Adapting Approximate Entropy as a Health Indicator of Rotating Machinery for Estimation of Remaining Useful Life

Cody Walker[1]

[1] *University of Tennessee Knoxville, Knoxville, TN, 37996, US*
cwalke42@vols.utk.edu

ABSTRACT

Despite significant attention to online health monitoring and prognostics of bearings, many common health indicators are not sensitive to early stages of degradation. This research investigates the use of approximate entropy (ApEn), previously developed for fault diagnostics, as a health indicator for prognostics. ApEn quantifies the regularity of a signal; as bearings degrade, the frequency content of vibration signals changes and affects the ApEn as the vibration becomes more chaotic. Early results suggest ApEn supports earlier degradation detection and more predictable progression from fault to failure. This research focuses on optimizing parameters of the ApEn calculation to provide guidance across a variety of bearing types, sizes, and geometries in both steady-state and transient operation.

1. PROBLEM STATEMENT

Bearings are a common failure point for most rotating equipment. Predicting the presence of a bearing fault can be done with a variety of methods including: Fourier analysis, wavelets, and Hilbert transforms (Lebold, McClintic, Campbell, Byington, & Maynard, 2000). However, the remaining useful life (RUL) of a bearing is difficult to predict without a clear indication of the bearing's current health. For bearings in particular, the transition from a healthy state to a failed state may seem abrupt depending on the health indicator. The goal of my dissertation is develop a health indicator that can gives earlier indication of fault initiation and better fault progression characteristics.

2. EXPECTED CONTRIBUTIONS

My research focuses on improving the Approximate Entropy (ApEn) statistic from a diagnostic tool into a health indicator for prognostics. ApEn measures the regularity of a signal. As a bearing degrades, there is an increase in the number of fre-

quency components. This will decrease the regularity of the signal thus increasing the ApEn value. Regular or repetitive signals have low ApEn values, while irregular or noisy signals have high ApEn values (Caesarendra, Kosasih, Tieu, & Moodie, 2013). This feature looks to be a promising health indicator, because it has an early fault detection time and trends both intuitively and favorably as the fault progresses.

The first step in calculating the ApEn is to convert the input signal into a phase space. This is done as the first step in several other nonlinear, feature-extraction methods including: largest Lyapunov exponent, correlation dimension, and fractal dimension (Williams, 1998). The input signal is time series data, such as $Y = (y_1, y_2, ..., y_N)$. The time series Y then would be reconstructed into the phase space X following the form seen in equation 1

$$X = \begin{pmatrix} y_1 & y_{1+J} & y_{1+2J} & \cdots & y_{1+(m-1)J} \\ y_2 & y_{2+J} & y_{2+2J} & \cdots & y_{2+(m-1)J} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_M & y_{M+J} & y_{M+2J} & \cdots & y_{M+(m-1)J} \end{pmatrix}, \quad (1)$$

where J represents the lag, m is the embedding dimension, and M is the number of reconstructed vectors. Both the lag, J, and the embedding dimension, m, are predetermined. The lag spaces out the time series data to try and capture underlying features of the data. The lag can be chosen based on autocorrelation-function indicators or simply via a guess-then-check method (Williams, 1998). The embedding dimension increases the length of each reconstructed vector, which has a direct computational cost. Small embedding dimensions are preferred for their expediency and simplicity of interpretation, so m=2 was chosen (Williams, 1998).

The maximum difference is calculated between the corresponding elements of two reconstructed vectors as seen in equation 2

$$d[X(i), X(j)] = \max_{k=1,2,...,m} (|x(i+k-1)-x(j+k-1)|), \quad (2)$$

where i = 1,2,...,N-m+1, j = 1,2,...,N-m+1, and N is the number of data points in the reconstructed vectors (Yan & Gao, 2007). For each vector, the similarity between that vector and every other vector is computed in equation 3

$$C_r^m(i) = \frac{1}{N-(m-1)} \sum_{j\neq i} \Theta\{r - d[X(i), X(j)]\}, \quad (3)$$

where $j \neq i$, r is a predetermined tolerance, and Θ is the Heaviside function. This similarity is a measure of pattern frequency within the phase space within a tolerance of r given a window length of m.

$$r = k * std(S), \quad (4)$$

The tolerance, r, is defined as equation 4 where k is a positive constant, std is the standard deviation, and S is a population of related time series data. The natural logarithm of each $C_r^m(i)$ is then averaged, seen in equation 5.

$$\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} ln(C_r^m(i)) \quad (5)$$

The embedding dimension is then increased by one and each step is repeated to calculate ϕ^{m+1}. The ApEn value is denoted by equation 6.

$$ApEn(m,r) = \phi^m(r) - \phi^{m+1}(r) \quad (6)$$

Previous literature utilized ApEn to differentiate bearing defects from unfaulted bearings (Caesarendra et al., 2013; Yan & Gao, 2007). My research looks to advance this feature for use as a prognostic parameter to predict the RUL for bearings. The tolerance, calculated in equation 4, is a key component in the ApEn formulation with a wide range of recommended values. My research looks to determine the ideal tolerance for prognostics using multi-objective optimization methods to maximize ApEn's prognostic metrics of monotonicity, trendability, and prognosability (Coble & Hines, 2009). This will ensure that not only can ApEn be used to differentiate between faulted from unfaulted components but also be used as a health indicator. Once the ideal tolerance has been chosen, prognostic models will be created, tested, and validated using real, run-to-failure data.

I will be examining bearings under both stationary and nonstationary operating conditions to study how this affects the ApEn. Nonstationary operating conditions are commonly seen during start-up, shutdown, and transients in operation. I plan to combat the changes in velocity during nonstationary operation with adaptive resampling. This technique converts nonstationary, time-series vibration data into the stationary angular domain through the use of a tachometer (Blough & Professor, n.d.).

A warning system will be created based on the approximate entropy of other features related to the health of the system, such as vibration root mean square (RMS) and kurtosis. Access to raw vibration data may be limited due to the large space requirements, so features of the data are stored in lieu of the raw data. The ApEn of these features may also be useful in determining the health of a system. Although the feature may not have crossed a failure threshold, the changes in the regularity of these features could provide earlier indications that the system is no longer in an unfaulted state. I am working to develop a system will notify the end user that changes are occurring within these variables, and inspection of the system is required.

The maximum difference in equation 2 is essentially a distance measurement. By changing the input from a simple time-series vector into a time-series matrix composed of multiple signals representing the state of the system, the regularity of multiple signals can be determined at once. This approach looks to combine multiple features to represent the system then use the ApEn algorithm to produce a single number to be interpreted.

3. RESEARCH PLAN

The research plan can be summarized into four steps:

1. Data collection and generation.
2. Optimization of ApEn tolerance.
3. Development of diagnostic and prognostic models.
4. Comparison to current best practices.

3.1. Work Performed

A small bearing test bed has been constructed to produce nonstationary, run-to-failure data for roller ball bearings. The test bed consists of a motor, the test bearing housing, a spring assembly, and an alternator. The motor is controlled by a variable frequency drive (VFD), so the velocity profile will vary in user-defined patterns. The spring assembly provides a radial load to the test bearing housing, while the alternator provides a load for the shaft to turn. Each test bearing is first pre-seeded with an outer race fault then installed into the bearing test bed. During operation, the spring assembly provides a radial load to the test bearing housing equating 85% of the bearing's dynamic load capacity to hasten the degradation process. The bearings are ran under a sinusoidal velocity profile. Many measurements are taken throughout each run

including vibration, temperature, acoustic, spring load, motor current, and shaft speed.

Run-to-failure data has been acquired from other sources as well. The NASA data repository hosts two sets of run-to-failure bearing data (J. Lee & Services, 2007). These data sets do not contain a large population of runs, but they have been studied by many other scientists. This will provide an opportunity to compare the diagnostic and prognostic models created in this research with models and methods used by others. Accelerated motor degradation testing was performed at the University of Tennessee Knoxville. This test was designed to produce failures in the stator wiring by simultaneously overheating and overloading the motors. Sixteen motors were ran to failure and will be examined to determine if ApEn is a suitable health indicator for not only bearings but also motors and rotating equipment in general.
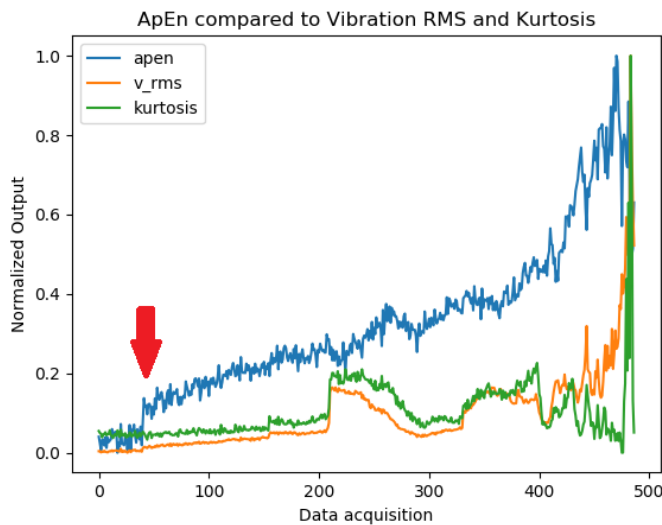

Figure 1. ApEn shows an earlier indication that a fault has occurred and a more monotonic trend than vibration RMS and kurtosis for a bearing outer race fault.

The ApEn threshold was first optimized to maximize the monotonicity of the ApEn value for a single bearing run using a basin-hopping technique. This threshold value will be used as the initial guess for future multi-objective optimization. In figure **??**, ApEn is compared to commonly used bearing health indicators, vibration RMS and kurtosis, for a bearing with an outer race fault. Each feature has been normalized, so they can be adequately compared. ApEn shows both an earlier indication of fault initiation, noted by the red arrow, and a more monotonic trend of damage progression.

3.2. Remaining Work

Work still remains on generating nonstationary, run-to-failure data using the small bearing lab. The ApEn will be calculated for a population to see if the same advantageous qualities seen during stationary operation is present during non-stationary operation. Multi-objective optimization will select the ideal ApEn tolerance based on the resulting ApEn's measures of monotonicity, trendability, and prognosibility.

Raw vibration data will be used as the input for the ApEn algorithm in prognostic models. A early-warning system will also be constructed using other features, such as vibration RMS and kurtosis, as the input for the ApEn algorithm. Changes in the regularity of these features may indict changes within the system itself. A multi-parameter input for the ApEn algorithm will be created to compare changes in the regularity of states within the system rather than comparing a single feature.

4. CONCLUSION

ApEn appears to be a promising health indicator with early fault detection and monotonic trending. This work hopes to push more maintenance strategies from preventative maintenance to condition-based maintenance through the use of more accurate and precise health indicators for rotating machinery.

REFERENCES

Blough, J. R., & Professor, A. (n.d.). *Adaptive Resampling-Transforming From the Time to the Angle Domain* (Tech. Rep.).

Caesarendra, W., Kosasih, B., Tieu, K., & Moodie, C. A. (2013). An application of nonlinear feature extraction - A case study for low speed slewing bearing condition monitoring and prognosis. *2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics: Mechatronics for Human Wellbeing, AIM 2013*, 1713–1718. doi: 10.1109/AIM.2013.6584344

Coble, J., & Hines, J. W. (2009). Identifying optimal prognostic parameters from data: a genetic algorithms approach. In *Annual conference of the prognostics and health management society, san diego, ca, september.*

J. Lee, G. Y. J. L., H. Qui, & Services, R. T. (2007). Bearing data set ims, university of cincinnati. "bearing data set", nasa ames prognostics data repository nasa ames research center, moffett field, ca.

Lebold, M., McClintic, K., Campbell, R., Byington, C., & Maynard, K. (2000). Review of vibration analysis methods for gearbox diagnostics and prognostics. In *Proceedings of the 54th meeting of the society for machinery failure prevention technology* (Vol. 634, p. 16).

Williams, G. (1998). *Chaos theory tamed* (Vol. 35) (No. 08). doi: 10.5860/CHOICE.35-4561

Yan, R., & Gao, R. X. (2007, feb). Approximate Entropy as a diagnostic tool for machine health monitoring. *Mechanical Systems and Signal Processing*, *21*. doi: 10.1016/J.YMSSP.2006.02.009