ANOMALY DETECTION OF MOTORS WITH FEATURE EMPHASIS USING ONLY NORMAL SOUNDS

Yumi Ono, Yoshifumi Onishi, Takafumi Koshinaka, Soichiro Takata, and Osamu Hoshuyama

NEC Corporation

ABSTRACT

This paper proposes an anomaly detection method for sound signals observed from motors in operation without using abnormal signals. It is based on feature emphasis and effectively detects anomalies that appear in a small subset of features. To emphasize the features, the method optimally estimates the contribution rates of various features to the dissimilarity score between an observed signal and the distribution of normal signals. We report here our evaluation of the method using sound data observed from PCs and fans in operation. The evaluation demonstrates that the proposed method emphasizes a small subset of narrow frequency ranges of sounds and that it achieves an error reduction rate of up to 76%.

***Index Terms*—** Anomaly Detection, Feature Emphasis, Fault Diagnosis, Fault Detection

1. INTRODUCTION

Sensor signals such as sounds and vibrations that are observed from machines in operation are often used to diagnose machine faults in factory acceptance tests or during machine maintenance. The most primitive and fundamental method involves having experts manually examine the signals. For example, they listen carefully to the sound and based on their experience, judge it to be a machine failure if it is peculiar. This method requires extensive experience and great care.

To automate this method, anomaly detection systems have been developed from accumulated know-how [1]. They were first developed for plants such as oil and gas plants [2][3]. These systems effectively detect anomalies but require both normal and abnormal signals to examine the properties of abnormal signals before operation. However, abnormal signals are rare and hence difficult to collect. Therefore, it is useful to detect anomalies without using abnormal signals.

Methods to detect anomalies without using abnormal signals have also been proposed for various machines such as spacecrafts [4][5], aircrafts [6], space shuttles [7][8], bearings and couplings of rotating machines [9], and turbine rotors [10]. These methods learn rules that capture the normal behavior or a stochastic model of the normal signals. The observed signal that deviates from the rules or the model is regarded as abnormal. They manually select features that show the anomaly. However, the features are difficult to select when the anomaly appears in various features depending on the observed signals. Furthermore, the detection accuracy decreases if an anomaly appears in a small subset of features.

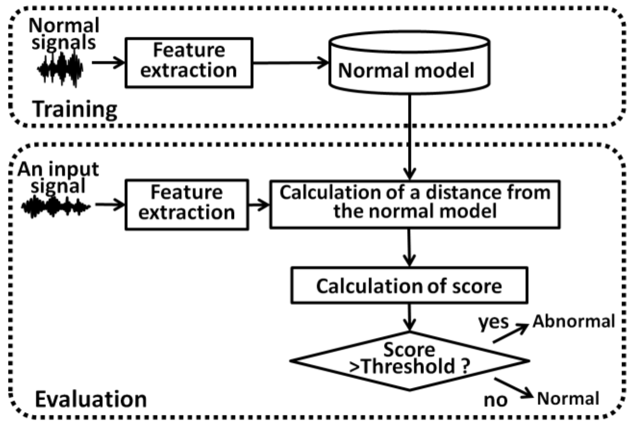

Figure 1. Flow of our anomaly detection system.

This paper proposes a method based on feature emphasis to detect abnormal sounds of motors with multi-sound sources. The anomaly appears in various features, which are the log amplitudes of sound frequencies, depending on the sound source that shows the anomaly. The features are optimally emphasized for each observed signal using the normal signals.

2. ANOMALY DETECTION METHOD

A typical method to detect anomalies only from normal signals is to create a model of normal signals and detect outliers from the model [11][12]. An input signal is regarded as abnormal when a dissimilarity score S_0, derived from the distance between features of the input signal and the model distribution of normal signals, exceeds a predefined threshold. The simplest definition of score S_0 is

$$S_0 = \frac{1}{N}\sum_{i=1}^{N} D_i , \tag{1}$$

where D_i is the distance of the ith feature (i=1, 2,.., N) and N is the number of features. The method treats all the features equally so that the accuracy decreases if an anomaly appears in a small subset of features.

In order to approach the problem, we propose a method to detect anomalies efficiently by emphasizing the small subset of features in which the abnormal signals appear. Fig. 1 shows a flow chart of our anomaly detection system. First, various features are extracted from signals of normal rotors, and a normal model, which is the distribution of normal signals, is trained from the features. Second, evaluation data are examined and determined to be normal or abnormal. The same features are extracted from the data, and the distance from the normal model of each feature is calculated. Then a newly defined score S_α is also calculated from distance D_i and finally, the input signal is detected to be abnormal if S_α exceeds a predefined threshold θ.

We define S_α as a generalized measure of S_0 in order to emphasize the features in which the abnormal signals appear. The score of each feature $s_{\alpha,i}$ and the total score S_α are

$$s_{\alpha,i} = F_{\alpha,i} D_i \,, \tag{2}$$

$$S_\alpha = \sum_{i=1}^{N} s_{\alpha,i} \,. \tag{3}$$

The emphasizing function $F_{\alpha,i}$ satisfies

$$F_{\alpha,i} = F_\alpha(D_i) = {D_i}^\alpha \Big/ \sum_{j=1}^{N} {D_j}^\alpha \,, \tag{4}$$

where index α is a constant that controls the contribution rate of each feature to score S_α . The function $F_{\alpha,i}$ emphasizes the sensitivity of the total score S_α to feature i of which D_i is high. Note that S_α with α=0 is used in the conventional method, whereas S_α with α>0 is used in the proposed method. In fact, S_α equals S_0 in (1) when α=0.

Index α is optimally determined to be α^* so that the score dispersion of the normal signals represents a minimum value. We determine α in that way to enlarge the difference between the score distribution of normal signals and that of the abnormal ones without using abnormal signals. Index α^* is given by

$$\alpha^* = \underset{\alpha}{\operatorname{argmin}} \frac{\sigma_\alpha}{\bar{S}_\alpha} \,, \tag{5}$$

$$\sigma_\alpha = \sqrt{\frac{1}{M} \sum_{m=1}^{M} \{S_{\alpha,m} - \bar{S}_\alpha\}^2} \,, \tag{6}$$

where σ_α is the standard deviation of the normal signals' scores, $\bar{S}_\alpha$ is the average score, $S_{\alpha,m}$ is the score of normal signal m, and M is the number of normal signals for training.

Here we rewrite S_0 and S_α using α-norm, which is a generalized length in a vector space, and discuss their meanings. Mathematically, the α-norm of an N-dimensional vector is

$$\| \mathbf{D} \|_\alpha = \left(\sum_{i=1}^{N} D_i^\alpha \right)^{1/\alpha} . \tag{7}$$

Using (7), we rewrite S_0 and S_α as

$$S_0 = \frac{1}{N} \| \mathbf{D} \|_1 \,, \tag{8}$$

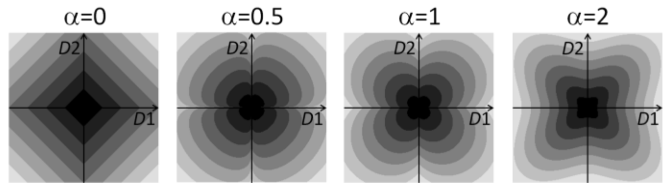

Figure 2. Contours of S_α.

$$S_\alpha = \frac{\| \mathbf{D} \|_{\alpha+1}{}^{\alpha+1}}{\| \mathbf{D} \|_\alpha{}^\alpha} \,. \tag{9}$$

Let us consider a case of two-dimensional features for simplicity. Fig. 2 displays two-dimensional contour plots of S_α when α is 0, 0.5, 1, and 2. The light-colored regions have larger S_α values than the dark regions. The contour of S_0 has a diamond shape, whereas those of $S_{0.5}$, S_1 and S_2 have flower-like shapes. This means that S_α when α is 0.5, 1, or 2 exaggerates peculiar features more than S_0 does. Thus, S_α with α>0 has the desired properties to detect an anomaly that appears in a small subset of features.

3. EXPERIMENT USING SOUND DATA OF PERSONAL COMPUTERS

An anomaly detection experiment was conducted using sound data observed from personal computers (PCs) assuming a test of silence before shipment. Each PC has three sound sources: the hard disc drive (HDD), fan, and optical disk drive (ODD). If the sound sources that made a louder sound could be detected, we could block shipments of defective PCs that do not meet a certain noise standard. Our method is worth applying in this situation because in actual situations, the abnormal data are rarely produced and are hard to obtain.

3.1. Experimental Conditions

Sound data were obtained with a piezoelectric control vibration sensor attached to operating PCs. The frequency range was 10 Hz–10 kHz. All PCs were of the same model. Each data sample is called an "event" hereinafter. Each event was 20 seconds long. Log amplitudes of spectra were obtained as features. The number of features N was 3197. Of the 20 total events, 13 normal events were used as training data, while 3 normal events and 4 abnormal events were used as evaluation data. The normal events of the training data and that of the evaluation data were alternated with each other. The Mahalanobis distance was applied to D_i defined as

$$D_i = \sqrt{(x_i - \mu_i)^T V_i^{-1} (x_i - \mu_i)} \,, \tag{10}$$

where x_i is a feature of the evaluation data, and μ_i and V_i are respectively the average and the variance of the ith feature, which is extracted from normal signals. Threshold θ was determined to be the score of the event that ranks the top 20% of all normal events.

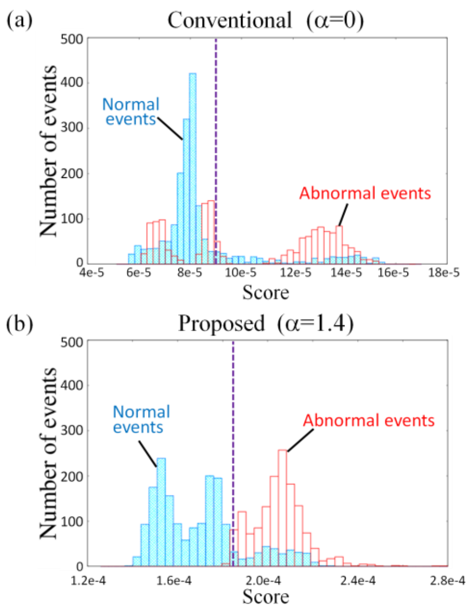

Figure 3. Score histograms.

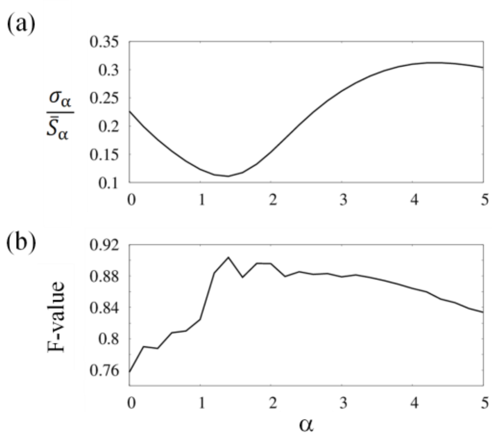

Figure 4. Standard deviation and F-value versus α.

3.2. Experimental Results

The accuracy of our anomaly detection system was evaluated with an F-value, which is the harmonic average of recall and precision. Recall refers to the ratio of events detected to be abnormal among all abnormal events. Precision refers to the ratio of abnormal events among the events detected to be abnormal. Table 1 shows the recall, precision, and F-value when α=0 and $\alpha=\alpha^*$=1.4. The F-value increased from 0.626 to 0.900 with the proposed method. Thus, the error reduction rate is 73.3%.

Figs. 3a and 3b show histograms of score S_α when α=0 and α=1.4, respectively. The shaded blue bars are the normal events, and the unshaded red bars are the abnormal events. The vertical dashed lines indicate threshold θ. The events on the right side of each line are identified as abnormal, and those on the left side are identified as normal.

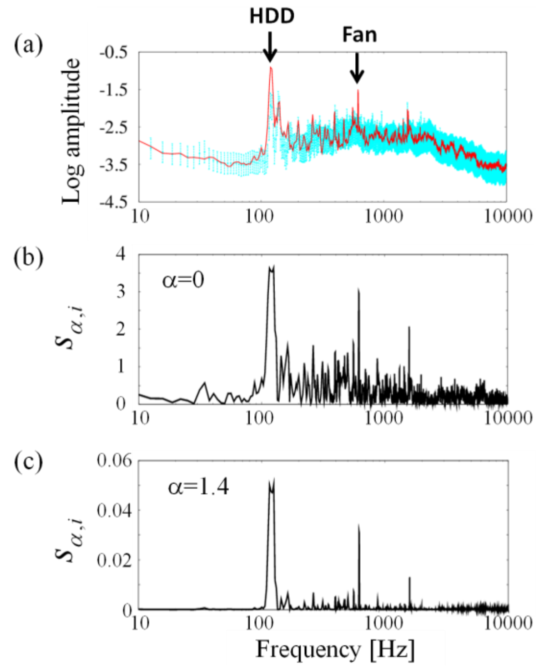

Figure 5. Abnormal event example.

There are many abnormal events that are mistakenly identified as normal in Fig. 3a, whereas such events hardly exist in Fig. 3b. This indicates that the proposed method significantly improved the detection accuracy.

Note that 20% of normal events must be identified as abnormal in both figures because threshold θ is defined as the score of the event that ranks the top 20% of all normal events. Naturally, as θ is set to be larger, the number of normal events identified as abnormal decreases, but the recall also decreases. Therefore, the choice of θ depends on the user's priority of these two properties.

Here we explain the relationship between α and F-value. Fig. 4a plots $\sigma_\alpha / \bar{S}_\alpha$ and Fig. 4b plots the F-value as a function of α. The value of α (=1.4 in the present case) that gives the minimum value of $\sigma_\alpha / \bar{S}_\alpha$, which is α^* given by (5), indeed gives the maximum F-value.

The spectrum and the scores of an abnormal event are illustrated in Fig. 5. In Fig. 5a, the red curve is a spectrum of the event, and the blue lines indicate the average spectrum and the error bars (average ± standard deviation) of the normal events. The amplitude of the event exceeds those of normal ones at 120 Hz and 620 Hz, which are the respective sound frequencies of the HDD and the fan, and this implies that the HDD and the fan generate abnormally loud sounds. Figs. 5b and 5c show the scores $s_{\alpha,i}$, where α=0 and α=1.4, respectively. The scores at 120 Hz and 620 Hz are more emphasized than those of other frequencies in Fig. 5c compared to those in Fig. 5b. Thus, those frequencies were emphasized properly by using the proposed method. In other abnormal events, the frequencies that show the anomaly were also emphasized properly even when the number of

	α=0 (conventional)	$\alpha=\alpha^{*}$=1.4 (proposed)
recall	0.554	0.996
precision	0.719	0.821
F-value	0.626	0.900

Table 1. Detection accuracy of experiment using PC sound data.

	α=0 (conventional)	$\alpha=\alpha^{*}$=0.2 (proposed)
recall	0.917	1.000
precision	0.976	0.974
F-value	0.946	0.987

Table 2. Detection accuracy of experiment using fan sound data.

sound sources generating abnormally loud sounds was different from that of the event in Fig. 5.

4. EXPERIMENT USING SOUND DATA OF FANS

Another experiment was conducted with communication devices, each of which had a single sound source consisting of a cooling fan. Sound data of cooling fans in operation were used for the experiment assuming a fault prediction test. A fan making an abnormal sound is liable to break down when electrical corrosion occurs for various reasons. Some accidents can be prevented if the anomaly sound is detected before breakdown.

4.1. Experimental Conditions

Sound data of fans in operation were obtained using various microphones. All fans were of the same model. The data samples were 4-15 seconds long and were obtained in different places; therefore, the sound level varied. To minimize the difference in sound levels, the power level was normalized so that the average amplitude equaled unity. The frequency range was 10 Hz–3500 Hz and there were 81 features N. Of the 37 total events, 23 normal events were used as training data while 3 normal events and 11 abnormal events were used as evaluation data. The normal events of the training data and that of the evaluation data were alternated with each other. The Mahalanobis distance defined in (10) was applied to D_i . Threshold θ was determined to be the score of the event that ranks the top 10% of all normal events.

4.2. Experimental Results

The accuracy of the anomaly detection was evaluated with the F-value indicated in Table 2. The F-value increased from 0.946 to 0.987, depending on the increase in recall, which was equal to 1.0. Thus, the error reduction rate is 75.9%.

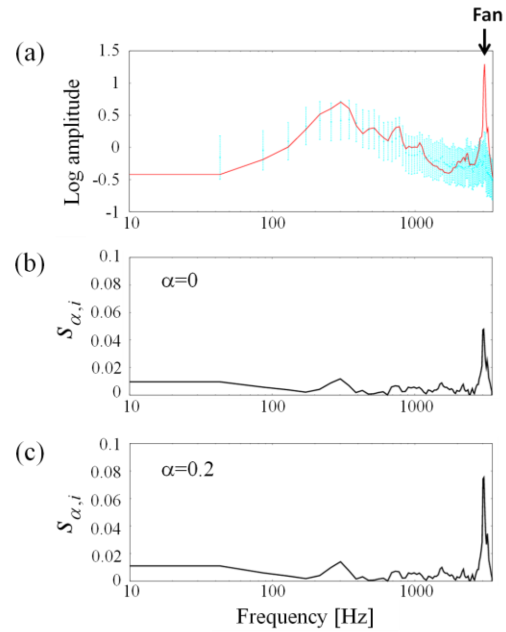

Figure 6. Abnormal event example.

One of the abnormal event examples is displayed in Fig. 6. The spectrum of the abnormal event is plotted as a red curve in Fig. 6a. The average spectrum of normal events and error bars (average ± standard deviation) are also indicated in the figure with blue lines. The red curve largely exceeds the blue line at 3300 Hz, which is the characteristic frequency of a fan with electrical corrosion. Figs. 6b and 6c show score $s_{\alpha,i}$, where α=0 and α=0.2, respectively. The score at 3300 Hz is the largest in the two figures. This property is more emphasized in Fig. 6c. Thus, an anomaly of an internal fan can be detected with high accuracy by using the proposed method.

5. SUMMARY

We presented an anomaly detection method for sound signals of rotors based on feature emphasis without using abnormal signals. To emphasize the features that show the anomaly, this method optimally determines the contributing rate of each feature to the dissimilarity score between an observed signal and the distribution of normal signals. Experiments were conducted using sound data observed from PCs and fans through vibration sensors and microphones, respectively. The error reduction rates were 73% for PC data and 76% for fan data. Our future tasks include conducting further experiments using a large number of various machine data.

Acknowledgements
We thank Takeshi Zenko and Yasutaka Nishii for providing the sound data of fans. We are also grateful to Shinichi Ando for his helpful suggestions.

6. REFERENCES

[1] L. H. Chiang, E. L. Russel, and R. D. Braatz, *Fault Detection and Diagnosis in Industrial Systems*, Springer-Verlag, London, 2001.

[2] R. A. Collacott, *Mechanical fault Diagnosis and Condition Monitoring*, Chapman & Hall, London, 1977.

[3] M. P. Boyce, C. B. Meher-Homji, and B. Wooldridge, "Condition Monitoring of Aeroderivative Gas Turbines," *Gas Turbine and Aeroengine Congress and Exposition*, Canada, ASME Paper No.89-GT-36, 1989.

[4] T. Yairi, Y. Kato, and K. Hori, "Fault Detection by Mining Association Rules from House-keeping data," Proceedings of the International Symposium on Artificial Intelligence, Robotics and Automation in Space, 2001.

[5] R. A. Martin, M. Schwabacher, N. Oza, and A. Srivastava, "Comparison of unsupervised anomaly detection methods for systems health management using space shuttle main engine data," Proceedings of the Joint Army Navy NASA Air Force Conference on Propulsion, Denver, 2007.

[6] P. Grabill, T. Brotherton, B. Branhof, J. Berry, and L. Grant, "Rotor Smoothing and Vibration Monitoring Results for the US Army VMEP," Proceedings of the American Helicopter Society 59th Annual Forum, 2009.

[7] A. Khatkhate, S. Gupta, A. Ray, and R. Patankar, "Anomaly detection in flexible mechanical couplings via symbolic time series analysis," *Journal of sound and vibration*, vol. 311, issues 3-5, pp. 608-622, 2008.

[8] H. Park, R. Mackey, M. James, M. Zak, M. Kynard, J. Sebghati, and W. Greene, "Analysis of Space Shuttle Main Engine Data Using Beacon-based Exception Analysis for Multi-Missions," Proceedings of the IEEE Aerospace Conference, New York, vol. 6, pp. 2835-2844, 2002.

[9] D. L. Iverson, "Inductive System Health Monitoring," Proceedings of the International Conference on Artificial Intelligence, IC-AI '04, CSREA Press, Las Vegas, 2004.

[10] S. E. Guttormsson, R. J. Marks, M. A. El-Sharkawi, and I. Kerszenbaum, "Elliptical novelty detection of excited running rotors," *IEEE Transactions on Energy Conversion*, vol. 14, no. 1, 1999.

[11] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey," *ACM Computing Surveys*, vol. 41, no. 3, article 15, 2009.

[12] M. Markou and S. Singh, "Novelty detection: a review-part 1: Statistical approaches," *Signal Processing*, vol. 83, no. 12, 2481-2497, 2003.