Fast Calculation Method of Abnormality Degree for Real Time Abnormality Detection in Vehicle Equipment

Minoru KONDO Yo SAKAIDANI
Drive Systems Laboratory, Vehicle Control Technology Division

Vibration monitoring is effective for early detection of equipment failure. In the vibration monitoring system proposed in this paper, abnormality detection is performed by applying the nearest neighbor method (NN) to the octave band analysis results of vibration. However, the NN requires a long calculation time and is not suitable for detecting abnormalities in real time. Therefore, applying the One Class Support Vector Machine (OCSVM) to abnormality detection was considered. In this paper, the OCSVM was applied to actual vibration data, and the calculation time was compared with those of the NN. The result shows that the calculation time is significantly reduced compared to the NN approach.

Keywords: vibration analysis in octave bands, condition monitoring, machine learning

1. Introduction

Traction motors and reduction gears are used in EMUs. In diesel cars, rotating machines such as diesel engines, transmissions, reduction gears, and propulsion shafts are used for driving. The ability to detect abnormalities in any of these vehicle parts early enough to prevent failure can help improve the reliability and safety of the railways. In general, it is known that vibration monitoring is effective for monitoring the condition of equipment. Therefore, RTRI is developing vibration monitoring systems for vehicle equipment [1-6].

During this development, in order to deal with complicated vehicle equipment vibrations, an abnormality detection system was proposed that applies machine learning to the octave band analysis results of vibration, and confirmed the effectiveness of the system in abnormality simulation tests. The proposed system assumes that vibration measurement and octave band analysis are performed on the train, the results are transmitted to the ground via wireless network and analyzed by a computer on the ground to detect abnormalities using a method based on the nearest neighbor method.

However, when abnormality detection results need to be flagged rapidly or in order to reduce the amount of wireless communication to the ground, real-time abnormality detection on a train becomes necessary. The method based on the nearest neighbor method is not suitable for on-board abnormality detection because the calculation time required for abnormality detection increases as training data increases. Therefore, we studied the application of OCSVM (One Class Support Vector Machine) [7] as a method for detecting abnormalities on a train.

OCSVM is a well-known machine learning method for abnormality detection, and is known to be effective for a range of problems. In OCSVM, abnormality detection boundaries are created using only a small part of the training data called support vectors, which should make it possible to speed up abnormality detection calculations. This paper reports on the results of applying OCSVM to actual vibration data, comparing the calculation results and calculation time of the abnormality degree with the conventional method, and verifying its effectiveness.

2. Condition monitoring system using vibration for traction machines

2.1 Overview of the system configuration and operation

Firstly, the outline of the target condition monitoring system is described. Figure 1 shows a configuration example of the condition monitoring system. The main components of the condition monitoring system are a "condition monitor" mounted on the train and a "diagnosis program" built into a computer on the ground. In this system, octave band analysis is performed while the vibration of the device is constantly measured by the condition monitor.

In the system proposed so far, it is assumed that an abnormality is detected by transmitting the octave band analysis result (vibration data) to the ground and analyzing the vibration data by the diagnosis program. Although the amount of data is greatly reduced by performing octave band analysis, a train with a large number of devices, pro-
duces vast volumes of data. For example, when performing 1/12 octave band analysis per second for vibrations from 10 Hz to 10 kHz, more than 100 numerical values are generated per second per vibration measurement point.

If abnormality detection can be performed in real time in the condition monitor installed on each vehicle of the train and only the result can be transmitted, the amount of communication data will be greatly reduced, and the system can be easily implemented.

Therefore, this paper describes an attempt to build a system that detects abnormalities in the condition monitor in real-time.

2.2 Model training process

In the abnormality detection method used in the present system, abnormality is detected by comparing measured vibration data with normal condition vibration data (training data). In order to detect abnormalities in the condition monitor, it is necessary to extract the training data from acquired vibration data which need to be transferred from the condition monitor. In addition, when using OCSVM, it is necessary to extract support vectors from training data and transfer the support vectors to the condition monitor.

In this system, as shown in Fig. 1, vibration data is recorded on a storage medium such as an SD card in the condition monitor, and the storage medium is removed during periodic inspections and the vibration data is transferred to the computer on the ground. Then, the vibration data is put into the diagnosis program to calculate support vectors, and the result is recorded in a storage medium and returned to the condition monitor on the train. Then, real-time abnormality detection is performed using the support vectors.

By doing this, it is possible to perform abnormality detection based on past normal data without communicating a large amount of vibration data.

3. Abnormality detection method

3.1 Abnormality detection by one class classification

In the abnormality detection method in this paper, normal condition vibration data are given as training data, and then it is determined whether the vibration data to be diagnosed (test data) correspond to normal vibration using a one-class classification machine learning method.

In a multiclass classification problem, a datum is classified into one of the predetermined classes. To detect abnormalities using this method, it is necessary to prepare training data to define both normal and abnormal classes in advance. However, in general, the frequency of vehicle device failure is very low, and acquiring failure condition vibration data in advance is difficult, thereby also making it difficult to prepare “abnormal” training data.

On the other hand, in the one-class classification, training data are provided only for one class, and a datum is classified into that class or another. With this method, an abnormality can be detected even if only “normal” data are prepared, and thus it is considered suitable for detecting a vehicle device abnormality.

There are several one-class classification methods in the literature, but in the development so far, a method based on the nearest neighbor method has been used. However, as mentioned above, the problem with the nearest neighbor method is that the memory and calculation time required for abnormality detection increase as training data volume rises. In order to detect abnormalities in real-time on a train, OCSVM with a small calculation load at the time of abnormality detection is considered suitable. Therefore, this paper compares these two methods.

3.2 Nearest neighbor method

Firstly, an abnormality detection method using the nearest neighbor method is described.

In this method, as shown in Fig. 2, the vibration measurement result is divided into pieces of waveform for every one second, and octave band analysis is performed for each piece of waveform. Next, principal component analysis (PCA) and whitening are performed as preprocessing on the analysis result.

In the PCA, components in the direction along which the data are widely distributed are extracted as principal components. Whitening is a process of subtracting the average value of all training data and dividing it by the standard deviation for each principal component. In this paper, the first to sixth principal components are used based on a past study [1].

When abnormality detection is performed by the proposed method based on the nearest neighbor method, each preprocessed datum corresponds to the coordinates of one point in the multidimensional space, and the distance between the test data (point X) and each training data is calculated, and a value obtained by subtracting 1 from a value obtained by dividing the distance from the closest training data (point A) by the reference distance is defined as the degree of abnormality [3]. In this case, when the calculated distance exceeds the reference distance, the abnormality...
degree becomes a positive value, so that the abnormality can be determined based on whether the abnormality degree is positive or negative. The reference distance is set so that 99% of the training data is judged to be normal when the training data itself is evaluated. However, this method is greatly degraded if point A is actually one of the abnormal data accidentally mixed into the normal data. Therefore, the same calculation is performed up to \( k_{\text{NN}} \)-th closest training data, and the averaged value is used as the abnormal degree. This predefined number \( k_{\text{NN}} \) is called the number of nearest neighbors, and in this paper it is calculated as \( k_{\text{NN}} = 4 \). The degree of abnormality defined according to the above concept is represented by (1).

\[
f(\bar{x}) = \frac{1}{k_{\text{NN}}} \sum_{i=1}^{k_{\text{NN}}} \frac{[\bar{x} - \text{NN}_{i}(\bar{x})]}{d_i} - 1 \tag{1}
\]

where, \( \bar{x} \) is the test datum, \( \text{NN}_{i}(\bar{x}) \) is the \( i \)-th closest training datum to \( \bar{x} \), and \( d_i \) is the reference distance for \( i \)-th closest data. An open source machine learning library [8] was used to create a program that performs these processes.

### 3.3 Overview of OCSVN

The next sections overview an abnormality detection method using OCSVM.

A support vector machine is a method originally developed for two-class classification problems, and OCSVM applies this method to one-class classification problems. When using OCSVM as the abnormality detection method in our system, OCSVM is applied to the data after performing the above-described PCA and whitening. In OCSVM, as shown in Fig. 3, a set of peripheral data of training data are selected, and a decision boundary which surrounds most of the training data while passing through the peripheral data is created, and abnormalities are determined at the decision boundary. The data selected to create the decision boundary are called support vectors. The support vectors are located on or outside the decision boundary.

The decision boundary is expressed as the contour line of a function using the support vectors as parameters and the coordinates of a point in the feature space as variables. Therefore, the value of the function can be used for calculating the degree of abnormality.

OCSVM is often used in combination with the kernel method, and our program also uses the kernel method. In the kernel method, when calculating an inner product between vectors, the value of a kernel function is used as an inner product instead of a normal inner product. The kernel function used here was the standard Gaussian kernel. The Gaussian kernel is expressed by the following equation.

\[
k(\tilde{x}, \tilde{y}) = e^{-\gamma ||\tilde{x} - \tilde{y}||^2} \tag{2}
\]

where, \( k(\tilde{x}, \tilde{y}) \) is a kernel function, and \( \tilde{x} \) and \( \tilde{y} \) are vectors such as test data. Since \( \gamma \) is a parameter and needs to be adjusted to an appropriate value, a value adjusted in advance was used.

The function \( f(\bar{x}) \) for calculating the degree of abnormality is given by the following equation using a kernel function.

\[
f(\bar{x}) = -\frac{\sum \alpha k(\tilde{x}_i, \bar{x}) - \rho}{\rho} \tag{3}
\]

where, \( \rho = \sum \alpha k(\tilde{x}_i, \tilde{x}_i) \), and \( \tilde{x}_i, \alpha_i \), and \( \rho \) are the support vector, the corresponding coefficient, and the intercept.

This function \( f(\bar{x}) \) adds a negative sign to the value obtained by dividing the value of the decision function normally used in OCSVM with the intercept so that the degree of abnormality is in the range of 1 to -1 and if the abnormality is determined to be abnormal, the degree of abnormality becomes a positive value.

These parameters are determined based on the training data and are calculated as solutions to the minimization problem of (4).

\[
\min \frac{1}{2} \sum_{i=1}^{\nu} \alpha_i \alpha_j k(\tilde{x}_i, \tilde{x}_j) \quad 0 \leq \alpha_i \leq \frac{1}{\nu l} \sum_{i=1}^{\nu} \alpha_i = 1 \tag{4}
\]

where, \( \nu \) is a parameter taking a value between zero and one, and \( l \) is the amount of training data.

With the above calculations, the function \( f(\bar{x}) \) can be defined. However, as shown in the calculation process, there are two parameters \( \nu \) and \( \gamma \) that need to be set to appropriate values in advance.

### 3.4 OCSVM parameter settings

OCSVM parameters \( \nu \) and \( \gamma \) have a large effect on the abnormality detection performance, so they need to be appropriate values. If the abnormal data are available at the time of setting, the parameters can be optimized easily by actually evaluating the abnormality detection performance using the abnormal data. However, in the case of a condition monitoring system, abnormal data is often not obtained at the starting stage. Therefore, it is necessary to optimize the OCSVM parameters using only normal data.

Among \( \nu \) and \( \gamma \), \( \nu \) has the characteristics of being the upper bound on the ratio of abnormal data included in the training data (theorem 1) and the lower bound on the ratio of support vectors (theorem 2). Therefore, from the viewpoint of the theorem 1, in the case of training data that does not include abnormal data, the value of \( \nu \) should be set to a sufficiently small value. However, if the value of \( \nu \) is fixed, when the number of the training data is very large, the number of support vectors becomes very large due to the theorem 2, and the calculation time required for calculating the degree of abnormality increases. Therefore, firstly, the lower limit value of the number of support vectors is determined, and the value obtained by dividing the lower limit value by the amount of training data \( l \) is used as the value of \( \nu \). As the points of the number of dimensions \( n \) of the space are required to specify the hyperplane in the multidimensional space, the dimensionality \( n \) can be used as the lower bound of the number of support vectors. Then, if the dimensionality \( n \) of the training data is specified, the
value of $v$ is determined as follows.

$$v = \frac{n}{l} \tag{5}$$

Note that what is determined by the value of $v$ is the lower limit of the number of support vectors, and the final number of support vectors increases depending on training data.

On the other hand, in the case of $\gamma$, if the value is too large, this leads to the possibility of overfitting to the training data and misclassifying normal data as abnormal. Conversely, if the value of $\gamma$ is too small, there is insufficient fitting to the training data, and increasing the risk of the possibility of misclassifying abnormal data as normal. Therefore, as a method of setting the value of $\gamma$ to an appropriate value, Xiao et al. [9] proposed a method in which the verification data are generated from the training data and adjusting $\gamma$ to an appropriate value (DTL method). In the DTL method, verification data is created based on a concept similar to that of the nearest neighbor method. Therefore, adjustment is performed so that a result similar to the nearest neighbor method is obtained. Therefore, in this paper, $\gamma$ is determined by the DTL method.

### 3.5 Diagnosis of failure types by abnormality detection for each frequency band

As described above, by applying the one-class classification to the octave band analysis result, it is considered that a versatile abnormality detection method capable of detecting various failures of various devices can be realized. However, since all conditions other than normal are determined to be abnormal, there is a problem that the type of the abnormality cannot be determined, and it is difficult to determine a counterplan after the abnormality is detected.

On the other hand, in the vibration inspection of various types of equipment, it has long been practiced to clarify the frequency of abnormal vibration by frequency analysis and to estimate the cause of the abnormal vibration from the frequency. Applying this knowledge to traction equipment for railway vehicles, the frequency of abnormal vibration can be roughly classified into three frequency bands, high frequency, medium frequency, and low frequency as shown in Table 1.

Therefore, as shown in Fig. 4, the results of the octave band analysis are divided into three frequency bands, and abnormalities are detected for each set of data in each frequency band. Table 1 suggests that simplified diagnosis should be possible [3].

The abnormality simulation test described later uses 1/12 octave band analysis. Therefore, low frequency band is below 96 Hz, medium frequency band is from 102 Hz to 970 Hz, and high frequency band is above 1030 Hz, although the boundaries of the three frequency bands are 100 Hz and 1 kHz as described above. In the verification by the abnormality simulation test described in the next section, the results of abnormality detection for each of these three frequency bands are shown.

### Table 1 Type of abnormality and corresponding inspection

| Frequency band | Type of abnormality | Type of inspection |
|----------------|---------------------|--------------------|
| Low frequency  | Under 100 Hz        | Appearance check   |
|                |                     | Dimension measurement |
|                | Unbalance           |                     |
|                | Misalignment        | Hammering inspection |
|                | Looseness           | lubricant check     |
| Middle frequency | 100 Hz – 1 kHz   | Overhaul inspection |
|                | Looseness           |                     |
|                | Wear & collision    |                     |
| High frequency | Over 1 kHz          | Lubricant check     |
|                | Wear & collision    | Overhaul inspection |
|                | Cracks              |                     |

![Fig. 4 Division of octave band analysis result](image)

### 4. Verification of the proposed method by abnormality simulation test on a traction motor

#### 4.1 Abnormality simulation test

In order to compare the nearest neighbor method with OCSVM, vibration data was used from an abnormality simulation test [5] which was carried out for a bearing abnormality of a traction motor of a high speed train. In this abnormality simulation test, each condition monitoring item such as vibration was measured as operating the traction motor at a fixed condition. Figure 5 shows the test motor and the test bearing.

In the abnormality simulation test, an abnormal bearing was used in the test motor in order to simulate the abnormality. The inner ring of the abnormal bearing was machined to form a flat patch as shown in Fig. 5. The vibration was measured by attaching vibration acceleration sensors to the output side and the non-output side on the
test motor. The vibration acceleration sensor is a one-axis vibration acceleration sensor, and the sensor is attached so as to measure vibration in the direction of the rotation axis.

In the test, a load test and a no-load test were performed for both the abnormal bearing and the normal bearing. In the load test, the test motor was operated at the operating points shown in Fig. 6, and vibrations at each of those times were measured. In the no-load test, vibrations were measured by rotating without a load at the same rotation speed as in the load test. The measurement was performed for 180 seconds for each operating point. For normal bearings, tests were performed twice to acquire training data and test data, and the obtained vibration data was randomly sampled and split into training data and test data.

4.2 Comparison of abnormality detection results

To confirm that OCSVM detected abnormalities in the same way as the nearest neighbor method, OCSVM was applied to the abnormality simulation test results of the traction motor with bearing abnormality. Since vibrations on the non-output side of the traction motor were monitored in the assumed condition monitoring system, the results of abnormal vibration detection on the non-output side are discussed. Figure 7 shows a comparison between the results of abnormality detection by the nearest neighbor method and the OCSVM.

The chart in Fig. 7 shows the calculation result of abnormality occurrence rates indicating the ratio of data classified as abnormal for each condition. Since various vibrations occur even in a normal bearing, the occurrence of false positive is not completely eliminated in the nearest neighbor method or OCSVM. Therefore, the abnormality occurrence rate within a certain period of time serves as an index for abnormality detection.

According to Fig. 7, in the case of a normal bearing, both methods judged the vibration data to be normal for almost all data, the occurrence of false positive was small, and the abnormality occurrence rate was less than 8%. Regarding the abnormal bearing, in both cases, an abnormality was detected in high frequency components under almost all conditions, and depending on the condition, an abnormality was detected in some data even in a medium frequency component.

For high-frequency components, there were few differences between OCSVM and the nearest neighbor method depending on the conditions, such as the low abnormality occurrence rate at a full load of 50 km/h in the nearest neighbor method and almost no abnormalities detected at a full load of 200 km/h for the OCSVM. As a result, the results of the nearest neighbor method and OCSVM were almost the same, and it was confirmed that the same abnormality detection was possible.

4.3 Comparison of calculation time

The purpose of applying OCSVM is to speed up abnormality detection and enable real-time abnormality detection in the condition monitor. Therefore, the impact on shortening calculation time for abnormality detection using OCSVM was evaluated.

Fig. 5  Tested traction motor and bearing

Fig. 6  Operating point during load test

Fig. 7  Comparison of abnormality detection results (Upper: Nearest neighbor, lower: OCSVM)
During the evaluation, calculations were performed on a PC, and the elapsed time at that time was measured. However, simultaneous running of multiple programs on computers affects the execution time of the program. Therefore, in order to reduce the effect as much as possible, the same calculation was performed several times and the shortest calculation time was used, because it was assumed it reflected the case when execution was least affected by other programs being run at the same time.

The evaluation was performed for the nearest neighbor method and OCSVM. However, there are several methods for speeding up the calculation in the nearest neighbor method, and they are available in the machine learning library [8] used for abnormality detection. Among these methods Ball Tree is used as the default method. However, in the case of implementing the program in a condition monitor, the alternative is to use a method based on Brute force, which is easy to implement due to its simple algorithm. Therefore, for the nearest neighbor method, calculations were performed both using ‘Brute force’ and ‘Ball tree’. Figure 8 shows that when the nearest neighbor method (Brute force) is used as a reference, the nearest neighbor method (Ball tree) has a calculation time of about 40% and OCSVM has a calculation time of about 5% of the reference. In addition, while the calculation time of the nearest neighbor method varies depending on conditions, the calculation time of OCSVM was a fixed value. This confirms that OCSVM is an abnormality detection method suitable for real-time detection in a condition monitor.

![Comparison of calculation time](image)

**Fig. 8 Comparison of calculation time**

5. Conclusions

This paper describes the study of a method using OCSVM as an abnormality detection method to detect abnormalities in real time in a condition monitoring system that monitors the vibration of vehicle equipment and applies an on-board abnormality detection method based on machine learning. Abnormality detection performance and calculation speed were evaluated using vibration data obtained by the abnormality simulation test.

As a result, it was confirmed that OCSVM detects abnormalities in the same way as the nearest neighbor method and that the calculation time can be significantly reduced compared to the nearest neighbor method.

**References**

[1] M. Kondo, S. Manabe, T. Takashige, H. Kanno, “Traction Diesel Engine Anomaly Detection Using Vibration Analysis in Octave Bands,” *QR of RTRI*, Vol.57, No.2, May 2016.

[2] M. Kondo, T. Takashige, “Abnormality Detection for Auxiliary Drive Shafts on Diesel Cars Using Vibration Condition Monitoring,” *QR of RTRI*, Vol.59, No.1, Feb 2018.

[3] M. Kondo, “Machine diagnosis for railway vehicle equipment using anomaly detection method with octave band vibration spectra,” *Transactions of the JSME*, Article ID 17-00594, May 2018 (in Japanese).

[4] M. Kondo, “Prototype Selection Method for Vehicle Condition Monitoring using Machine Learning,” *IEEE Transactions on Industry Applications*, Vol. 139, No. 2, pp.199-205, Feb 01, 2019 (in Japanese).

[5] Y. Sakaidani, M. Kondo, “Bearing Fault Detection for Railway Traction Motors through Leakage Current,” *XIII International Conference on Electrical Machines (ICEM)*, pp.1768-1774, Sept. 2018.

[6] K. Nishitani and M.Kondo, “Application of a condition monitoring system through vibration to railway vehicles in operation,” *Japanese Railway Engineering*, 203, pp.12-14, Jan. 2019.

[7] Schölkopf, B., Williamson, R., Smola, A., Shawe-Taylor, J., & Platt, J., “Support Vector Method for Novelty Detection,” *Advances in Neural Information Processing Systems*, 12, pp.582-588, 1999.

[8] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E., “Scikit-learn: Machine learning in Python,” *The Journal of Machine Learning Research*, vol. 12, pp.2825-2830, 2011.

[9] Y. Xiao, H. Wang, L. Zhang, and W. Xu, “Two methods of selecting Gaussian kernel parameters for one-class SVM and their application to fault detection,” *Knowledge-Based Systems*, Vol. 59, pp. 75-84, 2014.

**Authors**

Minoru KONDO, Dr. Eng.
Head of Drive Systems Laboratory, Vehicle Control Technology Division
Research Areas: Condition Monitoring, Traction Systems

Yo SAKAIDANI, Dr. Eng.
Researcher, Drive Systems Laboratory, Vehicle Control Technology Division
Research Areas: Traction Motors