Anomaly detection approach for fan monitoring –
an industrial case study

Antoni Lis1,2, Ziemowit Dworakowski1 and Piotr Czubak1
1AGH University of Science and Technology, Poland,
{alis@agh.edu.pl, zdw@agh.edu.pl, czubak@agh.edu.pl}
2Elmodis Sp. z o.o., Poland, antoni.lis@elmodis.com

Abstract

The key issue in monitoring rotary machinery is the necessity to verify efficiency of the developed methods under practical, industrial conditions while maintaining repeatability of the verification procedure. This research is aimed at comparing data obtained from five similar over-hung fans for the purpose of damage detection algorithms development. The fans are installed next to each other and are used in the same industrial process. The similarity of machinery allows for meaningful comparison of the proposed methods and estimation of expected repeatability of results. Condition monitoring procedures described in this paper are based on acceleration signals, which were gathered with piezoelectric accelerometers. All of the sensors were positioned at analogous places on monitored fans. Feature extraction methods described in the paper focus on statistic approach to time waveforms analysis along with signal processing techniques based on Fourier Transform. Features calculated from signals are used as inputs to anomaly detection (AD) algorithms. The data are gathered not only for the intact state of machinery, but also include a case of breakdown and analyses of maintenance induced changes. The authors evaluate various approaches to damage detection with respect to their reliability and practical applications.

Keywords: Anomaly detection, condition monitoring, vibrodiagnostics, bearing damage, broadband features

1. Introduction

Nowadays condition monitoring industry is facing the data revolution associated with concept of industry 4.0. As a result, the idea of periodic vibration measurements acquired with a handhold analyzers is slowly getting out of date. On-line monitoring devices continuously taking measurements and storing them in cloud services are becoming a new standard in maintenance practice. Vast amounts of data which are being collected have to be correctly analyzed to assess the condition of machines under monitoring. The basic analysis to be conducted on gathered data is a comparison of measured values to limits (warnings and alerts) provided by norms. If the monitoring system is capable of capturing time-domain signals, then more sophisticated types of analysis can be performed [1]. Collection of vibration signals as a function of time allows to carry out in-depth time and frequency domain examinations. These methods are important to distinguish between different kind of damages such as imbalance, misalignment, bearing fault etc. [2]. Next, features can be extracted, from the signals,
and used as an input to machine-learning systems for the purpose of machine state classification [3].

This paper presents AD algorithm based on nearest-neighbour method [3] applied for practical, industrial condition monitoring case. Data showed in the paper include both normal and damage condition of monitored fans. Trend and frequency analysis for damaged state is presented. Strengths and limitations of both anomaly detection and trend analysis are discussed. The main contribution of this work lies in comparison of an AD approach with classic well-established methods in practical case for which the damage is present in the structure throughout the entire data acquisition period.

The paper is organized as follows. In section 2 machines and acquisition settings are presented. Section 3 presents, condition monitoring approach based on norms, followed by signal processing part including time and frequency domain analyses. Both approaches are presented on analysis of DE bearing damage on one of the fans. Section 4 is a study of an anomaly detection algorithm applied to the collected data. Finally, section 5 sums up and concludes the paper.

2. Industrial background and data acquisition settings

Data used in this paper are acquired with on-line monitoring system installed on five fans that participate in the same manufacturing process. Each fan consists of an overhung rotor, with 12 blade impeller being driven by a 22 kW, two pole induction motor. All motors are powered directly from supply line with 50 Hz frequency. The process is stationary so the rotational speed is constant and nominal speed equals to 1465 RPM. Fans work continuously 24 hours a day. A fan rotor is mounted directly on the motor shaft which is supported by the rolling-ball-element bearings positioned inside housing on drive end (DE) and non drive-end (NDE). A schematic view of the system is presented in fig. 1

![Figure 1. Schematic drawing of a monitored fan](image)

Vibration signals analyzed in the paper are measured with piezoelectric accelerometers (CTC AC102-1A) with sampling frequency of 31.25 kHz. The measurements are done only in one, radial direction (see fig. 1). Sensors are positioned at analogous places at all five motors which allow for meaningful comparison of the acquired data. Due to design of the machines it was impossible to mount sensors directly at the bearings. Each measurement consist of a 5-second-long acceleration signal collected with an online-
monitoring system. The machines are vibroisolated. Data analyzed in this article were gathered during time span of 4 months in which one of the DE bearings of fan #4 was severely damaged and replaced.

3. Trend, time and frequency domain analysis of overhung fan

3.1 Trend analysis with norm-based thresholds

In this paragraph the fan #4 DE bearing breakdown case will be described. The velocity RMS (Root Mean Square) value as a function of time is presented in fig. 3. Each point in a diagram is calculated as a RMS value from 5-second-long acceleration signal integrated to velocity (in frequency domain). Warning and alarm levels in the diagram are taken from the ISO 10816-3:2009 norm. A schematic overview of a diagnostic path is depicted in fig. 2

![Figure 2. Diagnostic path of trend and norm – based analysis](image)

Fig. 3 includes data acquired for developing damage, moment of failure and state after a new bearing was installed. Since a condition monitoring system was installed with the damage already present in the bearing, there is no data acquired for the bearing prior to damage occurrence.

Shape of velocity RMS curve in fig. 3 rises and falls significantly before the final breakdown. This kind of behavior of the RMS value before the failure of rolling element bearing is known in the literature [6-7].

![Figure 3. Velocity RMS value measured on a MV4 fan in a period of 13 days during development of damage in DE bearing. Warning and alert levels are taken from the norm.](image)
A drastic change in velocity RMS amplitude is a strong premise of increasing damage in the machine. Vibration magnitude and its changes are straightforward indicators that immediate maintenance intervention is required.

The 25% change in magnitude, suggested in the norm to take action, is met as a condition in first day of data collection. After reaching an alarm level it took only one day for a fan to stop because of complete damage of a DE bearing. The bearing was seized up so that spinning of a shaft was disabled.

Features calculated from acceleration signal can be utilized for bearing damage recognition with higher efficiency than using a velocity RMS signal.

In fig. 4 trends of both acceleration RMS and acceleration PP (Peak to Peak) features are shown. Acceleration amplifies high frequencies which are essential in recognizing bearing faults [1]. Thresholds for these are not standardized by the norm. There are, however, severity charts which can be used for interpretation of their absolute magnitude (see [2]). The red and yellow points in the fig. 4 show when a velocity signal is in warning (yellow) and alert (red) range from the ISO 10816-3:2009. Trend lines in the fig. 4 can be applied for establishing thresholds for both of the features for this and other four fans as well.

Acceleration waveforms are presented in fig. 5. The time domain function of acceleration represents the last measurement before a breakdown (left) and one after the bearing replacement (right).
3.2 Frequency domain analysis

In order to investigate the bearing failure in more detail, the analysis in frequency domain was performed. The data presented in fig. 6 come from the same period of time as data presented in sec. 3.1. In fig. 6 (a) waterfall diagram presents frequency spectra for all the signals acquired for fan #4. Signs of damage are visible from the beginning of acquisition period. During first 7 days energy in acceleration spectra is spread mostly in range between 0 and around 5000 Hz. Later, the whole frequency range exhibits rise of amplitudes. As the fault progresses, more and more frequencies are being excited. After bearing replacement high frequency regions of the spectrum decrease in amplitude notably. To analyze this change more precisely (b) diagram is presented. In fig. 6 (b) a result of subtraction between the last spectrum before bearing replacement and first after it is shown.

![Figure 6. Waterfall diagram presenting development of DE bearing damage (a), plot of result of subtraction of the last signal before and first after the bearing replacement (b)](image)

The acceleration level dropped down significantly after the replacement of the DE bearing. Especially frequencies higher than 5 kHz which are most likely resonances of the bearing (see [8]) fell down drastically (between 50 to 80 dB). Low frequencies related to rotational speed got higher amplitudes than before due to the return to the intact condition of the machine.

3.3 Order analysis of velocity spectra

To study the bearing damage more closely, another type of frequency domain analysis was conducted. In fig. 7 there is a waterfall diagram of velocity spectra in which one of the axes is organized in orders of rotational speed. Ball pass frequency of an outer race (BPFO), for DE bearing, is pointed at the diagram with an arrow. BPFO (see [8]) is a dominant frequency in a spectrum with strong second and fourth harmonic.
Figure 7. Order analysis in a form of waterfall diagram based on velocity spectra’s for first 30 orders

The data presented in fig. 7 are result of integration of an acceleration signal in frequency domain. The method used for this calculation is presented in [9]. This procedure can be interpreted as a low pass frequency filtering of an acceleration signal. It is to be mentioned that the amplitude of the first order vibration before the replacement was 5 times smaller than the amplitude for BPFO frequency. After the bearing replacement the rotational speed is dominant in the velocity spectrum. BPFO frequency and its harmonics are dominant in an initial part of a time-order spectrum. After the bearing replacement they are no longer visible, which is consistent with expectations.

4. Anomaly detection

4.1 Theory and background of anomaly detection

A basic principle of operation of the AD used for fault detection requires acquisition of normal data, that is, data that was gathered on a machine without any faults [3-5]. The algorithm then selects a region in feature space that represent typical behavior of the machine. New data are checked against this region - if they do not fit inside the region it is assumed that machine entered a novel state which, if all possible operational conditions are accounted for in a training set, is a good indication of damage. An advantage of AD method lies in a variety of features, which can be used to detect an anomaly. Moreover, novelty detection approaches does not require examples of faults in training data to work as intended. A drawback, on the other hand, lies in necessity to select thresholds based only on examples of intact data which, in some cases, may lead to overcautious or insensitive system.

The scheme of AD-based diagnostic path for a fan is shown in fig. 8. Acceleration data are collected by an on-line monitoring system described in section 2. After specified period of time gathered data are used to define a normal range of fan's operation. Subsequently, the threshold for anomaly detection can be proposed. Features extracted, in this case from acceleration signals are points in multidimensional feature space [3].
There are several techniques that can be used to decide whether a new point is going to be classified as a anomaly [10]. Technique applied in this paper is based on the nearest-neighbor method. Distance between the features is calculated using the Euclidean metric. Thresholds for the purpose of this article are set to a value that is a result of summation of mean and three times the standard deviation of all inter-sample distances in training database. This comes from an assumption that a distribution of data in a training database is close to normal. Used features are listed in the Table 2.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Acceleration RMS</td>
</tr>
<tr>
<td>Time</td>
<td>Acceleration kurtosis</td>
</tr>
<tr>
<td>Time</td>
<td>Acceleration peak to peak</td>
</tr>
<tr>
<td>Frequency</td>
<td>Velocity RMS</td>
</tr>
</tbody>
</table>

The acceleration RMS represents the energy of a raw signal that amplifies high frequencies. According to [11] Kurtosis is used to reflect the peakiness of a signal, which allow for detection of impulses and measurement of the divergence from a fundamental Gaussian distribution. Peak-to-peak is a feature that amplifies variations of amplitude. Velocity RMS is a form of low pass filtering of a acceleration signal, which amplifies frequencies characteristic for basic rotor dynamics faults like imbalance, misalignment, bearing frequencies etc.

4.1 Anomaly detection results

Results for normalized distances in four dimensional space for five fans are presented in the fig. 9. There are two diagrams for every fan. Each diagram on the left-hand side presents training data and ones on the right-hand side presents testing data. The horizontal red dashed line represents in both training and testing data an anomaly threshold. Collection of data was done within 4 months with an on-line condition monitoring system. Data for each fan were gathered with a specified for each machine time interval. Data collection settings are described in section 2.
Figure 9. Training and testing data for each fan.
Data for fan #4 include damage example described in section 3. Even though the training dataset for fan #4 was collected while the damage was already present in the bearing, the percentage of anomalies above the threshold in testing dataset reached 90%. Thus, the damage condition of the fan #4 was recognized correctly despite lack of training on an intact machine.

Condition of fans #1, #2, and #3 during a data collection time didn’t show any sign of existing fault. The percentage of anomalies above the threshold among those three fans in testing database is the highest for fan #1 and it is equal to 4.84%.

Normalized distance in training dataset for fan #5 increased significantly. The percentage of anomalies above the threshold reached 44%. That is the highest result among not damaged fans. There was no damage reported by the maintenance crew during data collection, but the machine state has changed indicating presence of a yet undiscovered damage.

Bar chart in fig. 10 presents percentage of detected anomalies in testing dataset for all of five monitored fans.

![Figure 10. Percent of anomalies in testing datasets for all the monitored fans](image)

### 5. Conclusions

In this paper two approaches to vibration-based condition monitoring were presented: a classic approach including analyses of time and frequency representations of vibration signals and novelty detection based on a vector of broadband features. The experimental evaluation encompassed five over-hung fans. Both trend and anomaly detection analysis enabled correct detection of a severe damage case. The most important finding of the article lies in fact that the anomaly detection method, despite training on an already-damaged case, was able to detect damage with high confidence. This observation contrasts with a typical novelty detection approach in which it is assumed that condition of the machine under monitoring is intact during training data acquisition. The detection was possible due to the fact that developing damage not only caused an increase in feature levels (visible from the beginning of a monitoring period) but also rendered them to be fluctuating. As a result, features used as inputs to the system were entering novel regions of feature space with each consecutive stage of damage development.
Acknowledgements

The work presented in this article was supported by the National Centre for Research and Development in Poland under the project no. POIR.01.02.00-00/16. This research was made possible due to data provided by R&D department of Elmodis company.

References