

Hybrid Prediction Based on BP Neural Network and Markov Chain

Guofeng Liu, Shaobin Huang, Xiufeng Piao, Yuan Cheng

College of Computer Science and Technology
Harbin Engineering University
Harbin, China
liuguofeng22@163.com

Abstract—At present, there are many kinds of prediction methods, how to play the advantages of different prediction methods and take advantage of useful information contained in the data have become the focus of research. This paper starts from the studies of combination prediction method. First, it described the current research status of combination prediction methods. Second, it analyzed the characteristics and applicability of different prediction methods. Finally, it proposed a hybrid prediction method based on BP neural network and Markov chain, and applied it to the prediction of personal payment in endowment insurance fund. The experimental results show that the proposed method is effective and feasible.

Keywords—hybrid prediction; BP neural network; markov chain; endowment insurance

I. INTRODUCTION

Recently, in order to utilize the collected data effectively, and provide better services for decision-making and strategy development, the prediction techniques arise spontaneously. In the practice of prediction, there are many different prediction methods can be used to solve the same problem, but the accuracy of prediction results often have great differences. Therefore, in order to improve the accuracy of prediction, the problem of how to improve existing models and how to propose more applicability method have become the focus of the research. Combination prediction method can overcome the shortcoming of single prediction method, and provide a higher applicability, so it is widely used in the prediction field.

Combination prediction is also known as comprehensive prediction; it combined different prediction methods by weighting to achieve the purpose of optimization. Combination prediction [1] was proposed by J. M. Bates and C. W. J. Granger in 1969. Because it overcomes the shortcomings of traditional prediction methods, and can make full use of the information provided by the data, so it got high degree of attention from prediction academic community. At present, there are many achievements of combination prediction. Bunn proposed a winning method [2] in 1975, in his combination prediction model, the weight of single prediction method was viewed as the smallest probability of absolute error in the next prediction [3]. In the beginning of 90's, Granger and Teräsvirta found the accuracy of nonlinear combination prediction is higher than linear combination prediction [4]. In China, Shuo Wang and

Xiaowo Tang discussed the prediction system from multiple angles [5]. To ensure the rationality of selected weights, Huayou Chen and Yisheng Xu used the significant test to determine the combination prediction weights [6]. Bo Zhong proposed a weight determination method based on rough set theory, and converted the problem of determining the weight coefficient into the evaluation problem of rough set attribute importance, it avoid massive calculations and subjective of traditional weights determining method [7]. Changxiu Cao proposed an automatic weighted combination prediction method. The overall idea is that the current weights depend on the error change rate of the last time weights [8]. Jingrong Dong studied the prediction model further, and proposed the nonlinear combination prediction method based on wavelet transform, the experimental results show that the method has a high value to deal with uncertain nonlinear systems [9]. Dehong An and Wenxiu Han improved the optimal weighting determined method, and proposed a optimal weighting determined method based on the BP neural network model, and applied it to load prediction, the results show that the proposed method is feasible [10]. Bo Hu and Jianmin Liu established the level analysis model for freight volume influencing factors, then used the combination prediction method to predict the freight volume from 2005 to 2006, and achieved better prediction results [11].

This paper started from the research of existing combination prediction methods. First, it introduced the current status of combination prediction method. Second, it described the characteristics and applicability of single prediction method, and proposed a hybrid prediction method. The method utilized the nonlinear mapping ability of BP neural network to integrate the result of single prediction method, and used the Markov chain prediction method to process results of BP neural network, including narrow the prediction range and improve accuracy. Finally, the experiments on real world data and application on the personal payment of endowment insurance indicated the effectiveness of the proposed method.

II. THE CHARACTERISTICS AND APPLICABILITY OF EXISTING PREDICTION METHODS

A. Regression Prediction Method

The whole idea of regression prediction method is to find the regression equation between independent variable and dependent variable according to the historical data variation,

and use the regression equation to predict. In the model, the dependent variable and independent variable should be synchronized in time domain. That is the predicted value of dependent variable should be estimated by synchronized independent variable. At present, the simple linear regression prediction method is the most commonly used forms of regression prediction. It established the regression equation as form of $y = bx + a$, for a given time series, and the coefficients a, b can be computed by the least square method, and then we can use the obtained regression equation to predict.

Compared with other prediction methods, regression prediction method is a mature technology, the prediction process is relatively simple, and it can consider the correlation of predicted object and relevant factors. But its drawback is that all data will be treated no distinction, and considered the influence degree of data to the predicted object are same. In addition, the main factors selection and variable factors quantization are difficulties in the present study. When sample size is large and prediction range is short, the regression prediction method will obtain a better performance.

B. Exponential Curve Prediction Method

Exponential curve prediction method is a kind of trend extrapolation prediction method. When the predicted object presents an upward or downward trend with time change, and without obvious seasonal fluctuations, if we can find a curve to describe the trend of the predicted object, then it is reasonable to believe that the trend is likely to continue into the future. The general form of exponential curve prediction equation is $y = a + b^x$, in order to calculate parameters a and b , the equation needs linear processing, that is $\ln(y_i) = \ln a + \ln b^x$, and uses the least square method to estimate the parameters a and b .

Exponential curve prediction method belongs to nonlinear prediction method, when the time series obeys exponential distribution, and shows a rising trend, the accuracy of prediction results is high. While for the random fluctuations and seasonal data, the accuracy of prediction results is low.

C. Gray Prediction Method

Gray prediction method is originated from the gray system theory. the essence of the gray system theory is to change the chaotic raw data into strong regularity data through certain processing means and researches. Because of the method have many advantages, such as less data needed, simple operation, and without considering of data distribution and transformation, so it widely used in economic prediction. When the time series obeys exponential distribution, the gray prediction method can obtain higher prediction accuracy.

Compared with other prediction methods, gray prediction also has some limitations. First, the development coefficient and background value influenced the prediction accuracy. The method will fail if the development coefficient is too large, and the accuracy of prediction will be improved through selecting the recent data as background value [13-14].

Second, because it is constructed by accumulated data, if the original data is irregular or more randomness, the accuracy of method will be influenced. Finally, it is not suitable for long-term prediction. Therefore, the selection of training data sets should be given adequate attention.

D. BP Neural Network Prediction Method

Because the BP neural network has good ability of nonlinear fitting, it is widely used in prediction field [15-16]. The whole idea of the BP neural network algorithm is as follows: the training data accesses to the network from input layer, after processed by hidden layer, it will transfer to the output layer, and output the result. If there exists an error between the output result and actual value, the BP neural network will enter back-propagation phase, the weight of each layer will be adjusted, the forward propagation will repeat, until the difference between output result and actual result falls in the specified error range.

The advantages of BP neural network are as follows: first, it has adaptive and self-learning ability. Second, it has high parallel. Finally, it has nonlinear mapping and good fault-tolerant ability. Although the BP neural network prediction method can effectively achieve nonlinear mapping between input and output, it also has some drawbacks, such as slow convergence, fixed line of reasoning, there is no theoretical basis about the number of hidden nodes, needs larger training data sets, may appear over-fitting, etc.

E. Markov Chain Prediction Method

Markov chain is a special random process, and its character is that the time series should meet the "no after-effects" property, that is the future state of the system is only related with its current state, rather than its previous state. This feature makes the Markov prediction method can make a better prediction for time series with strong stochastic volatility. Therefore, it is widely used in prediction field.

When Markov prediction method is used to predict, the key is to calculate the state transition matrix, because the matrix revealed the internal law when different states convert between each other, it can be used to solve the next step state of the system. Compared with other prediction methods, its application is not very wide, because the time series must meet the Markov properties.

III. THE HYBRID PREDICTION METHOD BASED ON BP NEURAL NETWORK AND MARKOV CHAIN

According to the analysis of different prediction methods, it can be seen that each method has their own merit and limitation. So it is unrealistic to solve the practical problems effectively through any single one of the prediction methods. For example, for an exponential distribution time series, people can easily make the following judgments, namely the accuracy of prediction results will be improved by choosing the exponential curve prediction method or gray prediction method. In contrast, the accuracy of prediction results will reduce if we choose regression prediction method. But, when the time series presents significant volatility, it is clear that the exponential curve prediction method and gray prediction method will also have larger errors.

In order to play the advantages of different prediction methods and make full use of the information provided by the data, and solve the complex issues which can not be solved by a single prediction method, this paper proposed a hybrid prediction method based on BP neural network and Markov chain, the overall structure of the hybrid prediction shown in figure 1.

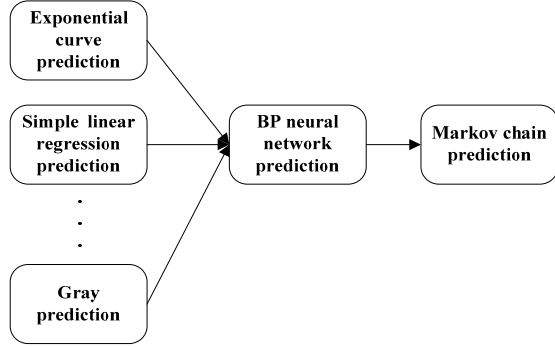


Figure 1. A HYBRID PREDICTION METHOD BASED ON BP NEURAL NETWORK AND MARKOV CHAIN

The whole idea of the method is as follows: first, it uses different prediction methods to fit the sample data, and chooses methods with higher precision as a part of hybrid prediction method. Second, it uses BP neural network to integrate the result of individual prediction method. Because of BP neural network is well known as its adaptive and self-learning ability, therefore, we can make full use of it to determine the coefficient of different prediction methods. Finally, it uses Markov prediction method to estimate the relative error of BP neural network prediction method, and determines the final prediction result.

According to Section 2, BP neural network prediction method is suitable for the case of larger sample size, when the sample size is small and has significant volatility, although it presents a high fitting ability, the generalization ability will reduce. In order to improve the adaptability of method and accuracy of results, this paper uses Markov chain for further processing. Markov chain prediction method can be expressed as a general form $X(n)=X(0)P^n$, where $X(n)$ is the state probability vector at time n , $X(0)$ is the initial state probability vector, and P is a state transition matrix, when $n=1$, P is called once state transition matrix. The key of using Markov chain to predict is the division of state and the solving of P . In this paper, the relative error between the result of BP neural network and actual value will be divided according to certain interval, the interval of an error will be determined and then the state transition matrix will be computed. We will calculate the transition probability between two states as formula (1).

$$P_{ST} = N_{ST} / N_S \quad (1)$$

Where N_{ST} is sample size from relative error state S to the next state T , and N_S is sample size when error state is S .

If the current state of the system is S , observing the state transition probability from S to other states in P , and selecting the state corresponding to the maximum probability value as the next state of the system. Use the median of an interval corresponding to the state as the next year prediction error probability. It can be computed by following formula.

$$H = H_0 / (1-e) \quad (2)$$

Where H_0 is a predicted value of BP neural network, H is a predicted value of Markov chain, and e is the median of an interval corresponding to the certain state.

IV. EXPERIMENT AND ANALYSIS

In order to verify the generality and effectiveness of the proposed method, we choose the personal payment in somewhere endowment insurance fund as experimental data. Figure 2 shows the monthly personal payment from 2006 to 2007 in somewhere social security department. (Data source: CNAO IN CENTER)

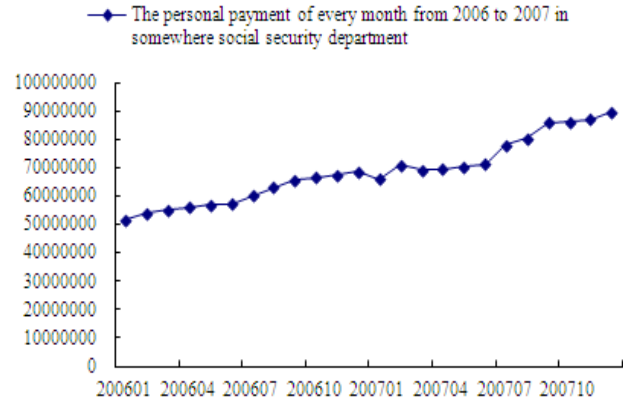


Figure 2. THE MONTHLY PERSONAL PAYMENT FROM 2006 TO 2007

Figure 2 shows, on the whole, the data presents a rising trend and with volatility. Different prediction methods were used to fit the actual value, the fitting results were shown in figure 3.

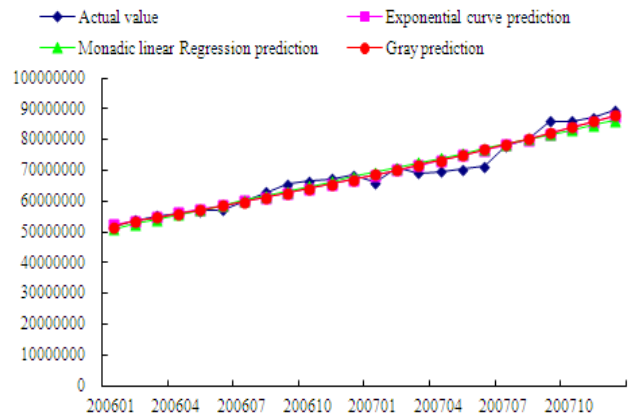


Figure 3. THE RESULTS OF SINGLE PREDICTION METHOD

Figure 3 shows, the fitting results of exponential curve prediction method and gray prediction method are essentially coincident, and the whole accuracy is higher than simple linear regression method, but in some significant volatility data points, the fitting values present a large deviation, such as April and May, 2007, and so on. So, the prediction results are not satisfactory.

In order to investigate the effect of simple linear regression prediction results on hybrid prediction results, we integrated the prediction results of different methods using BP neural network, and design the following two experimental schemas: (1) In the three-layer structure, there are three units in input layer, ten units in hidden layer, and one unit in output layer, where the units in input layer represent simple linear regression, exponential curve and grey prediction methods used in this paper respectively. (2) In the three-layer structure, there are two units in input layer, seven units in hidden layer, and one unit in output layer, where the two units in input layer represent two prediction methods having better fitting results, i.e. exponential curve and grey prediction methods. And the fitting results of these two experimental schemas are shown in figure 4 and 5.

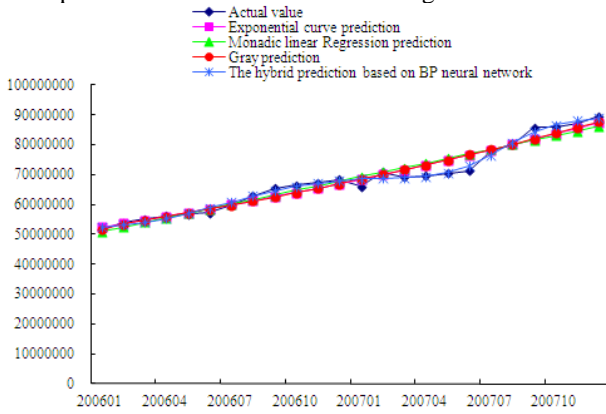


Figure 4. THE RESULTS OF HYBRID PREDICTION METHOD BASED ON BP NEURAL NETWORK (1)

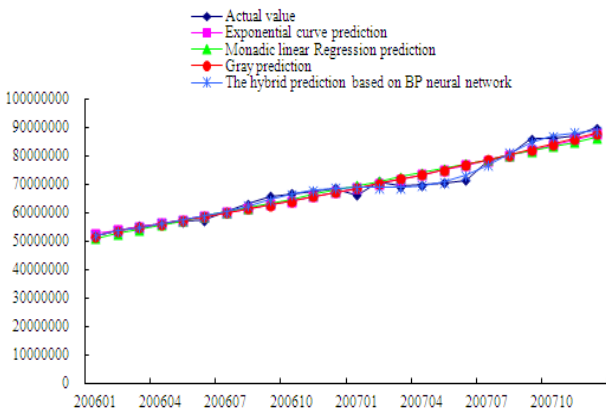


Figure 5. THE RESULTS OF HYBRID PREDICTION METHOD BASED ON BP NEURAL NETWORK (2)

Figure 4 and 5 shows, both of the experimental schemas have good performance, but in some data points, the accuracy of experimental schema 2 is higher than 1. It is mainly because the error of simple linear regression prediction method is larger than other methods, so it impacts the accuracy of hybrid prediction method.

Next, we use the hybrid prediction method to predict the amount of individual contributions in this area from January to April, 2008. At the same time, we also use single prediction method to predict, and the comparison between prediction results of different methods are shown in Table I.

Table I shows, the whole prediction accuracy has been improved effectively, through combining single prediction method based on BP neural network, because it overcomes the limitation of single prediction method, and combines their advantages. In order to further improve the accuracy of prediction, in this paper, the results of BP neural network prediction method will be further processed by Markov chain prediction method. According to the formula shown in Section 4, we can calculate the state transition probability P , and calculate the prediction values based on P . The results were shown in Table II.

It can be seen that the overall accuracy of prediction is further improved, so the method proposed in this paper is feasible.

V. CONCLUSION

According to the limitation of single prediction methods, such as accuracy is not high and the applicability is not strong, etc., this paper achieved the following works: 1. It analyzed the combination prediction research status at home and abroad; 2. It discussed the characteristics and applicability of single prediction method; 3. It proposed a hybrid prediction method based on BP neural network and Markov chain; 4. It validated the proposed method using real data sets, and compared it with single prediction method. The results show that the accuracy of proposed method is better than other single prediction method.

ACKNOWLEDGMENT

This work is sponsored by National Key Project of Scientific and Technical Supporting Programs under grant number 2009BAH42B02 and 2012BAH08B02, the National Natural Science Foundation of China under grant number 60873038 and the Fundamental Research Funds for the Central Universities under grant number HEUCF100603. Some research results of this paper come from National Audit Data Center, and it obtained by the first author.

REFERENCES

- [1] J. M. Bates and C. W. J. Granger, "The combination of forecasts," *Operational Research Quarterly*, vol. 20, 1969, pp.451-468.
- [2] D. W. Bunn, "A bayesian approach to the linear combination of forecasts," *Operational Research Quarterly*, vol.26, 1975, pp.325-329.
- [3] D. W. Bunn, "Statistical efficiency in the linear combination of forecasts," *International Journal of Forecasting*, 1985, pp. 151-163.
- [4] C. W.J. Granger and T. Teräsvirta, "Modelling Nonlinear Economic Relationships," Oxford University Press, 1993.

- [5] Shou Wang and Xiaowo Tang, "Research on combination forecasting system," Scientific Management Research, vol.18, 2000, pp.42-45.
- [6] Huayou Chen and Yisheng Xu, "Weight estimation and test of significance about combination forecasting," Operations Research and Management Science, vol.9, 2000, pp.75-78.
- [7] BO Zhong, Zhi Xiao and Jiaqi Zhou, "Determination to weighting coefficient of combination forecast based on rough set theory," Journal of Chongqing University, 2002, vol.25, pp.127-130.
- [8] Xiaowo Tang, Jing Wang and Changxiu Cao, "A new combined forecasting method based on method of fuzzy adaptive variable weight," Journal of University of Electronic Science and Technology of China, 1997, pp.47-53.
- [9] Jingrong Dong, "Research on nonlinear combination forecasting method based on wavelet network," Journal Of Systems Engineering, vol.15, 2000, pp.383-388.
- [10] Dehong An, Wenxiu Han and Yihong Yue, "Improved combination forecast method and its application in shor-term load forecasting of a power system," Systems Engineering and Electronics, vol.26, 2004, pp.842-844.
- [11] Bo Hu and Jianmin Liu, "Radial bases function neural network model for freight volume forecast based on combination forecast method," Computer and Communications, vol.24, 2006, pp.34-40.
- [12] M. A. Gebers, R. C. Peck, "Using traffic conviction correlates to identify high accident-risk drivers," Accident Analysis and Prevention, Vol. 6, pp.903-912, 2003.
- [13] Guanjun Tan, "The structure method application of background value in gray system GM(1,1) model," Systems Engineering Theory and Practice, vol.20, 2000, pp.98-103.
- [14] Dahai Zhang, Shifang Jiang and Kaiquan Shi, "Theoretical defect of grey prediction formula and its improvement," Systems Engineering Theory and Practice, vol.22, 2002, pp.1-3.
- [15] Mian Xing, "Research on combined grey neural network model of seasonal forecast," Systems Engineering Theory and Practice, 2001, pp.31-35.
- [16] Qian Liu and Lirong Xiong, "Predictive model of water requirement of paddy rice based on artificial neural network," Journal of Huazhong AgriculturalUniversity, vol.26, 2007, pp.885-887.

TABLE I. COMPARISON OF DIFFERENT PREDICTION RESULTS

Year	The hybrid prediction based on BP neural network		Simple linear regression prediction		Exponential curve prediction		Gray prediction		Actual value
	Prediction value	Relative Error (%)	Prediction value	Relative Error (%)	Prediction value	Relative Error (%)	Prediction value	Relative Error (%)	
200801	89811845	-0.23	87726355	2.10	89485201	0.13	89792631	-0.21	89603848
200802	91345362	0.94	89259279	3.21	91503409	0.77	91839806	0.41	92214658
200803	93177056	-0.74	90792202	1.84	93567135	-1.17	93933655	-1.56	92489915
200804	94934300	0.18	92325125	2.93	95677405	-0.60	96075241	-1.02	95107821
200805	96304986	1.57	93858049	4.07	97835270	0.01	98265653	-0.43	97841422
200806	97207375	0.05	95390972	1.92	100041801	-2.87	100506004	-3.35	97252889

TABLE II. COMPARISON BETWEEN HYBRID PREDICTION BASED ON BP NEURAL NETWORK AND THE METHOD PROPOSED IN THIS PAPER

Year	The hybrid prediction based on BP neural network		The method proposed in this paper		Actual value
	Prediction value	Relative Error (%)	Prediction value	Relative Error (%)	
200801	89811845	-0.23	88913727	0.77	89603848
200802	91345362	0.94	92258817	-0.05	92214658
200803	93177056	-0.74	92245285	0.27	92489915
200804	94934300	0.18	95883643	-0.82	95107821
200805	96304986	1.57	97268036	0.59	97841422
200806	97207375	0.05	98179449	-0.95	97252889