World Academy of Science, Engineering and Technology
International Journal of Mechanical and Mechatronics Engineering
Vol:2, No:9, 2008

Bearing Fault Feature Extraction by Recurrence Quantification Analysis

V. G. Rajesh, and M. V. Rajesh

***Abstract*—**In rotating machinery one of the critical components that is prone to premature failure is the rolling bearing. Consequently, early warning of an imminent bearing failure is much critical to the safety and reliability of any high speed rotating machines. This study is concerned with the application of Recurrence Quantification Analysis (RQA) in fault detection of rolling element bearings in rotating machinery. Based on the results from this study it is reported that the RQA variable, percent determinism, is sensitive to the type of fault investigated and therefore can provide useful information on bearing damage in rolling element bearings.

***Keywords*—**Bearing fault detection, machine vibrations, nonlinear time series analysis, recurrence quantification analysis.

I. Introduction

BEARINGS are among the most important machine components in the vast majority of rotating machines and exigent demands are made upon their carrying capacity and reliability. Generally, a rolling bearing cannot rotate for ever. It often works well in non-ideal conditions, but sometimes minor problems cause bearings to fail quickly and mysteriously without any notable warning. The bearing failures are mainly resulted from excessive wear or damage in rolling ball elements as well as in the inner/outer races of the bearing. Presently real-time condition monitoring systems for bearing systems often fail to provide sufficient time between warnings and on the other hand, inaccurate interpretation of operational conditions may result in false alarms and associated unnecessary costs and downtime [1].

Traditionally, the detection of faults has become possible by comparing the sensitive features of signals from sensors in the machinery while running in normal and faulty conditions. This method of the detection of faults has showed considerable success and several techniques have been developed. The use of vibration signals is quite common in the field of condition monitoring of rotating machinery. Analyzing the vibration signals directly in the time domain is one among the simplest and cheapest diagnosis approaches [2]. However, as the damage increase, the vibration signal becomes more random and the temporary statistical values reduce to more like that of normal bearing levels. This is the most important shortcoming of this approach [3]. In the frequency domain approach the major frequency components of vibration signals and their amplitudes are used for trending purposes [4]. One of the drawbacks of frequency-domain approaches is that they require the bearing defect frequencies to be known or pre-estimated. The time-frequency domain approach use both time and frequency information allowing for the transient features, such as impacts. However, this approach fails to analyze the continuously smooth signal.

In this paper, the powerful method of RQA is used to study and characterize the experimental sensor signals generated during the normal and faulty states of the bearing under study. The study has been carried out with an objective to discriminate these signals as due to a good or faulty bearing on the basis of the calculated sensitive RQA variables.

II. Experimental Setup And Data Acquisition Systems

Two arrays of experiments, one with a good bearing and the other with a defective bearing, containing 10 trials in each array are conducted. A total of 20 numbers of experiments are conducted and the corresponding acquisition and recordings of sensor signals representing time history of bearing vibration signature during operations are done. Here a faulty bearing refers to a bearing with damage induced in one of the eight balls. The level of damage induced in the bearing is relatively large in order to provide a fault comparison.

A. Test Rig

The study of fault detection for the bearing is carried out in a specially designed test rig. It consists of an induction motor of 1410 rpm and a shaft of uniform cross section. The shaft passes through two deep grooved ball bearings (6305 Type) which reside within two Plummer blocks separated by a distance. One end of this shaft is connected to a single spherical load of 1.1 kg and to the other end is rigidly attached a pulley. The shaft is driven by the motor at the designated speed using a v-belt. The entire arrangement is rigidly bolted on to the test bench keeping the axes of the shaft as well as motor horizontal. The first Plummer block provides support to the shaft near the pulley and affords stability to the whole test bench whereas the second Plummer block provides support near to the load. During experiments only the second Plummer block is opened to change the bearings and analyze for defects while the first is maintained intact.

V. G. Rajesh is with the Mechanical Engineering Department, Model Engineering College, Cochin, Kerala, 682021, India (phone: 0944-722-8784; fax: +91-484-2577379; e-mail: rvgmec@gmail.com).

M. V. Rajesh is with the Electronics Engineering Department, Model Engineering College, Cochin, Kerala, 682021, India (e-mail: m.v.rajesh@hotmail.com).

B. Data acquisition

An ADXL-150 accelerometer sensor is mounted horizontally on top of the second Plummer block to pick up the vibrations in the bearing during operations. The resulting output of the accelerometer is amplified and passed through an anti-aliasing filter having a cut-off frequency 1 kHz before digitization. The analog voltage signal from the filter is sent to DAC NI PCI 6221 through NI SHC68-68-EPM and SCB 68 for converting it to the digital domain. The sampling rate for this signal is fixed at 10 kHz. The digitized data is recorded in the PC hard drive using NI Lab VIEW.

Vibration data is recorded continuously for 30 sec duration during every trial of the experiments and from each of which 25000 data points representing a 2.5 sec duration vibration are randomly selected and analyzed. Fig. 1 and Fig. 2 show the time history of the sampled vibration sensor signals from the system using a good bearing and a faulty bearing.

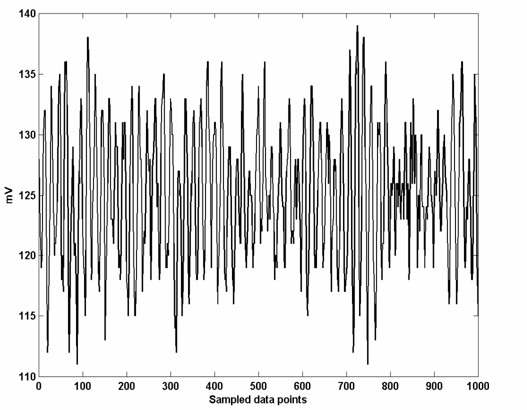

Fig. 1 Time series of recorded vibration signals-Good bearing

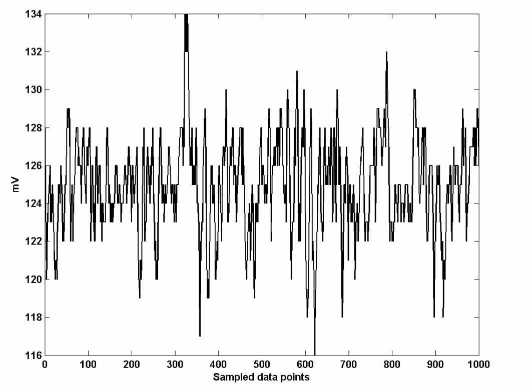

Fig. 2 Time series of recorded vibration signals-Faulty bearing

III. Non Linear Analysis and Methods

Most often, in signal analysis, the amplitude distribution of the signal is analyzed and various statistical moments are used as characteristics. The nonlinear time series analysis (NTSA) approach is basically different from the statistical one, in the respect that it can overcome inherent limits of the traditional linear and statistical tools. Despite its wide range of applications, NTSA suffers from the problems of non-stationarity of the measured time series data, which may lead to pitfalls that may invalidate the analysis. This can be overcome to a very great extent by the recurrence plots (RPs) and the RQA. In 1987, Eckmann et al. [5] introduced the concept of RPs that can visualize the recurrence behavior of the phase space trajectory of dynamical systems. Subsequently, the recurrence quantification analysis (RQA) was developed by Zbilut and Webber Jr. [6], [7] and extended with new measures of complexity by Marwan et al. [8]. The basic idea behind RQA is the identification of recurrence of local data points in a reconstructed phase-space. This method of analysis is found to be independent of various limiting constraints like the data size, it's stationarity and assumptions regarding the statistical distribution of data. Thus, it is ideally suited for analyzing experimental signals which are characterized by non-stationarity and noise.

In the following sections our approach is described; based on phase space reconstruction, the recurrence plot and the recurrence quantification analysis.

A. Phase Space Reconstruction

Takens [9] proved a theorem that is the firm basis of the methodology of delays. Since one variable only is measured (the usual case in an experiment) the delay coordinate approach is used in the present analysis. Given a time series $x(j)=x(1), x(2), x(3),.........x(N)$ we define points $X(i)$ in an m- dimensional state space as

$$X(i)=[x(i),\ x(i+\tau),\ x(i+2\tau),.......x(i+(m-1)\tau)] \qquad (1)$$

for $i=1,2,3,....,N-(m-1)\tau$ where i are time indices, τ, a time lag and m is called the embedding dimension. Time evolution of $X(i)$ is called a trajectory of the system, and the space, which this trajectory evolves in, is called the phase space.

B. Selecting the Minimum Embedding Dimension

The embedding dimension is the minimum dimension at which the reconstructed attractor can be considered completely unfolded. This parameter is usually estimated by the method of False Nearest Neighbors (FNN) given by Abarbanel [10].

By checking the neighborhood of points embedded in projection manifolds of increasing dimension, the algorithm eliminates 'false neighbors'[11]. A natural criterion for catching embedding errors is that the increase in distance between two neighbored points is large when going from dimension m to $m+1$. This criterion is stated by designating as a false nearest neighbor any neighbor for which the following is valid.

$$\left[\frac{R_{m+1}^2(i,i_r)-R_m^2(i,i_r)}{R_m^2(i,i_r)}\right]^{1/2}=\frac{|X(i+m\tau)-X(i_r+m\tau)|}{R_m(i,i_r)}>R_{tol} \qquad (2)$$

Here i and i_r are the times corresponding to the neighbor and the reference point, respectively. R_m and R_{m+1} denotes the distance in phase space with embedding dimension m and $m+1$ respectively, and R_{tol} is the tolerance threshold. For the present analysis the embedding dimension corresponding to the lowest value of FNN is selected.

C. Selecting the Time Lag

To choose the time lag, τ, we use the non linear correlation function of average mutual information (AMI). Fraser et. al [12] establishes that delay corresponds to the first local minimum of the average mutual information function $I(\tau)$ which is defined as follows.

$$I(\tau) = \sum P(X(i), X(i+\tau)) \log_2 \left[\frac{P(X(i), X(i+\tau))}{P(X(i))P(X(i+\tau))} \right] \quad (3)$$

where $P(X(i))$, is the probability of the measure $X(i)$, $P(X(i+\tau))$ is the probability of the measure $X(i+\tau)$ and $P(X(i), X(i+\tau))$ is the joint probability of the measure of $X(i)$ and $X(i+\tau)$. Plotting $I(\tau)$ versus τ makes it possible to identify the best value for the time delay, this is related to the first local minimum.

The values for time lag, τ, and embedding dimension m for the good bearing and faulty bearing have been calculated following the AMI and FNN methods and are shown in Table I. The values of time lag as well as the embedding dimension of the sensor signals for the two types of bearings differ. Since the work is aimed at monitoring of bearing failure that takes place during its operation, phase space reconstruction using two different sets of values is avoided here. Instead, they are chosen from the representative values of the faulty bearing which is found to take higher embedding dimension.

TABLE I
PHASE SPACE RECONSTRUCTION PARAMETERS

Bearing Type	Time Lag	Embedding Dimension
Good bearing	4	9
Faulty bearing	5	11

D. Recurrence Plots based Analysis

A recurrence plot (RP) is a way to visually investigate the multi dimensional phase space trajectory through a two-dimensional representation [8]. Recurrence of states of the system, in the meaning that states are arbitrarily close after some time, is a well-known property of deterministic dynamical systems and is typical for nonlinear or chaotic systems. An RP is derived from the distance plot, which is a symmetric NxN matrix where a point (i, j) represents some distance between coordinates $X(i)$ and $X(j)$ on the phase space trajectory. Thresholding the distance plot at a certain cut-off value transforms it into an RP which shows all the recurrent points as black spots.

$$RP(i, j) = \Theta(\varepsilon - \|X(i) - X(j)\|) \quad (4)$$

where $i, j = 1,.,.,.,N$, ε is a cut-off distance, $\|\bullet\|$ is some norm and $\Theta(\bullet)$ is the Heaviside function [13].

In the present analysis recurrence plots are constructed applying L_2 norm in distance calculations. The threshold ε is chosen by analyzing the measure of recurrence point density [14] as percentage of maximum distance (Table II). Again, as followed and due to reasons assumed in phase reconstruction, we use the threshold ε values obtained for the faulty bearing (27 for vibration) as representative values for RQA estimation.

TABLE II
CALCULATED VALUES FOR THE THRESHOLD

Bearing Type	Threshold ε
Good bearing	22
Faulty bearing	27

Since the RPs itself does not contain any visually appreciable quantitative information we utilize the RQA approach for the purpose in the present study.

E. Recurrence Quantification Analysis

The RQA is a tool based on the statistical description of the parallel lines distribution among the RP [9]. Measures of complexity are defined using the recurrence point density and diagonal line structures in the recurrence plot. These measures provide a qualitative description of the dynamics underlying the time series that is studied. In the original definition Eckman et al [5] used a fixed number of neighbors for determining recurrences. In the present analysis we use a fixed value for the threshold ε due to which the RP is symmetric across the central diagonal, called the line of identity (LOI). Attention is focused on the diagonal and vertical structures in the RP since from those stem the recurrence variables or quantifications. As the recurrence plot is symmetrical across the central diagonal, all quantitative feature extractions take place within the upper triangle in the RP [14], excluding the long diagonal (which provides no unique information) and lower triangle (which provides only redundant information).

We can derive eight statistical values from a RP using RQA. The first value is **Percent Recurrence**, quantifies the percentage of recurrent points falling within the specified radius. The second variable is **Percent Determinism** and measures the percentage of recurrent points that are contained in lines parallel to the main diagonal of the RP, which are known as deterministic lines. A deterministic line is defined if it contains a predefined minimum number of recurrence points. It represents a measure of predictability of the system. The third recurrence variable is **Linemax**, which is simply the length of the longest diagonal line segment in the plot, excluding the main diagonal line of identity .The fourth variable value is called **Entropy** and it refers to the Shannon entropy of the distribution probability of the diagonal lines length. Entropy is a measure of signal complexity and is calibrated in units of bits/bin and is calculated by binning the deterministic lines according to their length. The fifth statistical value is the **Trend** which is used to detect non-stationarity in the data. The trend essentially measures how quickly the RP pales away from the main diagonal and can be utilized as a measure of stationarity. For the detection of chaos-chaos transitions, Marwan et al. [10] introduced other two additional RQA variables, the **Percent Laminarity** and **Trapping Time**, in which attention is focused on vertical line structures and black patches. Percent Laminarity is analogous to percent determinism except that it measures the percentage of recurrent points comprising vertical line structures rather than diagonal line structures. Trapping time on the other hand is the average length of vertical line structures. It represents the average time in which the system is "trapped" in a specific state. The eighth recurrence variable is **Vmax**, which is simply the length of the longest diagonal line segment in the plot. This variable is analogous to the standard measure Linemax.

IV. Episodic Recurrence Quantification Analysis

A Recurrence Quantification Episodic test is conducted on the full length of sample data sets (Fig. 3). Here an epoch is designed to have a width of 512 data points and is made moving giving a 128 point data shift resulting in a total of 191 epochs. From the tests it was found that only one among the eight RQA variables; the percent determinism shows constancy of value over the whole length of data. Moreover, as shown in Fig. 3, there is a wide separation between the means of the percent determinism values of the two data sets suggesting of two distinct dynamics. This is explained by the source of the data; the upper graph in each is from the system using good bearing whereas the lower graph is from the system using faulty bearing.

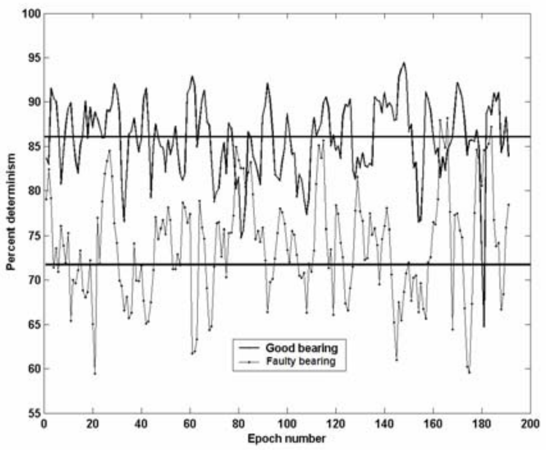

Fig. 3 Episodic Recurrence Analysis of test signals

Finally, it is to be noted here that the above tests are conducted with constant input parameter values for both the data sets; good as well as faulty bearing. As a check, we have examined the effects on RQA variables if the calculated input parameter values were used in RQA. Since the representative values of faulty bearing test data have been used as the constant input parameters, it is sufficient to analyze the good bearing signal data only, but with the calculated input parameter values for it. The RQA results indicates a similar trend here too, justifying the assumption to use constant input parameter values for online detection, as it is found to have no trade-off in using instantly calculated values for the input parameters epoch by epoch.

V. Conclusion

Wide separation between the mean values of RQA variable, Percent Determinism representing the two conditions under study suggest that RQA can be an efficient tool in analyzing time series related to fault detection in bearings. The definite advantage being that the RQA features can be extracted very easily from a noisy or non stationary time series signals which often is a challenge in mechanical systems signal processing. Also, since the data size and computational resource requirements are not large or intensive in comparison to the existing condition monitoring methods, the RQA based approaches proves to be a cost effective alternative. All these factors make RQA an attractive feature extraction methodology suitable for deployment in real-time condition monitoring systems involving roller bearings for quick retrieval of information with regard to bearing fault from sensor signals.

Acknowledgment

The authors thank Aju, Hariharan, Mathew and Mathew for their technical support. They also thank the Institute of Human Resources Development, Trivandrum and the Model Engineering College, Cochin for their facilities. They also thank the NPIU for providing financial support under the TEQIP Programme.

References

[1] F.K. Choy, J. Zhou, M.J. Braun, and L.Wang, "Vibration monitoring and damage quantification of faulty ball bearings," *Transactions of the ASME, Journal of Tribology*, vol. 127, pp. 776-783, Oct 2005.

[2] Antoni and R.B Randall, "Differential diagnosis of gear and bearing faults," *Transactions of the ASME, Journal of Vibration and Acoustics*, , vol. 124(2), pp. 165-171, 2002.

[3] T. Xinmin, D. Baoxiang, and X. Yong, "Bearings fault diagnosis based on HMM and fractal dimensions spectrum," in *Proc. of the 2007 IEEE Int. Conf. on Mechatronics and Automation*, China, 2007, pp. 1671–1676.

[4] K.R. Al-Balushi and B. Samanta, "Gear fault diagnosis using energy-based features of acoustics emission signals" *in Proc. of the I. Mech. E. Part 1, Journal of Systems and Control Engineering*, vol. 216(3), 2002, pp. 249-263.

[5] J.P. Eckmann, S.O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *Europhysics Letters*, vol. 5, pp. 973–977, 1987

[6] J.P. Zbilut and C.L. Webber Jr., "Embeddings and delays as derived from quantification of recurrence plots," *Physics Letters A*, vol. 171 (3–4), pp. 199–203, 1992.

[7] C.L. Webber Jr. and J.P. Zbilut, "Dynamical assessment of physiological systems and states using recurrence plot strategies," *Journal of Applied Physiology*, vol. 76-2, pp. 965-973, 1994.

[8] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, "Recurrence Plot Based Measures of Complexity and its Application to

Heart Rate Variability Data," *Physical Review E,* vol. 66(2), 026702, 2002.

[9] F. Takens, "Detecting strange attractors in turbulence," Dynamical Systems and Turbulence, Warwick, Lecture Notes in Mathematics 898, Springer-Verlag, pp. 366-81, 1981.

[10] H. D. I. Abarbanel, Analysis of the observed chaotic data, Springer-Verlag, Telos, 1996

[11] M.B. Kennel, R. Brown, and H.D.I. Abarbanel, "Determining embedding Dimension from phase-space reconstruction using a geometrical construction", Physical Review A, vol. 25, pp. 3403-3411, 1992.

[12] A.M. Fraser and H.L. Swinney, "Independent Coordinates for strange attractors from mutual information," Physical Review A, vol. 33, pp. 1134-1140, 1986.

[13] N. Marwan, "Encounters with Neighbors - Current developments of concepts based on recurrence plots and their applications", Ph.D. Thesis, University of Potsdam, ISBN 3-00-012347-4, 2003.

[14] C.L. Webber Jr, and J.P. Zbilut, "Recurrence quantification analysis of nonlinear dynamical systems," Tutorials in contemporary nonlinear methods for the behavioral sciences, (Chapter 2, pp. 26-94), M.A. Riley, G. Van Orden, eds. Retrieved December 1, 2004.