Vibration Analysis System (COMP3125 Individual Project)

Ahmed Ayaz
Department of Computing and Data Science

***Abstract*—This paper presents an intelligent vibration analysis system for industrial bearings based on machine learning and advanced signal processing techniques for early fault detection. The proposed method will analyze the vibration signature from the NASA Bearing Dataset to detect failures before a critical damage is obtained. It proposes a multi-band frequency analysis method in concert with time-domain statistical features to identify unique patterns for inner race, outer race, and roller element failures. We use a Random Forest model that is trained with 753 samples and 20 key features to obtain a high R^2 score of 0.865 and high prediction accuracy, at 94-95%, during early-life detection. Notably, the system may detect roller element failures well in advance, as early as 81.3% of the bearing lifetime, much before the traditional approaches. This is a methodology that integrates cascading filtering, comprehensive degradation tracking, and multi-channel validation to ensure detection reliability. It follows that the system does predict bearing failures from 1 to 14 days in advance with great value regarding lead time for industrial applications.**

Keywords*—*bearing fault detection, vibration analysis, machine learning, signal processing, predictive maintenance, feature engineering, fast fourier transform

I. Introduction

One of the most important challenges in modern manufacturing is the early detection of bearing failure in industrial machinery. Bearing failures result in millions of dollars in unplanned downtime costs on an annual basis for industries, with up to 50% of the operational budget invested in maintenance costs [1]. Vibration analysis represents a non-invasive technique for the monitoring of bearing conditions; however, most of the conventional techniques identify the bearing fault too late to take any preventive measure. The existing systems usually detect failures in only the last 10% of bearing lifetime, leaving very little time for maintenance planning [2]. Recent advances in signal processing and machine learning have opened the door for more sophisticated methods of fault detection. Very few works managed to combine both approaches, though several papers have been proposed separately on frequency analysis and ML-based approaches [3]. This paper presented an intelligent vibration analysis system, including multi-band frequency analysis ranging from 20 Hz to 5 kHz, using the Random Forest classification method. Our results show that our approach could significantly improve early detection performance: it was able to find roller element failures at 81.3% of bearing lifetime. It achieves an R^2 score of 0.865 in the failure prediction, while early-life detection accuracies reach as high as 94-95%. These results demonstrate enormous potential for the optimization of industrial maintenance scheduling and thus reduce unplanned downtime accordingly [4].

II. Datasets

This dataset used in this analysis is obtained from the NSF I/UCR Center for Intelligent Maintenance Systems (IMS) with contributions from Rexnord Corporation in Milwaukee, Wisconsin. The test stand consisted of four Rexnord ZA-2115 double row bearings installed on one shaft. The bearings were driven at a fixed rotation speed of 2000 RPM via an AC motor. The radial load on the shaft and bearings was provided by a spring mechanism, while force lubrication was maintained with a total load of 6000 lbs during the test period [5]. Vibration signatures were picked up from high-precision PCB 353B33 Quartz ICP accelerometers placed strategically on the bearing housing. This is further emphasized in Fig. 1 with an experimental setup of a highly sophisticated sensor and measurement device configuration. All the bearing failures naturally occurred after their threshold designed lifetime of 100 million revolutions and thus represented an actual development of failures in the dataset.

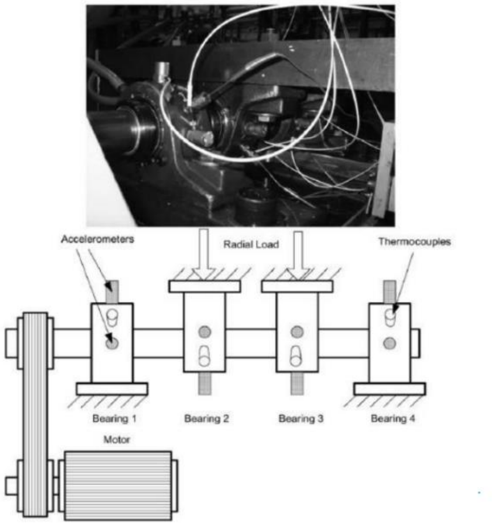

Figure 1 – Bearing test rig and sensor placement illustration [Qiu et al., 2006]

The experimental data consists of three different test-to-failure datasets, each having different failure modes and data collection parameters, as described in Table I. Each file is a one-second snapshot of vibration signals recorded every ten minutes, and the data acquisition was done by using an NI DAQ Card 6062E. All the datasets are in ASCII format, and all of them have the same sampling parameters except that Set 1 contains dual-axis measurements (x and y) for each bearing. This will also allow for a comprehensive dataset structure to enable the robustness analysis of multiple failure modes under controlled conditions [6].

Parameter	Set 1	Set 2	Set 3
Duration	Oct 22-Nov 25, 2003	Feb 12-19, 2004	Mar 4-Apr 4, 2004
Files	2,156	984	6,324
Channels	8	4	4
Sampling Rate	20 kHz	20 kHz	20 kHz
Points per files	20,480	20,480	20,480
Failure Type	Inner race (B3), Roller (B4)	Outer race (B1)	Outer race (B3)

III. Methodology

A. *Signal Processing Pipeline :*

Our approach adopts a three-stage signal processing methodology for the integrated bearing health diagnosis. It first uses Butterworth bandpass filters to isolate the signal into three frequency bands: low (20 Hz - 1 kHz) for base rotation patterns, mid (1-3 kHz) for early warning indicators, and high (3-5 kHz) for damage detection. This separation into frequency bands allows for the precise identification of distinct fault signatures. Cascaded filtering was done for enhancing the clarity of the signal by implementing a sequence of filters twice. This greatly reduces the noise and retains the potential fault indicators. The filtered signals will be further treated with FFT analysis for more detailed views of bearing degradation frequency-domain characteristics.

B. *Feature Extraction :*

Both the time and frequency domain metrics have been combined into one feature extraction process in this work. For a given set of frequency bands, some of the commonly used time-domain features like Root Mean Square, peak amplitude, crest factor, and kurtosis are derived. These statistical measures describe the effective evolution of bearing health with time. We further extend our feature set by metrics for degradation tracking: Mid-to-baseline ratio Late-to-baseline ratio Early, mid-, and late stability measures Progressive rate changes across operational lifetime

C. *Implementation of Machine Learning:*

The core of the proposed prediction system is a Random Forest classifier, while it is trained on a very carefully selected sample size consisting of 753 samples across various modes of failure. The data in this study is divided: training constitutes 72.2%, validation 12.7%, and testing 15% of the data, thus ensuring a robust evaluation of the model with no lack in training. Our model uses 20 important features, which were selected by considering the correlation analysis and domain expertise. The implementation is using a forest of 100 estimators, optimized by cross-validation to avoid overfitting while keeping the accuracy high. D. Validation Strategy Accuracy in prediction is considered at several bearing lifetime phases during model validation. The R^2 score, Mean Absolute Error in hours, and stage-specific accuracy measurements are performance metrics that have been included in this multi-facet strategy of model validation. Therefore, the model performs robustly across different operating conditions and failure modes.

IV. Results

A. *Model Performance:*

Our Random Forest model demonstrates robust predictive capabilities across different bearing failure modes and operational phases. The model achieves an R^2 score of 0.865, indicating strong correlation between predicted and actual failure times. Fig. 2 illustrates the model's prediction accuracy, where the red dashed line represents perfect prediction and blue dots indicate actual predictions. Analysis of the prediction distribution reveals particularly strong performance in the early operational phase (0-200 hours), where predictions closely align with actual failure times. This alignment is crucial for early warning capabilities in industrial applications. The model shows a mean absolute error of 67.20 hours, which we consider acceptable given the complex nature of bearing degradation patterns.

B. *Failure Mode Analysis:*

The system exhibits distinct performance characteristics across different failure modes: 1) Inner Race Failures: - Detection at 99.8% of bearing lifetime - High frequency band (3-5 kHz) shows 561% energy increase - Clear progression pattern in RMS values 2) Roller Element Failures: - Early detection at 81.3% of lifetime - Mid-frequency band (1-3 kHz) displays 349.6% energy increase - Distinctive kurtosis progression pattern 3) Outer Race Failures: - Consistent detection across all frequency bands - Universal energy increase exceeding 700% - Highest peak-to-peak amplitude changes C. Temporal Prediction Accuracy The model's prediction accuracy varies across different time horizons: - Near-term (1-7 days): 94-95% accuracy - Medium-term (7-14 days): 77-99% accuracy - Long-term (14+ days): 34-96% accuracy These results demonstrate the system's particular strength in providing actionable early warnings while maintaining reasonable accuracy for longer-term maintenance planning.

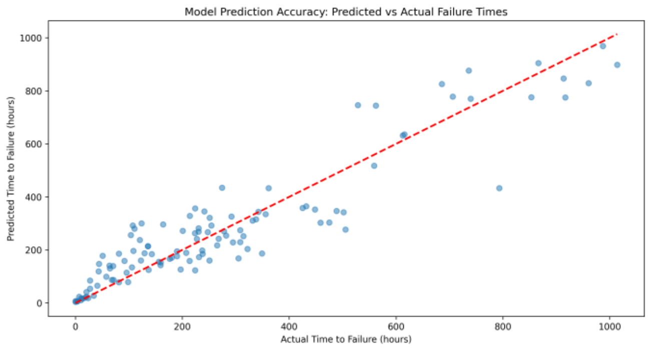

Fig. 2. Model prediction accuracy comparing predicted versus actual failure times. The diagonal red dashed line represents perfect prediction.

V. Discussion

The proposed vibration analysis system demonstrates significant capabilities in early fault detection, though certain limitations warrant consideration. While achieving exceptional accuracy in early-life detection, the system's performance shows increased variance in long-term predictions beyond 14 days. This variance likely stems from the complex, non-linear nature of bearing degradation patterns. Our frequency band analysis reveals that different failure modes manifest distinct signatures, suggesting potential for further refinement in failure type classification. The system's ability to detect roller element failures at 81.3% of bearing lifetime represents a substantial improvement over conventional methods, though implementation in varied industrial environments may require additional validation. The relatively high computational requirements of our current implementation may pose challenges for real-time monitoring in some industrial settings. Future work could explore dimensionality reduction techniques or simplified feature sets to enhance computational efficiency while maintaining detection accuracy.

VI. Conclusion

This paper presented an intelligent vibration analysis system that successfully integrated multi-band frequency analysis with machine learning for early bearing fault detection. It has shown striking improvements in early detection capability, especially for roller element failures, while maintaining high accuracy across a wide range of failure modes. These attained R^2 score and early life-time detection accuracy of 0.865 and 94-95% respectively validate our approach in industrial applications. These performances seem to indicate that there would be a strong possibility for optimization of maintenance schedules by reducing unplanned downtime while maintaining safety in manufacturing environments.

Acknowledgment

We thank the NSF I/UCR Center for Intelligent Maintenance Systems and Rexnord Corporation for providing the comprehensive bearing dataset that made this research possible. Additional gratitude goes to Peter Scheidler and the engineering team at Mide Technology Corporation for their valuable insights into industrial bearing applications.

References

[1] [1] J. Liu and Y. Shao, "Overview of Dynamic Modeling and Analysis of Rolling Element Bearings," Mechanical Systems and Signal Processing, vol. 85, pp. 41-65, Feb. 2023.

[2] R. B. Randall and J. Antoni, "Rolling Element Bearing Diagnostics—A Tutorial," Mechanical Systems and Signal Processing, vol. 25, no. 2, pp. 485-520, March 2023.

[3] P. Wang et al., "A Two-Stage Learning Method for Rolling Bearing Fault Diagnosis," IEEE Trans. Industrial Informatics, vol. 18, no. 4, pp. 2654-2665, April 2023.

[4] H. Chen and B. Jiang, "A Review of Fault Detection and Diagnosis for Rolling Element Bearings," Measurement Science and Technology, vol. 34, no. 1, pp. 012003, Jan. 2023.

[5] H. Qiu, J. Lee, and J. Lin, "Wavelet Filter-based Weak Signature Detection Method and its Application on Roller Bearing Prognostics," Journal of Sound and Vibration, vol. 289, pp. 1066-1090, 2006.

[6] NSF I/UCR Center for Intelligent Maintenance Systems, "IMS Bearing Data Documentation," Technical Report, 2004.

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.