Available online at www.sciencedirect.com

ScienceDirect

Procedia Earth and Planetary Science 1 (2009) 1418–1424

Procedia Earth and Planetary Science

www.elsevier.com/locate/procedia

The 6th International Conference on Mining Science & Technology

Safety evaluation of human accidents in coal mine based on ant colony optimization and SVM

Jin Zhu*, Ma Xiao-ping

The School of Information and Electrical Engineering, China University of Mining &Technology, Xuzhou 221116, China

Abstract

The man-made causes have manifested great significance in the safety evaluation of coal mine. Because factors which have been affecting human accidents are quite various and, their interaction are so highly complicated that statistical analysis can hardly be carried out, this paper has made elaborate analysis of the leading factors through four aspects: employees' qualities, organization management, facilities factors and working environment. Based upon those specific analysis, by converting feature selection of intricate multi-factors to parallel combinatorial optimization and global optimization, evaluative indices of human accident at coal mines and relevant feature extraction method have been come up with in this paper. After effective feature selection, a safety evaluation model of human accidents based upon SVM is established which can availably tackle the problem of "hard to evaluate accurately with limited samples of human accidents in coal mine". Evaluation and analysis on human accidents data of ventilating system of Caitun Mine proves that the proposed approach is effective and practicable.

Keywords: ant colony optimization; SVM; human accidents; safety evaluation

1. Introduction

Coal mine accidents occur frequently in our country. According to the statistics, more than 91% of coal mine accidents are directly or indirectly rooted in human accidents, which therefore have become the most crucial factors seriously affecting the system safety [1]. However, for the sake of various kinds of leading causes of human accidents and highly complicated interactions of them, as well as the special traits of coal mine industry, the unsubstantial basement of safety system engineering, and the difficulties of collecting statistical data accurately, it is rather limited to evaluate human accidents purely through traditional safety system engineering methods (e.g. SCL-Safety Check List, FTA-Fault Tree Analysis, AHP- Analytic Hierarchy Process, etc).

Ant Colony Optimization (ACO) is a novel evolutionary algorithm, which has been successfully applied to various NP-hard combinatorial optimization problems. It possesses many desirable merits, such as positive feedback, distributed computing and parallelism, etc.. In this paper, factors resulting in human accidents in coal mine were analyzed in details at first. And then, the evaluative indices and extractive methods of relevant feature factors in human accidents based upon ACO were presented by converting feature selection of intricate multi-factors to parallel combinational and global optimization. Through valid feature selection, the structural complexities were reduced and the disturbance imposed by redundant factors on system evaluation was dispelled. Finally, this paper established a SVM-based evaluation model of human accidents with limited samples, conducive to reinforcing safety management and promoting safety production in coal mine.

2. Feature selection of human accidents based on ACO

2.1. Human accidents in coal mine

Coal mine production, whose environment is complicated and geologic condition is atrocious, has always been threatened by

* Corresponding author. Tel.: +86-15852499056.
E-mail address: jin_jz@126.com.

1878-5220/09/$– See front matter © 2009 Published by Elsevier B.V.
doi:10.1016/j.proeps.2009.09.219

natural disasters such as gas, flood, fire, grime and roof fall. Meanwhile, coal mine accidents, whose influencing factors are much closely related to both the exploitation environment and the artificial factors, are also quite reticular and embodied in the interactions of human beings, substance and environments. By analyzing a great number of accident cases, the safety experts propose that artificial factors contribute most to safety production in coal mines [2].

Presently, there is no consensus made concerning the definition of human accidents. Scholars at home and abroad have their own definitions from distinct points of view. Zeng Weihua, from the perspective of human mind, considers human accidents mainly due to the artificial mistakes, whether intended or unconscious [3]. Swain points out there are two reasons leading to human accidents from intrinsic and extrinsic angle: one is the unsuitable designing of working conditions exceeding the individual capabilities; another is the individual impertinent actions [4]. Thus, human accidents involve not only staff's intended or unconscious mistakes, but also mistakes caused by environment which surpasses individual capabilities and qualities, along with human distorted actions aroused by external conditions. According to the various origins and bounds of human accidents in coal mines, we plot out the leading factors from four aspects: individual factors, facilities, environment, organization and administration. The structure is presented as follows.

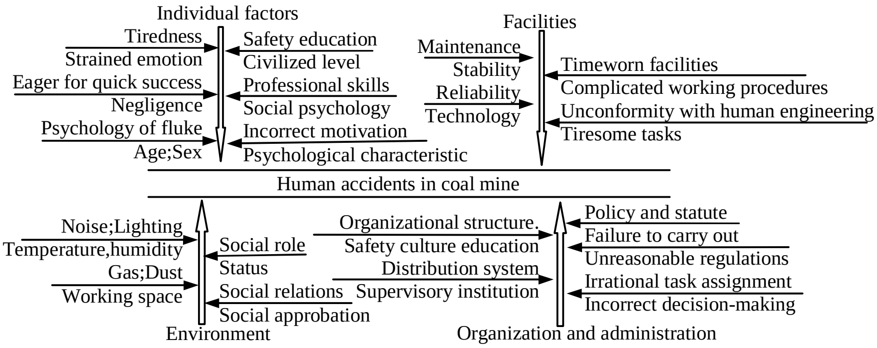

Fig. 1. Structure of leading human accidents in coal mine

2.2. Feature selection based on ACO

Coal mine production system is a complex dynamic system under special circumstance, where influencing factors correlate to each other with entanglement. It can be seen from Fig.1, that the state of coal mine system is affected by various different factors whose acting bounds and degrees are barely the same. Some factors have intensively uncertain connotations and extensions on which quantitative statistical analyses are hardly taken, hence, the veracity, validity and objectivity of system safety evaluation has been influenced to a high degree. So it is necessary to eliminate the redundant factors and to extract key factors. The evolvement of system state in coal mine is originated from a colony effect on transformation of inner factors, an effect which can be embodied through either the effects of each factor imposed on system or the transformative relationships among factors. According to the Gray Relational Analysis, each couple factors maintain the relationships by calculating the geometry similarity of their own transformative sequence: the more similar, the closer relationship they have, vice versa [5]. Since every factor keeps certain relationship with the others, the solution of the problem of the system state evolvement could be converted into finding out a feature model in which all factors maintain maximal relationships with each other. Obviously, this is a typical combinatorial optimization problem can be solved by ACO, considering its' various advantages on this issue (For more detail on ACO, refer to [6], et al.). So, based on ACO, we extract leading factors from human accidents and construct feature model where all the factors have maximal relationships affecting current system. This is helpful to effectively reduce the complicacy of system configuration and largely facilitate system safety evaluation.

Considering n influencing factors in system, m samples of every factor within specific time series make up a $n \times m$ transformative sequent matrix

$$\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nm} \end{bmatrix}.$$

In order to eliminate the dimensional effects among factors, we impose normalization on transformative sequent matrix, obtaining the following matrix

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix},$$

where, $x_{ij} = c_{ij} / (\frac{1}{m}\sum_{k}^{m} c_{ik})$,for i=1,2,…,n, j=1,2,…,m. The absolute difference value matrix of the ith factor's transformative sequence against the others is

$$\begin{bmatrix} \Delta x_{11} & \Delta x_{12} & \cdots & \Delta x_{1m} \\ \Delta x_{21} & \Delta x_{22} & \cdots & \Delta x_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \Delta x_{n1} & \Delta x_{n2} & \cdots & \Delta x_{nm} \end{bmatrix},$$

where, $\Delta x_{ij} = |x_{ik} - x_{jk}|$,for i=1,2,…,n, j=1,2,…,m, k=1,2,…,m. Then, the relationship coefficient matrix is given by

$$\begin{bmatrix} \Delta' x_{11} & \Delta' x_{12} & \cdots & \Delta' x_{1m} \\ \Delta' x_{21} & \Delta' x_{22} & \cdots & \Delta' x_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \Delta' x_{n1} & \Delta' x_{n2} & \cdots & \Delta' x_{nm} \end{bmatrix},$$

in which $\Delta' x_{ij} = \Delta x_{ij} / (\min(\Delta x) + \max(\Delta x))$, $\min(\Delta x)$ and $\max(\Delta x)$ respectively indicate the maximal and minimal value in the ith factor's relationship coefficient matrix. The relationship between factor i and factor j is

$$d(i,j) = \sqrt{\sum_k (\Delta' x_{ik} - \Delta' x_{jk})}\,, \tag{1}$$

$d(i,j)$ refers to similarity between the transformative sequence of factor i and factor j, symbolizing the relationship among factors. The smaller $d(i,j)$ is, the more closer relationship factor i has with the jth factor. And then, the relationships between each couple of factors are calculated sequently. Regarding each influencing factor as a city node, the relationship between each couple of nodes as one edge, then we can get a complete connecting graph. Since the system state lies on the relationships among every factor as well as the contributions of their own, the process of finding the feature factors set which decides system state is transformed into searching an influencing factor model with the maximal contribution and relationship in graph. Based on the principles of ACO algorithms on TSP, considering M ants that each one of them is corresponding to a city, each ant uses state transition rule to probabilistically choose the next factor node, according to the contribution , relationship and pheromone trail between current node i and candidate nodes. Edges with higher pheromone trail and heuristic information will be selected in large probability. The pheromone trail value is updated after each iteration to avoid a too rapid convergence towards a sub-optimal region. The process is iterated until the tour counter reaches the maximum or all ants make the same tour. By employing this positive feedback mechanism, therefore, the evolution towards global optimization can be guaranteed. An ant k in node i chooses the city j to move to at t time by applying the following probabilistic formula

$$P_{ij}^{k} = \begin{cases} \dfrac{\varepsilon_{ij}\tau_{ij}^{\alpha}(t)\eta_{ij}^{\beta}(t)}{\sum_{v \in allowed_k} \varepsilon_{iv}\tau_{iv}^{\alpha}(t)\eta_{iv}^{\beta}(t)} & j \in allowed_k \\ 0 & \text{otherwise} \end{cases}, \tag{2}$$

where, $\varepsilon_{ij} = \|x_j\| \big/ \sum_{v \in allowed_k} \|x_v\|$ represents contribution imposed on system state by factor j which is connected to factor i. $allowed_k$ is the set of factors that haven't yet been visited by ant k when in city i. $\tau_{ij}^{\alpha}(t)$ signifies the amount of pheromone trail on edge(i,j) at time t, whereas $\eta_{ij}^{\beta}(t)$ = $d(i,j)$ denotes the relationship on edge(i,j). α and β are parameters that control the relative importance of pheromone trail versus relationship. After each iteration, the pheromone trail on edge(i,j) is updated according to the following formula

$$\tau_{ij}^{new} = (1-\rho)\tau_{ij}^{old} + \rho\Delta\tau_{ij}\,, \tag{3}$$

$$\Delta\tau_{ij} = \begin{cases} Q/L_b & (i,j) \in G \\ 0 & \text{otherwise} \end{cases}, \tag{4}$$

in which Q is a constant and L_b is tour length of the iteration-best solution. G is the factor node set belonging to iteration-best solution. $\rho \in (0,1)$ is a coefficient representing evaporation of the trail. $\Delta\tau_{ij}$ is the pheromone trail accumulated on edge(i,j) in this iteration. After iterative counter reaches maximum, all ants have completed a tour, thus the global best solution L_g is found. We calculate contributions of each edge to the system state on global best path: $g(i,j) = L_g/[\mathrm{H}d(i,j)]$, H is a constant. The most influencing factor set can be selected with more than a given value. By selecting all the nodes, whose $g(i,j)$ are more than a given value, we constitute the most influential feature factor set upon current system state.

2.3. Description of algorithmic framework

In the real implement of our algorithm, we limit the range of possible pheromone trails on each solution component to an interval[$\tau_{\min}$,$\tau_{\max}$] to avoid unlimited accumulation of the pheromone trails, as has been adopted in MMAS(Max Min Ant

System). Meanwhile, the pheromone trails is initialized to τ_{max} in order to achieve a higher exploration of solutions at the start of the algorithm. Algorithmic skeleton for ACO-based upon feature selection is listed as follows:

Step 1. Initialize: We regard each influencing factor in current system as a city node, computing the relationships between every couple of nodes by Eq. (1), randomly selecting a node as the starting point, and using a greedier search to locate a preferable node sequence. Then we make τ_{max}=D/L_0 the upper pheromone trail limits, where D is a constant and L_0 denotes length of the preferable node sequence that could ever be found. All pheromone trails are initialized to τ_{max}, and then, set lower pheromone trail limit: $\tau_{min}=\tau_{max}/n$. M ants are placed on M randomly chosen cities. Set tour counter NC=0 and the default value of parameters: α, β, ρ, etc.

Step 2. At time t, the kth ant currently located at city i moves to city j with a probability:

$$\begin{cases} \arg\max_{v \in allowed_k} \varepsilon_{iv}\tau_{iv}^{\alpha}(t)\eta_{iv}^{\beta}(t) & , \quad q \le q_0 \\ \text{Eq.(2)} & , \quad \text{otherwise} \end{cases}, \qquad (5)$$

where, q is a variable chosen randomly with uniform probability [0,1], $q_0 \in (0,1)$ is a parameter. We apply the roulette wheel selection to choose next factor node.

Step 3. After selecting a node, all ants go to Step 2. to continue their search, repeating those processes until their taboo lists are full, namely, all nodes have been visited.

Step 4. Global solution is replaced by the better solution that ants have found, through comparing the iteration-best solution with the existed global-best solution. τ_{max} and τ_{min} are recomputed similarly as mentioned above in Step 1. Eq.(3) is used to update global pheromone. Then, we check whether pheromone trails on each solution component are within the limitation: $[\tau_{min}, \tau_{max}]$, if $\tau_{ij}>\tau_{max}$ then $\tau_{ij}=\tau_{max}$; otherwise, if $\tau_{ij}<\tau_{min}$ then $\tau_{ij}=\tau_{min}$.

Step 5. NC= NC+1.The taboo lists are initialized. The previous process is iterated until the tour counter reaches the maximum.

Step 6. We select all factor nodes with $g(i,j)$ that are more than a given value as the leading factors model in system.

3. Safety evaluation model of human accidents based on SVM

Feature selection based on ACO can effectively condense influencing factors in human accidents. However, generally, the referential samples are quite rare since the data of human accidents are hardly analyzed. It is rather difficult to guarantee the accurate evaluation of human accidents.SVM, which is developed by Vapnik according to structural risk minimization principle from statistical learning theory, is a novel machine learning method and has such advantages as it can handle pattern recognition of high dimensional data, small samples and non-linear feature vectors. Therefore, this paper proposes a safety evaluation model of human accidents based on SVM (For more detail on SVM, refer to[7][8],et al.).

Given a set of feature factors X_k, k=1,2,…,N, of human accidents which have the most influence on coal mine safety extracted by ACO, we firstly adopt safety check list to assess every factor, in which the evaluative grades are divided into 5 levels: **best** $d_{k1}(u_1=5)$, **better** $d_{k2}(u_2=4)$, **general** $d_{k3}(u_3=3)$, **worse** $d_{k4}(u_4=2)$, **worst** $d_{k5}(u_5=1)$. d_{ki} is the total votes of factor k cast in level i,k=1,2,…,N. The term in brackets represents points of corresponding level, that is, if one expert votes for a level then he/she will add relevant points to this level. We invite 10 experts to conduct the evaluation. According to the distinct function of each influencing factor, we employ AHP to compute their weight coefficient w_k,k=1,2,…, N. The frame of safety check list is as follows.

Table 1. Safety check list

Items	Weight	Levels of evaluation				
		worst	**worse**	**general**	**better**	**best**
Item 1	w_1	d_{15}	d_{14}	d_{13}	d_{12}	d_{11}
Item 2	w_2	d_{25}	d_{24}	d_{23}	d_{22}	d_{21}
…	…	…	…	…	…	…
Item N	w_N	d_{N5}	d_{N4}	d_{N3}	d_{N2}	d_{N1}

After the evaluation by all experts, we get a 1×N evaluative score matrix $X = [x_1\ x_2\ \ldots\ x_N]$,where $x_i=w_i\sum_{k}^{N} d_{ik}u_k$ is the relevant score of respective factors, $x_i \in w_i[10,50]$. Without loss of generality, we select score matrixes as training data when all factors belong to level “**worst**”, “**worse**”, “**better**”, “**best**” respectively: X_1,X_2,X_3,X_4. If the corresponding level better than(≥) “**general**”, then y_i=1,otherwise y_i=-1. Then, we get training data set$(X_1,y_1),(X_2,y_2),\ldots(X_4,y_4)$. The process of evaluating system safety turns to firstly finding out separating hyperplane which could divide system states into evaluative level space, viz.$WX+b=0, W\in R^N, b\in R$. The decision function is

$$f(X)= WX+b \qquad (6)$$

$$\text{s.t. } y_i(WX_i+b) \geq 1-\xi_i \quad i=1,2,\ldots,4,$$

where $\xi_i \geq 0$ are slack variables that deal with errors in evaluating misclassifications. The evaluative classification problem can be posed as

$$\min\Phi(W)=(\|W\|^2+C\sum_i \xi_i), \tag{7}$$

in which C is a user-specified positive parameter, controlling the tradeoff between classification violation and margin maximization. By using the Lagrangian optimization approach, Eq. (7) can be presented as

$$f(X)=\sum_i^l a_i y_i K(x, x_i)+b \quad , \tag{8}$$

α_i is Lagrangian multipliers, l is the number of support vectors, and $K(x,x_i)$is kernel function. When a sample is situated on the hyperplane, $f(X) = 0$. The more the sample distributes toward level "**best**", the bigger $f(X)$ will be, and $f(X)>0$. The more it distributes toward level "**worst**", the less $f(X)$ will be, and $f(X)<0$. Consequently, the value and sign of function f can reflect current safety state in system. The safety degree can quantitatively be judged from distance between the sample and hyperplane.

4. Experiments and analysis

4.1. Feature selection of human accidents

We investigated the effect of our feature selection strategy on accident data [5] with 30 influencing factors from 2006.1 to 2006.7 of certain coal mines, including: "mistakes in memory", "weak discernment", "decision-making", "social environment", "skills" , "organizational culture", "immature technology" and so on, as presented in the following table.

Table 2. Accident data from 2006.1 to 2006.7

Factors\Month	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.
mistakes in memory c_1	2	0	0	1	0	1	2
weak discernment c_2	0	0	0	0	0	0	1
decision-making c_3	16	4	4	6	6	13	19
social environment c_4	1	1	1	1	1	1	1
skills c_5	7	1	1	2	4	5	9
organizational culture c_6	6	5	3	2	2	5	8
…	…	…	…	…	…	…	…
immature technology c_{30}	1	1	1	1	1	1	1

Considering all accident data as a 30×7 transformative sequent matrix, we make normalization and calculate relationship $d(i,j)$ (i=1,2,…,30, j=1,2,…,30) between factor c_i and c_j according to Section 2.2. Then, a complete graph of 30 nodes is set up, every node representing an influencing factor in human accidents. The distance between node i and node j is assigned as relationship $d(i,j)$.

The parameters for ACO are chosen as follows: the initialized number of ants is equal to 30, as many as the number of factor nodes. Meanwhile, we initialize the pheromone trails to τ_{max}, and set α =1,β =5,ρ =0.5,H = 1000. All the tests have been carried out in 3000 cycles per trail, and then the best solution is obtained through ten trials. Finally, the global solution is L_g = 567. By calculating and ranking the contribution of each edge$(i,j)\in L_g$, we gain top 10 leading factors. Results are shown in Table 3

Table 3. Contributions of leading factors to the system

Factor	Contribution
1.decision-making	0.7692
2.organizational culture	0.2857
3.unreasonable tasks assignment	0.2857
4.inadequate maintenance	0.1408
5.disobedient emotions and habits etc.	0.1370
6.failure to supervise	0.0943
7.inaccurate training methods	0.0629
8.regulations	0.0515
9.skills	0.0334
10.social environment	0.0236

From the above table we can see, decision-making and organizational factors, which contribute mostly to the current system state, are the major influencing factors in human accidents. Therefore the principal countermeasures are constitutional improvement, enhancing safety concepts of decision-makers and constituting favorable safety cultural atmosphere. Next, safety inspection and supervision, training and education, along with unhealthy habits would also influence the system safety remarkably, which necessitate the enhancement of safety inspection and supervision mechanism as well as intensification of safety ideology educations. Finally, individual skills and social environment play relatively weak roles, however, from the perspective of system safety, coal mine corporations still need to improve the miner's qualities, reinforce their professional skills training and create a comfortable working environment for them.

4.2. Safety evaluation of human accidents

After analyzing feature factor vectors of human accidents extracted by ACO, we select 6 representative influencing factors: organizational institution, administrant system, safety management scheme, safety inspection, safety investment and crew's qualities, which have relatively important impacts on coal mine safety, and then we design the safety check list, using safety evaluation model based upon SVM to value human accidents of ventilating system on Caitun mine. Assuming that the following table is score of evaluative level of each factor in system assessed by all experts [9],

Table 4. Safety check list

Items X	Weights W	Levels of Evaluation				
		worst	**worse**	**general**	**better**	**best**
organizational institution x_1	0.200	1	2	3	4	0
crew's qualities x_2	0.088	0	3	3	4	0
safety investment x_3	0.064	0	3	4	2	1
safety management scheme x_4	0.215	2	4	4	0	0
administrant system x_5	0.215	0	3	2	3	2
safety inspection x_6	0.215	0	2	4	3	2

then, the evaluative score matrix of current system is X_C = [6.000 2.728 1.984 4.730 7.310 7.095]. We respectively select score matrix whose factors all belong to one level, ranging from "**worst**", "**worse**", "**better**" to "**best**", as training samples: X_1 = [2.00 0.88 0.64 2.15 2.15 2.15], X_2 = [4.00 1.76 1.28 4.30 4.30 4.30], X_3 = [8.00 3.52 2.56 8.60 8.60 8.60], X_4 = [10.00 4.40 3.20 10.75 10.75 10.75], with which compose the training data set$(X_1,-1),(X_2, -1),(X_3,1),(X_4, 1)$. Then evaluative model established in Section 3 is employed to train SVM. Experiments are completed on C-SVM with the commonly used RBF kernel. As a result, 4 SVs are found out: $SV_1 = X_1, SV_2 = X_2$, $SV_3 = X_3$, $SV_4 = X_4$, and $f(X_1)= -1.0$, $f(X_2)= -0.856$, $f(X_3)=0.856$, $f(X_4)=1.0, f(X_C)=0.0164$. Since $f(X_C)>0$ and $f(X_C)<f(X_3)$, the current system state is between "**general**" and "**better**", yet far away from level "**better**". This is consistent with the known facts in the coal mine. It can be seen from the comparison of score matrix X_C and X_3 that the limitation of current system is mainly embodied in 4 aspects: organizational institution, administrant system, safety management scheme, safety inspection. Hence, it is necessary for us to perfect organizational institution, ameliorate safety management scheme, enhance the degree of safety inspection and improve the crew's qualities.

5. Conclusions

In this paper, we make effective features extraction of influencing factors through Ant Colony Optimization, which confirms that the administrant and organizational factors are leading ingredients in human accidents. Besides, we propose a safety evaluation model of human accidents based on Support Vector Machine, providing a newly researching direction in safety evaluation. Experiments on coal mine data set have confirmed that our method is effective and practicable. Additionally, we should notice that safety production in coal mines is a long-term system project. Settling the disadvantages detected in safety evaluation timely and effectively is crucial to coal mine safety production.

Acknowledgements

Financial support for this work, provided by the Natural Science Foundation of Jiangsu Province, is gratefully acknowledged.

References

[1] J. Chen, Q.G. Chao, R.Z. Li, Forecast and Evaluation of Human Error in Coal Mine Safety Production. Mining Safety and Environmental Protection. 34 (2007) 78-81.
[2] X.Z. Zhang, Safety Evaluation on Management System of Coal Mine Ventilating System Based on AHP. Zhong Zhou Coal Mine. 1 (2007) 76-81.
[3] L.L. Yan, S.M. Wei, Study on Mechanism of Man-Cause accidents and its prevention measures. Shan Xi Coal Mine. 1 (2004) 14-16.
[4] L.J. Hu, J.H. Chen, Identification of Key Factors of Human Error for Coal Mining Safety. Journal of University of South China (Social Science Edition). 6 (2007) 31-34.

[5] T. Chen, Study on the Organization Factors of Human Errors in the Coal Mine Accidents, Master dissertation. Xi'an University of Science&Technology. 2008. (In Chinese)

[6] M. Dorigo, L.M. Gambardella, Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem. IEEE Transactions on Evolutionary Computation. 1 (1997) 53-65.

[7] V.N. Vapnik, An overview of statistical learning theory. IEEE Transactions on Neural Networks. 10 (1999) 988-999.

[8] V.N. Vapnik, Statistical Learning Theory.2nd ed.. New York:Springer-Verlag.1999.

[9] Z.Q. Zhou, The Security, Stability and reliability Appraisal and Application Study of The Ventilating System of Mine, Master dissertation. Liao Ning Technical University, 2003.