculation.

## III. WAVELET NEURAL NETWORK BASED ON RS ATTRIBUTE REDUCTION

The structure of WNN is shown in Fig. 1 where  $x_i (i = 1, \dots, I)$  is the input,  $y_k (k = 1, \dots, K)$  is the output,  $a_j, b_j (j = 1, \dots, H)$  is the dilation factor and translation factor respectively,  $w_{ji}$  is the connective weight between hidden

and input layers, and  $w_{kj}$  is the connective weight between output and hidden layers.

According to the RS performance, the realization steps of WNN method based on RS attribute reduction are as follows:

1. The input variable  $x_i$  are taken as the condition attributes, the output variable  $y_k$  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 the empirical formula

$$J \geq \sqrt{I + K} + a \quad \text{or} \quad J \geq \log_2 I \quad (3)$$

where  $a$  is a positive integer in [ 1, 10].

4. BP algorithm is used to train WNN and test the model.

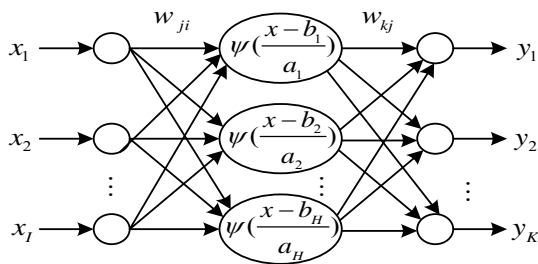


Fig. 1 Structure chart of WNN

IV. EXPERIMENT SIMULATION

SSE Composite Index daily closing index (China Shanghai stock market) can be selected as the research object. Research goal is to predict the following day variation movement T of SSE Composite Index. According to the size of the percent yield, 5 variation movements can be defined as

- Plunging (T = 1)  $R \in (-\infty, -2.0\%]$
- Falling (T = 2)  $R \in (-2.0\%, -0.5\%]$
- Invariability (T = 3)  $R \in (-0.5\%, 0.5\%]$
- Climbing (T = 4)  $R \in [0.5\%, 2.0\%]$
- Soaring (T = 5)  $R \in [2.0\%, \infty)$

where  $R$  denotes the percent yield, and can be defined by the following formula

$$R_t = (P_t - P_{t-1}) / P_{t-1} \times 100\% \quad (4)$$

where  $P_t$  is the price series.

The statistical results of 15 index feature are shown in

Table 1 that is mainly used to determine the state number of each index feature in the discretization process.

The specific prediction steps are as follows:

1. For SSE Composite Index closing index, the data on its fluctuation trend and corresponding feature in 2003 is separate into 7 sample sets on a yearly basis. Each sample set is divided into the training sample set and test sample set with 4:1 ratio, as shown in Table 1. 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 is 500.

2. Considering the training sample set in 2003, 15 index features in SSE Composite Index daily closing index can be taken as the condition attributes, and the following day variation movement T of SSE Composite Index can be taken as the decision attribute, establishing the initial decision table of dimension reduction for WNN input variables, as shown in Table 2, and the proposed discretization algorithm is used to automatically discretize the condition attributes in the decision table, as shown in Table 3.

TABLE I. STATISTICAL RESULT OF EACH FEATURE

ID	Index feature	Maximum	Minimum	Average	S.D.
(1)	Random %K	97.201	5.816	52.334	24.021
(2)	Random %D	93.160	9.308	52.302	20.733
(3)	W&R	100.000	0.000	47.620	33.419
(4)	MTM	318.570	-339.320	3.003	49.335
(5)	RSI	100.000	0.000	49.915	28.773
(6)	BBI	2081.753	376.727	1087.394	415.682
(7)	5BIAS	29.020	-16.731	0.102	3.172
(8)	10BIAS	53.285	-21.516	0.246	5.144
(9)	MB	128.827	-108.958	1.119	17.939
(10)	PSY	90.000	10.000	51.117	16.286
(11)	AR	473.016	10.875	96.524	56.412
(12)	BR	378.016	12.213	103.406	54.956
(13)	MV	363.553	1.393	44.394	42.289
(14)	TAPI	30422.995	213.594	4493.407	3567.456
(15)	VR	100.000	0.000	50.232	23.172

TABLE II. INITIAL DECISION TABLE OF DIMENSION REDUCTION FOR INPUT VARIABLES

Input	Sample ID				
	1	2	3	...	200
$x_1$	34.121	15.460	54.029	...	76.258
$x_2$	11.435	50.127	19.232	...	34.812
$x_3$	22.374	5.166	43.185	...	9.454
$x_4$	-11.027	-80.179	215.156	...	47.550
$x_5$	32.175	12.040	60.137	...	53.022
...	...	...	...	...	...
$x_{15}$	15.142	26.750	31.082	...	40.531
$y$	1	3	4	...	4

3. In light of reduction the result for condition attributes in Table 3, the number of WNN input layer nodes can be determined as 8. The corresponding index feature is (1), (2), (4), (5), (8), (11), (13) and (15) respectively. The number of WNN hidden layer nodes can be initially taken as 6 according to the formula (1).

TABLE III. DISCRETIZED DECISION TABLE OF DIMENSION REDUCTION FOR INPUT VARIABLES

Input	Sample ID				
	1	2	3	...	200
$x_1$	3	1	5	...	7
$x_2$	1	5	2	...	3
$x_3$	1	1	2	...	1
$x_4$	5	3	8	...	2
$x_5$	3	1	6	...	5
...	...	...	...	...	...
$x_{15}$	2	3	3	...	4
$y$	1	3	4	...	4

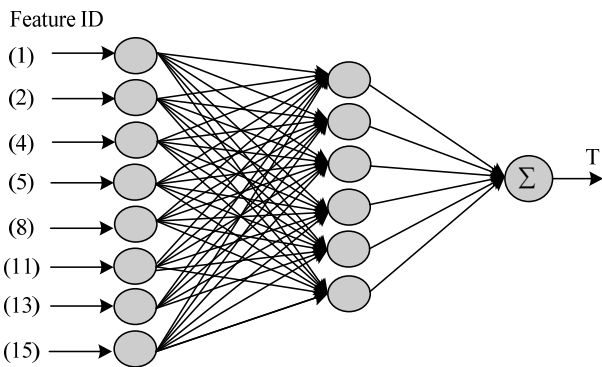


Fig. 2. WNN topological structure model

4. According to the modeling process, BP algorithm is used to train WNN with the inputs: (1), (2), (4), (5), (8), (11), (13) and (15) as shown in Fig. 2, and the stock price prediction model is established and the prediction result is obtained.

To promote better understanding of the performance for proposed model, RS-WNN model is compared with BP Neural Network (BPNN), GA Neural Network (GANN) and WNN model in the prediction results. For the convenience of comparison, the inputs of BPNN, GANN and WNN are 15 continuous features. The comparison of average prediction accuracy among BPNN, GANN, WNN and RS-WNN model are shown in Fig. 3.

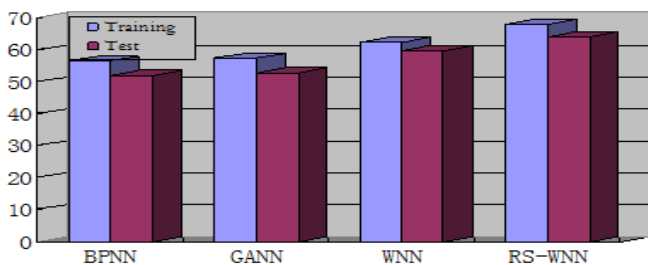


Fig. 3. Comparison of prediction accuracy for 4 models

It can be seen from Fig 3 that, compared with other 3 models, the proposed RS-WNN model has obvious advantages in network complexity, training and test accuracy. It shows that, using the attribute reduction ability adequately, RS can reduce the input dimensions of WNN and eliminate the redundant nodes in hidden layer, which can greatly simplifies the WNN model structure and enhances the performance of the prediction with average correct rate 64% within less training time, . Thus it provides a feasible and effective solution for the movement prediction of stock price.

V. CONCLUSION

For stock price prediction, an integrated prediction method based on RS and WNN is proposed. RS has an excellent attribute reduction capability. WNN is good at nonlinear system modeling. Combining with the both, it can reduce the input dimension and optimize the network topology structure for modeling and prediction of WNN, decreasing the calculation complexity with the improvement of stock price prediction accuracy. Simulations show that the proposed prediction method reduces the size of the network, simplifies the network structure, and has an improvement in contrast with simple WNN in the prediction effect.

REFERENCES

- [1] T. White, "Economic prediction using neural networks: the case of IBM daily stock returns," IEEE International Conference on Neural Networks, vol. 2, No. 6, 1988, pp. 451-458.
- [2] T. Kimoto, K. Asakawa, M. Yoda, et al, "Stock market prediction system with modular neural networks," 1990 IJCNN, International Joint Conference on Neural Networks, vol. 1, 1990, pp. 1-6.
- [3] N. Baba, M. Kozaki, "An intelligent forecasting system of stock price using neural networks," 1992IJCNN, International Joint Conference on Neural Networks, vol. 1, 1992, pp. 371-377.
- [4] M. H. Pesaran, M. A. Timermann, "Recursive modeling approach to predicting stock returns," Cambridge: University of Cambridge Press, 1999.
- [5] K. Kwong, "Financial forecasting using neural network or machine learning techniques," Thesis of Electrical Engineering, University of Queensland, 2001.
- [6] O. Dennis, M. Charles, "Neural network forecasts of Canadian stock returns using accounting ratios," International Journal of Forecasting, vol. 19, No. 3, 2003, pp. 453-465.
- [7] M. T. Philp, K. Paul, S. O. Choy, "Design and implementation of NNS for Hong Kong stock price forecasting," Engineering Applications of Artificial Intelligence, vol. 20, No. 4, 2007, pp. 453-461.
- [8] A. Leonel. Laboissiere, 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.
- [9] Q. H. Zhang, A. Benveniste, "Wavelet network," IEEE Transactions on Neural Networks, vol. 3, No. 6, 1992, pp. 889-898.
- [10] Q. H. Zhang, "Using wavelet network in nonparametric estimation," IEEE Transactions on Neural Networks, vol. 8, No. 2, 1997, pp. 227-236.
- [11] Z. Pawlak, "Rough sets," International Journal of Computer and Information Sciences, vol. 1, No. 11, 1982, pp. 341-356.
- [12] J. Zhao, G. Y. Wang, Z. F. Wu, H. Tang, et al, "An efficient approach to compute the feature core," Mini-Micro Systems, vol. 24, No. 11, 2003, pp. 1950-1953.