Study on prediction of sintering drum strength under small sample lacking information

Qiang Song[1*], Ai-min Wang[2]

[1]*Mechanical Engineering Department, Anyang Institute of Technology, Anyang 455000, Henan, China*

[2]*Computer and Information Engineering College, Anyang Normal University, Anyang 455000, Henan, China*

Received 1 March 2014, www.tsi.lv

Abstract

The paper provides a grey model and support vector machine algorithm and method for prediction of sinter drum strength based on the characteristics of large time delay, strong coupling, nonlinear, sintering process, put forward a kind of Combination forecasting model of drum strength based on grey model and support vector machine, the drum strength of sinter ore Laboratory values as output variables, the variables associated with the drum strength of sinter as input variables, using support vector machine powerful machine learning method and strong nonlinear fitting ability, so as to establish a stable, high precision of drum strength, the drum strength stronger generalization ability of the forecasting model, the method of the method has the high prediction accuracy, fast and convenient, and has great popularization and application value, and lay a good foundation for the green sintering technology of sintering.

Keywords: GM(1,1), LS-SVM, drum strength, Prediction

1 Introduction

The sintering process of iron and steel enterprises is the powdered iron material (such as Brazil, South Africa ore , India) with adding a certain proportion of the flux and fuel ignition and combustion, tiled in large sintering machine, produces a certain amount of fuel combustion with high temperature liquid phase, the other of un-melted particles bonded together, after cooling into porous block ore has a certain strength, as the blast furnace smelting of raw materials. The sinter is always the major raw materials for blast furnace at home and abroad, especially in China, the blast furnace sinter have accounted for more than 90%, the sinter output and quality directly affects the quality and quantity of indexes of iron-making and steelmaking. Therefore, the sintering production occupies an important position in China's iron and steel enterprises. At the same time, with China's accession to the WTO organization, the iron and steel industry in China has joined the ranks of international competition, adjusting the requires of the iron and steel industry structure and technological transformation, learn the advanced experience of foreign countries, sintering, iron-making raw material industry to accelerate the reduction of development have become an irreversible trend toward large-scale set, and it is imperative that domestic iron and steel enterprises to the international advanced iron and steel enterprises, to improve the control level of the existing sintering process to the advanced international level as soon as possible. From the control perspective, the sintering process are multivariable, nonlinear, large delay, strong coupling characteristics of complex controlled object, it is related to the temperature, pressure, speed, flow, a large number of physical parameters, including the complex process of physical change, chemical changes, and the distribution of gas in the solid material layer, the temperature field distribution etc. many aspects of the problem. Artificial traditional control method has been unable to meet the requirements of large sintering machine control; there is an urgent need for more accurate control method, stable to ensure normal operation of sintering production. Therefore, this paper uses the grey support vector machine algorithm to predict sinter tumbler strength, achieved satisfactory results.

1.1 SINTERING PROCESS

Sintering is a method that makes powdered materials (such as fine ore or preparation concentrate) into block mass under conditions involving incomplete fusion by heating to high temperature. Its production is sinter, which is irregular and porous. The following parts are usually included in sintering process: acceptance and storage of iron-containing raw materials, fuel and flux; crushing and screening of raw materials, fuel and flux; batching, mix-granulation, feeding, ignition and sintering of mix material; crushing, screening, cooling and size-stabilization of sinter. The flowchart is shown in Figure 1.

* ***Corresponding author*** e-mail: songqiang01@126.com

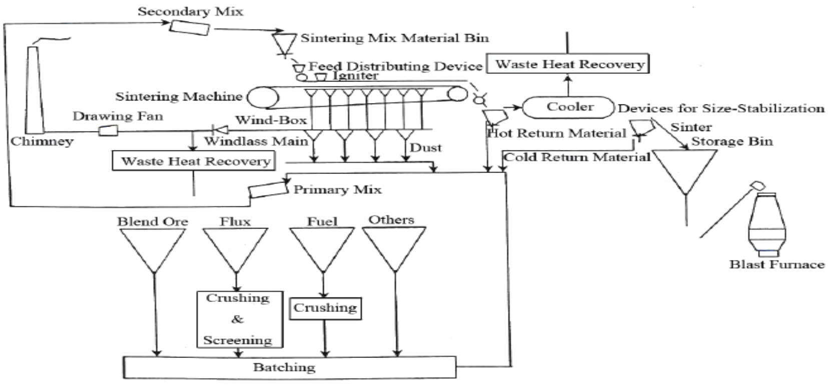

FIGURE 1 Sintering process

2 Grey model

2.1 GREY RESIDUAL ERROR CORRECTION MODEL

The term grey System is first published in a paper by Professor Deng Ju-Long, he presented the idea of "Control Problem of Unknown System" in his paper at the Sino-US Conference on Control System held in Shanghai in 1981. Presently, his paper on "Control Problem grey Systems" has been published in the Journal of System and Control Letters, nothing that the grey system theory is formal declaration on the international academia then. The basic concept of grey system theory is that all the nature for the existence of known information is white while unknown information is black, and the uncertainty information between the known (white) and the unknown (black) is grey. Grey system is mainly to dig out the nature of system under lack of information. It emphasizes information supplement to the system, and full use of white information have been identified through conducting systematic relational analysis and model construction. It makes the system state change from grey to white by prediction and decision methods to explore and understand the system [5-9]

A grey system is a system that is not completely known, i.e., the knowledge of the system is partially known and partially unknown. In recent years, grey models have been successfully employed in many prediction applications. The GM (1,1) model means a single differential equation model with a single variation. The modelling process is as follows: First of all, observed data are converted into new data series by a preliminary transformation called AGO (accumulated generating operation). Then a GM model based on the generated sequence is built, and then the prediction values are obtained by returning an AGO's level to the original level using IAGO (inverse accumulated generating operation).

A grey modelling algorithm is described as follows. (1) Suppose there is a set of discrete data that is unequal intervals as follows:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(\mathrm{n})) . \tag{1}$$

Accumulate the discrete data above once to get a new serial, that is

$$x^{(1)} = (x^{(1)}(1), x^{(2)}(2), ..., x^{(\mathrm{n})}(\mathrm{n})) . \tag{2}$$

(2) The GM (1,1) model can be constructed by establishing a first order differential equation for $x^{(1)}(\mathrm{t})$ as:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u . \tag{3}$$

The equation's general solution is: $x^{(1)}(t) = Ce^{-at} + \frac{\mu}{a}$

(3) The grey parameter a and u can be obtained by using the least square method:

$$\hat{\alpha} = \begin{bmatrix} a \\ u \end{bmatrix} = (B^T B)^{-1} B^T Y_N , \tag{4}$$

where $B = \begin{bmatrix} -0.5\left(x^{(1)}(2) + x^{(1)}(1)\right) & 1 \\ -0.5\left(x^{(1)}(3) + x^{(1)}(2)\right) & 1 \\ ... & ... \\ -0.5\left(x^{(1)}(n) + x^{(1)}(n-1)\right) & 1 \end{bmatrix}$,

$Y_N = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ ... \\ x^{(0)}(n) \end{bmatrix}$ grey parameters $\hat{\alpha}$ will be substituted into the time function, then:

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{u}{a} \right) e^{-ak} + \frac{u}{a} . \tag{5}$$

(4) Dealing $\hat{x}^{(1)}(k)$ for derivative and return to original equation then obtain

$$\hat{x}^{(0)}(k+1)=(1-e^{a})\left(x^{(0)}(1)-\frac{u}{a}\right)e^{-ak}. \quad (6)$$

(5) Calculating the difference of $x^{(0)}(k)$ and $\hat{x}^{(0)}(k)$ and the relative error $\varepsilon^{(0)}(k)=x^{(0)}(k)-\hat{x}^{(0)}(k)$.

The residual GM(1,1) model could be established to improve the predictive accuracy of the original GM(1,1) model. The modified prediction values can be obtained by adding the forecast values of the residual GM(1,1) model to the original $\hat{x}^{(0)}(k)$. However, the potency of the residual series depends on the number of data points with the same data, which is usually small when there are few observations. In these cases, the potency of the residual series with the same data may not be more than four, and a residual GM(1,1) model cannot be established. Here, we present an improved grey model to solve this problem.

2.2 RESIDUAL FORECASTING MODEL

To evaluate modelling performance, we should do synthetic test of goodness

$$C=\frac{s_2}{s_1}, \quad (7)$$

where $s_1=\frac{1}{n}\sum_{k=1}^{n}\left(x^{(0)}-\bar{x}^{(0)}\right)^2$, $s_2=\frac{1}{n}\sum_{k=1}^{n}\left(\varepsilon(k)-\bar{\varepsilon}\right)^2$ deviation between original data and estimating data:

$$\varepsilon^{(0)}=(\varepsilon(1),\varepsilon(2),...,\varepsilon(\mathrm{n}))=$$
$$(\hat{x}^{(0)}(1)-\hat{x}^{(0)}(1),\hat{x}^{(0)}(2)-\hat{x}^{(0)}(2)-,...,\hat{x}^{(0)}(\mathrm{n})-\hat{x}^{(0)}(\mathrm{n})).$$

$$P=P\{|\varepsilon(k)-\bar{\varepsilon}|<0.674581. \quad (8)$$

The precision grade of forecasting model can be seen in Table. 1.

Finally, applying the inverse accumulated generation operation (AGO), we then have prediction values $\hat{x}^{(0)}(k)=\hat{x}^{(1)}(k)-\hat{x}^{(1)}(k-1)$.

TABLE 1 Precision grade of forecasting model

Precision grade	P	C
Good	0.95≤p	C≤0.35
Qualified	0.80≤p<0.95	0.35<C≤0.5
Just	0.70≤p<0.80	0.5<C≤0.65
Unqualified	p<0.70	0.65<C

2.3 ESTABLISH THE NEURAL NETWORK MODEL

This paper will change the variables related to the research index to carry on the estimation respectively, which will obtain a few estimated values as the input of BP neural network and adopt a hidden layer that the transfer function is the sigmoid function, while the output of grey neural network is real examination values of the alkalinity, the model structure shows Figure 1.

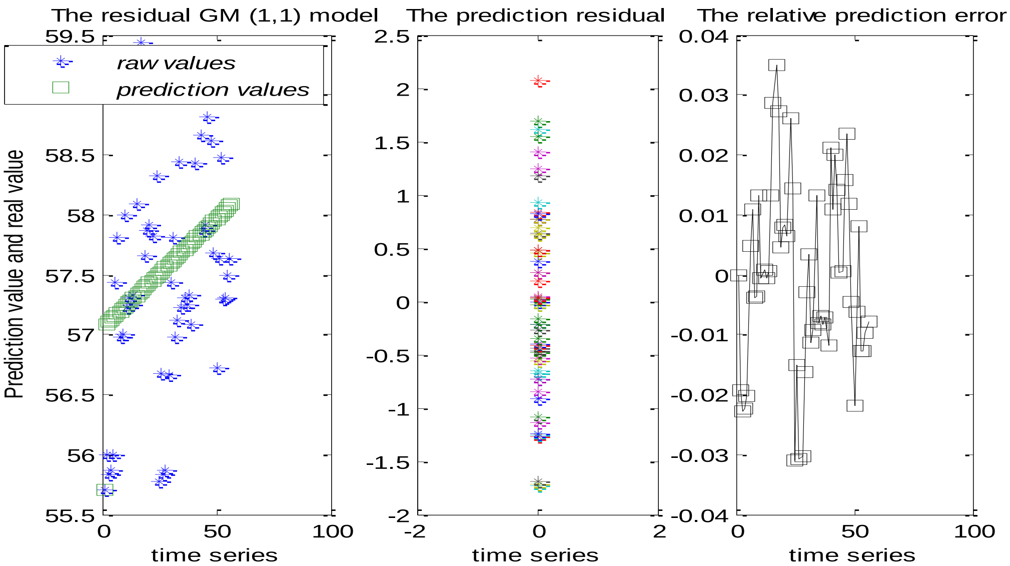

FIGURE 2 The prediction of Drum strength based on Grey residual error correction model

This paper adopts the grey neural network to predict the drum strength, aiming at this important output index, in the whole craft process, synthesizes the variables related to the drum strength make sure that ten important input variables as the input of GM(1,1), such as the layer thickness the trolley speed the first mixing water rate the mixing temperature the content of SiO2 in the mineral the content of CaO in the mineral the content of FeO in the mineral the second mixing water rate the proportion of CaO the proportion of Coal. This ten important variables(fifty-six datum) store in Excel database .In Matlab6.5,if we use the "import wizard", we may easily embed datum of Excel database into Matlab6.5,as long as we input the database name in Matlab window, we may employ the database respectively.

As can be seen, the grey theory to predict the residual correction, its accuracy is not improved, the grey model prediction accuracy of linear system is relatively high, but for the grey theory of nonlinear system, frequent changes in the still is not be desired, and therefore need to use other algorithms.

3 LS-SVM prediction model

3.1 LS-SVM PREDICTION model

The LS-SVM, evolved from the SVM, changes the inequality constraint of a SVM into all equality constraint and forces the sum of squared error(SSE) loss function to become an experience loss function of the training set [8]. Then the problem has become one of solving linear programming problems. This call be specifically described as follows. Given a training set $\{x_t, y_t\}_{t=1}^{N}$, with $x_t \in R^n$, $y_t \in R$, $x_t \in R^n$ is input vector of the first *t* samples, $y_t \in R$ is the desired output value of the first *t* corresponds to samples, N is the number of samples data, the problem of linear regression is to find a linear function y(x) that models the data. In feature space SVM models take the form:

$$y(x) = \omega^T \varphi(x) + \mathrm{b}\,, \tag{9}$$

where the nonlinear function mapping $\varphi(x): R^{\mathrm{n}} \to R^{n_{\mathrm{h}}}$ maps the high-dimensional space into the feature space.

Having comprehensively considered the complexity of function and fitting error, we can express the regression problem as the constrained optimization problem according to the structural risk minimization principle:

$$\min J(w,e) = \frac{1}{2} w^{\mathrm{T}} w + \gamma \frac{1}{2} \sum_{t=1}^{\mathrm{N}} e_t^2\,, \tag{10}$$

subject to the restrictive conditions, $y(x) = w^{\mathrm{T}} \varphi(x_t) + b + e_t$, for t = 1,…, N, where γ is margin parameter, and e_i is the slack variable for x_i.

In order to solve the above optimization problems, by changing the constrained problem into an unconstrained problem and introducing the Lagrange multipliers, we obtain the objective function:

$$L(w,b,e,\alpha) = J(w,e) - \sum_{\mathrm{t=1}}^{\mathrm{N}} \alpha_{\mathrm{t}} \{w^{\mathrm{T}} \varphi(x_t)\,, \tag{11}$$

where α_{t} is Lagrange multipliers. According to the optimal solution of Karush-Kuhn-Tucker(KKT) conditions, take the partial derivatives of (5) with respect to w, b and e respectively, and 1et them be zero, we obtain the optimal conditions as follows:

$$\begin{cases} \dfrac{\partial L}{\partial w} = 0 \to w = \sum\limits_{t=1}^{N} \alpha_t \varphi(x_t) \\ \dfrac{\partial L}{\partial b} = 0 \to \sum\limits_{t=1}^{N} \alpha_t = 0 \\ \dfrac{\partial L}{\partial e_t} = 0 \to \alpha_t = \gamma e_t \\ \dfrac{\partial L}{\partial \alpha_t} = 0 \to w^T \varphi_t + b + e_t - y_t = 0 \end{cases}. \tag{12}$$

After elimination of e_t and w, the equation can be expressed as a linear function group:

$$\begin{bmatrix} 0 & \boldsymbol{1}^T \\ \boldsymbol{1} & \varphi(x_t)^T \varphi(x_t) + \boldsymbol{D} \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}, \tag{13}$$

where $\boldsymbol{y} = [y_1, \cdots, y_N]$, $1 = [1, \cdots, 1]$, $\alpha = [\alpha_1, ..., \alpha_N]$, $\boldsymbol{D} = diag[\gamma_1, \cdots, \gamma_{\mathrm{N}}]$. Select $\gamma > 0$, and $\boldsymbol{\varphi} = \begin{bmatrix} 0 & \boldsymbol{1}^T \\ \boldsymbol{1} & \varphi(x_t)^T \varphi(x_t) + \boldsymbol{D} \end{bmatrix}$ guarantee matrix $\begin{bmatrix} \mathrm{b} \\ \boldsymbol{\alpha} \end{bmatrix} = \boldsymbol{\varphi}^{-1} \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}$.

Finally, the LS-SVM regression model can be expressed as

$$y(x) = \sum_{t=1}^{N} \alpha_t \exp\{-\| x - x_t \|_2^2 / 2\sigma^2\} + b\,, \tag{16}$$

where σ is a positive real constant. Note that in the case of RBF kernel function, one has only two additional turning parameters σ and γ, which is less than standard SVM [12].

This LS-SVM regression leads to solving a set of linear equations, which is for many application in different areas. Especially, the solution by solving a linear system is instead of quadratic programming. It can decrease the model algorithm complexity and shorten computing time greatly. The LS-SVM algorithm software package is run in MATLAB 7.0.1 software.

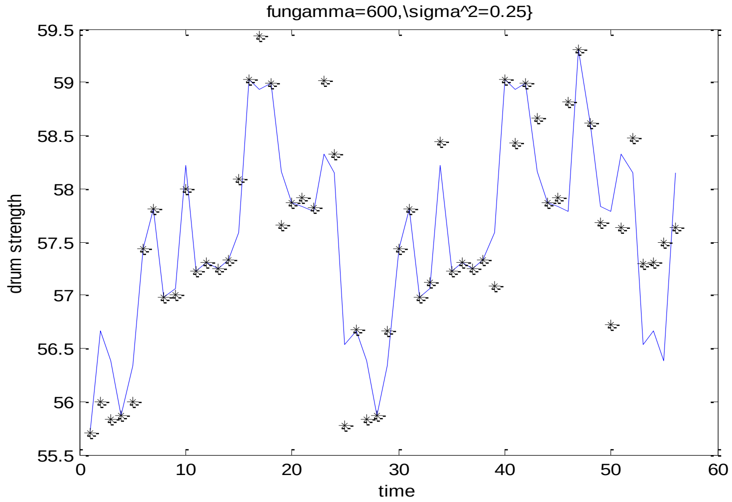

FIGURE 3 The prediction of Drum strength based on LS-SVM

3.2 COMBINATION FORECASTING MODEL.

We could see from the example that the difference value between forecasting value and practical value is large if the single model GM (1, 1) is employed to forecast. In order to make the forecasting result as close as the real value, literature [8] provide a combined grey neural network model. This kind of model could make the forecasting value as the input sample (learning sample) of the neural network, and make the real value as the target sample of the neural network. Adopting suitable structure and training on the neural network, then we could get a series of authority value threshold value, which correspond some corresponding crunodes [13 - 16].

This kind of combined method mainly use the characteristics such as three BP neural network layers carries the network which contains at least one concealed layer of S style and one linear output layer could approach any rational function. Through training, we endow the neural network the ability of simulating the relation between sequence data and sequence respectively. Meanwhile, literature provides the way that compensating the modelling error by neural network and blur logic, and doing some study training of network using heredity arithmetic. Thus, we bring forward a new combined model of grey neural network synthesizing the literature.

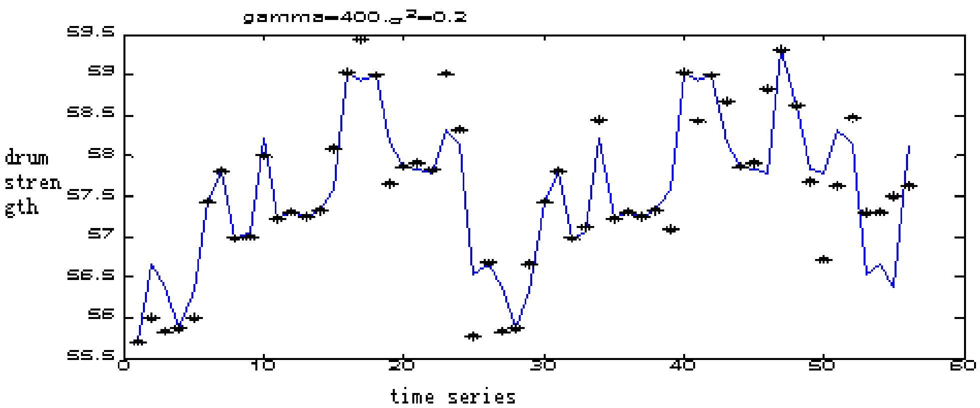

FIGURE 4 The prediction of Drum strength based on Combination

TABLE 2 The forecasting results of the three models

Model name	Mean relative error	Mean absolute error	Mean square error
Grey residual error correction model	2.09	0.65	2.888
LS-SVM	0.89	-0.042	9.02e-004
Combination forecasting model	0.03727	0.03269	8.66e-006

Figure.2-4 show the forecasting results of the three models respectively, respectively from Figure.2 ,it can be seen that although Grey residual error correction model has the better forecasting precision in drum strength the accuracy starts to decline obviously in some data points. In figure.3, the forecasting curve of LS-SVM is more close to observed curve than the curve of Grey residual error correction model from Figure.4, it can be seen that

LS-SVM as the highest forecasting accuracy. This combination method is better than the single method, which can get optimal combination prediction model from Table 1 it can be seen that three models present quite satisfactory forecasting results. By comparing the Mean relative error, mean absolute error and Mean square error of Combination forecasting model is smaller than that of Grey residual error correction model and LS-SVM. Moreover, combination-forecasting model has higher precise prediction than grey residual error correction

4 Conclusion

The original GM(1,1) model is a model with a group of differential equations adapted for variance of parameters, and it is a powerful forecasting model, especially when the number of observations is not large. In this paper, we have applied an improved grey GM(1,1) model by using a technique that combines residual modification with ls-svm. Our study results show that this method can yield more accurate results than the original GM(1,1) model and also solve problems resulting from having too few data, which may lead the same data residuals lower than four and violate the necessary condition of setting up a GM(1,1) model. The improved grey models were then applied to predict the drum strength in sintering process. Finally, through this study, our Combination forecasting model, is an appropriate forecasting method to yield more accurate results than the original GM(1,1) model and LS-SVM. In short, the Combination forecasting model is effective with the advantages of high precision, less requirement of samples and simple calculation. The grey neural network will be greatly applied and extended in future. However, we should admit that the combined grey neural network model is not perfect both in theory and practice. We will continue to do further study and some discussion based on the grey neural network model in future.

Acknowledgments

This thesis is greatly indebted to all who have instructed us in the course of our studies for the Master's degree, including professors WANG Ai-min, whose excellent lectures help brighten and broaden our vision to a great extent. In particular, we would like to extend our heart-felt thanks to our supervisor, professor YANG Cheng-zhi, for his meticulous instruction in both my study and life. We owe professor YANG Cheng-zhi the choice of the subject and the rectification of language errors in the thesis. Furthermore, our thanks are expressed to our roommates GUO Yi-jun and FENG Meng who help us with great encouragement, and their rich technology knowledge gives me great assistance in enlarging my point of view.

References

[1] FAN Xiao-hui, WANG Hai-dong 2002 *Mathematical model and Artificial Intelligence of sintering process* Central South University Press *(in Chinese)*

[2] Guo Wei-qiang, Yan Fei, Han Ning 2012 Wildfire Smoke Detection using Dynamic Features and Neural Network *JDCTA* **6**(11) 61-8 *(in Chinese)*

[3] SHI Meifeng, HE Zhongshi, LIU Xin, HUANG Meiyan 2012 A Distance Weight Object Tracking Method based on Combining Mean Shift and GM(1,1) *JDCTA* **6**(1) 318-25 *(in Chinese)*

[4] Zhipiao Liu, Qibo Sun, Shangguang Wang, Fangchun Yang 2012 The Performance Prediction of Cloud Service via JOGM(1,1) Model *AISS* **4**(5) 70-7 *(in Chinese)*

[5] Rui-lin Xu, Xin Xu, Xing-zhe Hou, Bo Zhu, Min-you Chen 2012 A Prediction Model for Wind Farm Power Generation based on Genetic-neural Network *JCIT* **7**(14) 11-9 *(in Chinese)*

[6] Deng J L 2002 *Grey Forecasting and grey Decision* Press of Hua Zhong University of Science and Technology *(in Chinese)*

[7] Rong-fei Ma 2011 Adaptive Error Control Mechanism based on GM(1,1) model and Cross-layer Design for Video Streaming over Wireless *JDCTA* **5**(8) 215-21 *(in Chinese)*

[8] Ching-Lien Huang, Yung Hui Chen, Tian-Long John Wan 2012 Optimization of Data Mining in Dynamic Environments based on a Component Search Neural Network Algorithm *JCIT* **7**(7) 216-23 *(in Chinese)*

[9] Shaio Yan Huang , An-An Chiu, Bau-Chi Wang 2012 Applying Intellectual Capital on Financial Distress Prediction Model in Taiwan Information Technology and Electronic Industry *IJACT* **4**(8) 270-80 *(in Chinese)*

[10] Xiaodong Wu, Han Wu, Guoqing Han, Yongsheng An 2012 Predicting the Solubility of Sulfur in Hydrogen Sulfide Using a Back-Propagation Neural Network *IJACT* **4**(8) 281-7 *(in Chinese)*

[11] Han-Chen Huang 2012 Using Artificial Neural Networks to Predict Restaurant Industry Service Recovery *IJACT_CST* **4**(10) 315-21 *(in Chinese)*

[12] Fuwei Zhang, Qin Su 2012 Application of Wavelet Neural Network into the Sustainable Development of Power Market *AISS* **4**(10) 269-77 *(in Chinese)*

[13] NIU Dong-xiao, LV Hai-tao, ZHANG Yun-yun 2008 Combination method of mid-long term load forecasting based on support vector machine within the Bayesian evidence framework *Journal of North China Electric Power University* **2008**(6) 62-6 *(in Chinese)*

Authors

Song Qiang, born in Xintai city of Shandong province

Current position, grades: Associate professor of Anyang Institute of Technology.
University studies: Kunming University of science and technology in 2006.
Scientific interest: areas of intelligence control theory and rough set theory.

Wang Ai-min, born in 1957

Current position, grades: professors of Anyang normal college.
University studies: doctor's degree from Wuhan institute of technology.
Scientific interest: SVM and rough set of intelligent decision-making research.