Statistical Approach to Diagnostic Rules for Various Malfunctions of Journal Bearing System Using Fisher Discriminant Analysis

Byungchul Jeon[1], Joonha Jung[1], Byeng D. Youn[1], Yeonwhan Kim[2], and Yong-Chae Bae[2]

[1]*Department of Mechanical and Aerospace Engineering, Seoul National University, Seoul, 151-742, Republic of Korea*

puurni@empas.com
reallibero@gmail.com
bdyoun@snu.ac.kr

[2]*Power Generation Laboratory, KEPCO Research Institute, Daejeon, 305-760, Republic of Korea*

ywkim.kepri@gmail.com
baeych@kepco.co.kr

ABSTRACT

This research is focused on developing an efficient fault diagnosis procedure for a journal bearing system. Vibration data of journal bearing rotor simulator under four conditions (i.e. a normal condition and three anomaly conditions including unbalance, rubbing and misalignment) was used to develop the algorithm. In order to improve diagnostic performance, cycle based time-domain features and frequency-domain features were extracted after resampling process being applied to the raw vibration data. Then, the optimal feature selection was accomplished by mixture of random combination performance test and Fisher Discriminant Ratio (FDR) criteria. After selecting optimal features, Fisher Discriminant Analysis (FDA) algorithm classified each abnormal conditions mentioned above. To end with, the result of classification is evaluated and verified.

1. INTRODUCTION

The modern machineries widely deployed in manufacturing sectors and power plant facilities have rotors as a core part. Naturally, bearings supporting the rotors frequently fail to perform their designed responsibility due to various reasons. Failure in bearings may affect the entire system to deteriorate or cause stopover of the system since it incorporates high energy. This also can generate casualties or damages when the counter measures are not held in suitable time (Yaguo Lei, He, & Zi, 2008). To maintain performance of the rotating machineries and to prevent the catastrophe of having casualties and economic loss, numerous attempts have been made to diagnose the faults in their initial states.

Vibration data is one of the reliable parameters that efficiently represents the performance of machineries, and it is widely used to define the health status of systems (Gupta, 1997; Yaguo Lei, He, Zi, & Chen, 2008; Ocak, Loparo, & Discenzo, 2007). However, without proper signal processing techniques and knowledge on vibration, the data itself does not denote any information of health status. Though sometimes even when processing has been done properly, lack of knowledge hinders the successful diagnosis. Therefore, the need for setting up a reliable diagnose algorithm without any help from experts has been steadily increasing (Jardine, Lin, & Banjevic, 2006; Y. Lei, He, Zi, & Hu, 2007; Wong, Jack, & Nandi, 2006).

In response to the request, an automatic diagnosis algorithm implementing Artificial Neural Network (ANN) was developed (Chen & Mo, 2004; Li, Chow, Tipsuwan, & Hung, 2000; Samanta & Al-Balushi, 2003). Vibration data were acquired from both the normal and abnormal bearing system, and from the data time-domain features or frequency-domain features were extracted, which were used as an input for ANN. ANN diagnosed the system as normal or abnormal upon those features. In addition, features from wavelet analysis in time-frequency domain facilitated constructing ANN based diagnosis (Al-Raheem & Abdul-Karem, 2010; Han, Yang, Choi, & Kim, 2006; Sanz, Perera, & Huerta, 2007; Yang, Han, & An, 2004). Rather than piling more features, study on selecting effective features such as genetic algorithm took a part in this process (Han et al., 2006). Many fault diagnosis algorithms based on ANN have been introduced. However the limited use of ANN, which require certain amount of data, had led a way to other machine learning (Ahmadi, Moosavian, & Khazaee, 2012).

In order to overcome the limitation in ANN, Support Vector Machine (SVM) based algorithms were suggested. Since SVM is a linear classifier for two-class problems, its use has been limited to linearly separable data sets. However, with the invention of the kernels and other techniques as well, SVM has gained popularity among researchers (Huo-Ching & Yann-Chang, 2012; Yang, Han, & Hwang, 2005). Often, ANN and SVM were used individually to compare performance of each algorithm, whereas others tried to combine these two method to generate more reliable diagnosis algorithm (Salahshoor, Kordestani, & Khoshro, 2010; Samanta, Al-Balushi, & Al-Araimi, 2003).

Fisher Discriminant Analysis (FDA) is another widely used machine learning technique. The basic principle of FDA is similar to that of SVM, but then FDA utilizes the scatter of data rather than the data itself. The advantage of using scatter over data lies in computational efficiency. Specifically, for large multi-class data set, FDA can save its resources while SVM wastes resources finding the optimal vector. The performance difference of FDA and SVM depends on the data set, which does not show much difference in this research. Thus, FDA was chosen as the main classifying algorithm.

In this research, advanced fault diagnosis algorithm for journal bearing system has been developed. Advanced algorithm can be attributed to the features extracted from vibration per cycle while other researches have extracted features for certain amount of time. ‘A cycle’ method allows to identify the fault characteristics of the vibration signal more thoroughly. To achieve features per cycle, data were resampled before being extracted. Then, extracted features numbered more than 50, which needed dimensional reduction. In addition, not only the features incorporating cycle characteristics of vibration but also average and standard deviation of multiple cycles can represent the faults clearly. Features selection method by Fisher Discriminant Ratio (FDR) and random combination of features has been applied.

Through the paper, the following section will cover the type of features extracted from the test-bed. Then, in section three feature extraction procedures will be clearly stated, and in section four, feature selection method will be revealed. Finally, the result of the classification will be discussed.

2. EXPERIMENTAL SETUP AND DATA ACQUISITION

2.1. Experimental Setup

The RK4 rotor kit of GE Bently Nevada was used as a journal bearing rotor system for implementing anomaly conditions. This experimental apparatus is shown in Figure 1. Rotor shaft with a disc of 800g supported by two journal bearings were tested. Two shafts were connected by a flexible coupling to acquire more reliable data. The vibration data was acquired from the middle of the test-bed, close to the point where the abnormal conditions were induced. Among several anomaly conditions of rotor systems, three kinds of abnormal conditions, unbalance, rubbing, misalignment, were induced to the test-bed.

For unbalance test, a small amount of weight has been injected in the disc. Rubbing test was done by a rubbing screw to make partial rub on shaft. Additional misalignment device with ball bearing shifted the shaft up & downward to produce misaligned shaft data. In addition to those conditions, normal data was set as a reference.

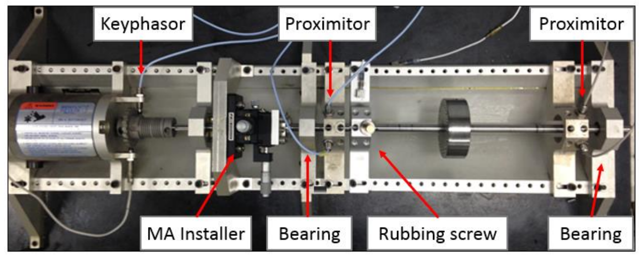

Figure 1. RK4 test-bed

2.2. Vibration Data Acquisition

To achieve consistent and reliable data sets, weight balancing procedure preceded the actual experiment. Unlike a ball bearing system or a roller bearing system, a journal bearing system shows relatively simple sinusoidal wave. Even the slightest alteration of the settings result in a big change of the waves. For example, improper disc joining practice will cause differences in the signal. Therefore, among various candidates, vibration amplitude and phase have been selected to represent the initial state of the system. So as to have consistent amplitude and phase throughout the whole data sets, balancing procedure preceded every experiment to make the system fit into the same amplitude and phase. This preceded action gives reliability to compare with the other data sets.

After the balancing procedure is done, vibration data for four conditions can be achieved from the proximity probe installed between the journal bearings. Two points on the shaft, just beside the bearings have been chosen, and at each point, two probes are mounted at a right angle to receive voltage signal. Both the relative and absolute displacement between the sensor and the shaft can be measured. In addition to the time-based vibration signals of each sensor, shaft centerline orbit could be tracked via vibration signals of two probes mounted in a right angle. The phase information can be obtained through the keyphasor signal which prints once-per-revolution pulse to provide a precise timing measurement. This keyphasor signal enables us to dissect the signal into a cycle unit, which will be discussed in section 4. Vibration signals of proximity sensor were acquired by the rate of 4,000 samples/s via NI DAQ 4432. Each normal and abnormal state has been repeated three times, and for each case, data was obtained for 60 second long at 3600 rpm.

3. FEATURE TYPES

The vibration data itself may show the difference among abnormal conditions graphically. Specifically, for journal bearing systems, modified sinusoidal wave of vibration undeniably proves that the system is not in a normal state. However, all data cannot be analyzed manually due to its large size. It will take tremendous amount of resources if all the data is processed by humans, while losing the reliability due to human factors. It is no doubt that quantified indicator is required to precisely diagnose malfunctions and to utilize the automatic system which can process big data in a short period of time.

Statistical parameters were defined for features of vibration data, the quantified indicator of vibration. Some features were extracted for every cycle, while others were extracted for number of cycles. Whether a cycle or few cycles, unit for features must be defined considering the statistical definition and implication.

3.1. Time-domain Features

Time-domain indicates statistical features from the pre-defined period of vibration data. Maximum, root-mean-square, kurtosis and more features are extracted from every rotation. Also, mean and deviation for every rotation in 60 cycles at 3600 rpm, are defined as each features.

The first three features in Table 1 represents the vibration amplitude. In other words, they can be indicator of kinetic energy of the system. The next five features form skewness to entropy can be interpreted as indicators of shape of the wave. Upper/lower bound and AR coefficient represents distribution characteristics and signal changes over time. Especially, the information of orbit can be gathered via proximity probe mounted at a right angle. The mean and variance of each time-domain feature for 60 cycles are also adopted as features for anomaly diagnostics. Table 1 lists the features of time-domain.

3.2. Frequency-domain Features

Features in frequency-domain also implies important characteristics of vibration signals as much as time-domain features. All the frequency features are based on the power spectrum for one-second long data. Power spectrum itself shows distribution of the frequency elements, but needs to be quantified just like the vibration data.

Five features were defined in this paper. The definition of frequency center (FC), root mean square frequency (RMSF), and root variance frequency (RVF) are stated in Table 2. (Wei, Guo, Jia, Liu, & Yuan, 2013; Yang & Widodo, 2009).

$s(f)$ denotes the power spectrum of signal, so that according to the definition FC and RMSF show alteration in position change of main frequencies, RVF describes the convergence of the spectrum power. Additionally, two more

Table 1. Time-domain features

Features	Description
Maximum	$\mathrm{Max}(X_i)$
Mean absolute	$\mathrm{Mean}(\lvert X_i \rvert)$
RMS	$\sqrt{\frac{\sum {X_i}^2}{N}}$
Skewness	$\frac{\sum (X_i - \bar{X})^3}{(N-1)s^3}$
Kurtosis	$\frac{\sum (X_i - \bar{X})^4}{(N-1)s^4}$
Crest factor	$\frac{X_{peak}}{X_{rms}}$
Shape factor	$\frac{X_{rms}}{\mathrm{Mean}(\lvert X_i \rvert)}$
Impulse factor	$\frac{\mathrm{Max}(X_i)}{\mathrm{Mean}(\lvert X_i \rvert)}$
Entropy	$-\sum p_i \times \log p_i$
Upper bound	$\mathrm{Max}(X_i) + \frac{\mathrm{Max}(X_i) - \mathrm{Min}(X_i)}{2(N-1)}$
Lower bound	$\mathrm{Min}(X_i) - \frac{\mathrm{Max}(X_i) - \mathrm{Min}(X_i)}{2(N-1)}$
AR Coefficient	Auto regressive coefficient(1st to 8th)
Effective orbit radius(1x, total)	$\frac{\sum (X_i^2 + Y_i^2)}{N}$
Aspect ratio of 1x orbit	$\frac{\text{Minor Axis}}{\text{Major Axis}}$

Table 2. Frequency-domain features

Features	Description
FC	$\frac{\int f \times s(f) df}{\int s(f) df}$
RMSF	$\left[\frac{\int f^2 \times s(f) df}{\int s(f) df}\right]^{1/2}$
RVF	$\left[\frac{\int (f - FC)^2 \times s(f) df}{\int s(f) df}\right]^{1/2}$
2X / 1X	$\sqrt{\frac{s(f_{2X})}{s(f_{1X})}}$
(Total-1X) / 1X	$\frac{\left[\int \sqrt{s(f)} df - \sqrt{s(f_{1X})}\right]}{\sqrt{s(f_{1X})}}$

features regarding the ratio between the main and other frequency components are introduced as in the last two rows in Table 2. $\sqrt{s(f_{1X})}$ and $\sqrt{s(f_{2X})}$ indicates the magnitude of 1X and 2X component of vibration signal, respectively.

4. STATISTICAL ANOMALY DETECTION METHODS

4.1. Feature Extraction

In this research, 47 features have been set as the candidate parameters for anomaly diagnosis of above mentioned conditions. Both time-domain and frequency-domain features are extracted for one rotation or/and one second. As stated in previous section, the raw vibration data should be segmented to maintain consistency of the features.

4.1.1. Preprocessing for Feature Extraction

The fundamental frequency of journal bearing systems dominates other frequencies. Naturally, the sub-harmonic frequencies as well as super-harmonic frequencies were often utilized in traditional diagnosis algorithm (Randall & Antoni, 2011). In this study, the test-bed used here shows typical journal bearing characteristics, so that features are extracted based on cycles. Feature extraction unit differs according to feature types, some use one rotation while others use multiple rotations. For either of the case, keyphasor signal must be implemented to segment the signal into exact cycles. The sampling rate, 4,000 samples per second, creates unevenly distributed sample points per cycle at target speed of 3,600 rpm, as in the Figure 2(b). And even if the sampling rate has altered to multiples of speed, the rotating speed cannot be controlled at exactly 3,600 rpm, which makes resampling process inevitable.

Resampling process enables the signal to have same number of data points per cycle. For example, in Figure 2(c), resampled signal shows eight points per cycle. With the given sampling rate and the target speed, signal was resampled to have 64 points per cycle starting from the keyphasor signal to the next keyphasor signal. Intervals between data points were set by equivalent rotation angle difference, so as to have same data points even when the rpm changes. The resampled signal can now be used to extract features in accordance with the same criteria.

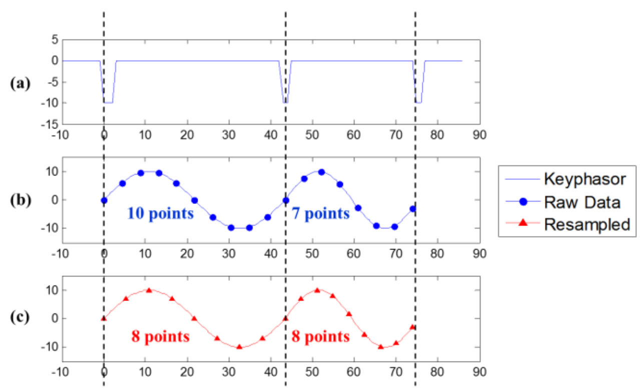

Figure 2. Resampling procedure (a) Keyphasor signal (b) raw Signal (c) resampled Signal

4.1.2. Cycle based Feature Extraction

As stated in section 3, time-domain features are extracted based on a cycle or several cycles. Features from certain period of time are universally used in developing fault diagnosis. However, considering the fact that fundamental frequency dominates in the journal bearing system, and the sensitivity that journal bearing sinusoidal waveforms have, one rotation of a signal would regard significant amount of information. If features are extracted one second without applying resampling process, for example, the particular information on a sinusoidal wave fades away as it is averaged with other non-particular information. This is the reason we are focusing on the cycle based features for journal bearing. Simultaneously, features related to valuable information such as the trend being shifted to other states are extracted from 60 cycle data. Widely scattered features of a cycle will grant a large variation, which itself can be an independent feature. Therefore, time-domain features are statistically described by the mean and variance terms of time-domain features.

On the other hand, for frequency-domain features, it is desirable to extract 60 cycle based features. The longer the signal is acquired, the higher resolution of FFT result can be achieved. Since the target speed is 60 rev/sec, extreme high frequency components are not required. Rather sub-harmonic frequencies or super-harmonic frequencies are required.

So far, from raw vibration data 47 features are extracted. Among the 47 extracted ones, a few features have been chosen to check whether they are able to separate the malfunctions clearly. As presented in Figure 3, health classes can be unclearly or clearly separated depending on selection of a key feature set. In other words, it is sufficient to classify all health states if a key feature set is properly selected.

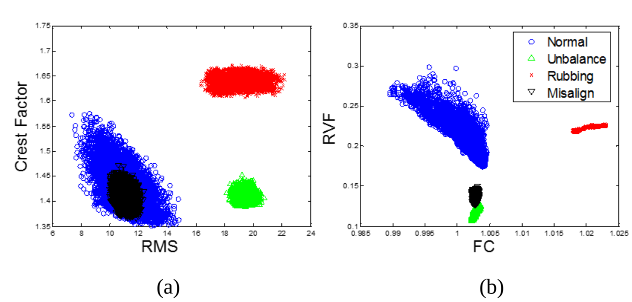

Figure 3. Graphical expression of features in (a) time-domain (b) frequency-domain

4.2. Feature Selection

Accuracy and computational efficiency are the two main factors that define the performance of the diagnosis algorithm. In view of those two points, the best feature sets are minimum number of features that produce good result. Minimizing the number of features would greatly contribute to reducing computational demands. Reducing the time and effort for computation may be very critical to some real-time diagnosis systems. Although real-time is not required, some features might hinder the characteristics of the data group which deteriorates algorithm performance. Therefore, many researches had been conducted solely on feature selection.

In this research, feature selection was accomplished by mixture of Fisher Discriminant Ratio ranking and random combination performance test.

4.2.1. FDR & Correlation Coefficient Ranking

FDR is a criterion that indicates separable ability for two-class data. In this research each abnormal conditions can be regarded as a class, as of universal terms. So high FDR value means that it can distinguish an abnormal condition from another condition. Its definition is in equation (1). The numerator shows that difference between mean of each class. In the denominator variance for each class data are summed to represent how well class data is congested. Specifically, two class data, whose mean difference is large, and which has small variance, FDR value for the feature will grant a high value (S. Theodoridis & Koutroumbas, 2008).

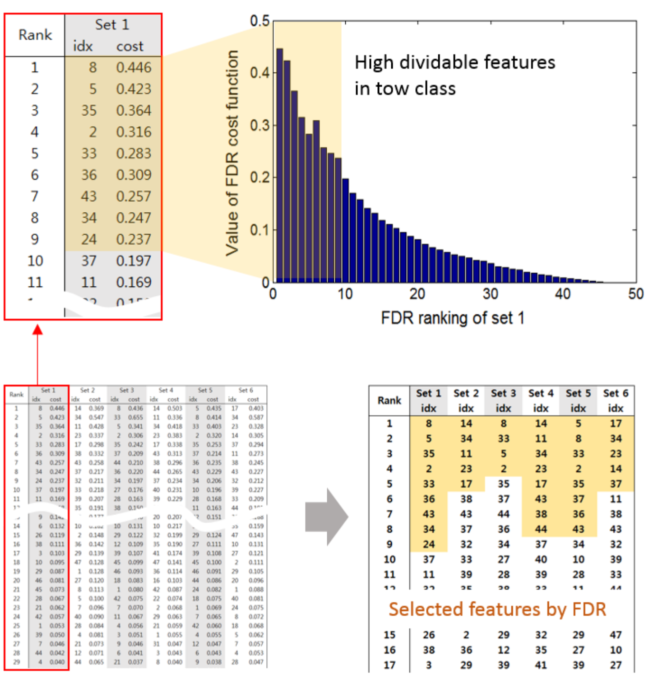

Figure 4. Feature selection using FDR

$$FDR = \frac{(\mu_i - \mu_j)^2}{\sigma_i^2 + \sigma_j^2} \qquad (1)$$

The explained FDR values will be derived for every feature, and also for every abnormal combination sets of two. In this research 47 features are extracted for four classes, so total of 47 x ${}_4C_2$ FDR values will be calculated.

However, FDR criteria does not take any consideration in reducing number of features. It only gives separable ability of individual features. Hence correlation coefficients between features are deliberated to obtain a cost function in equation (2). This cost function will sort out the features in a new ranking. The feature that used to have higher FDR value might be ranked very low in a new cost function ranking, and vice versa. The cost function can be used as a criteria for reducing the number of features.

$$i_k = \arg\max_{j} \left\{ a_1 C_j - \frac{a_2}{k-1} \sum_{r=1}^{k-1} \left| \rho_{i_r, j} \right| \right\} \qquad (2)$$

Yet, the combined FDR and correlation ranking is still based on two-class problems, which does not guarantee decent performance features for multi-class problems as well. To utilize in multi-class, random combination method is used.

4.2.2. Random Combination Test

To apply the feature rankings to multiple class problem, random combination of features are selected and evaluated by the performance of classification of training data set. First, the cost function value of feature rankings in section 4.2.1. is examined. Though its absolute values do not hold crucial meaning, they can be used as a rough measure for separable ability in each two-class sets. As shown Figure 4, for each two class combination set, features that have less than half of the maximum value of cost function are discarded as they have bad separable ability. Selected high

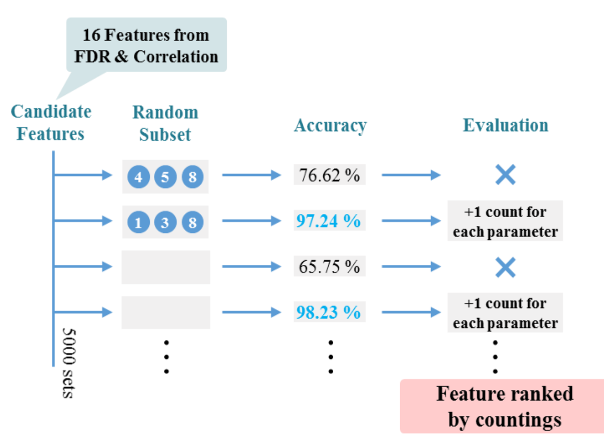

Figure 5. Random combination testing process

separable features represented in right bottom of Figure 4(the orange colored values). To find priorities among the selected high separable features, random combination test was applied. The brief process is shown Figure 5.

Random combination of features have tested 5000 times in this study. The occurrence of individual feature is accumulated when the prediction accuracy is above the threshold. The priority is ranked by the accumulated occurrence descending. The result in details will be described in section 5.1.

4.3. Classification – Fishier Discriminant Analysis

FDA (Fisher Discriminant Analysis) was used for a classification scheme. FDA classification algorithm is to find a hyper-plane, where projected data on to this plane maximizes the cost function, FDR(Welling, 2005).

In the two-class problem, hyper-plane becomes a single line, represented by **w**. Assuming that the data are projected, high FDR corresponds to the difference of mean value being far away and the variance of each class being as small. Finding the line **w** manually might be computationally demanding, but the maximum eigenvalue of $S_w^{-1}S_B$ matrix is proven to be the vector **w**, where S_B means covariance between classes, and S_W means covariance within the class.

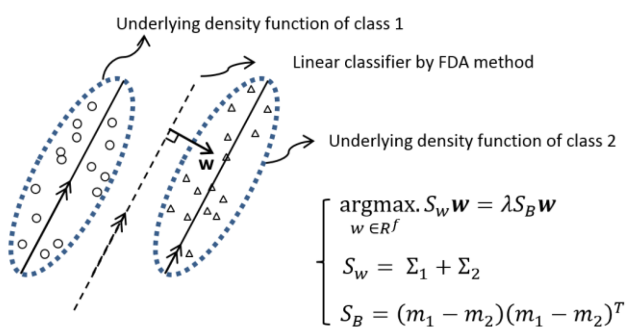

Figure 6. Fisher Discriminant Analysis for two class problem

For the multi-class problem, FDA criteria is substituted to other cost function J_3, and the rest are the same as the two-class problem(Sergios Theodoridis & Theodoridis, 2010).

$$J_3 = trace\{S_w^{-1} S_B\} \tag{3}$$

The overall procedure of developing classification starts with acquiring data sets. Other researches have used a part of one set for training, and the rest for testing. Conversely, this research acquired two sets of data, one for training and the other for testing. After both data sets were resampled and normalized, the defined features were extracted. Then, the features from training data set was used in feature selection process since there will be no testing data in real systems. Selected features of the training data were utilized to develop the classification model by FDA, and the three **w** vectors were derived. The selected features of testing data are classified with the **w** vectors.

5. RESULTS

This section can be divided into two parts. The first one will discuss the optimal selected features accomplished by feature selection process. The latter part will discuss the result of class prediction of testing data sets. The training set and testing set is listed in the Table 3.

Table 3. Training and testing data sets

Training Data				Testing Data			
Conditions		Data #	Features	Conditions		Data #	Features
		N	d			N	d
Normal	1	3600	47	Normal	1	3600	47
	2	3600	47		2	3600	47
	3	3600	47		3	3600	47
Unbalance	1	3600	47	Unbalance	1	3600	47
	2	3600	47		2	3600	47
	3	3600	47		3	3600	47
Rubbing	1	3600	47	Rubbing	1	3600	47
	2	3600	47		2	3600	47
	3	3600	47		3	3600	47
Misalign-ment	1	3600	47	Misalign-ment	1	3600	47
	2	3600	47		2	3600	47
	3	3600	47		3	3600	47

Before stating the result, data sets must be organized clearly. For feature selection and training the classifier, only training data sets were used. At classifier evaluation step, the testing data set was predicted using the classifier developed by training data sets.

5.1. Feature Selection Results

The main function of feature selection is to reduce the dimension to increase the efficiency of diagnosis algorithm. The test-bed vibration data had been transformed to time-domain and frequency-domain features. Total 47 candidate features were extracted to be used as an input in classification. However, 47 seemed heavy even for the simplest classification algorithm, because the number of data, or cycles, was quite large. At the same time, applying too much features in poor separability for anomaly diagnosis may lowering the efficiency of the classifier. When all 47 features are used, the class prediction for the training set leaves only 74.7% accuracy, because not all the features were capable of classifying the conditions. So, feature selection by mixed FDR and correlation coefficient criteria was performed. Features that had higher value than the half of the maximum cost function value had been selected. Through this mixed feature selection method, 16 features, almost one third of all 47 features, were recognized as valid parameters. These selected 16 parameter are same as the number listed in the x-axis of Figure 7.

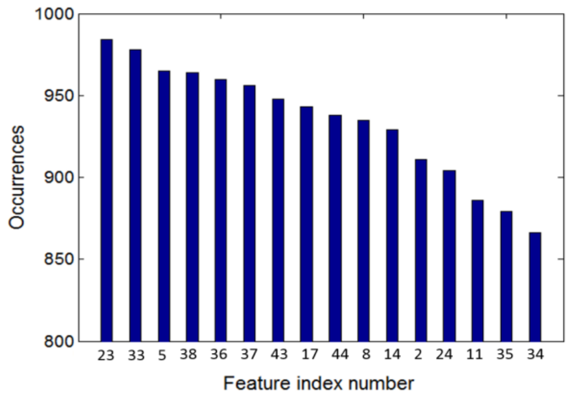

Figure 7. Occurrences of individual feature in random combination above threshold for accuracy

Among these 16 remaining features, three features were selected randomly for 5000 times to evaluate the performance of the combinations. Three was selected as the least number of parameters for classifying the four-class problem. Each feature combination of training data set in Table 3 were trained and tested. In order to acquire the optimal features, the threshold of prediction accuracy was used as 80%. The 80% criteria above is supposed to be reasonable in a sense that prediction accuracy using all 47 features yielding 74.7%, but further research needs to be done. Then, only the eligible feature combination scores the individual features as shown in Figure 5. The result produces a ranking list of 16 features, which are used in section 5.2 to predict the testing set classes. Through these selection process, optimal feature sets could be picked.

5.2. Classification Results

Before referring the classification result, the proposed feature extraction method in section 4.1. enhanced the consistency in features. Compared to the features from the previous studies, based on certain period of time, the proposed features showed separable ability more than twice as well as the previous ones.

With the improved features, FDR feature selection method was performed to find the optimal features for classification. The first step was to obtain the FDR & correlation ranking list which is based on only training sets. Then, feature combinations according to the list rankings were formed and classified the testing data set. Starting from the top three feature combinations, a next ranking feature was added each time after classification result was attained. The result is shown in Figure 8.

As it is displayed in the chart, all 16 feature combination does not yield good prediction result. Rather smaller number, from three to eleven features, gave 100% accurate prediction. In addition to the improved accuracy, computation time was saved greatly. The result more than

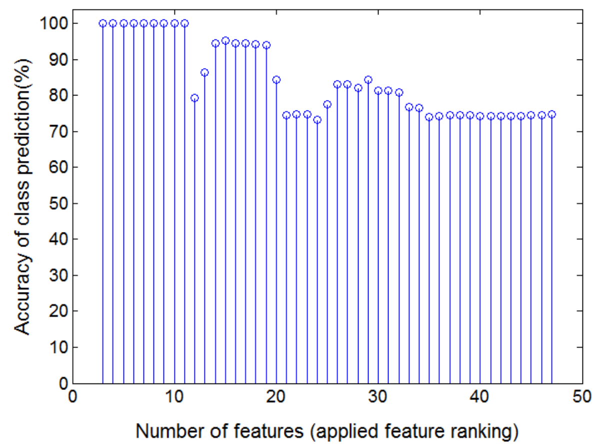

Figure 8. Classification accuracy by number of features

16 features have been achieved by adding left features after feature selection process. This chart insists that feature selection process was successful.

6. CONCLUSION

In this research, diagnose algorithm for four conditions of journal bearing systems has been developed. Two separate data sets were grouped as training set and testing set, respectively. Each of the condition was repeated three times and each test preceded the balancing procedure to enhance the reliability of the data sets. The initial vibration amplitude, indeed, had crucial effect in consistency. Considering the characteristics of a journal bearing system, features have been extracted based on a cycle or cycles after the proximitor signal was resampled. Keyphasor signal has made the resampling procedure possible, and that cycle segmentation became possible. Total of 47 Cycle based features are defined in time-domain and frequency-domain. Among those features 16 of them had been chosen to be effective parameter by FDR criteria and random combination performance test. This feature selection played key role in developing competent diagnosis algorithm with only three to eleven features being used. However, when choosing the features via random combination method, the accuracy threshold, which has not been studied deeply, plays key role. Further research must be conducted on this subject.

ACKNOWLEDGEMENT

This work was supported by the Power Generation & Electricity Delivery Core Technology Program of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) granted financial resource from the Ministry of Trade, Industry & Energy, Republic of Korea. (No. 2012101010001C)

REFERENCES

Ahmadi, H., Moosavian, A., & Khazaee, M. (2012). An appropriate approach for misalignment fault diagnosis based on feature selection and least square support vector machine.

Al-Raheem, K.F., & Abdul-Karem, W. (2010). Rolling bearing fault diagnostics using artificial neural networks based on laplace wavelet analysis. *International Journal of Engineering, Science and Technology, 2*(6).

Chen, C., & Mo, C. (2004). A method for intelligent fault diagnosis of rotating machinery. *Digital Signal Processing, 14*(3), 203-217.

Gupta, K. (1997). Vibration—a tool for machine diagnostics and condition monitoring. *Sadhana, 22*(3), 393-410.

Han, T., Yang, B.-S., Choi, W.-H., & Kim, J.-S. (2006). Fault diagnosis system of induction motors based on neural network and genetic algorithm using stator current signals. *International Journal of Rotating Machinery, 2006*.

Huo-Ching, S., & Yann-Chang, H. (2012, 4-6 June 2012). *Fault diagnosis of steam turbine-generator sets using evolutionary based support vector machine.* Paper presented at the Computer, Consumer and Control (IS3C), 2012 International Symposium on.

Jardine, A.K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing, 20*(7), 1483-1510. doi: DOI 10.1016/j.ymssp.2005.09.012

Lei, Y., He, Z., & Zi, Y. (2008). A new approach to intelligent fault diagnosis of rotating machinery. *Expert Systems with Applications, 35*(4), 1593-1600.

Lei, Y., He, Z., Zi, Y., & Chen, X. (2008). New clustering algorithm-based fault diagnosis using compensation distance evaluation technique. *Mechanical Systems and Signal Processing, 22*(2), 419-435.

Lei, Y., He, Z., Zi, Y., & Hu, Q. (2007). Fault diagnosis of rotating machinery based on multiple anfis combination with gas. *Mechanical Systems and Signal Processing, 21*(5), 2280-2294. doi: DOI 10.1016/j.ymssp.2006.11.003

Li, B., Chow, M.-Y., Tipsuwan, Y., & Hung, J.C. (2000). Neural-network-based motor rolling bearing fault diagnosis. *Industrial Electronics, IEEE Transactions on, 47*(5), 1060-1069.

Ocak, H., Loparo, K.A., & Discenzo, F.M. (2007). Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: A method for bearing prognostics. *Journal of sound and vibration, 302*(4), 951-961.

Randall, R.B., & Antoni, J. (2011). Rolling element bearing diagnostics—a tutorial. *Mechanical Systems and Signal Processing, 25*(2), 485-520.

Salahshoor, K., Kordestani, M., & Khoshro, M.S. (2010). Fault detection and diagnosis of an industrial steam turbine using fusion of svm (support vector machine) and anfis (adaptive neuro-fuzzy inference system) classifiers. *Energy, 35*(12), 5472-5482.

Samanta, B., & Al-Balushi, K. (2003). Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. *Mechanical systems and signal processing, 17*(2), 317-328.

Samanta, B., Al-Balushi, K., & Al-Araimi, S. (2003). Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. *Engineering Applications of Artificial Intelligence, 16*(7), 657-665.

Sanz, J., Perera, R., & Huerta, C. (2007). Fault diagnosis of rotating machinery based on auto-associative neural networks and wavelet transforms. *Journal of Sound and Vibration, 302*(4), 981-999.

Theodoridis, S., & Koutroumbas, K. (2008). *Pattern recognition*: Elsevier Science.

Theodoridis, S., & Theodoridis, S. (2010). *Introduction to pattern recognition : A matlab approach*. Burlington, MA: Academic Press.

Wei, X., Guo, K., Jia, L., Liu, G., & Yuan, M. (2013). Fault isolation of light rail vehicle suspension system based on ds evidence theory and improvement application case. *Journal of Intelligent Learning Systems & Applications, 5*(4).

Welling, M. (2005). Fisher linear discriminant analysis. *Department of Computer Science, University of Toronto*.

Wong, M., Jack, L., & Nandi, A. (2006). Modified self-organising map for automated novelty detection applied to vibration signal monitoring. *Mechanical Systems and Signal Processing, 20*(3), 593-610.

Yang, B.-S., Han, T., & An, J.L. (2004). Art–kohonen neural network for fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing, 18*(3), 645-657.

Yang, B.-S., Han, T., & Hwang, W.-W. (2005). Fault diagnosis of rotating machinery based on multi-class support vector machines. *Journal of Mechanical Science and Technology, 19*(3), 846-859.

Yang, B.-S., & Widodo, A. (2009). *Introduction of intelligent machine fault diagnosis and prognosis*: Nova Science Publishers.

BIOGRAPHIES

Byungchul Jeon received the B.S. degree of aerospace engineering from Republic of Korea Air Force Academy in 2001 and the M.S. degree of mechanical engineering from

Seoul National University in 2009. He is now a Ph. D. student at the School of Mechanical and Aerospace Engineering in Seoul National University. His current research is focused on the data driven Prognostics and Health Management for rotating machinery.

Joonha Jung received B.S. degree of mechanical engineering from Seoul National University (SNU) in 2012, and is involved in graduate School of Mechanical and Aerospace Engineering at SNU. He is currently doing research on Prognostics and Health Management for rotating machinery.

Byeng D. Youn received his Ph.D. degree from the department of Mechanical Engineering at the University of Iowa, Iowa City, IA, in 2001. He was a research associate at the University of Iowa (until 2004), an assistant professor in Michigan Technical University (until 2007), and an assistant professor in the University of Maryland College Park (until 2010). Currently, he is an associate professor at the School of Mechanical and Aerospace Engineering at Seoul National University, Republic of Korea. His research is dedicated to well-balanced experimental and simulation studies of system analysis and design, and he is currently exploring three research avenues: system risk-based design, prognostics and health management (PHM), and energy harvester design. Dr. Youn's research and educational dedication has led to: six notable awards, including the ISSMO/Springer Prize for the Best Young Scientist in 2005 from the International Society of Structural and Multidisciplinary Optimization (ISSMO), and more than 100 publications in the area of system-risk-based design, PHM and energy harvester design.

Yeon-Whan Kim received his BS and MS degree from the department of Mechanical Engineering from Ajou University in 1987 and 1990, and his Ph.D. degree from the department of Mechanical Design Engineering at Chungnam National University in 2008, Korea. He is a Principal Researcher at Power Generation Laboratory in KEPCO Research Institute. His research is dedicated to machinery vibration analysis, noise analysis, mechanism identification on flow induced vibration, health condition monitoring and diagnostics in power plants. His current research is focused on prognostics and health management of rotating machinery in 1000MW class thermal power plant.

Yong-Chae Bae received his BS, MS and Ph.D. degrees from the department of Mechanical Engineering from Chonnam National University, Republic of Korea, respectively. He is a Mechanical Engineering Group Leader at Power Generation Laboratory in KEPCO Research Institute. His research interests are vibration analysis, health monitoring and diagnostics in power plants.