

## Study on Electronic Banking Risk Warning based on Comprehensive Optimized Gray Theory

Tong Wang

*School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing, China*  
*wangtong@bupt.edu.cn*

### **Abstract**

*General warning method of electronic banks generates based on clear detection of relevant data. However, this method is mostly a reflection of electronic banks' risks in the past which is a manifestation of historical data that can only be able to analyze problems that have already occurred but can not give a better future anticipation. Therefore early warning system established in this paper is based on the data detection system and we use an optimized gray early warning method to predict the indicators of the risk of electronic banks. Based on the idea of using descending cumulation to change traditional gray 1-AGO sequence, we use weaken buffer operators to deal with the original data sequence and then use genetic algorithm to estimate a and b--important parameters of background values. Eventually the optimized gray GM (1,1) model prediction method generate and it can effectively forecast the risk profile of electronic banks. Then analyze the results, further read and dig the hidden meaning, give a series of practical conclusions for the development, risk prevention and control of electronic banks. Decision making depends on the degree of risk in the future. Then we can take effective measures to cope with the arrival of electronic bank crisis that may arise.*

**Keywords:** *Electronic bank, Risk warning, Gray model, Genetic optimization*

### **1. Introduction**

Electronic banking in China is still an emerging business. It is usually simply operated. It's convenient, quick and always ready. So it can counter the operating pressure of bank branches significantly. Electronic banking is becoming a major transaction mode of almost all banks. Meanwhile, it has an important impact on the corporate image of the commercial banks. It is also an important criterion for judging the strength of each bank's competitiveness. In virtue of electronic banking having its inherent mode of operation and unique operating environment, when people enjoy the convenience it brings, they may also come across a variety of risks. It is the core element that how to use scientific methods to warn the risks of electronic banking in the rapid development of electronic banking so that we can prevent and control effectively.

Domestic scholars have analyzed the studies of electronic banking risks from different perspectives. Li De described the development of electronic banking systematically. Electronic banks developed from the banks of traditional concept. So the e-banking not only has a traditional banking business risks but also has some specific business risks such as the risk of network operation, risks related to computer technology and legal risks in the network operation. Through a series of cases, he got that the risk is mainly due to the electronic banking network system instability. Many factors of instability lead to the possibility of a major accident. Eventually he gave some market signals which can help us to effectively identify the risks of electronic banking [1]. Yue Yiding *et al.*, classified the major risks of electronic banking and gave them a comprehensive comparison. They are technology risk, business risk and operation risk. Meanwhile, they also did related

researches on the business risk and compared the traditional banking risk. with electronic banking risks [2]. Yin Long considered that developing countries such as China can not completely copy the electronic banking supervision system of developed countries otherwise we can not strengthen the risk control of electronic banking under the existing conditions and we can not effectively establish or optimize our risk monitoring system. As we know, networking technology, computer hardware technology, the existing conditions, development mode of banks and even people's attitude on the acceptance of electronic banks in China are all different from these in developed countries. Therefore only after we strengthen the development of computer hardware and software of our country we can effectively improve the electronic banking regulatory system [3]. Zhang Qiang also came up with some valid suggestions on strengthening our electronic banking risk monitoring. He believed that electronic banking should be directly under the management of financial institutions and we need to establish strict access system for e-banking as well as strengthen security protection of e-banking operations platform etc. [4]. Wei Tielin believed that electronic banks have played an important role in China's overall economic development. However, the development of electronic banks is still short and there is little systematic theoretical study. Therefore the impact on economic development may be good but also can be bad. He thought that currently the most important for China is to plan for the long-term development strategy, mode of operation, risk control of electronic banks [5]. Zhao Liang held that risk management is an important problem in development of electronic banks. In order to better resolve these problems, we need to strengthen the national communication with developed countries and get good experience of risk management. Thereby we can increase the overall level of China's electronic banks' regulatory risk [6]. Sun Sen thought that prevention should be carried out from both internal and external aspects so that we can effectively improve the overall preparedness level of electronic banking. From the bank's internal point of view, they should strengthen their risk management; from the outside point of view, the financial supervision departments should strengthen the bank's risk monitoring electronic means [7]. Dai Wenjin held the point that for the risk management of electronic banking in financial industry, first we need to perfect the internal control system; then to foster highly qualified technical personnel compliance with electronic banking long-term development; finally to improve the cultural infrastructure bank business and build a healthy enterprise network cultural environment.

From the above analysis, we can know that after 30 years of reform and opening-up efforts, China has made a series of great progress in bank risk controlling especially the risk of electronic banks. However, due to the further development of information technology in recent years, some new phenomenon and problems generated in risk management. Therefore, we must further improve risk management of electronic banks based on the principle of keeping pace with the times. So we believe that study on risk control and early warning of electronic banks is very significant. In this paper, we propose a series of quantitative analysis method to control and early warning the risk profile of electronic banks. Thus we can provide an effective adjunct to bank regulators for more scientific e-banking risk management.

## 2. Establishment of Traditional Gray Model

Here we will introduce the most classic  $GM(1,1)$  model. The first "1" means one variable in the model and the second "1" means that there's one unknown number included in the model.

The main idea of the  $GM(1,1)$  model is explained as follows:

We set  $x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$  as the original amount and we can get that

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$$

$$\text{and } x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)$$

We describe the first-order linear ordinary differential equation as albinism differential equation of  $GM(1,1)$ .

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b$$

The differential equation is

$$x^{(0)}(k) + az^{(1)}(k) = b$$

$a$  is development coefficient,  $b$  is gray action.  $a, b$  are first-order parameter package of model  $GM(1,1)$ .

$$[a, b]^T = (B^T B)^{-1} B^T Y_n$$

And

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}$$

$$Y_n = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)]^T$$

The background value is

$$z^{(1)}(k+1) = \frac{1}{2}[x^{(1)}(k) + x^{(1)}(k+1)] \quad k = 1, 2, \dots, n-1.$$

The discrete solution of the  $GM(1,1)$  model is:

$$\hat{x}^{(1)}(k+1) = (x^{(0)}(1) - \frac{b}{a}) \cdot e^{-ak} + \frac{b}{a}$$

The restore value is:

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) = (1 - e^{-a})(x^{(0)}(1) - \frac{b}{a}) \cdot e^{-ak}, \quad k = 1, 2, \dots, n$$

Suppose that  $\bar{x}^{(r)}(k)$  is the fitted value of  $x^{(r)}(k)$ .

We define that  $q^{(0)}(k) = x^{(0)}(k) - \bar{x}^{(r)}(k)$  is the salvage of  $x^{(0)}$  at  $k$ .

$$S_1^2 = \frac{1}{n} \sum_{i=1}^n (x^{(0)}(i) - \bar{u}_1)^2$$

$$S_2^2 = \frac{1}{n} \sum_{j=1}^n (\bar{x}^{(0)}(i) - \bar{u}_2)^2$$

$\bar{u}_1$  is the mean of the original sequence,  $\bar{u}_2$  is the mean of the predicted values. We define the posteriori error ratio as:

$$C = \frac{S_2}{S_1}$$

### 3. Study on RiskPortfolio Prediction of Electronic Banks based on the Improved GM(1,1) Model

#### 3.1. Basic Principle of Buffer Operator

**Definition 1:** Suppose that  $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$  is a set of actual time series vectors, the observed behavior data sequence can be expressed by the actual time series vector group as follows:

$$\begin{aligned} X^{(0)} &= (x(1), x(2), \dots, x(n)) \\ &= (x^{(0)}(1) + \varepsilon_1, x^{(0)}(2) + \varepsilon_2, \dots, x^{(0)}(n) + \varepsilon_n) \end{aligned}$$

the  $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$  is the error term and  $x$  is perturbation vector.

If we want to start from the perturbation vector, simulate the change trend of the real behavior sequence correctly and have an in-depth understanding, we must reduce the

intervention of error term in the system. Otherwise, using time series set  $x$  directly to model and predict model can not reflect real variation of the group and lead to a too large error in the prediction so that we lost the practical significance of prediction.

Because of errors, and the error will be presented in a stochastic volatility trend, so the quantitative predictions differs very far from people's subjective qualitative analysis conclusion and even cases of drawing further apart occur frequently. Therefore, seeking to combine quantitative analysis with qualitative prediction and preprocessing of data in a way of "cleaning" is an important way to improve the accuracy of the prediction.

**Definition 2:** Set the behavioral data columns of the system as  $X = (x(1), x(2), \dots, x(n))$ . We have three forms as follows:

If  $\forall k = 2, 3, \dots, n$ ,  $x(k) - x(k-1) > 0$ , we call it monotonically increasing behavior data columns;

If  $\forall k = 2, 3, \dots, n$ ,  $x(k) - x(k-1) < 0$ , we call it monotonically decreasing behavior data columns;

If  $\exists k, k \in \{2, 3, \dots, n\}$ ,  $x(k) - x(k-1) > 0$ ,  $x(k) - x(k-1) < 0$ , we call it random behavior data columns.

Set  $M = \max\{x(k) | k = 1, 2, \dots, n\}$ ,  $m = \min\{x(k) | k = 1, 2, \dots, n\}$ , then we can use  $M - m$  to indicate the amplitude of behavior data columns.

**Definition 3:** Let  $x$  be behavior data columns of the system,  $D$  is the operator acting on  $x$ . Record the sequence after  $x$  acted on by  $D$  as  $xD = (x(1)d, x(2)d, \dots, x(n)d)$  and we call  $D$  sequence operators,  $xD$  first-order operator acting sequence.

Analogously, if  $D_1, D_2, D_3$  are sequence operators, then we call  $D_1, D_2$  second-order operators.  $xD_1D_2 = (x(1)d_1d_2, x(2)d_1d_2, \dots, x(n)d_1d_2)$  are called second-order operators acting sequence.

### 3.2. GM (1,1) Model and its Optimization

The basic optimization idea of GM (1,1) model in this chapter is:

First use the idea of reverse accumulation to change traditional gray 1-AGO sequence; then use weaken buffer operators to preprocess the original data sequence  $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ ; then use genetic algorithm to estimate a and b — important parameters of the background values. Now we get the ultimate form of optimized gray GM (1,1) model and forecasted analyze the indexes of the risk evaluation system in the fourth chapter.

Next we will first use the idea of reverse accumulation to change traditional gray 1-AGO sequence.

Let  $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$  be a set of quasi smooth sequence data, accumulate it reversely and get a new data column  $\tilde{x}^{(0)} = (\tilde{x}^{(0)}(1), \tilde{x}^{(0)}(2), \dots, \tilde{x}^{(0)}(n))$  which is the 1-DAGO sequence of  $\tilde{x}^{(0)}$ :

$$\tilde{x}^{(1)}(k) = \sum_{i=1}^k \tilde{x}^{(0)}(i) \quad k = 1, \dots, n \quad i = 1, \dots, n \quad (4)$$

Now we call  $\tilde{z}^{(1)} = (\tilde{z}^{(1)}(2), \tilde{z}^{(1)}(3), \dots, \tilde{z}^{(1)}(n))$  the proximate mean generation sequence of  $\tilde{x}^{(1)}$  which has been reversely accumulated:

$$\tilde{z}^{(1)}(k) = \frac{1}{2}(\tilde{x}^{(1)}(k) + \tilde{x}^{(1)}(k-1)) \quad k = 2, 3, \dots, n$$

The basic form of the GM (1,1) model after reverse accumulation is different from that of the traditional GM (1,1) model. We use the following formula to express:

$$-\tilde{x}^{(0)}(k-1) + a\tilde{z}^{(1)}(k) = b, \quad k = 2, 3, \dots, n \quad (5)$$

$-a$  is called development coefficient,  $b$  is called gray action.

Set  $\hat{a}$  as the parameter to be estimated vector and  $\hat{a} = (a, b)^T$ , then the least squares estimated parameter list of gray differential equation (2) meets:

$$\hat{a} = (B^T B)^{-1} B^T Y_n$$

and

$$Y = \begin{bmatrix} -x^{(0)}(2) \\ -x^{(0)}(3) \\ \vdots \\ -x^{(0)}(n) \end{bmatrix}, \quad B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad (6)$$

Secondly, we use the weaken buffer operator to deal with the initial sequence  $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$  beforehand. We construct two kinds of weaken buffer operator and use them to reset the initial sequence respectively, and then use the mean to form a new sequence.

According to relevant content 3.1, we define the following operators:

**Operator 1:**

Make the initial sequence

$$X = (x(1), x(2), \dots, x(n))$$

Calculate by a weakening buffer and get

$$XD = (x(1)d, x(2)d, \dots, x(n)d)$$

We have

$$x_1(k)d = \frac{kx(k) + (k+1)x(k+1) + \dots + nx(n)}{(n+k)(n-k+1)/2}; \quad k = 1, 2, \dots, n$$

**Operator 2:**

Make the initial sequence

$$X = (x(1), x(2), \dots, x(n)) .$$

The sequence is non-negative which means

$$x(k) > 0, k = 1, 2, \dots, n .$$

Calculate by a weakening buffer and get

$$XD = (x(1)d, x(2)d, \dots, x(n)d) .$$

We have

$$x_2(k)d = [(x(k) \times x(k+1) \times \dots \times x(n))^{n-k+1}]^{\frac{1}{n-k+1}} = [\prod_{i=k}^n x(i)]^{\frac{1}{n-k+1}}; \quad k = 1, 2, \dots, n$$

Then use the formula  $x(k)d = \frac{1}{2}x_1(k)d + \frac{1}{2}x_2(k)d$  and get the certain initial sequence.

**Operator 3:**

We use the genetic algorithm to estimate the parameters and background value a and b. The specific algorithm steps are as follows:

Step1: Binary encoding.

First we should fix the range of development factor a and gray action b. Then encode a and b to binary format based on genetic algorithm and we can get the initial population which the genetic algorithm needs.

Step2: To construct the fitness function.

The termination condition of parameter optimization is decided by the value of fitness function based on our real-time judgement. If the fitness value is in line with our requirements, we can consider the prediction accuracy of the gray model meets our requirements. Then you can terminate computing the genetic algorithm optimization. After the decoding process, the a and b we get are the optimal parameter values. Then you can use the sets of optimal parameters to model gray prediction model. And if the value does not meet the requirements, go to step 4.

For the fitness value of genetic algorithm, the following formula can be used as a measure of the standard:

$$MSE = \frac{1}{N} \sum (x_i - \hat{x}_i)^2 \quad (7)$$

Now we will describe the symbols of the formula.  $x_i$  is the actual value,  $\hat{x}_i$  is the predicted value,  $N$  the overall number of samples. The calculation result is less, the better accuracy the genetic algorithm has.

Step3: To calculate the fitness function.

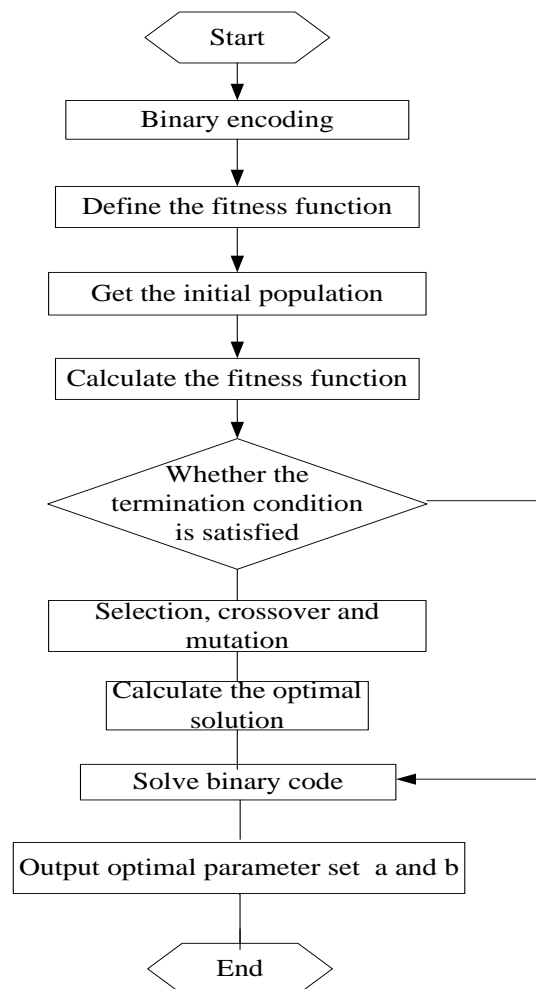
On the specific application of this algorithm in electronic bank risk assessment system, because of its industry characteristics—preference small degree of risk, and therefore we can set the parameters as follows:

Set the size of the population (*e.g.*,  $M = 100$ ); terminate evolution generation ( $T = 30$ ); use relatively large crossover probability ( $P_c = 0.95$ ); a relatively small mutation probability ( $P_m = 0.001$ ).

Step4: Selection, crossover and mutation.

Obtain offspring populations through genetic operator selection, crossover operator, mutation operator. Then go to the Step2 to judge the fitness value size.

According to the algorithm above, we can get an estimation value of the background's parameters a and b. Process of the algorithm above is shown by Figure 1 below:



**Figure 1. Parameter Optimization Flowchart of Genetic Algorithm**

Now we have got the optimized initial sequence as well as the estimated optimal parameter of background value. We can do the final prediction by the following formula:

The new value into the background gray prediction model there

$$\hat{x}(t) = [\tilde{x}^{(1)}(1) - \frac{b}{a}]e^{-at} + \frac{b}{a} \quad (8)$$

The time response sequence of GM(1,1) model—  $\tilde{x}^{(0)}(k) + a\tilde{z}^{(1)}(k) = b$  is:

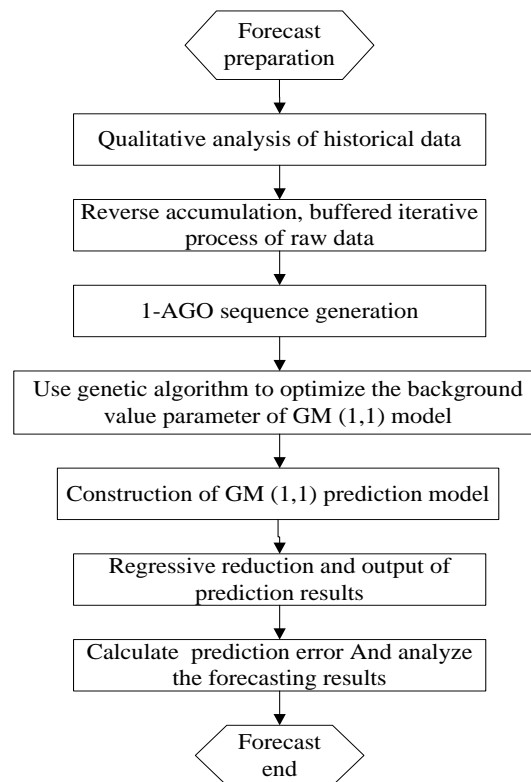
$$\hat{x}(k+1) = [\tilde{x}^{(1)}(1) - \frac{b}{a}]e^{-ak} + \frac{b}{a}, \quad k = 1, 2, \dots, n \quad (9)$$

Restore Value

$$\begin{aligned} \hat{x}^{(0)}(k+1) &= a^{(1)}\hat{x}^{(1)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \\ &= (1 - e^{-a})[\tilde{x}^{(0)}(1) - \frac{b}{a}]e^{-ak}, \quad k = 1, 2, \dots, n \end{aligned} \quad (10)$$

Above is the prediction equation.

In summary, use buffer operator and the background value to improve the optimum gray model and build a new process of prediction model. The flow diagram is shown in Figure 2:



**Figure 2. The Forecasting Process of Comprehensive Optimized Gray Model**

#### 4. Numerical Example and Result Analysis

Next we will give the comments on the 17 important indexes given by experts. They are given with the change of time and the unit is half year. Whereas the length of the paper limited we will only display four of them.







Then according to the comments in Table 1, we get the normalized three-level index evaluation matrix  $L = (L_{ij})$ . Multiply three-level index evaluation matrix and three-level index comprehensive weights to have three-level index evaluation results. On the basis of  $E_i^j = T_i^j \cdot V$ , the  $V = \{90, 80, 70, 60, 50\}$ , we can calculate the risk assessment indexes of the 17 two-level index layers from 2009 to 2010. The  $i$  means indicators content and  $i = 1, 2, \dots, 17$ . They are information risk, network security risk, system operation risk, network technology selection risk, business risk, operation risk, human risk, organization risk, management risk, control risk, the risk of running liability, the risk of conflict abroad, legal risk, credit risk, country risk, reputational risk and transaction risk.  $j$  means the time and  $j = 1, 2, \dots, 6$ . They are the first half of 2009, the second half of 2009, the first half of 2010, the second half of 2010, the first half of 2011 and the second half of 2011. For instance, the  $E_3^4$  means the system operation risk assessment index of electrical bank in the second half of 2010. The higher is the index, the safer is the system.

**Table 2. The Risk Assessment Index Table**

	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
$E_1^j$	81.703	82	84.851	84.561	85.703	89.280
$E_2^j$	80.652	82.438	84.032	82.000	84.000	88.652
$E_3^j$	82.331	82.000	84.000	83.851	84.830	88.000
$E_4^j$	82.614	80.693	85.141	82.000	84.000	88.000
$E_5^j$	82.661	82.911	84.854	87.145	84.855	89.517
$E_6^j$	81.232	82.000	84.000	86.927	85.232	88.000
$E_7^j$	82.236	82.407	84.636	86.565	84.000	88.734
$E_8^j$	80.239	82.000	84.470	82.000	84.000	88.239
$E_9^j$	80.000	83.382	81.379	82.000	84.000	88.000
$E_{10}^j$	81.540	82.687	82.626	82.682	84.000	89.313
$E_{11}^j$	82.470	82.304	85.530	85.548	85.674	90.000
$E_{12}^j$	83.212	81.736	84.788	85.896	84.394	90.000
$E_{13}^j$	84.000	81.526	83.746	87.052	84.254	90.000
$E_{14}^j$	81.616	82.941	86.384	85.113	86.941	90.000
$E_{15}^j$	84.000	81.580	83.169	87.160	84.831	90.000
$E_{16}^j$	81.834	82.941	86.167	85.549	86.167	90.000
$E_{17}^j$	81.666	83.347	86.694	85.306	86.694	90.000

We get  $a$  and  $b$  which are the background value parameters of optimized gray forecasting model. According to the degree of their impact on the modeling results, we use MATLAB programming to find and determine the final parameters which are 0.042314 and 14217.5023 respectively.

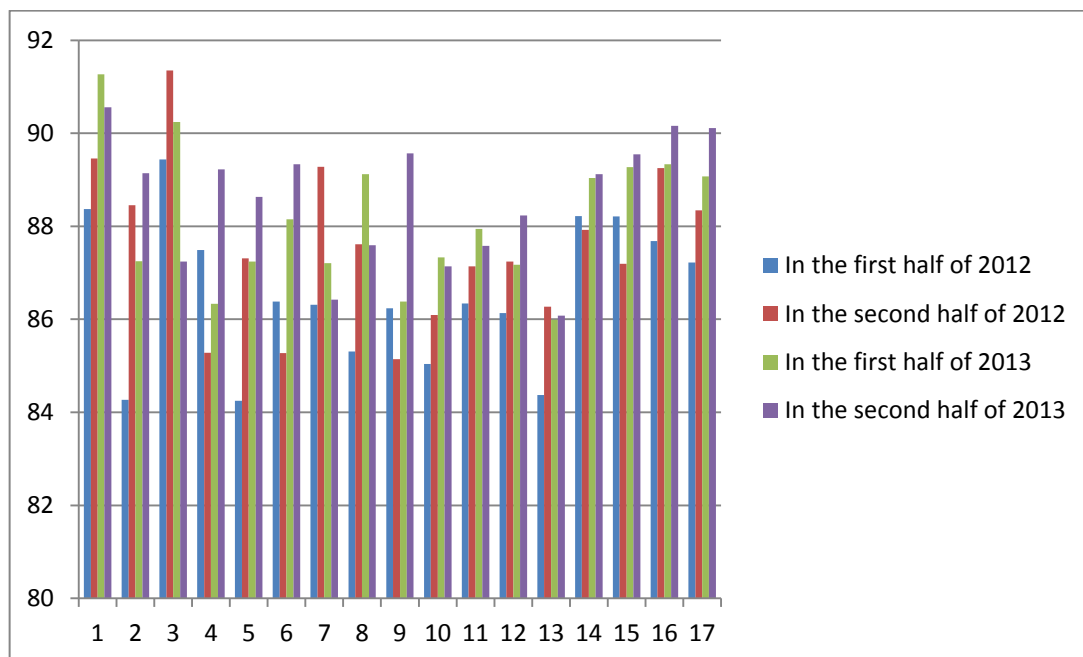
We use the above risk indexes as the sample data. The gray forecasting model in this chapter will be used to forecast the the risk assessment indexes of the next four periods: the first half of 2012, the second half of 2012, the first half of 2013 and the second half of 2013.

The table below is for the prediction results.

**Table 3. The Prediction Results**

	The first half of 2012	The second half of 2012	The first half of 2013	The second half of 2013
$E_1^j$	88.37	89.46	91.27	90.56
$E_2^j$	84.27	88.45	87.25	89.14
$E_3^j$	89.44	91.35	90.24	87.24
$E_4^j$	87.49	85.28	86.33	89.22
$E_5^j$	84.25	87.31	87.24	88.63
$E_6^j$	86.38	85.27	88.15	89.33
$E_7^j$	86.31	89.28	87.21	86.42
$E_8^j$	85.31	87.61	89.12	87.59
$E_9^j$	86.24	85.14	86.38	89.57
$E_{10}^j$	85.04	86.09	87.33	87.14
$E_{11}^j$	86.34	87.14	87.94	87.58
$E_{12}^j$	86.13	87.24	87.17	88.23
$E_{13}^j$	84.37	86.27	85.99	86.08
$E_{14}^j$	88.22	87.92	89.04	89.12
$E_{15}^j$	88.21	87.19	89.27	89.55
$E_{16}^j$	87.68	89.25	89.33	90.16
$E_{17}^j$	87.22	88.34	89.07	90.11

By the above table, we can get each predicted value. If we also give four groups data of 2012 and 2013 expert comments, then review the results in a similar calculation process according to Table 1. The result will be used as a risk assessment index, so that we can compare the actual values with the predicted values. The results of prediction are shown as follows:



**Figure 3. Schematic Predictions**

From the trend, we can get that for electronic banks, there are 17 important indicators and they are information risk, network security risk, system operation risk, network technology selection risk, business risk, operation risk, human risk, organization risk, management risk, control risk, the risk of running liability, the risk of conflict abroad, legal risk, credit risk, country risk, reputational risk and transaction risk. We can find the risks change over time as a large increase in electronic banks network security risks, electronic banks risk and risk management. While the control risks, the risk of run liability and conflict abroad transform relatively insignificantly. In the way of amplitude, the network technology selection risk and human risk are more prominent.

## References

- [1] L. De, "Risk and supervision of network bank", Beijing: Financial Theory and Practice, vol. 4, (2001).
- [2] D. Y. Yue and Q. T. Wu, "Network Finance", Nanjing: Southeast University Press, (2005).
- [3] Y. Long, "A Study on the Development and the Supervision of Internet-banking in China", Beijing: Journal of finance, vol. 1, (2001).
- [4] Q. Zhang, "The Latent Issues of Network Banking and Financial Supervision", Hunan: Journal of Hunan University (Social Science) , vol. 6, (2001).
- [5] T. L. Wei, "Several problems in the development of Internet Bank in China", Social Sciences In Hubei, vol. 8, (2003).
- [6] L. Zhao, "Supervision and Management on Cross-national Activities of Net work Banks", Hunan: Journal of Hunan Padio and Febevision University, vol. 4, (2004).
- [7] S. Sen, "Internet Bank", China Financial Publishing House, vol. 1, (2004).
- [8] W. J. Dai, "Prevention and management of Internet-banking risks", Shanghai: Shanghai Finance, vol. 1, (2005).