The Technique of Gas Disaster Information Feature Extraction based on Rough Set Theory

Hui Li
Department of Computer Engineering, Huai Hai institute of Technology, Lianyungang, China
E-mai:shufanzs@126.com

Shu Zhang
Shool of Business, Huai Hai institute of Technology, Lianyungang, China
E-mail: shufanzs@sina.com

Xia Wang
Department of Computer Engineering, Huai Hai institute of Technology, Lianyungang, China
E-mai:lijunmagic@yahoo.com.cn

***Abstract*—Gas accident has become the main contradiction which constrains the safety production of coal mine. It is one of the import issues in the field of coal mine safety to identify and analyze the dangerous factors that lead gas accidents, and establish an effective early-warming support system of gas in coal mine at present. In view of the characteristics of coal mine gas disaster, a high efficient gas disaster feature extraction algorithm based on rough set is proposed, which includes two phases: The algorithm refine the gas disaster information matrix using dimensionality reduction, then uses the entropy and maximum entropy to establish data mining model of gas disaster prediction. The effectiveness and practicality of rough set theory in the prediction of gas disaster and feature extraction was confirmed through practical application.**

***Index Terms*— gas outburst, feature extraction, rough set, gas disaster, gas prediction**

I. INTRODUCTION

The disaster caused by mine gas outburst is extremely serious, so its prediction has been widespread concerned around the world. Gas disasters, especially the outburst of gas, are very dangerous and complex phenomena in coal mines. Outbursts are hazardous through the mechanical effects of particle ejection and by asphyxiation from the gas produced. The violence of an outburst has frequently tossed miners back several meters from the face of a heading. In a number of cases the dislodged coal, which frequently consists of small particles, has engulfed the operators, preventing them from escaping while the released gas asphyxiated them. So it is necessary and significant to develop a useful, accurate method to monitor the potential warning signs of outbursts, and use an intelligent theory to predict the disaster in advance.

Manuscript received May 25, 2012; revised June 1, 2012; accepted July 1, 2012.

Supported by the technology open fund project of Jiangsu Province R&D Institute of Marine Resources (No. JSIMR11B12).

Corresponding author: Hui Li, email: shufanzs@126.com

The coal bed conditions vary a great deal in China, and a coal and gas outburst is the synthetic result of stress, gas pressure, the physical mechanics of coal and so on. With the development of computer technology and artificial intelligence, the method for predicting the mine methane has also been greatly developed, such as the chaotic time series prediction[1-4], gray relational analysis[5-7], and neural network prediction and so on[8-9]. The most existing methods for prediction of gas delay outburst adopt the BP neural network [10-13]. However, BP neural network is difficult to select parameters, neuron nodes in hidden layer etc., seriously affecting the performance of BP. To improve the performance of neural network, people proposed many types of neural networks, such as RBF neural network, wavelet network [14-15], fuzzy neural networks [16-17]. Although the wavelet basis with excellent local time-frequency characteristics and multi-scale analysis inherits from wavelet analysis adapts to describe short-term, high-frequency signal and the non-uniform sampling data [18], how to select parameters of the radial wavelet-based network as well as the structure of network and training is a challenging task. Most methods determine the parameters of wavelet basis in the framework of wavelet according to the distribution of samples. Then AIC (Alaikes information Criterion) integrating with other algorithms, such as genetic algorithms, Gram-Schmit algorithm, the gradient descent etc., adjust the parameters of WN [19]. However, there are some problems which exist in these methods, such as improper initial parameters, long training time, learning criteria based on experience risk minimization. After analyzing all existing methods of outburst prediction, we conclude that the intelligent method of patterning the human brain is the only potent approach which enables us to consider the multiple associated factors and make a precise prediction.

The effective method of preventing gas disasters is to effectively extract the characteristics of gas disasters in a lot of gas monitoring information, early recognition and find, capture gas hazard information, provide the basis to

doi:10.4304/jcp.8.4.983-989

prevent disasters. Feature extraction is the key technology of completing the judgment and identifying the gas disasters. Mine gas disaster feature extraction and recognition technology is one of the main techniques in the field of coal mine safety [20].

Because the correct identification of gas hazard information is based on the effective feature extraction, the more sophisticated feature extraction of the characteristic parameters set in the gas disasters, the higher probability of the gas disasters recognition. This paper proposes a novel gas disaster feature extraction algorithm based on rough set. The main idea of our method can be divided into two steps: Firstly, we optimize the gas disaster information matrix through attribution reduction; secondly, we construct the feature extraction model of gas disaster information using the entropy and max entropy of information entropy.

II. Preliminary Knowlegment

A. Principles used in Coal and Gas Outburst Focasting System

From the present situation of research on coal and gas outburst, there is insufficient information about the development of coal and gas outburst process; its variation rules are unclear to us, so it is difficult to analyze quantitatively. Outburst prediction process is a typical "lack of information" and "information uncertainty" system; it is a so-called grey system.

(1) Entirety principle

Subsystem or elements constitute a system according to a certain combination. New combination effects emerge. The article is committed to get a positive synergy. Specific to the coal and gas outburst prediction, avoid various prediction methods "in their own ways"; all of them are combined as an organic whole.

(2) Optimization theory

Even if the system has the same elements, different structures may achieve different functions. Therefore, for the established system, research on the principles of labor division and cooperation what internal factors accorded are important. It is a way to enhance system function and performance. In a later section, we will analyze how the various forecasting methods get labor division and cooperation to achieve the optimal requirements.

(3) Dynamic principle

In the process of coal and gas outburst prediction, we will encounter new situations, such as application of new prediction method or new mining methods, greater changes of mining environment, and so on. This requires that the system is capable of evolving.

(4) Feedback principle

The output information will be sent back to the input and get impacts on it. Feedback makes causes and consequences interaction to complete a common function. Feedback produces dynamic unification of the thing itself with the surrounding environment. In the process of coal and gas outburst prediction, the prediction results are obtained through the prediction methods working on the coal mine environment. The feedback prediction effects can be used to amend coal and gas outburst spatiotemporal forecasting system. According to the feedback prediction results, adjust the forecast focus and dynamical schedule forecasting equipment and human resources.

B. Issue Description and Related Works

Through the analysis of the monitoring data of coal gas, the basic feature can be expressed by the matrix that is the original features. Let m denote the number of gas hazard information acquisition channel, and each channel chooses n sample data, so form a $m \times n$ monitoring data, the matrix can be represented as follows:

$$R = (x_{ij}) \quad (\mathrm{i} = 1,2,...\mathrm{m};\ \mathrm{j} = 1,2,...\mathrm{n})$$

The data need to be preprocessed before conduct data reasoning. Usually, The data and knowledge need to be expressed by the information table, and construct the gas disaster features knowledge base. We will map the gas hazard information matrix to an incomplete information system, the concept of incomplete information system is as following:

Formally, a data set or an information system is a quadruple $S = < U, A, V, f >$,where U is a non-empty finite set of objects, called a universe, A is a non-empty finite set of features, V is the union of feature domains, $f: U \times A \to V$ is an information function ,which make $\forall X \in U, a_i \in A, f(X, a_i) \in V_a$ .We can split set A of features into two subsets: $C \subset A$ and $D = A - C$, conditional set of features and decision features ,respectively. Information system can be written briefly IS=(U,A).

Following we redefine and interpret the concept of information system in the rough set. Given an information system $S = < U, A, V, f >$, $U = \{x_1, x_2, ..., x_m\}$ denotes the universe of the gas disaster (object), for any subset $X \subseteq U$ called a concept or domain in U; $A = C \cup D$ is the attribution set, subset C and subset D denote the condition attribute and decision attribute respectively; V_{a_i} denote the value scope of attribute a_i ,V denote the union of attribute value; $f: U \times A \to V$ denote a information function, and assign the attribute value to each x in U .

The task of feature extraction and selection is to discover the most valid feature of classification. Sometimes specified ration evidence is needed to measure the validity of feature classification. There are many kinds of probability when mapped from high dimension feature space to low dimension feature space. A comparing standard is essential for selecting which mapping transformation is more efficient. Moreover, after choosing low dimension feature, such association probability of features is not unique. So a comparing norm is also needed to evaluate which association is the optimized classification.

The feature extraction method based on Euclid distance has its advantages but it cannot depict different types' approximate distribution of probability. It can not exactly show different types of overlapping modes. Therefore, it cannot directly get in touch with the error ratio. Feature extraction method based on approximate evidence should be taken into consideration.

In the time domain, outburst prediction is divided into two parts: one is early comprehensive forecast and the other is real-time forecast. Multiple indexes method [21] is used in the early comprehensive forecast. The prediction method used in the real-time forecast is as follows: use change law of gas signals, index of drilling cutting's quantity, drilling gas inrush initial velocity method and risk degree of outburst warning [22]. In the space domain, the prediction is divided into three parts: district prediction, local forecast and point prediction. Multiple indexes method is used in district prediction. Local prediction adopts multiple indexes method and prediction using change law of gas signals. Point prediction adopts index of drilling cutting's quantity, drilling gas inrush initial velocity method and risk degree of outburst warning.

Sun Jian et al. [23] propose a fuzzy neural network risk prediction model for rock burst trained with the improved BP algorithm based on the typical rock burst data. This method is an improvement of comprehensive index judgment and multi-index judgment with fuzzy mathematics. Practical engineering applications in Sanhejian Coal Mines indicate that this method is not only precise and simple, but also intelligent, with the predicted results well agreeing with the practical conditions.

Xinyu Li [24] propose an adaptive wavelet networks algorithm for prediction of gas outburst. First, adaptive clustering algorithm is first used to determine initial parameters of wavelet network according to the results of the clustering. Then genetic algorithm and SVM-RFE is adopted to tune the structure of the wavelet network and adjust the network parameters to improve generalization performance.

Hua Fu et al. [25] established a coal and gas outburst spatiotemporal forecasting system by using system engineering theory, combined with the current mine production conditions and based on the coal and gas outburst composite hypothesis. This system can guide forecasting work schedule, optimize prediction technologies, carry out step-by-step prediction and eliminate hazard hierarchically. From the point of view of application, the proposed system improves the prediction efficiency and accuracy.

Tiezhu Qiao el al. [26] proposed the weighted LS-SWM to improve sparse and robustness and its time series prediction model is used to analysis short-time mine working face gas emission. Under MATLAB2009b environment, using LS-SVM1.7 toolbox, specific algorithm model is established, further model is verified by Hebi 10th 1113 mine and gas outburst working face time series data

III. The Gas Disaser Feature Extraction Algorithm

A. The Dimension Deduction of Feature Space

It is well known that an information system or a decision table may usually have irrelevant and superfluous knowledge, from which it is inconvenient for us to get concise and meaningful decision. To acquire brief decision rules from inconsistent decision systems, attribute reduction is needed. In order to extract the most effective feature for recognizing the gas disaster informing among the gas disaster information feature set, we need to optimize the gas disaster information feature matrix. The attribute deduction is a good way. Following, we discuss the core attribute selection combing the rough set and attribute deduction.

In this paper, we present reduction algorithms based on the principle of Skowron's discernibility matrix. The information in the information table (also called decision table) relevant to attribute discriminate are concentrated in a matrix (called Discernibility Matrix) in such method, and to calculate the core attribute through the discernibility matrix. Our core attribute selection algorithm is based on the discernibility matrix, and we also study the other attribute combination as well as the core attribute in the discernibility matrix, and utilize disjunctive normal form to conduct attribute deduction.

The discernibility matrix is defined as:

Definition 1: Given an information system $IS=(U,A)$, $U=\{x_1,...,x_n\}$ we can split set A of features into two subsets: conditional set of features $C=\{c_1,...,c_m\}$ and decision features D: A=C∪D. Let $c_i(x_j)$ and $D(x_j)$ express the value of data point x_j on conditional attribute set C and decision attribute set D respectively. The value of every element in the discernibility matrix is defined as follows:

$$M_{i,j}=\begin{cases}0, & D(x_i)=D(x_j)\\ -1, & \forall c\in C, c(x_i)=c(x_j), D(x_i)\neq D(x_j)\\ \{c\in C: c(x_i)\neq c(x_j)\}, & D(x_i)\neq D(x_j)\end{cases}$$
$$i,j=1,...,n \tag{1}$$

This matrix states that the element value is 0 when the decision attribute is the same; the element value is different attribute combination when the decision attribute is different but can be distinguished by some condition attribute; the element value is negative one when the decision attribute is different but the condition attribute is the same which states that the data is wrong or the condition attribute is insufficient.

Considering the discernibility matrix contains all the attribute discernible information in the decision table, so as well as the core attribute, the value attribute should be acquired from the matrix element which the number of attribute combination is not one. Supposing the information table have two attribute combination as well as C_0, which is referred $t_{11},...,t_{1e}$ and $t_{21},...,t_{2k}$. To construct the expression: $P=(t_{11}\vee...\vee t_{1e})\wedge(t_{21}\vee...\vee t_{2k})$

and the attribute combination denoted by such a expression can distinguish all the decision in the original decision table. If the information table also has more attribute combination besides C_0 , the processing method may be deduced by analogy. Owing to the disjunctive normal form is composed of multiple conjunctive normal form, which attribute combination that should be used may be determined according to requirement and this attribute combination combine with the core attribute constitute the optimal attribute deduction under a given requirement. In this paper, we adopt the simplest attribute combination to serve as the finally result of attribute deduction. From the above discussion, we will propose an attribute reduction algorithm as follows:

Algorithm 1:The algorithm of attribute deduction

Input: decision table $IS = (U, A, V, f)$, where $A = C \bigcup D$ is attribute set;

Output: a set Ω of attribute set after deduction。

(1) Calculate discernibility matrix MD of decision table;

(2) Add the attribute with the number of attribute combination is 1 to the core attribute set C_0.

(3) Add the core attribute to attribute set Ω after attribute deduction, that is $\Omega = C_0$.

(4) Find out all the attribute combination Q not including core attribute, that is $Q = \{B_i \mid B_i \bigcap \Omega = \Phi, i = 1,..., s\}$;

(5) The attribute combination Q is denoted by the format of conjunctive normal form, that is: $P = \wedge\{\vee B_i, i = 1,..., s\}_{k=1,...,m}$;

(6) Transform P to disjunctive normal form.

(7) Select the simplest attribute combination and add it to set Ω ;

(8) Output attribute set Ω .

B. The Gas Disater Information Feature Extraction based on Maximum Entropy

The basic idea of the maximum entropy principle is: when we only master the distribution of the unknown part of knowledge, we should choose the probability distribution with the maximum entropy and satisfy the known knowledge. That is, To choose a statistical model, which satisfies all the known facts and not to do any hypothesis to the unknown facts. This principle reflected in our estimates of parameters $p(y|x)$, where x denotes the gas disaster happened condition, y denotes the gas hazard events, so the joint probability of x and y note for $p(x, y)$.

Feature selection is based on the sample data, and the sample data can be expressed by $(x_1,y_1),(x_2,y_2),\cdots,(x_n, y_n)$.where xi denotes decision attribute, y_i denotes categorical attribute used as class label provided by the expert. The training data can be expressed by using maximum likelihood estimation, that is:

$$\bar{p}(x, y) \equiv \frac{freq(x, y)}{\sum_{x,y} freq(x, y)} \tag{2}$$

Where $freq(x, y)$ denotes the number of times that (x, y) occurs in the sample.

Definition 1. feature function(or function for short): It is usually a binary value function $f(x,y) \to \{0,1\}$, for example, To the gas hazard information feature extraction problem, we may define the characteristic function as follows:

$$f(x, y) = \begin{cases} 1 & (\text{x} = \text{gas outburst feature}) \wedge \\ & (y = \text{abnormal vibration}) \\ 0 & \text{other types of gas disaster} \end{cases} \tag{3}$$

The expected value of feature function f_i with respect to the empirical distribution $\bar{p}(x, y)$ is exactly the statistic we are interested in. We denote this expected value by

$$E_{\bar{p}} f_i = \sum_{x,y} \bar{p}(x, y) f_i(x, y) \tag{4}$$

When we discover a statistic that we feel is useful, we can acknowledge its importance by requiring that our model accord with it. We do this by constraining the expected value that the model assigns to the corresponding feature function. The expected value of feature function f_i with respect to the model distribution $p(y|x)$ is

$$E_p f_i = \sum_{x,y} \bar{p}(x) p(y|x) f_i(x, y) \tag{5}$$

Where $\bar{p}(x)$ is the empirical distribution of x in the training sample. We constrain this expected value to be the same as the expected value of f_i in the training sample. That is, we require

$$E_{\bar{p}} f_i = E_p f_i \tag{6}$$

Combining (4),(5) and (6) yields the more explicit equation

$$\sum_{x,y} \bar{p}(x) p(y|x) f_i(x, y) = \sum_{x,y} \bar{p}(x, y) f_i(x, y) \tag{7}$$

We call the requirement (6) a constraint equation or simply a constraint. By restricting attention to those models $p(y|x)$ for which (6) holds, we are eliminating from consideration those models which do not agree with the training sample on how often the output of the process should exhibit the feature f .

To sum up so far, we now have a means of representing statistical phenomena inherent in a sample of data (namely, $E_{\bar{p}} f_i$), and also a means of requiring that our model of the process exhibit these phenomena(namely, $E_{\bar{p}} f_i = E_p f_i$).

Suppose that we are given n feature function f_i (i=1,2,…,n), so the model belong to the model set subject to the constraint can be formally defined as follows:

$$P=\{p \mid E_{\bar{p}} f_i = E_p f_i \text{ , i}=1,2,\wedge,\text{n}\} \quad (8)$$

There are many model meet the constraint condition, the goal is to produce model with the most uniform distribution model under the constraint set. Among the models which meet the constraint, the maximum entropy philosophy dictates that we select the distribution which is most uniform. But now we face a question what does "uniform" mean?

A mathematical measure of the uniformity of a conditional distribution $p(y|x)$ is provided by the conditional entropy

$$H(p) \equiv -\sum_{x,y} \bar{p}(x)p(y|x)\log p(y|x) \quad (9)$$

So, we may choose the model with maximum entropy as our goal model p^*, which is formally represented as follows:

$$p^*(y \mid x) = \arg\max\{-\sum(\bar{p}(x)p(y|x))\log(\bar{p}(x)p(y|x))\} \quad (10)$$

To address the general problem, we apply the method of Lagrange multipliers form the theory of constrained optimization.

$$p^*(y \mid x) = \frac{1}{Z(x)}\exp(\sum_i \lambda_i f_i(x,y)) \quad (11)$$

Where $Z(x)$ is a normalizing constant:

$$Z(x) = \sum_y \exp(\sum_{i-1}^{n} \lambda_i f_i(x,y)) \quad (12)$$

Where λ_i is feature parameters denoted the importance of every feature function. Through studying in the training set to get value, we can get the optimal solution and complete the construction of maximum entropy model. The value of λ_i and $p(y|x)$ can be got through a iterative algorithm, and the relevant steps are outlined here:

Algorithm 2. computing the parameters of λ_i and $p(y|x)$

Input: Feature functions $f_1, f_2,..., f_n$;empirical distribution $\bar{p}(x,y)$

Output: Optimal parameter values λ_i^* ; optimal model p_λ^*

① Start with λ_i =0 for all i∈ {1, 2,…, n}

② For i=1 to n

Let $\Delta\lambda_i$ be the solution to

$$\sum_{x,y} \bar{p}(x)p(y|x)f_i(x,y)\exp(\Delta\lambda_i f^{\#}(x,y)) = E_{\bar{p}} f_i$$

And draw the result: $\Delta\lambda_i = \frac{1}{M}\log\frac{E_{\bar{p}} f_i}{E_{P_\lambda} f_i}$

Where $M = \sum_{i=1}^{n} f_i(x,y)$;

③ $\lambda_i \leftarrow \lambda_i + \Delta\lambda_i$;

④ Go to step ② if not all the λ_i have converged very carefully check a printed copy.

IV. Experiments And Results

A. Experiment Data

In order to verify the validity of the algorithm proposed in this paper, we experiment t he feature extraction model based on the maximum entropy principle. The monitoring data research in our experiment collected from the ZhangShuangLou mining area of XuZhou coal mining group, among them the data of coal bed gas content, pressure, temperature, humidity, coal, electric interface and constant density of monitor are collected from the daily monitoring data of mining area.

In our experiments, we chose twenty-eight parameters : GAS-D(DaHang gas concentration), GAS-H(Mining face of gas concentration),GAS-J(Heading face of gas concentration), WIN-C(Wheel-dreven wind speed at the bottom of a well), TE-M(Temperature coal seam), TE-H(Mining face temperature), HU-M(Coal water content), HU-H(Mining face humidity), WIN-F(Face wind speed), Q-WJ(Q value), GAS-P(Coal bed gas pressure), SW-E(Acoustic emission frequency), WIN-H(Mining face wind speeds), KC-MTD(Ming methods), KC-SP(Mining speed), GAS-BIN(The gas initial loose speed). The whole data set totally have 2403 records including the incomplete records, and 1320 records of full target.

B. Experiment of Feature Extraction

We chose the gas concentration from the gas monitoring data set as the experiment object, which includes two gas concentration function of GAS-J and GAS-H. Supposing they meet Gaussian distribution, and distribution density respectively satisfy $f(x_1) \sim N(0,1)$, $f(x_2) \sim N(0,4^2)$ 。 Now, define $f(x)$ as the both of weighted mix, namely:

$$f(x) = (1-p)f(x_1) + pf(x_2)$$

Where p denotes the rate of pollution of the Standard of Gaussian distribution. According to the occurred condition of gas disaster and expert experience, we set p =0.3. Using algorithm 1 to create model for $f(x)$, the experiment results are: λ_0 =-1.4746, λ_1 = -0.1014, λ_2=-0.1758, λ_3=0.003, λ_4=0.0043,and the Fitting function is:

$$f(x) = \exp(-1.4746 - 0.1014x - 0.1758x^2 + 0.003x^3 + 0.0043x^4)$$

Comparing expected gas concentration distribution used maximum entropy model and theory distribution of gas concentration, as the results shown in figure 1, we can see that the approximation effect applying analytic fitting distribution function using the maximum entropy principle and the theory distribution is good.

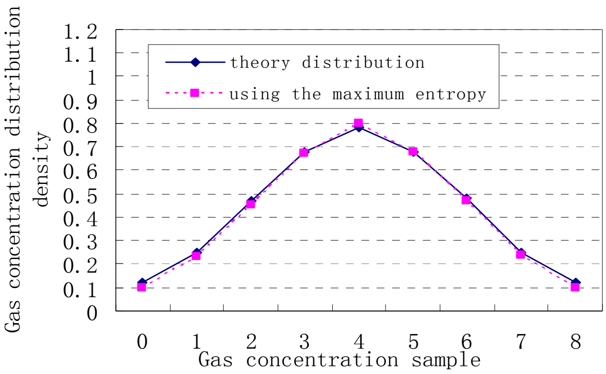

Figure 1. Maximum Entropy Model and Theory Distribution of Gas Concentration

We chose the TE-M as another experiment object, and the experimental results as shown in figure 2 shows. In the above figure, the abscissa denotes the time; y-coordinate denotes gas pressure value.

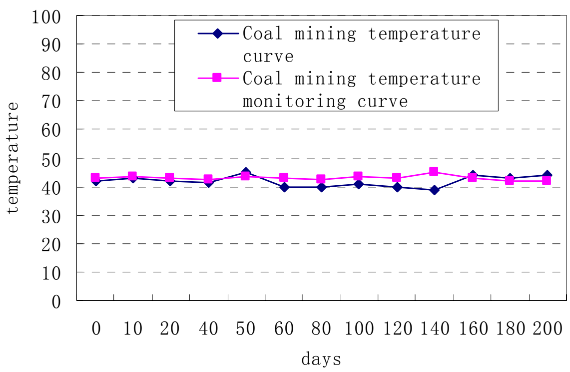

Figure 2. Coal Mining Temperature Curve and Coal Mining Temperature Monitoring Curve

C. *The Model Validation Experiment*

To choose a new data set to test the effectiveness of the maximum entropy model proposed in this paper, and the standard error is as follows:

$$E = \frac{\sum_{K=1}^{n} [y_e(k) - y_m(k)]^2}{2}$$

Where $y_e(k)$ denotes the experiment output, $y_m(k)$ denotes the actual output, n denotes the testing number. The accuracy of the prediction model based on the maximum entropy principle can reach 0.85 through the simulation,, the simulation of the error curve as shown in figure 3 :

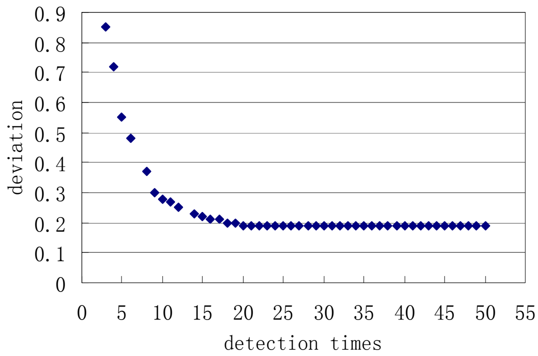

Figure 3. The validity check experiment of the maximum entropy model

V. Conclusion

This paper combine the singular value and maximum entropy principle, which is used in coal mine gas feature extraction, so as to realize the coal gas risk evaluation prediction. The research in this paper of rich warning system reliability evaluation, which provides effective decision support for the coal mine accident prevention and rescue work to provide. The results from our experiments show that our method is feasible and effective. Prediction results show that using rough set and maximum entropy principle predict outburst risk is scientific and feasible. But as a relatively new forecasting method, maximum entropy still has some defects to improve at present. For example: how to select the core attribution to increase the prediction accuracy, selection of punish coefficient etc. Therefore, to do a more accurate prediction, there is still a need for further discussion and study.

Acknowledgment

The authors would like to thank to Liu Yanwei for providing the gas time-series data. The authors thank Ma Xiaoping for providing ideas and help during the writing of this paper. Financial support for this work, provided by the technology open fund project of Jiangsu Province R&D Institute of Marine Resources. No. JSIMR11B12, the open fund project of Jiangsu Key Laboratory of Marine Biotechnology the National Natural Science Foundation of China (No.60974126) and the Natural Science Foundation of Jiangsu Province (No.BK2009094) is gratefully acknowledged.

References

[1] M.Y. Qiao, J.J. Qiao, H. Tao and X.P. Ma, "Mine Workface Gas Emission Time-Series Fractal Properties", Information Engineering and Mechatronics, vol.138, no.1, pp.1287-1296, 2011.
[2] V.A. Gromov, A.N. Shulga, "Chaotic time series prediction with employment of ant colony optimization", Expert Systems with Applications, Vol.39, No.9, pp.8474-8478, 2012.

[3] R.Chandra, M.Zhang, "Cooperative coevolution of Elman recurrent neural networks for chaotic time series prediction", Neurocomputing, Vol.86, no.1, pp.116-123, 2012.
[4] W. Liang, L.B. Zhang, M.D. Wang, "The chaos differential evolution optimization algorithm and its application to support vector regression machine", Journal of Software, Vol.6, no.7, pp.1911-1918, 2011.
[5] C.J. Tzeng, Y.H. Lin, Y.K. Yang, M.C. Jeng, "Optimization of turning operations with multiple performance characteristics using the Taguchi method and Grey relational analysis", Journal of Materials Procedding Technology, Vol.209, no.19, pp.2753-2759.
[6] R. Rao, V. Yadave, "Multi-objective optimization of Nd: YAG laser cutting of thin superalloy sheet using grey relational analysis with entropy measurement", Optics & Laser Technology, Vol.41, No.8, pp.922-930, 2009.
[7] C.F. Jeffrey Kuo, T.L. Su, P.R. Jhang, C.Y. Huang, C.H. Chiu, "Using the Taguchi method and grey relational analysis to optimize the flat-plate collector process with multiple quality characteristics in solar energy collector manufacturing", Energy, Vol.36, no.5, pp.3554-3562, 2011.
[8] Y.Hu, X.Z Mo, X.Z. Zhang, "Intelligent analysis model for outsourced software project risk using constraint based Bayesian network", Journal of Software, Vol.7, no.2, pp.440-449, 2012.
[9] X.Y. Tang , Z.G. Wang, "Predicting Gas Emission Based on Combination of Grey Relational Analysis and Improved Fuzzy Neural Network ", Advances in Intelligent and Soft Computing, , Vol.169, no.1, pp.275-280,2012.
[10] H. Li, "Pattern recognition and prediction study of rock burst based on neural network", Journal of Coal Science and Engineering ,Vol.16, No.4, pp.347-351, 2010.
[11] W.Z Cui, C.C. Zhu, W.X. Bao, "Prediction of the chaotic time series using support vector machines for fuzzy rule-based modeling", Acta Physica Sinica ,vol.54,no.7, pp.2221–2230 , 2009.
[12] X.R. Luo, F. Yang, Y.T. Kan, A.R. Zhang, "Research on Real-Time Alarm Theory of Delayed Coal and Gas Outburst", Journal of China University of Mining and Technology, vol.37,no.1, pp.163–166, 2008.
[13] Y.Z. Miao, J.W. Zhang, H.X. Zhang, "Coal and Gas Outburst Prediction Combining a Neural Network with the Dempster-Shafter Evidence", Advances in Neural Networks, Vol.5264, no.1, pp. 822-829, 2008,
[14] Z.G. Wu, J.Z. Zhang, Y.B. Tao, Y.L. He, W.Q. Tao, "Application of artificial neural nework method for performance prediction of a gas cooler in a CO_2 heat pump", International Journal of Heat and Mass Transfer, Vol.51, no.1, pp.5459-5464, 2008.
[15] H.L. Wei, S.A. BILLINGS, y.f. Zhao, L.Z. Guo, "An adaptive wavelet neural network for spatio-temporal system identification", Neural Networks, Vol.23, no.10, pp.1286-1299, 2010.
[16] X.Y. Li, "An Adaptive Wavelet Networks Algorithm for Prediction of Gas Delay Outburst", Advances in Neural Networks. VOl.1553, pp. 1000-1006, 2009
[17] C.N. Fan, K.K. Li, W.L. Chan, W.Y. Yu, Z.N. Zhang, "Application of Wavelet Fuzzy Neural Network in Locating Single Line to Ground Fault (SLG) in Distribution Lines", Electrical Power and Energy Systems, vol. 29, pp. 497–503, 2007.
[18] Y.l. Chai, "Study of Coal Gas Outburst Prediction Based on FNN", electronics and signal processing, vol.97, pp.687-693, 2011.
[19] A. Eric, Rying, L.B. Griff, L.J. Chyi, " Focused Local Learning with Wavelet Neural Networks". IEEE Trans. on Neural Netw, pp.304–320, 2009.
[20] L.H. lv, "Complex Industrial Systems Based on Wavelet Network and Robust Estimation Modeling Study". Zhejiang University, pp.23-29, 2001.
[21] J.E. Liu, F.L. Zeng, Z.L. Guo, "Gas Outburst Risk Analysis based on Pattern Recognition of RS-SVM Model", procedia engineering, vol.29, no.1, pp.170-173, 2012.
[22] H.Q. Zhu, W.J. Chang, B. Zhang, "Different-source gas emission prediction model of working face based on BP artificial neural network and its application", Journal of China Coal Society, vol.32, no.5, pp.504-508, 2007.
[23] J. Sun, L.G. Wang, H.L. Zhang, "Application of fuzzy neural network in prediction the risk of rock burst", Procedia Earth and Planetary Science, vol.1, pp. 536-543, 2009.
[24] X.Y Li, "An adaptive wavelet networks algorithm for prediction of gas delay outburst", Advance in neural networks, vol.5553, pp.1000-1006 ,2009.
[25] F. Hua, X.L. Jing, M.G. Liang, "Application of computational intelligence platform in coal and gas outburst prediction", Journal of coal science & engineering, vol.18, no.1, pp. 49-54, 2012.
[26] T.Z. Qiao, M.Y. Qiao, "Mine working face gas prediction based on weighted LS-SVM", In Proceeding of the 3th international conference on Artificial intelligence and computational intelligence, pp. 369-647, 2011.

Hui Li was born in Lianyungang, JiangSu Province, P.R.China, in October 20, 1979. She received her BE degree in computer science and technology from the Yang Zhou University, in 2007. She is a master candidate in the School of Information & Electrical Engineering in China University of Mining & Technology. She is currently a teacher in Department of Computer Engineering, Huai Hai institute of Technology, Lianyungang, JiangSu Province, P.R.China.

Her current research interest includes data mining, information processing and intelligent computing. He has published more than 20 papers in journals and conferences.

Shu Zhang, born in 1979. Received her BE degree in computer science and technology from the Yang Zhou University, in 2007. He is currently a teacher in school of business, Huai Hai institute of Technology, Lianyungang, JiangSu Province, P.R.China. His current research interests include data mining and e-commerce.

Xia Wang, born in 1977. Received her BE degree in computer science and technology from the Yang Zhou University, in 2011. She is currently a teacher in Department of Computer Engineering, Huai Hai institute of Technology, Lianyungang, JiangSu Province, P.R.China. Her current research interests include data mining and information retrieval.