International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 14 (2015) pp 34861-34866

Remaining Life Time Prediction of Bearings Through Classification Using Decision Tree Algorithm

R.Satishkumar, V.Sugumaran

SMBS, VIT University, Chennai Campus, Chennai, India
cr_sathi@yahoo.co.in, sugumaran.v@vit.ac.in

Abstract
Bearings are widely used in rotatory machines which are running continuously for many hours. Knowing the remaining life time of the bearings will help to schedule the preventive maintenance program and to buy the necessary spare parts in advance. Researchers generally use regression techniques to assess the remaining life time of the bearings. This paper proposes a machine learning based classification approach to assess the remaining life time of bearings. The vibration signals were acquired on continuous basis from bearings operating rated speed and load conditions. The experiment was carried out on a new bearing until the bearing is damaged (not suitable for use). The time duration from start of the experiment to the end of the experiment is divided into five stages. From the sample signals of these stages, descriptive statistical features like mean, median, skewness, etc., were extracted. Out of 12 such features, the best performing five features were selected using decision tree. Then the selected features were used for training the classifiers to build a predictive model. The trained decision tree algorithm (classifier) was validated using distinct data. The classification performance is 95.48%.

Key words: Life time assessment, bearings, classification, vibration signals, machine learning.

1. Introduction

Bearing is one of the critical components in many industrial and automotive applications. As condition of the bearing directly affect the functionality of the machine in many applications, the condition monitoring of bearing is inevitable. Therefore numerous studies on fault diagnosis of bearing were reported in the last three decades. Some of them focus on the diagnosis of simulated faults, namely inner race faults, outer race faults, cage faults, etc,. While others attempt with naturally occurred faults. Many of the research works carried out on life time prediction are based on accelerated load tests to save time in conducting the experiments. As the phenomenon of degradation is different in accelerated load test than that of the rated load tests, there is clear need for conducting an experiment with rated load and speed. Few researchers have conducted the rated load tests to collect signals to build regression models for predicting the life time of bearings.

In many practical applications, it is very difficult to collect huge data on regular intervals is highly challenging due to want of resources, energy and money. Regression models may give better prediction of the life time; however in many practical applications, it is sufficient to know the approximate state of the bearing for maintenance purpose i.e., whether the bearing is in critical state or safe region. The present study proposes a new classification based approach for predicting the remaining life time of the bearings. Here, the brand new bearing is considered as to be in stage 1. After 1000 hours of running at rated speed and load conditions the bearing is considered to be in stage 2. Similarly, after 1250 and 1500 hours of running the bearing is said to be in stage 3 and stage 4 respectively. Stage 5 represents the damaged bearing condition approximately (1800 hours of running).

The vibration signals corresponding to the above stages were collected and there from statically features were extracted. From the extracted features a data model was built using J48 decision tree classification algorithm.

2. Literature survey

Palmgren and Lundberg [1-3] have laid foundation in estimating the life time of the bearing. This has given way to establish the standard for predicting the life time of the bearings with respect to the loads and speed of the application [4-6].

A bearing is said to be damaged when it generates increased noise and vibration and also the temperature of the lubricant increases drastically. Apart from this the bearing damages will be reflected in the vibration signals. The spikes and the amplitude on the signals reveal the damages on the bearing. There are number of methods / techniques deployed in diagnosing the bearing faults. Further to this, there are more research works carried out on predicting the life time and the health of the bearing using the defect diagnosis methods.

As per the International Standard Organization (ISO) [7], failure prognostic refers to the “estimation of the time to failure and the risk for one or more existing and future failure modes”. It’s the art or act of predicting the future state on the basis of present signs and symptoms with respect to current state [8]. The effective deployment of prognosis system turns out to deliver an effective maintenance schedule. Bearing prognosis aims at anticipating the failure date by predicting the future health state of a bearing and its remaining useful life [9]. Most of the researches in past, related to condition based maintenance is generally focused on fault diagnostics. Failure prognostic is an upcoming or emerging research domain where new methods and tools are being developed. Palmgren [4] on his research paper demonstrated the impact for the life of bearings subject to radial and axial loads. He also derived the resultant of various types of loads that are changing over period of time**.**

A. Model-Based Prognostic Methods

The output of this method is purely based on defined analytical model. The models are defined in best possible ways to represent the systems degradation phenomenon and dynamic performance. Yu and Harris [10] have developed a model to precisely predict the life of bearings through stress based fatigue prognostic method. However, in reality, the environment is complex in nature and the degradation phenomenon is not consistent. Hence it is very complex to design a model and apparently the application of this method is limited.

B. Experience-Based Prognostic Methods

This method uses the real time data collected over period of time (breakdowns, maintenance data, operational data like scrap or rework etc) to predict the remaining useful life of the bearing or the system. This method is relatively simple as it uses the basic reliability functions only. Palmgren [4] in his papers has detailed on the relations with respect to the life of bearings and the applied load which is based on the experimental data. The results from this method are less precise when compared to other methods.

C. Data-Driven Methods

This method uses the captured online data with the accelerometers and sensors etc. and convert the same to information for use. For example the online vibration signals extracted from the system can be used to study the degradation of the system. There are number of methods and techniques like statistical method, Artificial Neural Network (ANN), Spall propagation model etc., to evaluate and predict the useful life of the bearings or system.

Shao and Nezu [11] and Gebraeel [12] have proposed artificial neural-network-based models for estimating the bearing life and bearing failures. These models can be designed to accommodate the variable operating conditions and quickly responds to the changes in the environments.

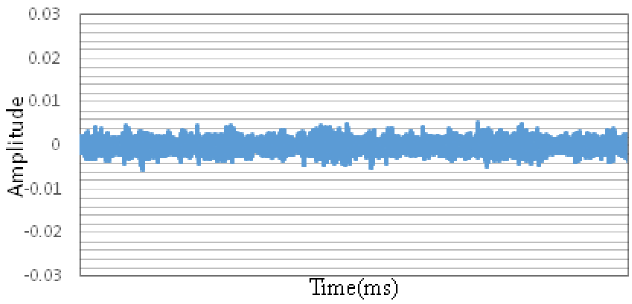
Data-Driven Methods are comparatively better than the rest of the methods as in real time environment developing a model and getting data is more complex and most of the time will not reflect the real environment. Also the prognostic results from this method are more precise since the online data reflects the system as it is.

3. Experimental setup and procedure

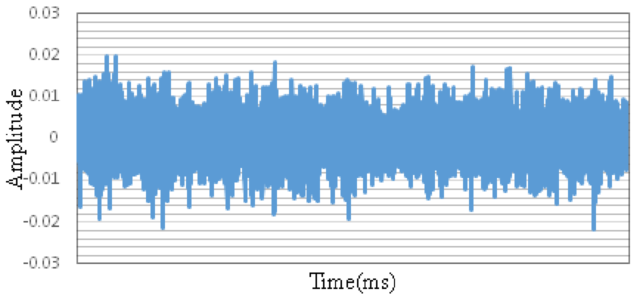
The entire experimental set-up is shown in Fig. 1. This set-up consists of a bearing, accelerometer, motor, DAQ card and lab view software loaded into a computer to acquire the vibration signals.

Fig. 1 Experimental setup

The bearings are mounted on to a shaft with Housing and in-turn connected to a variable speed motor. A few major parameters like speed, load, temperature, lubrication,. etc are simulated at real time conditions. The bearings are made to run to failure at variable speeds and variable axial loads. The vibration signals from the bearings are collected through the accelerometers placed on the housings and acquired via DAQ card. These signals are processed in the lab VIEW software which is connected to the computer. The experiment was run at real time conditions matching to the said bearing application. Adequate measures were taken to avoid bearing faults or damages due to fitments, misalignments etc. The experiments on each bearing were conducted till the bearing fails naturally. Of all the bearing monitoring data, vibration signal is more effective and suitable for reflecting bearing running condition. The amplitude of the vibration signals are monitored on daily basis for estimating the current health condition of the bearings. The vibration is small and smooth when the bearing is under normal condition. And the occurrence of bearing defect can cause fluctuation of vibration. During the process of degradation, the amplitude of vibration increases obviously as shown in the Fig. 2(a), 2(b). Thus vibration signal becomes the convenient variable to estimate the remaining useful life of the bearings.



2. (a) Brand new bearing



2. (b) Bearing after 1000 hours of running

4. Feature extraction

Descriptive statistical parameters such as mean, median, mode, sample variance, kurtosis, skewness, standard error, standard deviation, minimum, maximum, sum, and range were computed to serve as features. They are named as 'statistical features' here. Brief descriptions about the extracted features are given below.

(a) Standard error: Standard error is a measure of the amount of error in the prediction of *y* for an individual *x* in the regression, where *x* and *y* are the sample means and *'n'* is the sample size.

$$\text{Standard error of the predicted, } Y = \sqrt{\frac{1}{n-2}\left[\sum y-\bar{y}^{\,2} - \frac{\left[\sum x-\bar{x} \;\; y-\bar{y}\right]^2}{x-\bar{x}^{\,2}}\right]}$$

(b) Standard deviation: This is a measure of the effective energy or power content of the vibration signal. The following formula was used for computation of standard deviation.

$$\text{Standard Deviation } = \sqrt{\frac{\sum x^2 - \sum x^{\;2}}{n(n-1)}}$$

(c) Sample variance: It is variance of the signal points and the following formula was used for computation of sample variance.

$$\text{Sample Variance } = \frac{\sum x^2 - \sum x^{\;2}}{n(n-1)}$$

(d) Kurtosis: Kurtosis indicates the flatness or the spikiness of the signal. Its value is very low for normal condition of the bearing and high for faulty condition of the bearing a due to the spiky nature of the signal.

$$\text{Kurtosis} = \left\{\frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum\left(\frac{x_i-\bar{x}}{s}\right)^4\right\} - \frac{3(n-1)^2}{(n-2)(n-3)}$$

where 's' is the sample standard deviation.

(e) Skewness: Skewness characterizes the degree of asymmetry of a distribution around its mean. The following formula was used for computation of skewness.

$$\text{Skewness } = \frac{n}{n-1}\sum\left(\frac{x_i-\bar{x}}{s}\right)^3$$

(f) Range: It refers to the difference in maximum and minimum signal point values for a given signal.

(g) Minimum value: It refers to the minimum signal point value in a given signal. Therefore, it can be used to detect faulty signal condition.

(h) Maximum value: It refers to the maximum signal point value in a given signal.

(i) Sum: It is the sum of all feature values for each sample.

5. Feature selection

Feature selection is the process of selecting well contributing features among the available features. In the present study 12 descriptive statistical features have been extracted and out of 12 features the well contributing features have to be identified. Decision tree is used for feature selection. The feature which appears on the top of the decision tree is the best feature for the classification problem and coming down the next level, the next best features will be available. In this way a threshold is cut manually to select the number of features required for classification. All 12 features are ordered in the descending order of importance. Out of which only 8 features are appearing in the decision tree and rest of the 4 features are taken in the random order. An experiment was carried out by using only the top most features and classification accuracy is noted down, then top two features are selected and corresponding classification accuracy is noted. The process is continued till the classification accuracy for all the features are noted (refer Fig. 3). It is found that 5 features namely standard deviation, maximum, sum, skewness, kurtosis was found to be the best contributing features.

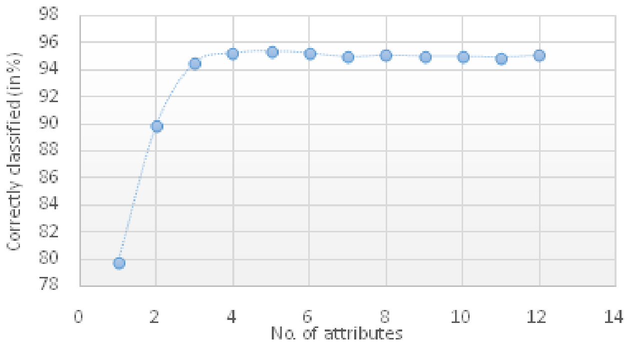


Fig. 3 Effect of Attributes

6. Classifier (J48 Decision Tree Algorithm)

A decision tree is related with ID3 (induction decision) consists of branches, one root, nodes and leaves. J48 algorithm is the one preferred for preparing Decision Trees. This methodology is made up of one root, number of leaves, nodes and branches. Branches refer to outcome of a decision and are represented by connecting lines. When the decision is numerical, the greater than branch is usually shown on the right and less than on the left. A node consists of one attribute. One branch refers to chain of nodes. Its path is from root to leaf. Oval shaped symbol represents node and rectangular shape depicts leaves. The algorithm identifies the good features for the purpose of classification from the given training data set and thus reduces the knowledge of domain to find out good features for solving classification problems.

The procedure of creating decision tree algorithm and performing the same for selecting features is summarized below:

i) The available feature set provides input to algorithm and decision tree is output.
ii) The decision tree possess leaf node, which expresses class label related with classes being classified.
iii) Branches of tree shows every single possible value of node from where they are originated.

7. Building and Testing of Predictive Model

The sound signals were recorded for normal and abnormal conditions. Totally 2500 samples were collected and classified into 5 different stages namely stage 1,stage 2, stage 3, stage 4 and stage 5 respectively. Stage 1 indicates that the bearing is in brand new condition. Stage 2 indicates the bearing is running at rated speed and load conditions after 1000 hours. Stage 3 and 4 indicates the bearing is running at rated speed and load conditions after 1250 and 1500 hours respectively. Stage 5 indicates that the bearing is in faulty condition. The statistical features explained in the section 4 will act as input to the algorithm. The paper deals with building predictive model using data available. In the present study, decision tree J48 algorithm is used for feature selection and classification. All 12 statistical features may not contribute equally for classification. Some features may contribute more compared to others. Irrelevant features sometimes may reduce the overall performance. Researchers have to extract all statistical features and have to select the good ones. Feature selection is done through decision tree.

First, only first best feature alone was used with decision tree and the classification accuracy was found. This happens to be standard deviation as it is the top node of the decision tree. Then, top two best features were used with decision tree and found the classification accuracy (standard deviation and maximum). This was repeated with top three features alone, top four features alone, till twelve features and the results were plotted in Fig. 3. It is evident from the graph that when the number of features is five the classification accuracy was highest (95.4 %). Any further increase in the number of features has adverse effect on classification performance. Hence, the number of features was fixed as five(standard deviation, maximum, sum, skewness, kurtosis).

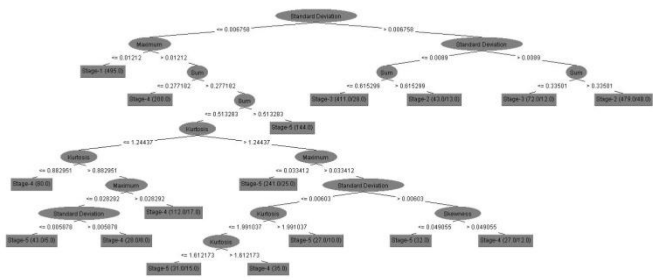


Fig. 4 Decision tree with selected five features

The decision tree depicted in Fig. 4 is built after feature selection. The structural information hidden in the dataset can be identified through decision tree and a set of 'if-then' rules can be formed.

(i) If the standard deviation value is less than or equal to 0.006758 and maximum less than or equal to 0.01212 then it is in stage-1
(ii) If the standard deviation is between 0.006758 and 0.0089 and sum is greater than 0.615299 then it is in stage-2
(iii) If the standard deviation is between 0.006758 and 0.0089 and sum is less than or equal to 0.615299 then it is in stage-3
(iv) If the standard deviation value is less than or equal to 0.006758, maximum value greater than 0.01212 and sum value less than or equal to 0.277182 then it is in stage-4
(v) If the standard deviation value is less than or equal to 0.006758, maximum value greater than 0.01212 and sum value is between 0.277182 and 0.513283 then it is in stage-5

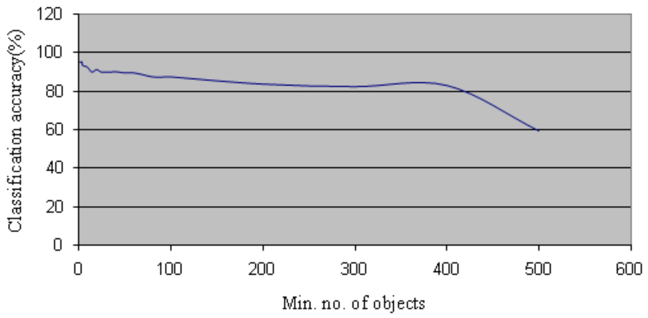


Fig. 5 Minimum no. of objects vs Classification accuracy

The number of objects required for forming a class was varied from 1 to 600 and found that when it was in 4, the algorithm are best classification accuracy of 94.4%(Fig. 5). When the number of data points is less the algorithm tends to over fit the data and when it is more the algorithm tends to generalize the built model. Hence, it is better to choose minimum number of objects to form a class as four. The confidence factor is varied from 0 and 1 and found that when confidence factor is 0.2 the algorithm gives best classification accuracy of 95.48%(Fig. 6)

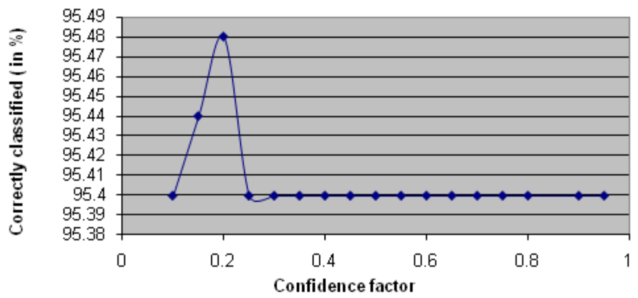


Fig. 6 Confidence factor vs Classification accuracy

Table 1 Detailed accuracy by class

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.996	0.001	0.998	0.996	0.997	0.998	Stage-1
	0.894	0.026	0.896	0.894	0.895	0.974	Stage-2
	0.896	0.027	0.892	0.896	0.894	0.978	Stage-3
	1	0.003	0.988	1	0.994	0.999	Stage-4
	0.988	0	1	0.988	0.994	0.995	Stage-5
Weighted average	0.955	0.011	0.955	0.955	0.955	0.989	

The detailed class wise accuracy of the algorithm can be found in Table 1. Out of the many terms used in Table 1, TP Rate and FP Rate are very important. 'TP rate' stands for true positive and 'FP rate' stands for false positive rate. True positive should be close to 1 and false positive rate should be close to 0 for better classification accuracy. In the present study, from Table 1, one can appreciate the closeness of 'TP rate' to '1' and 'FP rate' to '0'. The both values confirm that the built model is good one.

Table 2 Confusion matrix

a	b	c	d	e	Classified as
498	0	2	0	0	a = Stage–1
1	447	52	0	0	b = Stage-2
0	52	448	0	0	c = Stage-3
0	0	0	500	0	d = Stage-4
0	0	0	6	494	e = Stage-5

The classification accuracy of the decision tree algorithm can be found in Table 2. The interpretation of confusion matrix is as follows:

- The diagonal elements in the confusion matrix indicate the correctly classified instances.
- In the first row, the first element shows number of data points that belong to 'stage-1' is correctly classified as 'stage-1'.
- In the first row, the third element belongs to 'stage-1' is misclassified as 'stage-3'.

Similarly in all the rows, the diagonal elements represent correctly classified instances and the non diagonal elements were interpreted as misclassified instances. It should be noted that error percentage is less than 5% which is acceptable for many practical applications.

8. Conclusion

Bearing plays an important role in rotatory machines which runs continuously for many hours. The paper presented an algorithm based interpretation of vibrations signals for predicting the remaining life time of bearing through classification using decision tree algorithm. From acquired sound data, a model was built using data modelling technique. Decision tree J48 algorithm was used for feature selection and classification. The built model was tested with 10-fold cross-validation method and found its accuracy to be 95.48%. The error of about 5% is very much practically acceptable. Hence, the results of decision tree model can be practically used for predicting the remaining life time of bearings successfully.

9. Reference

[1] Palmgren, Ball and Roller Bearing Engineering, First edition, Translation by G. Palmgren and B. Ruley, SKF Industries, Inc., Philadelphia, PA, 1945.

[2] G. Lundberg and A. Paimgren, Dynamic Capacity of Rolling Bearings, Acta Polytechnica Mechanical

Engineering Series, vol.1, no. 3, Stockholm, Sweden, 1947.

[3] G. Lundberg and A. Palmgren, Dynamic Capacity of Rolling Bearings," ActaPolytechnica Mechanical Engineering Series, vol. 2, no. 4, Stockholm, Sweden, 1952.

[4] Palmgren, Die Lebensdauer von Kugellagern (The Service Life of Ball Bearings), Zeitschrift des Vereines Deutscherlngenieure, vol. 68, no. 14, pp. 339-341,1924,.

[5] Anon, Load Ratings and Fatigue Life for Ball Bearings, ANSI/AFBMA 9-1990, The Anti-Friction Bearing Manufacturers Association, Washington, DC, 1990.

[6] Anon, Rolling Bearings-Dynamic Load Ratios and Rating Life, ISO 281:1990(E), International Organization for Standardization, 1990.

[7] AFNOR, "Condition monitoring and diagnostics of machines – prognostics-part 1: General guidelines. NF ISO 13381-1," 2005.

[8] C.P. Henry, An historical view of the MFPT Society. In Proceedings of a Joint MFPT–JOAP–TSC Conference, Mobile, Alabama, pp. 3–15, 22–26 April 1996.

[9] D.A. Tobon-Mejia, K. Medjaher, N. Zerhouni and G. Tripot, A Mixture of Gaussians Hidden Markov Model for failure diagnostic and prognostic, 6th Annual IEEE Conference on Automation Science and Engineering (Case), Toronto, Canada, pp.338-343, 2010.

[10] W.K. Yu and T.A. Harris, A new stress-based fatigue life model for ball bearings, Tribology Transactions, vol. 44, no. 1, pp.11-18, 2001.

[11] Y. Shao and K. Nezu, Prognosis of remaining bearing life using neural networks, Proc. Inst. Mech. Eng., J. Syst. Control Eng.,pt. 1, vol. 214, no. 3, pp. 217–230, 2000.

[12] N. Gebraeel, M. Lawley, R. Liu and V. Parmeshwaran, Residual Life Predictions From Vibration-Based Degradation Signals:A Neural Network Approach, IEEE Transactions On Industrial Electronics, vol. 51, no. 3, pp. 694-700, 2004

.