International journal of COMADEM
Vol 20 No 4 (October 2017) 35-39

COMADEM

International Journal of Condition Monitoring and Diagnostic Engineering Management

Fault detection in bearings using autocorrelation

Rushit N. Shah[a], Michael H. Azarian[a]*, and Michael G. Pecht[a]

[a] *Center for Advanced Life Cycle Engineering, University of Maryland, College Park, MD 20742, USA*
* *Corresponding author. Tel.: +1 301-405-7555; email: mazarian@calce.umd.edu*

ABSTRACT

Autocorrelation is a special case of cross-correlation wherein a signal is correlated with a time-lagged version of itself – the resulting signal comprises only the periodic information from the original signal whilst reducing noise. This property of autocorrelation can be particularly useful in analysing bearing faults since vibration data from a bearing, with local faults/defects, consists of cyclostationary acceleration signals usually contaminated with noise from sensors and other environmental factors. This study introduces a method which provides early failure warning in rolling element bearings by applying an autocorrelation operation to vibration data. The Sequential Probability Ratio Test (SPRT) is used to detect anomalies indicative of incipient failure. The results from the autocorrelation analysis are compared with results from a simple moving-RMS analysis of the acceleration data. The developed method is shown to provide an earlier warning of failure than the RMS-based method. This method can detect early stages of degradation in bearings – which in turn allows earlier scheduling of maintenance and the avoidance of system failures.

Keywords: Autocorrelation; bearings; diagnostics; health management; prognostics; SPRT; anomaly detection

1. Introduction

Prognostics is a process that involves assessing the deviation of the present health state of a system, from its expected health state, in order to predict its remaining useful life (RUL) [1][2][3]. Prognostics methods may belong to one of two categories – physics-of-failure (PoF) methods or data-driven methods. PoF-based prognostics models utilize information related to a system's life cycle loading conditions, geometry, material properties and failure mechanisms in order to estimate the remaining useful life of the system [4]. While PoF-based prognostic methods can provide the extent of damage, assuming that the correct failure mechanism and loading conditions are known, these methods are unable to provide RUL when the failure mechanisms or the models describing those mechanisms are unavailable. Data-driven prognostic methods, on the other hand, make use of present and historical data collected from a system to statistically estimate the RUL of the system. These methods are not dependent on the foreknowledge of the geometry, material properties and failure mechanisms governing failure in a system, unlike PoF-based methods. Data-driven methods can describe complex patterns and trends in data, without requiring a specific failure model. Some examples of data-driven methods include neural networks (NN), support vector machines (SVM), linear discriminant analysis (LDA), random forest (RF), etc. This paper focuses on data-driven methods in order to estimate RUL of rolling-element bearings.

2. Bearing Fault Diagnostics

Bearings are load-bearing elements in machines, the smooth functioning of which proves to be critical in the failure-free operation of the system of which they are a part. Rolling element bearings, especially, are some of the most widely used components in mechanical systems and the failures of such systems are most frequently attributed to the failure of these bearings [5]. Owing to their criticality, early detection of faults in bearings is desired. Consequently, fault detection in bearings has been studied extensively. Vibration-based fault detection is the most common and reliable approach to this problem [6]. The emphasis on the vibration-based diagnostic techniques stems from the fact that a force impulse is generated every time a rolling element strikes a fault on a supporting bearing surface (inner race or outer race of the bearing). The series of impulses generated by such interactions between the fault and the rolling elements generally generates an amplitude modulated cyclostationary signal which can then be analyzed for fault signatures. The frequency spectrum of such a signal consists harmonics of resonance frequencies excited by the impulses and spaced at the bearing fault frequency. However, raw vibration signals are usually riddled with noise and disturbances produced by the system itself, as well as other external sources. In the domain of vibration-based fault detection, envelope analysis [7] is an established method of obtaining the bearing fault frequencies from the vibration signal. The general procedure for envelope analysis involves initial pre-processing using a high-pass filter to remove low-frequency noise and then inspecting the envelope spectrum of the resulting signal for bearing fault frequencies. Several improvements have been suggested over the original envelope analysis method. For instance, fast Fourier transform (FFT) is the most widely used technique to analyze spectra. Employing Hilbert transform for demodulation also allows effective extraction of the relevant section of the signal spectrum [8]. However, most methods in the domain of vibration-based fault detection focus on the raw vibration signal in the frequency domain. The frequency domain approach is suitable since extraction of fault frequencies provides detailed information about the nature and the location of the fault because fault frequencies are intrinsically related to the geometry of the bearing.

However, our study focuses on the early detection of a bearing fault (rather than gleaning information about location, nature or the extent of the fault) by examining the vibration signal time domain rather than the frequency domain. Analyses in the time domain involve extraction of various features from the temporal vibration signal such as mean, standard deviation, skewness, kurtosis, etc. Changes in these features over time are then used as indicators of health of the bearing. Advanced approaches in the time domain analysis involve fitting temporal signals to parametric time series models, followed by extraction of features from the parametric model. Popular methods of this type are the autoregressive (AR) model and autoregressive moving average (ARMA) model. Poyhonen et al. [9] fit vibration signals to an AR model and used the model coefficients as features for fault detection. The technique proposed by Garga et al. [10] uses AR modeling followed by dimensionality reduction. However, the AR and ARMA models are complex to model. Baydar et al. [11] studied the use of Principal Component Analysis (PCA) for analyzing time domain vibration signals from gears. For analyzing non-stationary signals, time-frequency domain analysis may be applied to such vibration signals. Time-frequency domain methods represent vibration signals in both time and frequency domain. The resulting distributions represent the energy of the signal and can help understand how the energy of the signal is distributed across different frequency bands over time. Popular time-frequency domain techniques include short-time Fourier transform [12] [13] and Wigner-Ville distribution [14] [15] [16] [17]. In this paper, a fast, new and computationally inexpensive method based on the autocorrelation function and SPRT is introduced. The rest of this paper is organized as follows. The next section provides theoretical background of autocorrelation and SPRT. Then, the developed bearing fault detection methodology is delineated - a new parameter called normalized autocorrelation function range (R_{range}) is then introduced and the methodology explained. Experimental results are provided to demonstrate the effectiveness of the developed method in the context of bearing fault detection. Finally, conclusions about the developed method are discussed. The next section discusses the autocorrelation function, normalized autocorrelation function and their properties. A theoretical background of SPRT is then provided. Finally, the method is tested using an open-source dataset and the results are compared with an equivalent fault-detection method based on the RMS of the vibration of the signal.

3. Autocorrelation function

Given a signal $f(t)$, the continuous autocorrelation function $R_{ff}(\tau)$ is defined as the continuous cross-correlation integral of $f(t)$ with itself, at lag τ :

$$\boldsymbol{R_{ff}(\tau) = \int_{-\infty}^{\infty} f(t)f(t-\tau)dt}$$

Alternatively, the autocorrelation function given above can be modified by including a normalizing factor to obtain the normalized autocorrelation function $R_{ff}(\tau)$:

$$\boldsymbol{R_{ff}(\tau) = \frac{\int_{-\infty}^{\infty} f(t)f(t-\tau)dt}{\sigma^2}}$$

where σ^2 is the variance of $f(t)$. The benefit of normalization is twofold. Firstly, it limits the range of autocorrelation function to $R_{ff}\epsilon[-1,1]$. Secondly, it provides a scale-free measure of similarity between different moments of the signal $f(t)$.

4. A New Health Indicator for Bearing Fault Detection

When a bearing is operating in a defect-free state, the vibration signal collected from the bearing is composed primarily of noise from the system. However, the vibration signal from a faulty bearing is composed of different periodic components. This is due to the periodic nature of the impulses generated when the rolling-element impacts a defect on the outer or the inner race of the bearing. This effect has a direct bearing on the autocorrelation function. As these periodic components appear in the vibration signal, the magnitude of correlation between different moments of the vibration signal increases. In other words, the measure of similarity between different regions of the vibration signal increases. To capture this change in the magnitude of the autocorrelation function, a new parameter called *normalized autocorrelation function range* (R_{range}) is defined for a given signal $f(t)$ as:

$$\boldsymbol{R_{range} = \max\left(R_{ff}(\tau)\right) - \min\left(R_{ff}(\tau)\right)}$$

5. Sequential Probability Ratio Test (SPRT)

The sequential probability ratio test (SPRT) is used to track any change in the magnitude of R_{range} which is then used as an indicator of the health of the bearing. SPRT is a statistical hypothesis test which determines whether the test data falls into the probability density distribution of the base line training data [18] [19]. The SPRT detects changes in the test data by comparing a null and alternative hypotheses. The null hypothesis is the case in which the test data adheres to a Gaussian distribution with a mean of 0 and a variance of σ^2 extracted from the training data, which represents the normal test data. There exist four alternative hypotheses - the positive/negative mean test, the normal variance test, and the inverse variance test. For the positive/negative mean test, the alternative hypotheses are that the test data follows a Gaussian distribution with a mean of +M or -M and a variance of σ^2, where M is a pre-determined system disturbance level. The SPRT index is the logarithm of the ratio of the probability that the alternative hypothesis is true to the probability that the null hypothesis is true. Given the null and alternative hypotheses, the SPRT index is given as:

$$\boldsymbol{SPRT = \frac{M}{\sigma^2}\sum_{k=1}^{n}\left(x_k - \frac{M}{2}\right)}$$

where x_k are successive observations of the test data. The SPRT index is continuously calculated and compared to the lower and upper threshold limits which are set by the user and determined by the desired level of sensitivity to detection. When the SPRT index is less than the lower threshold, it can be concluded that the test data is normal. When the SPRT index is greater than the upper boundary indicates that the test data is abnormal. The SPRT test can thus be used detect any deviation in the normalized autocorrelation function range from the base line data.

6. Methodology

Data from the bearing with outer race faults was analyzed for warning of an impending fault. The raw vibration signal in the time domain was used for the purpose of this analysis. The developed fault detection method is outlined in Fig. 2. The vibration signal is first divided into non-overlapping windows. The length of each window was fixed at 50000 samples. Thus, each window corresponds to a signal duration of 2.5 seconds. Selecting a window size of 50000 samples can be justified by the fact that the ball pass frequency for an outer race fault is ~236 Hz for this bearing. This implies that the interval between two consecutive impulses is ~1/236 = 0.0042 s. Since the sampling frequency is 20000 Hz, the interval between two consecutive impulses is ~86 samples. Hence, a window size of 50000 samples is sufficient to capture any fault related periodic components in the vibration signal. The normalized autocorrelation function was then computed for the first window. This autocorrelation function corresponding to the first window was then used to compute the normalized autocorrelation function range R_{ranae}. This value was recorded and saved. This constituted one iteration. This procedure was then repeated sequentially for every window, the R_{ranae} values for their corresponding windows were saved. After each iteration, the SPRT test was used to detect deviation in the R_{ranae} values. The first SPRT detection was used as the early warning of impending failure.

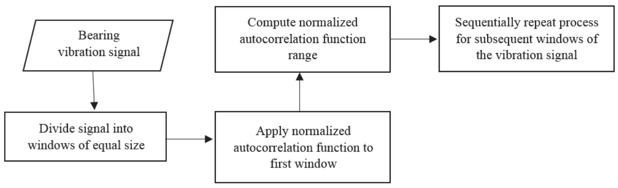

Figure 1. Flow chart of the developed fault detection method

7. Test setup and data collection

Bearing run-to-failure tests under constant load conditions were performed on a specially designed test rig, as shown in Fig.1, with bearing data from a data repository [20]. The bearing test rig has four test Rexnord ZA-2115 double row bearings mounted on a shaft that is driven by an alternating current. Table 1 lists the geometrical specifications of the bearings. The rotation speed was kept constant at 2000 rev/min., with a load of 6000 lb. being applied to the shaft and bearing radially.

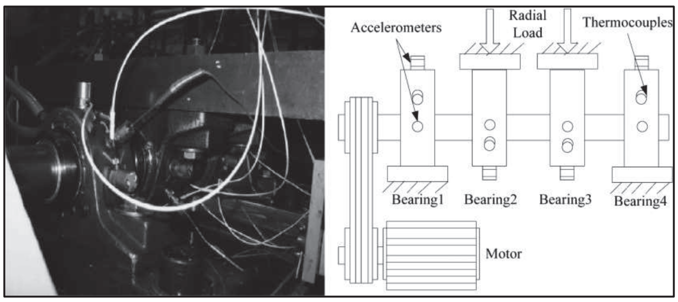

Figure 2 Test setup

Table 1. Geometrical specifications of the rolling element bearing

Number of rolling elements n	16
Diameter of rolling element d	8.4 mm
Pitch diameter of bearing D	71.5 mm
Contact angle γ	15.17°

A PCB 353B33 high-sensitivity quartz Integrated Circuits Piezoelectric (ICP) accelerometer installed on each bearing housing collected data at 20 kHz.

8. Results

The developed method was applied to the vibration signal corresponding to an outer race fault. Fig. 3 shows the normalized autocorrelation function range plotted as a function of the window number. Up to window number 200, the trend in R_{ranae} is uniform. The deviation in this trend is evident after window number 200. This deviation corresponds to the onset of the fault in the bearing. The process of damage propagation in a bearing involves onset of fault and damage aggravation followed by self-healing. Self-healing is the mechanism by which the spalls or cracks during damage propagation are smoothed out by continuous rolling contact which returns the bearing to a healthy or *near*-healthy state [21]. This mechanism continues until the bearing fails. In Fig. 3, the region between window numbers 300-350 corresponds to this self-healing mechanism. In order to automate the process of detection, SPRT was applied to this data. R_{ranae} values for the first 200 windows was used as the base line data since these windows correspond to the healthy section of the vibration signal from the bearing and the SPRT test is used to detect deviation from this base line data. It is expected that as the damage in the bearing accumulates, SPRT will indicate an alarm value of ‘1’, indicative of a fault and return to ‘0’ as the bearing self-heals. Hence, only the first SPRT detection is used as a fault indicator. For the developed method, the first detection was at window number 251. In order to compare this result with an equivalent method, the root mean squared (RMS) value of each window was used. Fig. 4 shows the RMS values for each window. For SPRT detection, the RMS values of first 200 windows were used as base line data. In this case, the first SPRT detection is observed at window number 308 i.e. once the bearing is already in the self-healing state [17].

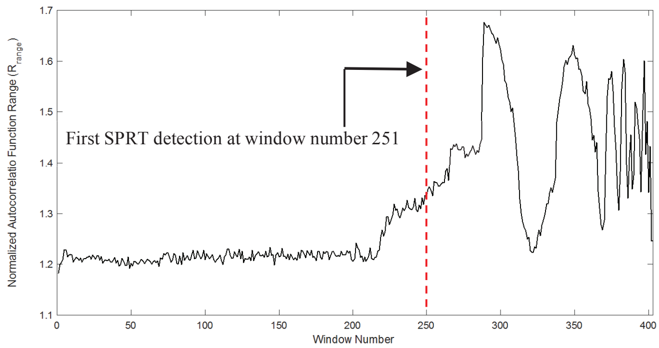

Figure 3. Normalized autocorrelation function range and first SPRT detection using the developed method.

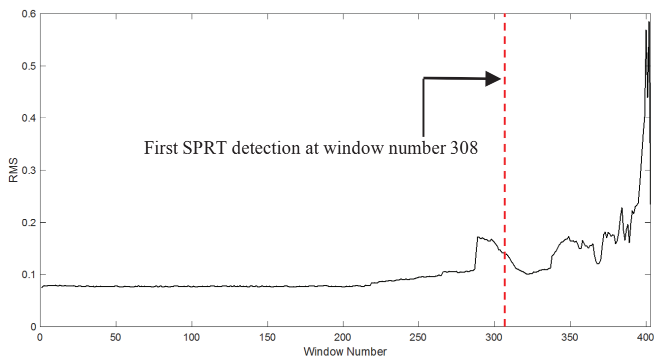

Figure 4. RMS-based fault detection method and first SPRT detection using this method.

The results show that the developed method can detected the fault (at window number 251) much early compared to the RMS-based method which detected the identical fault only after window number 308. The difference between the time of first detection of the fault, by the two methods is ~3 minutes.

9. Effect of window size

For the first study, a window size of 50000 samples was chosen. As discussed in the section on methodology, this window size was sufficient to capture any periodic impulse which may have occurred due to a fault in the bearing. To study the effect that selecting different window sizes has on the time of detection, the developed method was tested with seven different window sizes – 20000 to 80000 (at intervals of 10000 samples). For every window size, the percentage remaining useful life (% RUL) – time from first detection to bearing failure - was computed, for both, the developed method and the RMS-based method. The results are presented in Figure 5. As the window size was increased, the % RUL at the time of first detection decreased. Hence, a smaller window size is preferred. However, it must be ensured that the window size is large enough to capture the periodic impulses; i.e. the size of the window must be larger than the time between two impulses. Also, with increasing window size the decrease in % RUL (at time of first detection) was more significant in the case of the RMS-based method than the developed method. This demonstrates the robustness of the developed method. In every case, the developed method based on autocorrelation had smaller detection times for the fault than the RMS-based method. In the case where the window size was 20000 samples, the developed method was able to detect the fault with ~40% of the life of the bearing remaining. In the case where window size was 80000 samples, the developed method detected the fault with ~35% of the life of the bearing still remaining. In the same case, the RMS-based method provided a failure warning with only ~15% of the bearing life remaining.

10. Conclusion

Bearings play a critical role in many mechanical systems, and early detection of bearing faults is desired to avoid system failure and to schedule early maintenance. This paper presents such a method for rolling element bearings. The developed method utilizes the normalized autocorrelation function to compute a new scale-free parameter called normalized autocorrelation function range (R_{range}) which is used to characterize the state of the system. The sequential probability ratio test (SPRT) is then used to detect a fault based on the deviation of R_{range} from base line data.

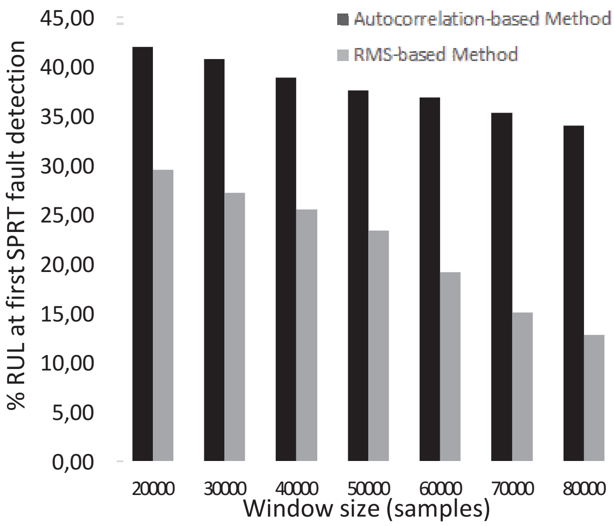

Figure 5. Effect of choosing different window sizes on the time of first SPRT fault detection

The developed method has been compared with an equivalent RMS-based method using the same vibration data. The method, although rudimentary, provides a promising way of performing prognostics for rolling-element bearings. In the set of experiments performed, the developed method consistently detected the fault in the bearing with ~35-40% bearing life remaining. Since this method involves computation of only one parameter for determining the health of the system, it is computationally inexpensive (it takes less than 5 seconds to execute for the bearing vibration data used in this study), and it can be easily integrated into larger system-level prognostics and health-management (PHM) routines.

References

1. Shunfeng and M. Pecht, "A Fusion Prognostics Method for Remaining Useful Life Prediction of Electronic Products," in IEEE International Conference on Automation Science and Engineering, 2009.
2. M. Pecht, Prognostics and Health Management of Electronics, New York, NY: Wiley-Interscience, 2008.
3. N. Vichare, "Prognostics and Health Management of Electronics," IEEE Transactions on Components and Packaging Technologies, vol. 29, no. 1, pp. 222-229, 2006.
4. S. Kumar, M. Torres and M. Pecht, "A Hybrid Prognostics Methodology for Electronics Systems," in Special Session on Computational Intelligence on Anomaly Detection, Hong Kong, 2008.
5. R. B. Randall and J. Antoni, "Rolling Element Bearing Diagnostics - A Tutorial," Mechanical Systems and Signal Processing, vol. 25, no. 2, pp. 485-520, 2011.
6. R. Golfashan and K. Y. Sanliturk, "Rolling Bearing Fault Diagnosis Using Recursive Autocorrelation and Autoregressive Analyses," in 23rd International Congress on Sound and Vibration, Athens, Greece, 2016.
7. M. S. Darlow, R. H. Badgley and G. W. Hogg, "Application of High Frequency Resonance Techniques for Bearing Diagnostics in Helicopter Gearboxes," Technical Report ,US Army Air Mobility Research and Development Laboratory, 1974.
8. F. Bonnardot, R. B. Randall, J. Antoni and F. Guillet, "Enhanced Unsupervised Noise Cancellation Using Angular Resampling for Planetary Bearing Fault Diagnosis.," International Journal of Acoustivs and Vibration, vol. 9, no. 2, pp. 51-60, 2004.
9. S. Poyhonen, P. Jover and H. Hyotyniemi, "Signal Processing of Vibrations for Condition Monitoring of an Induction Motor," in 2004, New York, ICSSP: First International Symposium on Control, Communications and Signal Processing.
10. A. K. Garga, B. T. Elverson and D. C. Lang, "AR Modeling with Dimension Reduction for Machinery Fault Classification," Critical Link: Diagnosis to Prognosis, Haymarket, pp. 299-308, 1997.
11. N. Baydar, Q. Chen, A. Ball and U. Kruger, "Detection of Incipient Tooth Defect in Helical Gears Using Multivariate Statistics," Mechanical Systems and Signal Processing, vol. 15, pp. 303-321, 2001.
12. W. J. Wang and P. D. McFadden, "Early Detection of Gear Failure by Vibration Analysis - I: Calculation of the Time-Frequency Distribution," Mechanical Systems and Signal Processing, pp. 193-203, 1993.
13. F. A. Andrade, I. Esat and M. N. M. Badi, "Gearbox Fault Detection Using Statistical Methods, Time-frequency Methods (STFT and Wigner-Ville Distribution) and Harmonic Wavelet - A Comparative Study," in Proceedings of the COMADEM, 99, 1999.
14. Q. Meng and L. Qu, "Rotating Machinery Fault Diagnosis Using Wigner Distribution," Mechanical Systems and Signal Processing, vol. 5, pp. 155-166, 1991.
15. M. C. Pan, H. Van Brussel, P. Sas and B. Verbeure, "Fault Diagnosis of Joint Backlash," Journal of Vibration and Acoustics, vol. 120, no. 1, pp. 13-24, 1998.
16. I. S. Koo and W. W. Kim, "The Development of Reactor Coolant Pump Vibration Monitoring and a Diagnostic System in the Nuclear Power Plant," ISA Transactions, vol. 39, pp. 309-316, 2000.

17. N. Baydar and A. Ball, "A Comparative Study of Acoustic and Vibration Signals in Detection of Gear Failures Using Wigner-Ville Distribution," Mechanical Systems and Signal Processing, vol. 15, pp. 1091-1107, 2001.
18. W. Lu and K. Gross, "Early Detection of Signal and Process Anomalies in Enterprise Computing Systems," in IEEE Int'l Conf. on Machine Learning and Applications, Las Vegas, NV, 2002.
19. K. Humenik and K. Gross, "Sequential Probability Ratio Tests for Reactor Signal Validation and Sensor Surveillance Applications," Nuclear Science and Engineering, vol. 105, 1990.
20. J. Lee, H. Qiu and G. Yu, Rexnord Technical Services: Bearing Data Set, IMS, Univ. Cincinnati: NASA Ames Prognostics Data Repository, NASA Ames, 2007.
21. T. Williams, X. Ribadeneira, S. Billington and T. Kurfess, "Rolling Element Bearing Diagnostics in Run-to-failure Lifetime Testing," Mechanical Systems and Signal Processing, vol. 15, pp. 979-993, 2001.