Bearing Race Faults Classification using Simulation-generated Training Data and Feature Free Methods

Cameron Sobie[1], Carina Freitas[1] and Mike Nicolai[1]

[1]*Siemens Industry Software N.V., 3000 Leuven, Belgium*
cameron.sobie@siemens.com

ABSTRACT

Rotating machinery is central to transportation and power generation, and maximizing its uptime while minimizing unplanned maintenance is important from safety and economic perspectives. Roller bearings are ubiquitous components in such machinery and are very often the cause of failures; thus, condition monitoring of roller bearings is a topic of key interest to many industries. Historically accomplished using human expertise and experience to estimate the condition of a machine from a limited set of signals and curated statistical features, machine learning has become an effective tool in bearing fault identification with modern computing and algorithmic advances. In this regard, two major challenges exist: first, the availability of in-service data from bearings containing faults is often rare or difficult to obtain, and moreover, for machines being deployed to the field for the first time, no such data exists. Secondly, the extracted statistical measures (features) must be chosen to maximize the classifier accuracy (or another metric), and their selection of these to maximize accuracy can be a challenging task. We directly address these challenges with two separate approaches. With validation against four experimental datasets, we show that in the absence of recorded in-service data, machine learning algorithms trained using simulated bearing vibration signals (i.e. simulation-driven machine learning) can classify bearing race faults with greater accuracy than those trained using data collected from other machines. Next, we propose convolutional neural networks and nearest-neighbor dynamic time warping (NNDTW) as statistical feature-free methods to detect bearing race faults using a signal processing pipeline based on angle synchronous averaging. We show that these methods offer superior accuracy from the simulation-driven perspective and can predict an in-service wind turbine fault over one month before failure.

1. INTRODUCTION

Roller bearings are fundamental components in rotating machinery such as wind turbines, electric machines, and transportation applications. The nature of their construction results in the bearing races being subject to cyclic loading with high stresses arising from Hertzian contact with the roller elements. As a result, bearing faults are expected under nominal loading and one can predict the fatigue life according to well-accepted standards; however, these standards describe probability distributions for lifetime and the fatigue life of an individual bearing varies according to its unique material properties and flaws. Therefore, it is important to detect minor flaws before a total failure occurs and unplanned maintenance is required.

Several measurement techniques such as linear accelerometers and acoustic arrays have been proposed to monitor bearings for faults, with linear accelerometers being favored. Many signal processing and feature extraction methodologies for bearing fault detection have been proposed. Time-series data can be directly analyzed to produce features differentiating between nominal and faulted bearings with measures such as root mean square (RMS), skewness, kurtosis, or crest factor (Camci, Medjaher, Zerhouni, & Nectoux, 2013). Preprocessing methods such as band-passing, calculating the envelope signal, and cepstrum calculation have also been proposed. The natural periodicity of rotating machines makes frequency analysis a natural approach, and each bearing fault (inner/outer race, roller element) has a characteristic frequency relative to the shaft frequency (assuming no slipping between the components). The power at these frequencies can be used as features in addition to other typical features such as spectral kurtosis. Fourier transform methods are limited to steady state operation and are not well-adapted to varying shaft speeds without additional processing. Modern time-frequency domain methods such as the Hilbert-Huang transform, and discrete wavelet transforms (DWT) can address these problems. Apart from transform methods, angle synchronous averaging (ASA) is a commonly used method to emphasize bearing defect signals and reduce noise (McFadden & Toozhy, 2000). This analysis method is resistant to varying shaft speeds and because it includes the bearing geometry, angle synchronous averages between different bearing geometries can be compared.

Machine learning algorithms have proven effective as a tool to move beyond manual bearing fault diagnosis. Statistical features have been shown to effectively determine bearing fault state for cross-validation between bearings in the same machine. For example, Rauber, de Assis Boldt, & Varejao (2015) created a feature pool including time-series, frequency domain, and wavelet transform statistics to successfully classify bearing faults with several different classifiers reaching 99% accuracy. More recently, one-dimensional convolutional neural networks applied to the bearing acceleration frequency spectrum have been proposed as a more effective method of bearing fault detections than statistical feature analysis with the additional benefit of removing the need for expert knowledge associated with feature extraction (Janssens, et al., 2016). This method shows promise, but in its proposed form was limited to stationary operating regimes as the input to the convolutional neural network is the power spectrum, which is inherently dependent on the operating speed and bearing configuration.

This work seeks to address major challenges facing machine learning based condition monitoring with the first deployment of a new machine such as, which data should be used to train the algorithm, and which features should be extracted to classify the fault state. We address these challenges simulating the vibration response of a defective bearing, and using this data to train machine learning algorithms to detect race faults. This methodology is validated and compared against data-driven methods using four different experimental datasets including one industrial application from a wind turbine. We address the second challenge by proposing two feature-free machine learning methods for bearing fault detection, and show that these algorithms outperform classical statistical methods in condition monitoring applications where prior in-service data containing faults does not exist.

2. METHODOLOGY

2.1. Signal processing and normalization

The signal processing in this work relies on calculating the envelope signal using a Hilbert transform and transforming time series data into the angle domain using angle synchronous averaging. Averaging several defect periods acts to reduce asynchronous noise and emphasize a characteristic vibration caused by a fault. The figures in this work average over 25 defect periods, which is observed to be optimal or nearly so for many algorithms and simulated/experimental data. Each accelerometer signal produces two ASAs (one per race) and each of these is considered by the classifiers. Figure 1 depicts typical ASAs for a bearing from the CWRU dataset.

Energy-based features are often proposed in literature and are effective when in-service data is available for a specific machine. However, this limits the robustness of the extracted features because energy-dependent features are correlated to the operating speed of the machine and are thus not robust against non-stationary conditions (or speeds that are not covered in experimental data) and therefore only energy-free features are considered and all ASAs are normalized to zero mean and unit variance before further processing.

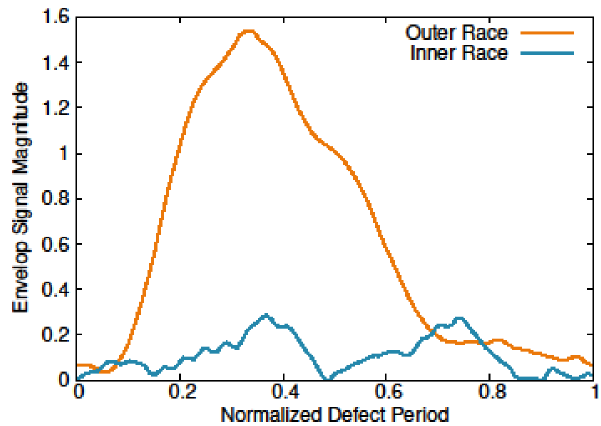

Figure 1. Angle synchronous averages for a bearing with an outer race defect.

2.2. Machine Learning

Fault classification from statistical features is performed here with the following algorithms: logistic regression, K-nearest neighbors, random forests (RF), support vector machines, and MLP neural networks. These algorithms have been researched extensively for bearing fault detection and details of their theoretical foundations, implementations and applications are widely available in the literature. Here, the hyper-parameters of these algorithms are optimized and the most accurate method in the data cross-validation case is selected.

Convolutional neural networks were proposed as a new machine learning method better adapted to analyze high dimensional data and data with inherent structure (e.g. images or time series data) than classical neural networks (LeCun & Bengio, 1995). While CNNs have been highly popularized by their state-of-the-art performance for image classification, one-dimensional applications have been more limited.

Dynamic time warping was first proposed as a method to compare time-sequences with the goal of improving spoken word recognition and is qualitatively an elastic distance measure between two sequences of ordered numbers. As a classification method, first nearest-neighbor dynamic time warping (NNDTW) has been shown to be effective at time series motif recognition matching at large data scales (Keogh & Ratanamahatana, 2005). In this work, NNDTW is used to detect race faults in roller bearings by comparing the

ASAs derived from experiments to simulation-generated ASAs, and the experimental signal is assigned the class of the nearest simulated signal.

2.3. Data Approaches

The model we adopt in this work is the one-dimensional 3-DOF model of Sassi, Badri & Thomas (2007), which is implemented as simulated using Siemens LMS Amesim (Siemens Industry Software NV, 2016). To summarize the model, the first oscillatory modes of the inner and outer races constitute two of the degrees of freedom, and the third is the roller element. The elastohydrodynamic film was initially modeled with as a spring-damper before a range of parameters was tested to explore the envelope of a typical dynamic bearing response. The roller element is assumed to be infinitely stiff. A race defect is modelled either with a prescribed force representing the roller element entering and exiting the defect. This information is used to create longer time series signals incorporating stochastic effects.

The experimental used in this work is composed of two publicly available datasets from the Case Western Reserve University (CWRU) and the Society for Machine Failure Prevention Technology (MFPT) as well as a dataset produced exclusively for this work using a machinery fault simulator (SQ). Each apparatus has a different bearing, loading, and speed configuration. All seeded-fault experimental data is considered here as an ensemble (except for the wind turbine data, considered in an independent section) and the accuracy scores reported are thus for the union of the CWRU, MFPT and SQ datasets.

The wind turbine high-speed bearing fault data is from a published dataset (Bechhoefer & Kingsley, 2009) that was collected from a commercial in-service 2MW wind turbine. Six seconds of data was captured per day for 50 days preceding the discovery of a crack in the inner race of the high speed bearing.

3. RESULTS

3.1. Classification of Seeded Faults

The statistical features derived from the ASA are the skewness, kurtosis, crest factor and margin factor, and from the ASA power spectrum, the skewness and kurtosis are used. Finally, including the energies of the first four discrete wavelet decompositions (using the Daubechies 2 wavelet) results in a total of twelve features. The extracted statistical features are normalized to zero mean and a maximum/minimum of ±1. The algorithms are optimized over their respective hyper-parameters, and the superior method was found to be a random forest classifier; a more detailed analysis is available in (Sobie, Submitted 01/2017). While statistical feature extraction attempts to encapsulate as much information as possible using carefully selected statistics and convolutional neural networks learn feature filters that capture the "shape" of a bearing acceleration signal associated to a fault, and dynamic time warping provides a direct distance measure between signals with no need for feature extraction or model training.

Figure 2 shows that both the CNN and NNDTW classifiers exceed the performance of the best statistical feature based classifier in simulation-driven performance, and in the case of the CNN, data-driven applications as well.

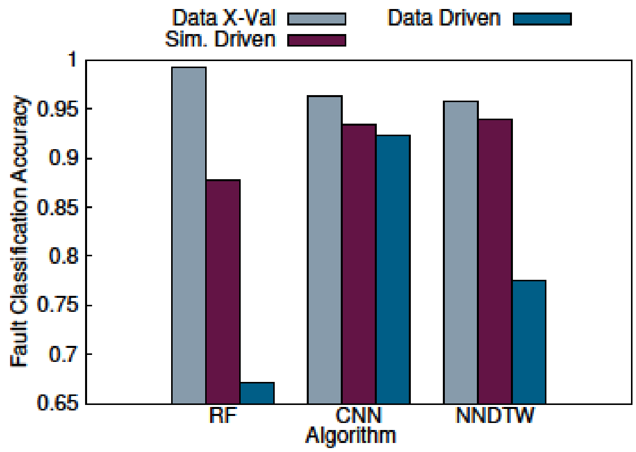

Figure 2. Seeded bearing fault classification accuracy.

Data cross-validation accuracy is above the statistical classifier average (not shown here) but lower than for random forests. NNDTW is particularly well suited to simulation-driven condition monitoring because the initial data that can be supplied through simulations can be augmented online with in-service fault data as it is recorded, whereas all other classifiers require retraining. NNDTW can also provide much greater insight into the fault compared to other methods, particularly in combination with the proposed simulation-driven methodology. Beyond the signal classification, all of the physical parameters associated to the nearest-neighbor match are available, which could give deeper insight into the defect nature and severity. Such insight is cannot be gained using a CNN or the statistical feature classifiers used here. However, it comes at a greater computational cost than the forward pass of a CNN.

3.2. Wind Turbine Industrial Application

The previous section establishes the validity and performance of simulation-driven machine learning, but the faults in the experimental datasets were seeded artificially rather than faults arising under normal loading. To further establish the effectiveness of the proposed method, an in-service fault from an industrial application is considered. A crack was found in the inner race of a high-speed bearing of a wind turbine, and fifty days of vibration data preceding the crack discovery have been made available. The data from each of the fifty days is angle synchronously averaged and the probability of a defect present is calculated using the algorithms. Figures 3 and 4 show that the probability of an outer race fault is consistently predicted below 20%, providing a baseline to which the inner race predictions can be compared.

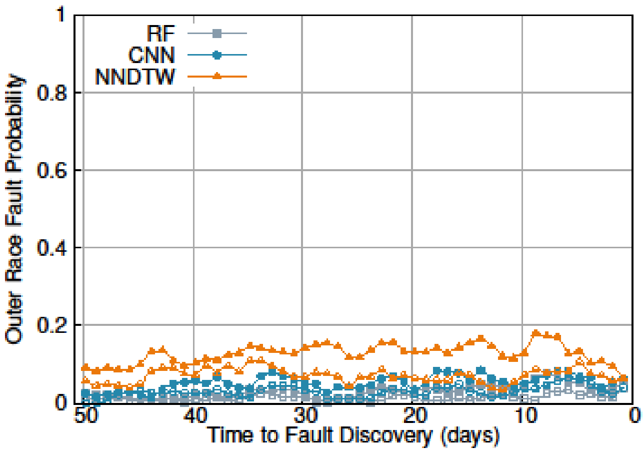

Figure 3. Wind turbine high speed bearing outer race fault probability.

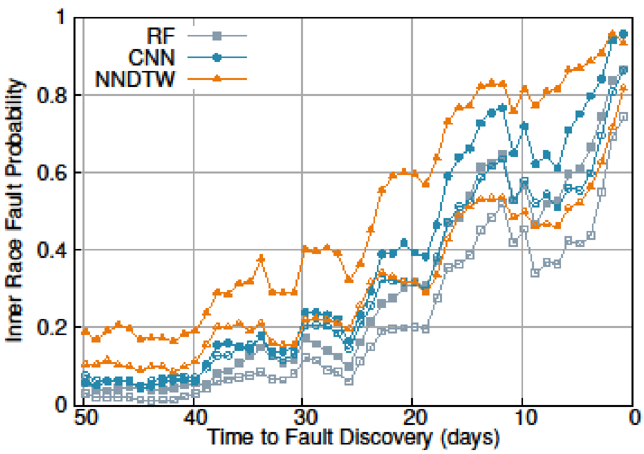

Figure 4. Wind turbine high speed bearing inner race fault probability. Filled/empty symbols represent simulation-/data-driven predictions.

The inner race fault state is steady for 10 days, after which it climbs to nearly 100% probability by the final day. All methods clearly detect the fault, and simulation-driven methods consistently predict defects with higher probability than data-driven methods.

4. CONCLUSION

Data-driven machine learning has received significant attention in the field of condition monitoring and failure prognosis, but its dependence on in-service failure data and the selection of useful features for unseen data presents challenges towards industrial adoption. We show that data generated with a simple bearing model with added stochastic effects can be used to train classifiers to accurately detect race faults across four experimental datasets. Simulated data cannot replace experimental data derived from the exact conditions under which monitoring is being performed; however, it provides a strong starting point for novel applications, to which in-service data can be added to further improve classifier performance. Two new applications of feature-free machine learning algorithms to fault identification proposed in this work are shown to match or outperform statistical feature based classifiers, and the novel application of NNDTW is of particular interest because new data can be included without algorithm re-training and it can give greater insight into the fault process as it identifies the most similar simulation.

ACKNOWLEDGEMENT

We would like to acknowledge the support of VLAIO and ITEA of the Reflexion project (#14035).

REFERENCES

Bechhoefer, E., & Kingsley, M. (2009). A review of time synchronous average algorithms. *Annual Conference of the Prognostics and Health Management Society*, (pp. 24-33). San Diego, CA.

Camci, F., Medjaher, K., Zerhouni, N., & Nectoux, P. (2013). Feature evaluation for effective bearing prognostics. *Quality and reliability engineering international*, 477-486.

Janssens, O., Slavkovikj, V., Vervisch, B., Stockman, K., Loccufier, M., Verstockt, S., . . . Van Hoecke, S. (2016). Convolutional neural network based fault detection for rotating machinery. *Journal of Sound and Vibration*, 331-345.

Keogh, E., & Ratanamahatana, C. A. (2005). Exact indexing of dynamic time warping. *Knowledge and information systems*, 358-386.

LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*.

McFadden, P., & Toozhy, M. (2000). Application of synchronous averaging to vibration monitoring of rolling element bearings. *Mechanical Systems and Signal Processing*, 891-906.

Rauber, T. W., de Assis Boldt, F., & Varejao, F. M. (2015). Heterogeneous feature models and feature selection applied to bearing fault diagnosis. *IEEE Transactions on Industrial Electronics*, 637-646.

Sassi, S., Badri, B., & Thomas, M. (2007). A numerical model to predict damaged bearing vibrations. *Journal of Vibration and Control*, 1603-1628.

Siemens Industry Software NV. (2016). Siemens LMS Imagine.Lab Amesim (Version 15).

Sobie, C. a. (Submitted 01/2017). Simulation-Driven Machine Learning: Bearing Fault Classification. *Submitted to Mechanical Systems and Signal Processing*.