

Study on Risk Classification of Goaf Based on RS-SVM

Haiping Yuan^{1,3}, Biaohua Le¹, Yubo Wang¹, Chenghao Chen¹, and Hanbing Bian^{1,2}

¹School of Civil Engineering, Hefei University of Technology, hefei anhui 230009, China

²Laboratoire de Génie Civil et géo-Environnement, Université de Lille, 5900 Lille, France

³State Engineering Laboratory of Highway Maintenance Technology, Changsha University of Science and Technology, Changsha, China

Abstract. According to the uncertainty and concealment of the risk of goaf, a risk classification model of goaf is constructed based on rough set (RS) knowledge and support vector machine (SVM) theory. In this paper, based on statistical analysis and measured data, nine parameters including mining method, empty area excavation depth, goaf height, maximum exposed area of empty area, maximum exposure height, maximum exposure span, pillar condition, empty volume and treatment rate are selected as the main influencing factors. The RS theory is used to reduce the sample, and SVM is compiled by Matlab. The one-to-one method is used to construct the binary classifier to realize the multi-class classification algorithm of goaf. Finally, a SVM model for evaluating the risk level of the goaf is obtained. The research shows that: based on RS theory, SVM has a good effect on the hazard classification of the goaf iron ore mine, and the difference with the actual situation is 13.3%. The research results have certain theoretical significance and guiding role for the safe mining of an iron mine in Eastern China.

1 Introduction

Because the mine uses the empty field method for mining, it leaves a large number of goafs. With the continuous consumption of mineral resources, the mine production activities are increasing, the mining depth is gradually deepening and the scale is gradually increasing. The goaf poses a more serious threat to the safety of the mine, and the phenomenon of rock stratum displacement and crack deformation is gradually jeopardized. The danger of the goaf has seriously affected the safety of the entire mine and has become an important factor in evaluating mine safety [1-3].

The risk assessment of goaf is still a key issue. At present, domestic and foreign scholars have done a lot of research on the risk of goaf. Chen Zhiqiang[4] established a two-level evaluation model for fuzzy comprehensive evaluation. According to the weight set and evaluation matrix, the fuzzy comprehensive evaluation of the goaf was completed according to the second-level evaluation model. Xu Xiaoding[5] constructed the pillar-rock beam support system in the plastic zone, and constructed the cusp catastrophe model of the support system by using energy theory, and obtained the instability judgment conditions related to the support system, so as to evaluate the stability of the goaf. Zhang Yaoping[6] used numerical simulation to simulate and predict the formation process of the empty area and the stability of the goaf. Hu Gaojian[7] established a three-dimensional discrete fracture network model of structural plane, studied the degree of cutting fracture of surrounding rock, and obtained the stability analysis result of complex goaf by using Mathews stability diagram method. Li Junjun[8] used adaptive genetic algorithm to globally optimize the initial weight and enthalpy of BP network, and predicted the risk level of goaf. Based on the theory of set pair analysis, Xiao Chao[9] constructed a pair of sampled measured index values that affect the stability of the goaf and the stability criteria of the goaf, and described the pair relationship through the relationship degree and the stability of the gab mine goaf classification. Wang Wei[10] established a stability evaluation index system for goaf, applied matter-element analysis theory, and established a matter-element extension model for the evaluation of goaf stability through steps of matter element structure and correlation function.

Due to the complex geological environment in the goaf and the uncertainties of the stability of the goaf, many studies on the stability of the goaf have certain limitations. In this paper, the rough set (RS) is used to reduce the sample attributes, and the influencing factors are selected based on the actual situation of an iron mine in Eastern China. The support vector machine (SVM) is used to classify the risk of the goaf, which is of great significance to mine safety.

2 Basic theory of RS-SVM

RS theory proposed by Pawlak et al[11-12] is a new mathematical tool to deal with the problems of ambiguity and inaccuracy. When dealing with redundant information, the reduction process has the following definition.

SVM is a new machine learning method proposed by Vapnik et al[13]. Based on statistical learning theory. The basic theory is as follows:

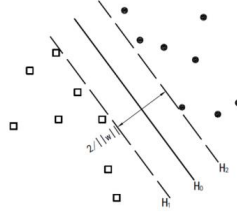


Figure 1 Optimal classification super-planar

In the case of linear separability, Suppose there is a training sample (x_i, y_i) , $x \in R^n, y = \{+1, -1\}, i = 1, 2, \dots, l$. As shown in Figure 1, the classification interval is $\frac{2}{\|w\|}$, making the interval maximum equivalent to $\frac{1}{2}\|w\|^2$ minimum.

Considering the existence of an error, using a slack variable $\xi > 0$, it is written as:

$$\min \frac{1}{2}\|w\|^2 + C \sum_{i=1}^l \xi_i \tag{1}$$

Where C is a normal number and the Lagrange function is established:

$$L = \frac{1}{2}\|w\|^2 + C \sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i (y_i (< w, x > + b) - 1 + \xi_i) - \sum_{i=1}^l \gamma_i \xi_i \tag{2}$$

In the nonlinear case, the partial derivative of the Lagrange function to the optimization target w, b, ξ is 0. According to the Kuhn-Tucker condition and the Mercer theorem, the kernel function $K(x_i, x_j)$ is selected to map the data to a high-dimensional feature space, and then in the high-dimensional Linear classification is performed in the feature space to obtain the dual form of the following quadratic programming:

$$\begin{cases} \max \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \sum_{i=1}^l \alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq C, i = 1, 2, \dots, l \end{cases} \tag{3}$$

Solving the optimal solution of α from the quadratic programming problem, and selecting the support vector of $\alpha \neq 0$ to obtain:

$$b = y_i - \sum_{i=1}^l \alpha_i^* y_i K(x_i, x_j) \tag{4}$$

Finally, the optimal classification function of SVM can be constructed:

$$f(x) = \text{sgn}(\sum_{i=1}^l \alpha_i y_i K(x_i, x) + b) \tag{5}$$

Where, the kernel functions is follow:

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \tag{6}$$

3 Engineering applications

3.1 Project overview

According to the current statistics, there is no less than 94 goafs in the upper plate area, F2 lower plate and east area of an iron mine in the east, including 168 sub-goafs. The total volume of the goafs is about 3.67 million m³, of which the eastern area is about 2.03 million cubic meters, accounting for 55.3% of the total volume of the empty area above 200m; most of the empty area is concentrated in the horizontal range of 200-300m, the total volume of the empty area is about 1.63 million m³, accounting for nearly 44.8% of the total volume of the empty area above 200m. The empty area is mainly formed by shallow hole mining method, and the shallow goaf accounts for nearly 75.8% of the total volume of the empty area above 200m.

3.2 Influencing factors of goaf

There are many factors affecting the stability of the empty area, such as engineering and hydrogeology, components of the ore block, depth of burial of the ore body, reasonable design of the pillars in the isolated pillar and the stope, vibration and impact of the blasting on the ore body and surrounding rock, The length of the formation of the goaf, the shape of the goaf, the exposed area, the exposed height, the thickness of the roof and the bottom layer, the treatment and strength of the empty area, the surrounding rock strength of the empty area, the scale and distribution of the ground surface and joint fissures, the development of geological structures such as faults, and so on. Moreover, the influence of each factor on the stability of goaf in a specific mine is quite different [14-15]. According to the data collected and mastered by an iron mine in the east of +200m and above, the evaluation criteria of the stability of the empty area include mining method (C₁), empty area excavation depth (C₂), goaf height (C₃), mining The maximum exposed area (C₄), the maximum exposure height (C₅), the maximum exposure span (C₆), the pillar condition (C₇), the volume of the goaf (C₈), and the treatment rate (C₉). SVM can't process qualitative indicators when learning sample data, so it needs to be quantified by qualitative indicators.

3.3 Selection of samples and kernel functions

At present, there are 168 sub-empty areas in an iron ore sub-space in the east. The training samples are randomly selected by stratified sampling. As shown in Table 1, the previous 33 sequences are used as training samples, and the rest are used as prediction samples. When selecting the kernel function, considering that the influencing factors of the goaf are numerous and the dispersion is large, the linear kernel function can not effectively classify the risk of the goaf; the Sigmoid kernel function can classify well, but the parameter value requirements Higher, and the effect is not ideal under small samples. The polynomial kernel function has global characteristics, while the RBF kernel function has a wide range of applications and a wide convergence domain [16]. In this paper, polynomial kernel function and RBF kernel function are used to predict the risk.

Table 2 Samples of goaf danger classification

Serial number	Excavation depth (m)	Mining method	Gob height	Maximum exposed area (m ²)	Maximum exposure height (m)	Maximum exposure span (m)	Pillar situation	Measured volume (m ³)	Governance rate	Actual level
1	166	1	12	8114.7	12	62	1	48684	1	I
2	336	1	14	2615	14	46	2	33085	1	I
3	308	1	15	2519	15	50	2	21721	1	I
4	160	2	17	3883	4.2	26	3	33805	3	II
.....										

3.4 Sample data processing and attribute reduction

The sample data size has a large impact on the SVM model, which leads to a decrease in the accuracy of the classification results. Therefore, in order to ensure the stability of the numerical operation and the convergence of the model network, and to eliminate the influence of the dimension, the sample data needs to be Normalized according to equation(7), namely:

$$S_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \tag{7}$$

Where: x_{ij} is the sample data before normalization; S_{ij} is the normalized sample data; $\min(x_j)$ is the minimum value in the measured data of the index; $\max(x_j)$ is the maximum value in the measured data of the indicator.

In the process of reduction using rough set theory, in order to narrow the information system and improve the clustering ability of the evaluation system, it is necessary to discretize the normalized data. This paper uses the

equidistant division method to discretize. The spatial dimension and information loss after attribute discretization should be minimized, therefore, take three breakpoints when discrete, and establish an information decision table.

Based on the information decision table, the redundant information in the data is removed, and the data dimension is reduced. The information decision table shows that the training samples 2 and 3, 16 and 18, 20 and 24, 23 and 28 overlap, thus deleting the duplicate Training samples 3, 18, 24, 28. At the same time, the attribute reduction is performed by the deletion method, and the reduction set is $\{C_2, C_5, C_8\}$.

3.5 Establishment of SVM multi-class classification model

At present, the common methods for multi-class classification by SVM include one-against-One, one-against- Rest, binary tree method and directed acyclic decision graph method[17], which are all based on SVM binary classification. In this paper, the SVM multi-class classifier is constructed by one-against-One. This method constructs $k(k - 1) / 2$ sub-classifiers in k -class training samples. Each sub-classifier only has two kinds of training samples in k class. Train between them, then use the voting method to examine the test samples one by one, and select the category with the most votes as the final output category of the classifier.

The RS-SVM multi-class classification model program is written based on matlab. The parameters are determined by the trial algorithm as follows: the kernel parameter of the polynomial kernel function $\gamma = 0.5$ and $q = 2$, the RBF kernel function $\sigma^2 = 0.3$, and the penalty parameter $C = 2.5$. The predicted results are shown in Figure 2.

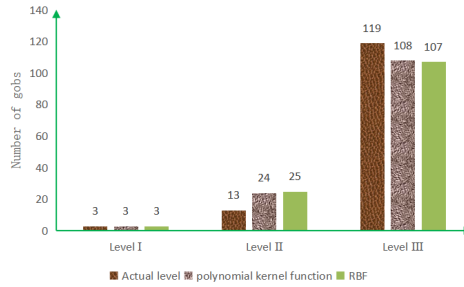


Figure 2 Comparison of risk classification of goaf

When using SVM for grading, the influence of the kernel function parameters and the penalty parameters is very obvious. The error of the final evaluation can be obtained by debugging the parameters, as shown in Figure 3 and Figure 4.

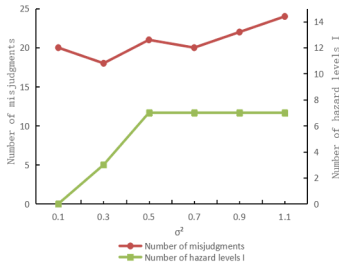


Figure 3 The effect of σ^2 ($C=2.5$)

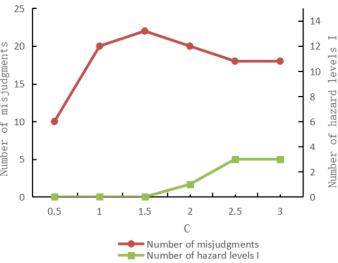


Figure 4 The effect of C ($\sigma^2=0.3$)

Using the above-mentioned RS-SVM multi-classification model to identify the risk of goaf, it is possible to comprehensively consider the risk of goaf and the nonlinear relationship between various factors, based on two kernel functions. Compared with the actual situation, the difference between the three is 13.3%. This shows that the support vector machine method, an intelligent learning method, can effectively guide the identification and evaluation of the risk of goaf in practice. It is a method of identification that is worthy of further study and popularization.

4. Conclusion

So far, the research on the risk of goaf is limited by the concealment and uncertainty of goaf. To this end, using the emerging intelligent learning method, through the learning and regular mining of limited data, seek out gofs. The nonlinear relationship between hazard and various influencing factors is as follows:

(1) Hazard identification based on support vector machine has good generalization ability in small samples. For complex systems such as goaf, the result of discrimination is reasonable and effective.

(2) RS theory can effectively reduce the number of samples, improve the training speed, and at the same time reduce the redundant attributes of the samples, so that the support vector machine analysis can be more accurate.

(3) In this paper, SVM is used to classify the risk of goaf. The appropriate nuclear parameters and penalty parameters can be accurate results, and the samples also need certain diversity and representativeness.

Acknowledgments

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research work was funded by the National Natural Science Foundation of China (51874112 and 51774107); the State Engineering Laboratory of Highway Maintenance Technology, Changsha University of Science and Technology (kfj170108).

References

1. H. T. Ma, N. W. Liu, Y. H. Wang, et al. Review on research status of controlling techniques for goaf disaster in metal mine[J]. *Journal of Safety Science and Technology*, 2014, 10(10):75-80
2. F. Q. Gong, X. B. Li, L. J. Dong, et al. Underground goaf risk evaluation based on uncertainty measurement theory[J]. *Chinese Journal of Rock Mechanics and Engineering*, 2008, 27(2):323-330
3. Q. M. Wang, B. G. Xu, S. H. Tang, et al. The actual state of mined-out areas in metallic and nonmetallic mines and the countermeasures for the treatment of the mined-out areas in China[J]. *MINING R&D*, 2009, 29(4):63-68
4. Z. Q. Chen, Y. M. Wang, S. M. Chen, et al. Stability analysis of goaf based on entropy weight method and fuzzy theory[J]. *MINING R&D*, 2019, 39(5):39-43
5. X. D. Xu, Y. J. Zhou, S. Pang. Analysis of catastrophic instability of plastic supporting system in old goaf of gypsum mine[J]. *Chinese Journal of Rock Mechanics and Engineering*, 2018, 37(11):2548-2555
6. Y. P. Zhang, P. Cao, H. P. Yuan, et al. Numerical simulation on stability of complicated goaf[J]. *Journal of Mining & Safety Engineering*, 2010, 27(2):233-238
7. G. J. Hu, T. H. Yang, Z. Q. Hu, et al. Mathews-stability-method-based multi-angle analysis and evaluation of mined-out zones stability[J]. *Journal of Mining & Safety Engineering*, 2017, 34(2):348-354
8. Y. J. Li, W. Q. Lin, Y. Chen. Evaluation on risk of goaf based on AGA-BP neural network[J]. *Journal of Safety Science and Technology*, 2015, 11(7):135-141
9. C. Xiao, H. C. Zheng, W. Z. Wu, et al. Stability evaluation of goaf in gypsum mine based on the set pair analysis[J]. *CHINA MINING MAGAZINE*, 2014, 23(10):107-114
10. W. Wang, Z. Q. Luo, L. X. Xiong, et al. Stability evaluation of goaf based on improved matter element extension model[J]. *Journal of Safety and Environment*, 2015, 15(1):21-25
11. Z. Pawlak. Rough sets[J]. *International Journal of Information and Computer Science*, 1982, 11(5):341-356
12. S. Z. Wu, P. Z. Yan. Attribute reduction algorithm on rough set information entropy and its application[J]. *Computer Engineering*, 2011, 37(7):56-61
13. S. F. Ding, B. J. Qi, H. Y. Tan. An overview on theory and algorithm of support vector machines[J]. *Journal of University of Electronic Science and Technology of China*, 2011, 40(1):2-10
14. L. X. Li, G. W. Wang. Sensitivity analysis of impact factors of goaf group stability and its application[J]. *MFTAI MINE*, 2018, 2018(2):27-34
15. Y. Feng, X. M. Wang, A. B. Cheng, et al. Method optimization of underground goaf risk evaluation[J]. *Journal of Central South University(Science and Technology)*, 2013, 44(7):2882-2888
16. H. Song, Y. Xue, L. G. Zhang. Research on kernel function selection simulation based on SVM classification[J]. *JISUANJI YU XIANDAIHUA*, 2011, 2011(8):133-136
17. Y. G. Liu, Z. H. You, L. P. Cao. A novel and quick SVM-based multiclass classifier[J]. *Pattern Recognition*, 2006, 39(11):2258-2264