Rock Burst Monitoring and Early Warning Based on Incremental Learning Method with SVM

Chunfang Wu, Ruisheng Jia and Tao Qiu
College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao, 266590, China

Abstract: The rock burst hazard is a common geological hazard. In this study, we investigate an approach for classification of rock burst situation. This study relies on support vector machine classifier which in the case of less prior knowledge, still has the ability of classification. First we describes the current research work on rock burst monitoring and early warning and reasons for the introduction of support vector machines and later propose support vector machines algorithm and its improvement strategies. The results illustrate that incremental learning method for support vector machine not only requires less prior knowledge, but also without affecting the performance at the same time and training time will be substantially reduced. The method for rock burst monitoring and early warning has exhibited remarkable detection and generalization performance.

Keywords: Classification, early warning, incremental learning method, monitoring, rock burst hazard, SVM classification

INTRODUCTION

Recent years, China’s coal mine accidents showed high momentum. Along with the increasing depth of coal mining and the complex of geological conditions, the rock burst hazards has become an important factor in triggering mine geological disaster. The reason that cause the rock burst hazard is extraordinarily complex (Pan et al., 2012; Qi et al., 2011). There are not only geological reasons, but also mining technology as well as people lack of awareness on the mechanism of rock burst and so on.

For the past few years, monitoring systems have been put into used and a significant amount of data have been obtained. Many researchers at home and abroad have attempted to apply time series analysis, neural networks (Gao et al., 2011) and gray model into detecting early warning rock burst areas. However, how to develop a comprehensive system of evaluation and how to control of this hazard is still the urgent subject of the research (Jiang et al., 2010).

With the rapid development of data mining technology, some coal mines have adopted further techniques in order to improve the performance (Song et al., 2010; Gao, 2004; Wang et al., 2004). In this study, support vector machine will be introduced into the field of rock burst monitoring which can be seen as a classification problem, that is, for a given sensor data: what kind of data is normal, what kind of data anomalies. By introducing the support vector machine to the field of rock burst monitoring, we can make the system in the case of a less priori knowledge, still has good generalization ability.

OVERVIEW OF SUPPORT VECTOR MACHINE CLASSIFIERS

SVMs based on the principle of structural risk minimization are a class of supervised learning algorithms (Xu et al., 2011). SVM algorithm is based on the statistical learning theory and the Vapnik-Chervonenkis (VC) dimension introduced by Vladimir Vapnik and Alexey Chervonenkis (Hang, 2000). Its essential idea is that low-dimensional points in space are mapped into a high dimensional space, so that they become linearly separable, then use the principle of linear division to determine the optimal separating hyperplane so that the points of the separate categories are divided by a clear gap that is as wide as possible (Vapnik, 1995). In high dimensional space, it is a linear division and in the original data space, which often happens that the points to discriminate are not linearly separable.

Consider the classification of two classes of patterns which are linearly separable. Suppose the training data \((x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)\), where \(x \in \mathbb{R}^n\) and \(y_i \in \{-1,1\}\) The decision function of the SVM has the form:

\[
f(x) = \text{sgn}\left(\sum_{i=1}^{l} y_i \alpha_i x \cdot x^T + b\right)
\]
where, \( y_i \) is either 1 or -1. \( x_i \) is the point to be classified, \( b \) is constant decided from training. Only a few \( \alpha_i \) will be greater than zero. The corresponding \( \alpha_i \) are exactly the support vectors (Farquad and Bose, 2012). When the training sets are non-linear, we can map the original sample space to a higher dimensional feature space. Since the inner product is in a relatively high-dimensional space which can’t easily lead to the dimensional disaster. The kernel function plays a key role for SVM in many such applications where the training data are not linearly separable:

\[
f(x) = \text{sgn}\left(\sum_{i=1}^{l} y_i \alpha_i K(x_i \cdot x) + b\right)
\]

The RBF function is usually considered as the most powerful. In the real model, the target classes are not linear separable, the process is usually done using the RBF kernel.

**Incremental learning method with SVM:** From the data set level, SVM classifier hyperplane location only associate with support vector machines, but not with other unrelated training samples (Takashi et al., 2007). That is, SVM generalization performance does not depend all the training data samples. Usually support vectors are only part of SVM training sample. Using this feature, we can greatly compress the useful data and reduce the training time (Rao et al., 2003). At the same time, it will not bring the problem of declining the classification accuracy.

Therefore, we introduce the incremental learning SVM method, assuming that the training samples is \( C \), specific steps are as follows:

- First, we extract a small set of samples \( S_0 \) from a large training set \( C \) which can guarantee a certain of classification accuracy and the cost of training time is acceptable. Then the initial classifier \( F_0 \) is obtained.
- \( S_0 \) is removed from the training set \( C \). The remainder is randomly divided into the same number of mutually exclusive subset1, subset2…subset\( n \).
- Use the initial classifier training \( F_0 \) to train subset1; misclassified sample set is obtained which is called Misclassification Training Samples (MTS).
- Extract support vector from \( S_0 \) and merge MTS as a new training sample set, then train with subset2.
- These cycles until the fault of all subsets of points are added to the initial training set and the final classification is obtained.

Algorithm thought is: Let initial hyperplane of \( F_0 \) is \( L \), for any subset of the sample \( x \), \( x \) to the hyperplane distance is:

\[
r = \frac{|f(x)|}{||w||}
\]

If \( x \) is misclassification, it indicates that classifier is not able to properly reflect misclassified sample information and should be absorbed into the training set. If \( F_0 \) can be correctly classified and in the \( 1-\varepsilon < |f(x)| < 1+\varepsilon \), then retain \( x \), i.e., \( x \) is closer to the support vector, \( \varepsilon \) is controllable threshold (Cauwenberghs and Poggio, 2001).

The model for monitoring and early warning of the rock burst:
The rock burst intrusion detection system consists of acoustic emission data preprocessing, incremental learning support vector machine and decision making system, as shown in Fig. 1.

The entire model can be divided into two processes: the training phase and the testing phase. The training applies the incremental support vector machine to train data in order to obtain the optimal separate hyperplane and its corresponding parameters. During the test phase, the data is processed as the input of the incremental support vector machine and then according to the above theory formula, the likelihood of rock burst is obtained.

**EXPERIMENTS**

In this study, we address feature parameters from acoustic emission sensor to determine possible rock burst hazard. Where rock burst is defined as two kinds of situation states: normal and dangerous. The input parameters of incremental support vector machine includes energy, wave hit (event) counts, big event rate of the acoustic emission, the total number of ringing, \( m \) value method (Li et al., 2008).

During the experiment, we select sound monitoring sensor data from a coal mine in May 2011 and preprocess through method of PCA and normalization. We extract 90% of sample set from the sound sensor as a training set and the remainder samples as a test set.
Table 1: Comparison of two methods

<table>
<thead>
<tr>
<th>σ^2</th>
<th>SVM-type</th>
<th>SV. total</th>
<th>t</th>
<th>Accuracy (%)</th>
<th>G-means (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0001</td>
<td>SVM-standard</td>
<td>825</td>
<td>253.74s</td>
<td>71.46</td>
<td>74.00</td>
</tr>
<tr>
<td></td>
<td>SVM-inc</td>
<td>678</td>
<td>183.54s</td>
<td>78.46</td>
<td>79.65</td>
</tr>
<tr>
<td>0.001</td>
<td>SVM-standard</td>
<td>810</td>
<td>260.93s</td>
<td>77.18</td>
<td>78.20</td>
</tr>
<tr>
<td></td>
<td>SVM-inc</td>
<td>658</td>
<td>173.66s</td>
<td>77.42</td>
<td>78.62</td>
</tr>
<tr>
<td>0.0001</td>
<td>SVM-standard</td>
<td>730</td>
<td>232.84s</td>
<td>77.48</td>
<td>79.36</td>
</tr>
<tr>
<td></td>
<td>SVM-inc</td>
<td>642</td>
<td>169.23s</td>
<td>78.56</td>
<td>80.42</td>
</tr>
<tr>
<td>0.00001</td>
<td>SVM-standard</td>
<td>699</td>
<td>242.53s</td>
<td>76.00</td>
<td>77.30</td>
</tr>
<tr>
<td></td>
<td>SVM-inc</td>
<td>631</td>
<td>156.85s</td>
<td>77.49</td>
<td>77.80</td>
</tr>
<tr>
<td>0.000002</td>
<td>SVM-standard</td>
<td>678</td>
<td>252.50s</td>
<td>75.60</td>
<td>75.80</td>
</tr>
<tr>
<td></td>
<td>SVM-inc</td>
<td>621</td>
<td>150.23s</td>
<td>76.86</td>
<td>77.04</td>
</tr>
</tbody>
</table>

Using the third part of the introduction of the incremental learning method, 10% of the training set is as the initial training set S0, the remaining part of the collection is divided into several disjoint small collections.

The results in Table 1 show comparison results between the standard SVM and incremental learning method SVM. (Which SVM-standard represents the standard support vector machines and SVM-inc represents the incremental learning support vector machine).

As can be seen from Table 1, compared with standard support vector machine, the incremental support vector machine learning methods can maintain accuracy and g-means slightly higher. As the number involved in the training of support vector machine reduces, the use of training time is less than 2/3 of the standard support vector machine. It proves that the incremental support vector machine has a shorter training time, higher accuracy and g-means. And it proves the algorithm is efficient. And when σ^2 = 0.0001, we can obtain the optimal hyperplane constructed with the best classification results, with σ^2 values continue to decrease although the algorithm reduces training time, it also brought the problem of declining accuracy rate due to decline in the number of support vectors.

**CONCLUSION**

Based on the support vector machine, we propose a strategy based on incremental learning method and apply it to the rock burst monitoring and warning areas. Conclusions are as follows:

- By preserving the original support vector and classifier misclassification, it can greatly reduce the number of support vectors, which can improve the classification accuracy and g-means at the same time and training the time is shortened.
- Simulation example results prove that the proposed incremental learning method support vector machine is validity, reliability, easy to implement and thus it provides a practical ideas and methods in the field of engineering practice.

**ACKNOWLEDGMENT**

In this study, the research was sponsored by the Nature Key Technology R and D Program of the Ministry of Science and Technology of P.R. China (No. 2012BAK04B06).

**REFERENCES**


