A Systematic Methodology for Gearbox Health Assessment and Fault Classification

Hassan Al-Atat[1], David Siegel[1], and Jay Lee[1]

[1] *NSF I/UCRC for Intelligent Maintenance Systems (IMS), University of Cincinnati, Cincinnati, Ohio, 45220*

atathf@mail.uc.edu
siegeldn@mail.uc.edu
jay.lee@uc.edu

ABSTRACT

A systematic methodology for gearbox health assessment and fault classification is developed and evaluated for 560 data sets of gearbox vibration data provided by the Prognostics and Health Management Society for the 2009 data challenge competition. A comprehensive set of signal processing and feature extraction methods are used to extract over 200 features, including features extracted from the raw time signal, time synchronous signal, wavelet decomposition signal, frequency domain spectrum, envelope spectrum, among others. A regime segmentation approach using the tachometer signal, a spectrum similarity metric, and gear mesh frequency peak information are used to segment the data by gear type, input shaft speed, and braking torque load. A health assessment method that finds the minimum feature vector sum in each regime is used to classify and find the 80 baseline healthy data sets. A fault diagnosis method based on a distance calculation from normal along with specific features correlated to different fault signatures is used to diagnosis specific faults. The fault diagnosis method is evaluated for the diagnosis of a gear tooth breakage; input shaft imbalance, bent shaft, bearing inner race defect, and bad key, and the method could be further extended for other faults as long as a set of features can be correlated with a known fault signature. Future work looks to further refine the distance calculation algorithm for fault diagnosis, as well as further evaluate other signal processing method such as the empirical mode decomposition to see if an improved set of features can be used to improve the fault diagnosis accuracy.

1. INTRODUCTION

Diagnosis and health assessment of rotary machinery using vibration signals from the machine has been a domain of interest for many years. Prior to a total failure, degradation and incipient level of damage in components of the machinery can demonstrate behavioral features hidden within the vibration signals. In practice, a machine with no faults will have a vibration signature, "normal" signature, based on its system dynamics and forces acting on the system. Different mechanical faults in different components will display different vibration signatures that can be differentiated from the "normal" signature, with the utilization of the proper signal processing techniques. One of the main challenges in diagnosis and health assessment of rotary machinery is the potentiality of the machine or equipment to operate in a multitude of regimes, and thus their vibration behavior will be different in each regime. Another, challenge is that the "normal" vibration signature from every operating regime might not be available for prior training. This paper proposes a multi-regime health assessment and fault-diagnosis systematic methodology for gearbox systems, which utilizes different techniques for signal processing, regime segmentation, baseline "normal" signature detection and fault-diagnosis. This methodology of developing a fault classifier without baseline data and for a system that operates under multiple regimes and loading conditions, although applied to a gearbox for this application; could be extended to other applications with proper adjustment of the selected features and regime identification method.

The 2009 Prognostics and Health Management Data Challenge (PHM 2009 Data Challenge), focused on developing and applying techniques in the area of fault classification; data collected from a generic gearbox was used to facilitate the data sets for the challenge. A schematic of the gearbox used for data collection is shown in Figure 1, note that the measured signals consisted of two accelerometer signals along with a tachometer signal. The gearbox contains several mechanical elements, including three shafts, 4 gears, and 6 bearings; the overall objective of the data challenge was to specify the condition of each of the mechanical components and specify the particular fault if it was not in the healthy state. For example, a particular bearing could be in the healthy state, or have an inner race, outer race, or ball defect. For each data set, a 45 line diagnostic output was to be specified that detailed the condition of each mechanical component based on the available vibration and speed signals.

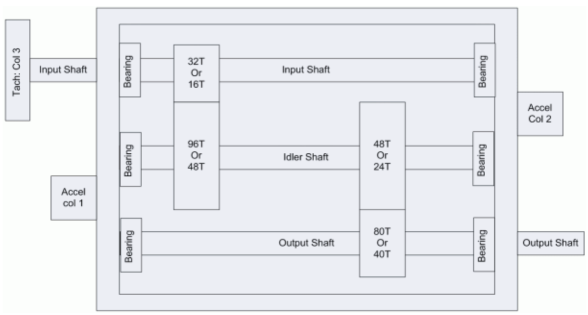

Figure 1: Schematic of Gearbox Used in PHM 2009 Challenge Data (PHM 2009 Data Challenge)

The data provided consisted of 560 data sets, in which the gearbox was tested under 5 different speeds, 2 different loads, and two different gear types. The 560 data sets consisted of data in which the gearbox was tested under different operating conditions as well as under different conditions of the mechanical components. No training data set was provided for the data challenge and fault detection was to be determined by analyzing the available vibration and speed signals and basing the diagnosis on the signature of the signal compared with the known fault signatures in the literature.

A picture of the experimental gearbox tested and used for data collection is shown in Figure 2. In this particular instance, the gearbox is shown with helical gears; however a set of spur gears was also tested and the spur gears consisted of twice the number of teeth of the helical gears.

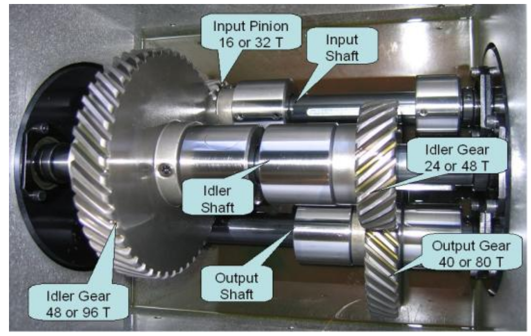

Figure 2: Inside View of Gearbox Used for Data Collection and Testing (PHM 2009 Data Challenge)

2. OVERALL METHODOLOGY TO GEARBOX FAULT DIAGNOSIS

The overall approach developed for gearbox fault diagnosis consists of several key steps in which the final output is specific diagnostic information for each mechanical component in the gearbox. For this particular application and data set, the inputs consisted of two vibration signals and a tachometer signal; the overall methodology shown in Figure 3 would require similar inputs since the fault diagnosis method is based on the vibration signals.

The initial key step requires the use of several signal processing and feature extraction methods to extract relevant information from the input vibration signals. By transforming the signal into frequency or time-frequency domains, relevant information that is correlated with particular gearbox faults can be extracted. This necessitates the use of several signal processing and feature extraction methods, since depending on the nature of the fault would indicate which signal processing method to use.

Regime segmentation allows for a fair comparison between the extracted feature sets; since the influence induced by operating the gearbox at different speeds or loads is reduced. By assessing the health and diagnosing the condition of each gearbox mechanical component for each regime; the influence of operating conditions is held constant and a change in particular features is only due to degradation of a particular component.

Health assessment consisted of using a specific set of features that are well correlated with overall gearbox health and using this feature set to assess the overall gearbox condition. For this particular application, a health assessment algorithm is used to find the data set with the minimum feature vector sum in each regime; and this is used to determine the baseline data sets.

Fault classification is the final step in the gearbox fault diagnosis methodology; the fault classification is triggered after health assessment, since only if the gearbox overall health has degraded does it necessitate further diagnosis to determine the particular problem. For each particular fault, features correlated with this failure signature are used to calculate the distance from each data set to the baseline data set in each regime, and this is used to calculate a probability of each fault based on the distance value from normal. The final diagnosis is dependent on inputs from the feature extraction step, regime segmentation, and health assessment calculation; this places much importance on the prior steps before the final diagnosis.

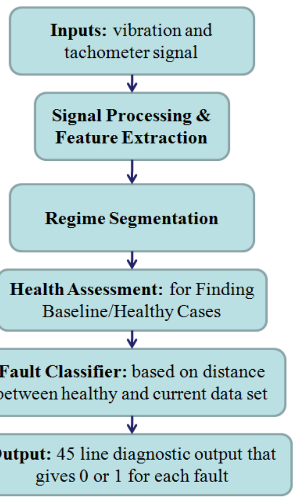

Figure 3: Flow Diagram for Gearbox Fault Diagnosis

3. SIGNAL PROCESSING AND FEATURE EXTRACTION

The health assessment and fault classification algorithms require the appropriate features as inputs in order to make the right assessment of the condition of the gearbox components as well as the level of damage. As described in the review by Samuel et al. (2005), this places much importance on extracting and selecting the most suitable set of features that are correlated to gearbox fault signatures as well as fault severity.

The signal processing and feature extraction methods used to extract a multitude of condition indicators from the vibration signals are presented; the health assessment and fault diagnosis section provide the details of the particular subset of features and algorithm used for assessing the gearbox health and providing the fault diagnosis information.

3.1 Time Domain Feature Extraction

Features from the raw time signal can be used to provide an overall understanding of the vibration level exhibited by the monitored gearbox as well as the distribution of the vibration data. The time domain feature values can be compared to a known baseline and this provides some level of assessment of gearbox condition but limited ability to diagnosis the particular fault. The root mean square value, defined for a sampled signal is given in Eq. (1) and provides an overall indicator of the vibration energy.

$$s_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} s_i^{\,2}} \qquad (1)$$

As mentioned by Decker et al. (2003), potential time domain features include the peak to peak vibration level, crest factor, and statistical measures such as kurtosis, among others; however for the development of this particular method, only the RMS value provided a useful feature from the raw vibration time signal.

The RMS value can only provide insight that a particular fault is occurring but insight on the exact gearbox component that has damage or what failure mode is occurring cannot be inferred only using this indicator. The RMS feature can be used in an overall health assessment algorithm along with other potential features; however it has limited used for providing specific diagnosis for mechanical systems comprised of several components such as a gearbox.

An example of the level of insight the RMS feature can provide is shown in Figure 4, in which the RMS value is 2.6 times higher for the output vibration signal for a gearbox with a gear in the idler shaft having a broken tooth compared to a healthy gearbox operating under the same load and speed settings.

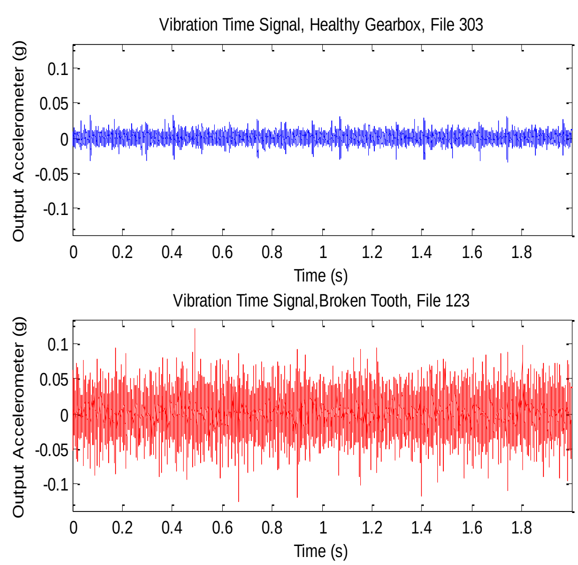

Figure 4: Vibration Time Signal for Healthy Gearbox and Gearbox with a Broken Gear Tooth

3.2 Time Synchronous Average Time Signal

By processing the tachometer signal, the vibration signal can be segmented into blocks for the duration of one revolution of the input, output or idler shaft, and averaging the signal for each block of data can highlight certain phenomena that are synchronous with the shaft rotation. Keller et al. (2003) mentioned several signal processing techniques for helicopter gearbox vibration analysis including further processing of the time synchronous signal by taking the Fast Fourier Transform and extracting peak information from particular orders of interest. However, there is still relevant information that can be extracted from the synchronous time signal including gear tooth problems related to impacts.

Impacts that occur repetitively for each shaft rotating could be an indication of a particular fault; for example a broken tooth would generate an impulsive impact once per revolution of its respective shaft and this can be more easily detected by analyzing the time synchronous average signal. As mentioned by Choy et al. (2004), the frequency spectrum obtained by performing the FFT on the time synchronous signal might not provide insight into the impact caused by this particular fault and the time or time-frequency domain is more appropriate way to analyze this particular signal. This is due to a pure impulse in the time domain containing broadband energy in the frequency domain. An example time synchronous signal is shown in Figure 5 for a data set that contains a broken gear tooth on an idler shaft gear, this particular fault has a clear signature due to the a periodic impact occurring.

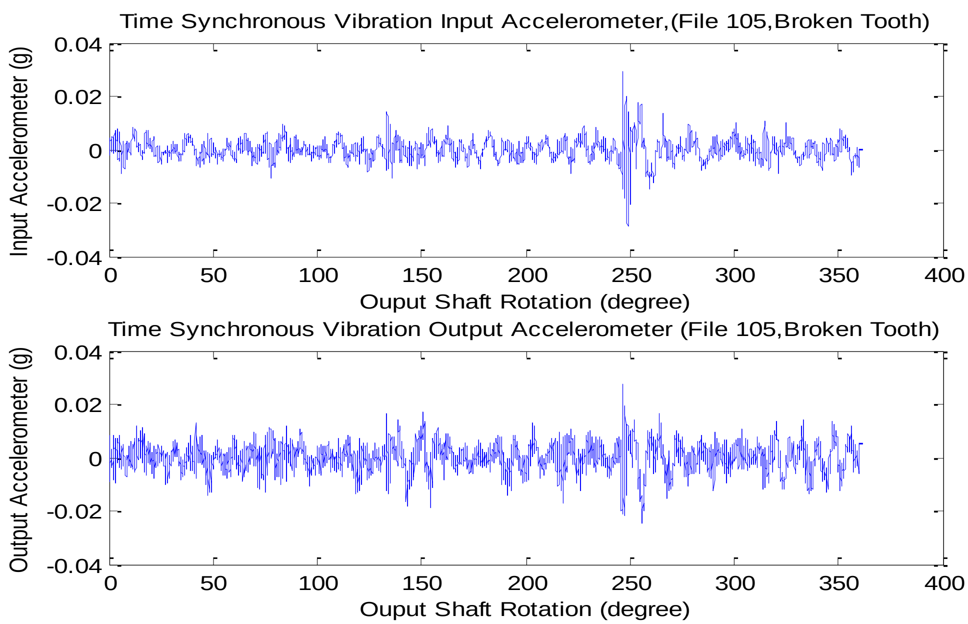

Figure 5: Time Synchronous Vibration Time Signal for Gearbox with a Gear that has a Broken Tooth

Both the peak to peak vibration and the energy operator feature can be used to process the synchronous time signal in an automated feature extraction routine to quantify this impact. The energy operator (EO) indicator used by Ma (1995), is a normalized kurtosis value defined by Eq. (2), where N is the number of data points in a sampled signal s.

$$EO = \frac{N^2 \sum_{i=1}^{N} (\Delta x_i - \Delta \bar{x})^4}{\left(\sum_{i=1}^{N} (\Delta x_i - \Delta \bar{x})^2 \right)^2} \qquad (2)$$

where: $\Delta x_i = s_{i+1}{}^2 - s_i{}^2$

For the particular example shown in Figure 5, this particular gearbox with a broken gear tooth had an energy operator value that is 3 times higher than the baseline data set case for the output accelerometer signal. The energy operator in the time synchronous average signal is able to characterize this type of impact due to a gear tooth breakage. The impact could also be associated with other gearbox faults related to bearings, bent shaft, or bad key; however quantifying the impact is important and more than one feature can be used to isolate a particular fault.

3.3 Synchronous Average Vibration Spectrum

As discussed by Grabill et al. (2001), the Fourier Transform of the time synchronous averaged signal can reveal periodic occurrences that are related to a particular shaft of interest; information related to shaft problems such as imbalance, as well as gear problems related to sidebands can be analyzed using the frequency domain spectrum. Figure 6 shows the frequency domain spectrum for a healthy gearbox and a gearbox with input shaft imbalance.

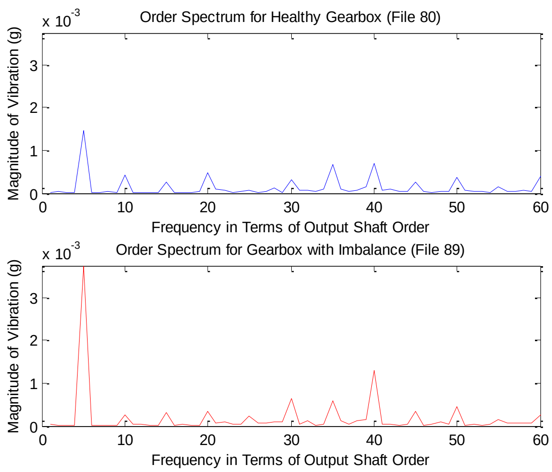

Figure 6: Frequency Spectrum for Healthy Gearbox and Gearbox with Imbalance Input Shaft

The frequency spectrum plot shows a much higher peak in the spectrum at 5X, which is the speed of the input shaft. For this particular example, the vibration from the input accelerometer at 5X is 2.5 times greater for the gearbox with imbalance in the input shaft compared to the healthy gearbox.

3.4 Continuous Wavelet Transform

For visual understanding of the impulse in the time synchronous signal, time-frequency method such as the continuous wavelet transform can reveal the broken gear tooth impact as described by Zheng et al. (2002). In this particular example, the continuous wavelet transform with a mother wavelet of Daubechies order 8 is used to process the time synchronous average signal and the impulse can be clearly seen in Figure 7. For faults that cause impacts or are transient in nature, the use of time-frequency signal processing methods such as the continuous wavelet transform can be used to provide a visual understanding of the particular fault that is occurring.

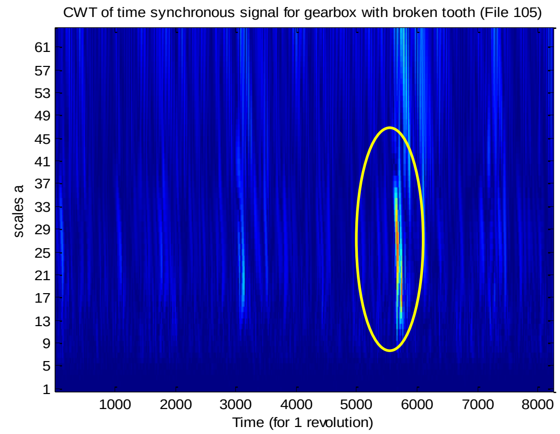

Figure 7: Continuous Wavelet Transform for Gearbox containing a gear with a broken tooth

3.5 Discrete Wavelet Transform

The vibration exhibited by a gearbox with a multitude of frequency components requires the use of advanced signal processing techniques to decompose the signal to isolate particular fault signatures more easily. The use of wavelet decomposition described by Peng et al (2003) is well suited for this particular task, in that the high frequency aspects of the signals denoted as the details can be isolated from the low frequency components. The gearbox vibration signal for a bent shaft shown in Figure 8 is an example in which the use of the wavelet decomposition technique provided a more robust feature set for fault diagnosis.

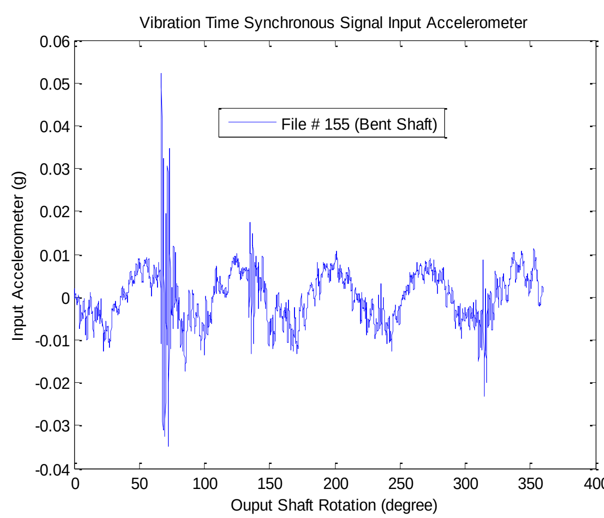

Figure 8: Vibration Time Synchronous Average Time Signal for Gearbox with Bent Input Shaft

The time synchronous vibration signal for a bent shaft condition shows an impulse impact along with a lower frequency harmonic component. The impact can be quantified by features such as kurtosis from the time synchronous signal; however the harmonic component is of particular interest since it is occurring at a frequency of 5 times the output shaft rotating speed which is the input shaft speed.

A wavelet decomposition of level 5 with a mother wavelet of Daubechies order 8 is used to further decompose the signal; the result in Figure 9 shows the approximation level 5 signal and the original signal after removing the approximation signal. The decomposition technique can be used to further analyze the harmonic component or the impact.

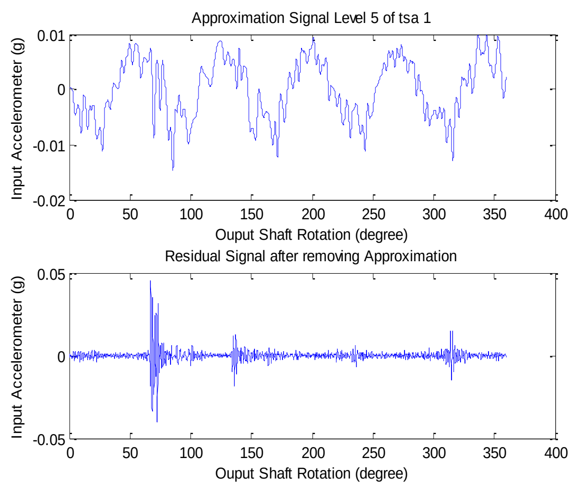

Figure 9: Wavelet Decomposition Signal for Bent Shaft Case

Further processing of the harmonic signal by taking the Fast Fourier Transform shown in Figure 10 can also be used to extract additional information.

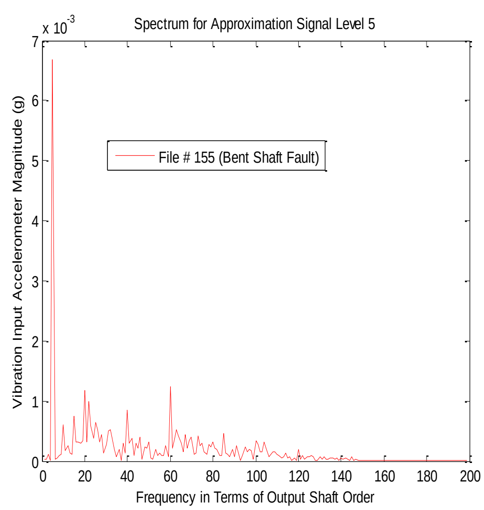

Figure 10: Frequency Spectrum of Harmonic Component after using wavelet decomposition

The frequency spectrum prior to decomposition would be difficult to interpret since the impulse would spread broadband energy; the frequency spectrum of the approximation signal provides much more relevant information related to the harmonic component associated with the input shaft speed. The peak at 5X divided by the next largest peak in the approximation signal spectrum was one of the features used to characterize the input bent shaft fault.

3.6 Spectral Kurtosis

Applying the kurtosis statistical measure for the raw vibration time signal is not necessarily suitable for detecting incipient damage and the use of a more localized way of capturing the transient nature of impulses generated from gears or bearings with early stages of damage are needed. The use of the spectral kurtosis by Antoni et al. (2006) has shown to be suitable solution for characterizing the transient impulsive type faults that occur for mechanical components; the spectral kurtosis technique has shown to be effectively used as a machine surveillance indicator as well as a way to select an optimum band pass filter for mechanical fault detection.

The overall procedure for computing the spectral kurtosis described by Antoni (2006) consists of taking the Short Time Fourier Transform (STFT) for a given block-size and computing the kurtosis statistical calculation across each spectral line; this provides a kurtosis value as a function of the frequency. The use of spectral kurtosis for the purpose of the gearbox health assessment and diagnostic method was to use the features from the spectral kurtosis calculation to characterize the overall health status of the gearbox.

A plot of the spectral kurtosis value as a function of frequency is shown in Figure 11 for a healthy gearbox and a gearbox with a broken tooth on an idler shaft gear. There is a noticeable difference in the kurtosis value at higher frequency and in particular for the output accelerometer in a frequency range of 10-20 KHz, the kurtosis value is much larger for the damaged gearbox.

For both the input and output vibration signal, three frequency bands (below 10 KHz, from 10 KHz-20 KHz and above 20 KHz) were used to calculate the sum of the kurtosis value in each frequency band and these 6 features were potential features that could be selected to determine the overall gearbox health.

In this particular example shown in Figure 11, the spectral kurtosis sum feature from 10 KHz-20 KHz for the output accelerometer was 9.5 times greater than the same feature for the healthy gearbox; this reaffirms the utility of the spectral kurtosis feature extraction method for gearbox healthy assessment. The spectral kurtosis band features provide an additional indicator for determining the overall health state of the gearbox but does not provide detail information on the particular fault that is occurring; further diagnosis requires additional features to isolate a particular gearbox fault.

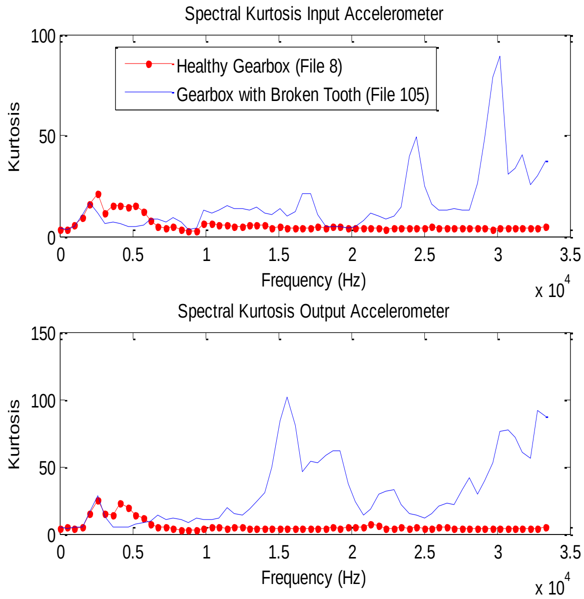

Figure 11: Spectral Kurtosis Plot for Healthy Gearbox and a Gearbox with Broken Gear Tooth

3.7 Envelope Spectrum for Bearing Fault Frequency Peak Information

The high frequency envelope frequency extraction method is a well established method for providing fault information for a particular bearing of interest. The overall signal processing procedure consists of band-pass filtering around an excited natural frequency, using the Hilbert Transform to calculate the envelope, and taking the Fourier Transform of the analytical signal. As described by Tse et al. (2001), the impacts caused by a bearing defect excite a few high frequency modes of the system and the bearing fault frequencies are amplitude modulated; by performing the band pass filtering and demodulating the signal, this characteristic fault information at the peaks can be extracted. For this particular application, a Chebyshev band pass filter was centered at 8950Hz with an upper frequency limit set at 9250Hz and a lower frequency limit at 8650Hz. For each particular bearing fault such as an inner race or outer race defect, there is a particular peak in the frequency domain that is representative of damage for this particular bearing failure mode. As mentioned in Li et al. (2000), from the specified bearing geometry, there is a set of equations that relate the bearing fault frequencies with the number of rolling elements, pitch diameter, contact angle, and diameter of the rolling elements. Figure 12 shows the envelope spectrum for the output accelerometer for both a healthy gearbox and a gearbox with an input shaft (output side) inner race bearing defect. The peak at 245Hz, which corresponds to the ball pass frequency inner race (BPFI), is clearly much higher for the gearbox with this particular inner race problem. In a similar manner the peaks at the other bearing fault frequencies can be extracted from the envelope spectrum and this subset of features can be used as inputs to classify the condition of the gearbox bearings.

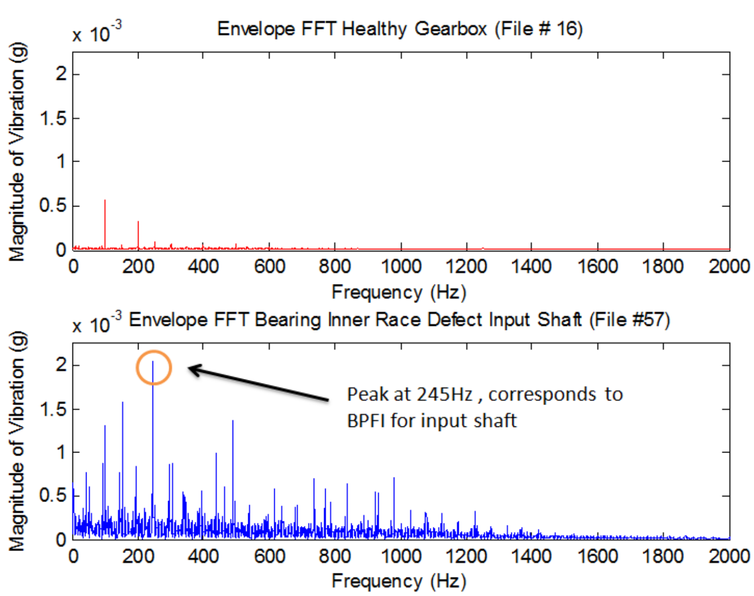

Figure 12: Envelope Spectrum for Healthy Gearbox and Gearbox with Inner Race Bearing Defect

3.8 Features Specifically for Gear Fault Diagnosis

Particular features have been specifically developed for monitoring the health of gearbox gears, and these specific gear condition indicators were extracted and considered for use in the health assessment and fault diagnosis algorithm. A more detailed description of the FM4, FMO, NB4, sideband level and index indicators are provided by Vecer et al. (2005), and a quick review of the potential use of each gear condition indicators is presented since the use of these indicators were incorporated into the gearbox health assessment and classification algorithm. Higher sidebands around the gear-mesh frequency can indicate wear or manufacturing error such as eccentricity, the sideband index is an average value of the sidebands for a particular gear mesh frequency and this feature was taken for each gear. The sideband level is a similar feature and is a ratio between the sum of the sidebands around a particular gear mesh frequency divided by the standard deviation in the time average synchronous signal.

The FM4 is taken from the residual signal, in which the gear-mesh harmonics and shaft harmonics are removed and the kurtosis is calculated for the residual signal; if one tooth is detective or damaged this feature should be greater than normal. The NB4 feature is also used to characterize gear damage and consist of band pass filtering around a particular gear mesh frequency and taking the envelope of the signal using the Hilbert Transform and calculating the kurtosis for the analytical signal. The FMO feature, also known as the zero order figure of merit, is calculated by taking the ratio between the peak to peak vibration levels for the time synchronous average signal divided by the sum of the gear mesh harmonics. These particular features were potential features used in the health assessment and fault classification method.

4. REGIME SEGMENTATION

In order to assess the gearbox condition and diagnosis particular faults, it is necessary to minimize the effect of operating variables. This allows for a fair comparison, because the operating effects would influence the features extracted from the vibration signal and higher feature values might only be due to loading or speed effects and not due to degradation in a particular gear, bearing or shaft component. The overall regime segmentation method is shown in Figure 13 and is used to segment each data set by gear type, load, and speed. The result of the regime segmentation procedure is 20 clusters, where each data set is segmented by load, speed, and gear type.

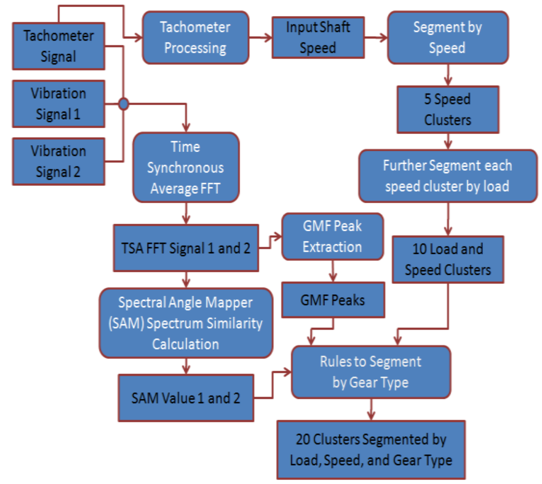

Figure 13: Regime Segmentation Flow Chart

4.1 Segment by Speed and Load

The data sets provided from the gearbox test-rigs were collected during different operating regime settings, including different input shaft speeds as well as two levels of applied braking torque load on the output shaft. By processing the tachometer square wave signal, the input shaft speed can be determined; for this particular experimental testing it is clearly observed that the input shaft speed was tested at 5 different speeds. A light or heavy braking load was applied to the output shaft, depending on the loading case. Further analysis of the speed signal in each regime showed two clusters, the cluster that had the slightly lower speed was due to a greater braking torque load being applied. Figure 14 shows the input shaft speed for the 45Hz input shaft speed cluster, the higher braking torque load causes a slight reduction in input shaft speed and allows for segmenting each data set into a high or light load regime.

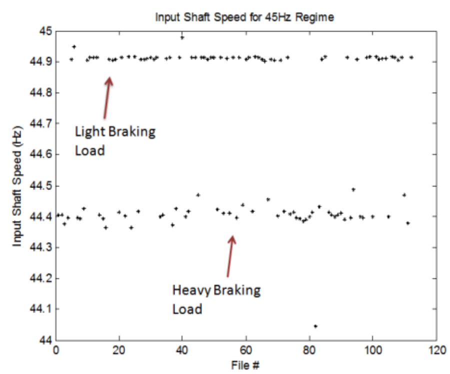

Figure 14: Example of Segmenting by Load

4.2 Similarity Spectrum Measure for Gear Type Segmentation

The gear mesh frequency is based on the shaft speed and the number of teeth on a particular gear; the helical gears used in this gearbox had a gear mesh frequency at 40 and 80 times the output shaft speed while the spur gears which had twice as many teeth had a gear mesh frequency at 80 and 160 times the output shaft speed. Using only information at the gear mesh frequency peaks at orders of 40, 80, and 160, was not enough information to segment the data sets by gear type. This is due to harmonics of the gear mesh frequency would coincide with the gear mesh frequency of the other gear type; additional information is necessary to segment the data set by gear type.

A similarity measured defined as the Spectral Angle Mapper has been used by Sheeley et al (2009) for current spectrum signals, and is incorporated into this gear type segmentation task. The time synchronous average spectrum between two data sets denoted by s_i and s_k is compared using the similarity measure shown in Eq. (3). This calculation is essentially a dot product calculation and is a value between 0 and 1, with 1 indicating a pair of spectrum that is closer in similarity.

$$\mathrm{SAM}(s_i, s_k) = \left(\frac{\mathrm{dot}(s_i, s_k)}{|s_i||s_k|} \right) \tag{3}$$

The use of the similarity measure for segmenting by gear type is outlined by the procedure listed below:

1. Find the data set in each of the 10 speed and load clusters that have the maximum value of the sum of the vibration spectrum peak at orders of 40 and 120 times the output shaft speed from the output accelerometer.
2. Calculate the similarity measure between this data set and all the other data sets in that regime, and take the 24 data sets that have the most similar spectrum as the helical gear type.
3. The other remaining data sets in that regime are given the label as spur gears.

By utilizing the information that corresponds to the gear mesh frequency peak and 3 times the gear mesh frequency peak for a helical gear mesh pair for the idler and output gear mesh pair, the data set with this maximum sum in each speed and load cluster is found. A data set in a particular regime that has a very similar spectrum would imply that it also contains helical gears. An example plot is shown in Figure 15 (a), in which 24 helical gears are clearly separable using this method.

This method for segmenting files by gear type was validated for a new labeled dataset published by PHM society for the same gearbox test-rig (PHM public Datasets, 2009). The labeled dataset consisted of 14 cases, in which 8 were from spur gear and 6 from helical gear. Each labeled case, was run in 5 different speeds (30Hz, 35 Hz, 40Hz, 45 Hz, and 50Hz) under two loads (high and low) and 2 replications. Figure 15 (b) shows the segmenting by gear type for the labeled data in one regime; however the method showed 100% classification for all other regimes.

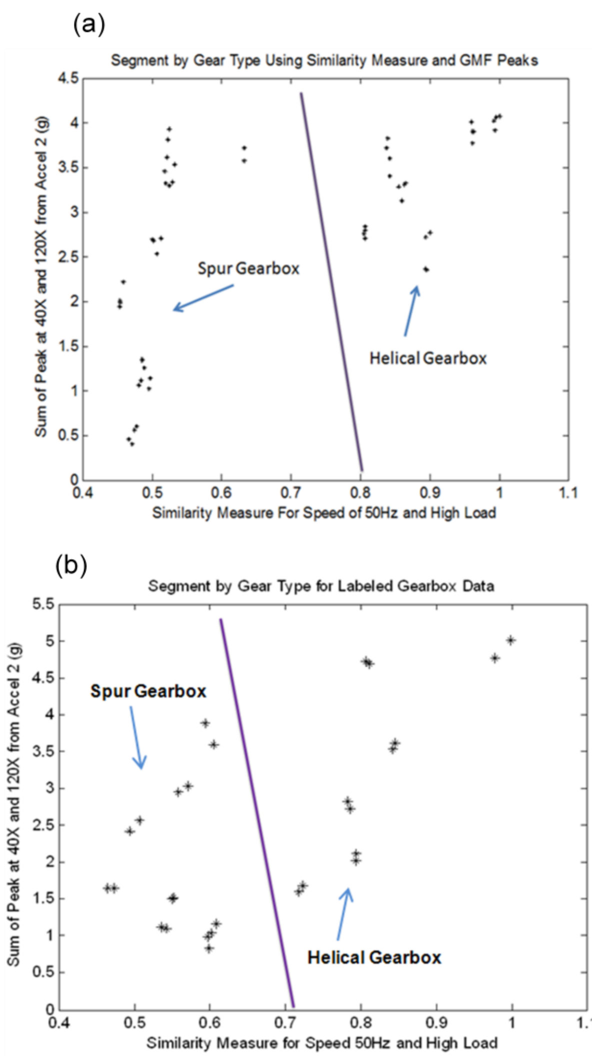

Figure 15: Example of Segmenting by Gear Type for 50Hz and high load regime a) for PHM data challenge 2009 dataset (PHM 2009 Data Challenge) b) validated result for labeled dataset (PHM Public Datasets, 2009)

5. GEARBOX HEALTH ASSESSMENT

An overall system health method or anomaly detection routine is used to provide an initial measure of diagnosis; it determines whether the system is in the normal state but not information on what fault is occurring. Although knowledge of the exact fault that is occurring is useful for reducing the maintenance cost related from logistics of ordering spare parts as well as reducing the labor time to determine the problem; only if the system health is degrading does it make sense to trigger a fault diagnosis classifier. In this particular instance, an overall system health calculation is done to determine the data sets in which the gearbox is in the baseline healthy state regarding all of its components; this baseline data set in each regime is later used by the fault diagnosis calculation.

5.1 Feature Set for Overall Gearbox Health

A list of the feature set is provided below and includes features related to overall vibration level, gear sideband information, features that are correlated to impacts from broken tooth, bearing fault frequency features, and features related to peaks due to shaft imbalance or other shaft problems.

1. RMS Value from raw time signal (input and output accelerometer).
2. Peak to Peak Level from time synchronous signal (input and output accelerometer).
3. Energy Operator from time synchronous average signal (input and output accelerometer).
4. Peak at 25X from input accelerometer from time synchronous average FFT.
5. Peak at 5X from input and output accelerometer from time synchronous average FFT.
6. Mean, max and sum of a set of features related to peaks corresponding to sidebands around gear mesh frequency.
7. Mean and sum of a set of the set of features related to sideband index and sideband level.
8. Mean of spectral kurtosis features.
9. Mean of a set of features related to bearing fault frequency peaks.
10. Max of a set of features related to peaks for shaft related problems (10X, 15X, 20X).

5.2 Health Assessment Calculation

For each of the 20 regimes segmented by gear type, load and input shaft speed, the following health assessment procedure was used to determine the overall gearbox health state as well as find the baseline data sets.

1. Normalize the feature set for health assessment so each feature has the same weight.
2. Calculate the sum squared of the feature vector for each file and store this value as the health value for this data set.
3. In each regime, find the file that has the minimum health value, this gearbox data set would be in the best health state since it has low vibration level, and features related to gear and bearing and shaft problems are all low values.
4. After finding the gearbox data set with the minimum health value in each regime, use the similarity measure to find the other 3 data sets that are most similar to this healthy one.
5. For each operating regime, there are 4 replications, so the similarity measure is just used to ensure that the other 3 data sets that are also healthy in each regime are included.

Overall, this method was able to find all 80 data sets for healthy gearbox; the baseline data sets were later used for designing the fault classification algorithm.

Baseline Data files	**Speed**	**Load**	**Gear Type**
14 , 428, 431, 490	30 Hz	Light	Spur
84 , 379 , 391 , 439	30 Hz	Light	Helical
175 , 190, 369 , 446	30 Hz	Heavy	Spur
72 , 287, 497, 531	30 Hz	Heavy	Helical
101 , 165, 412, 524	35 Hz	Light	Spur
70 , 380 , 420 , 548	35 Hz	Light	Helical
184 , 265 , 355, 543	35 Hz	Heavy	Spur
108 , 238 , 436 , 463	35 Hz	Heavy	Helical
44 , 209 , 322 , 441	40 Hz	Light	Spur
8 , 77, 182, 252	40 Hz	Light	Helical
42 , 95 , 113, 469	40 Hz	Heavy	Spur
233 , 297 , 320 , 444	40 Hz	Heavy	Helical
181, 222, 356, 462	45 Hz	Light	Spur
303, 504, 505, 519	45 Hz	Light	Helical
116 , 212, 350, 425	45 Hz	Heavy	Spur
29 , 128 , 186 , 452	45 Hz	Heavy	Helical
4 , 172, 324 , 347	50 Hz	Light	Spur
16 , 94, 194, 460	50 Hz	Light	Helical
80, 193, 404, 555	50 Hz	Heavy	Spur
376, 481, 526 , 536	50 Hz	Heavy	Helical

Table 1: Identified baseline data files in each regime

6. GEARBOX FAULT DIAGNOSIS

6.1 Fault Diagnosis Overall Method

The fault diagnosis process is responsible for identifying the location and kind of defects in each data set. There are different techniques that can be used for this purpose; with each technique having its own merits and drawbacks. Occam's razor principle indicates that "the simplest method" to model the problem should be preferred. In this case, the simplest diagnosis method is the rule-based diagnosis where each defect is diagnosed based on the values of some features exceeding specified thresholds. The drawback of rule-based is that the specification and selection of the thresholds will be very problem specific and cannot be generalized for other problems, especially since this particular application is dealing with physical defects. Although the vibration signature could usually indicate degradation and defects, it is quite difficult to specify thresholds for specific defects that can be generalized with a high level of certainty.

Given the limitations of the rule-based diagnosis model in finding generalized thresholds or rules, a more general approach should be pursued. The approach should be capable of providing a more general method that provides the desired level of accuracy regarding fault classification.

The overall proposed approach for gearbox diagnosis is a systematic regime-and-similarity-based and is summarized in Figure 16. The diagnosis will be performed in each regime separately and will be based on calculating the probability of defect for each of the defects (33 total defects) from the distance to the regime's baseline. For each kind of defect a specific set of features are selected (based on experience and established literature for specific faults); the selected features should be capable of distinguishing the fault; the feature set will be normalized (0 to 1 scale); the distance of the feature set of each data file to the baseline in each regime is calculated; the Probability of defect P(d) is calculated for each data file for each of the 33 defects; and finally the probability value will be used to determine the existence of each defect in the data set.

The proposed approach was used to detect 5 different defects: Broken tooth in idler shaft gear output side (Gear 3); input shaft imbalance; input shaft bent; output shaft bad key; and inner race defect for bearing on input shaft output side. For each defect, different features were selected based on experience and the established literature in which certain features are known to be correlated to defects or degradation of specific mechanical components. Other faults in other components can also be detected using the same methodology; only different features should be selected based on the particular fault signature.

For the specific application of diagnosis using vibration data, most defect features are expected to increase from normal level if the defect exists. But, if the defect does not exist, the specific features are expected to be either equal or within the variation of the normal set. Since there is not enough data to establish the variations of the normal set, any feature value below the normal feature value will be considered within the normal region of operation, and any value greater than the normal feature shall be considered outside the normal region and might be an indicator of a defect based on how far away it is from the normal region.

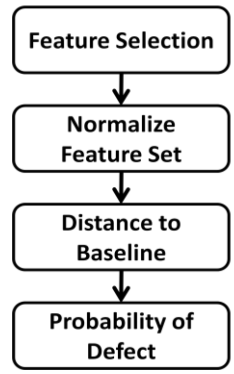

Figure 16: Overall diagnosis approach

The distance function used is a modified Euclidean distance function as shown in Eq. (4-5) as follows:

$$D(X,Y) = \left\{ \sum_{j=1}^{n} \left[\left(X_j - Y_j \right)^2 U\left(X_j - Y_j \right) \right] \right\}^{\frac{1}{2}} \tag{4}$$

$$\mathrm{U}(\mathrm{X}_j - \mathrm{Y}_j) = \begin{cases} 0 & \text{if } \mathrm{X}_j < \mathrm{Y}_j \\ 1 & \text{if } \mathrm{X}_j \geq \mathrm{Y}_j \end{cases} \tag{5}$$

Where D(X,Y) is the distance between the two feature vectors X and Y each of size n, and U(X-Y) is the unit step function as defined by Eq. (5). The modified distance function will be used to find the distance of each feature vector in a regime to the feature vector of the normal set. But, some data sets might have some features lower than the normal set, which can be explained by the variations in the normal regime operation. These data sets are within the normal variation for the specific feature. The variance in a particular vibration feature can be due to several factors related to sensor noise and for this application the noise on both accelerometers and the tachometer signal would have an influence on the feature values. The probability of defect calculation, by using a distance measure that combines information from multiple features and sources of information as well as not considering features that have values less than the baseline value accounts for some of the uncertainty in the feature values due to sensor noise.

The traditional Euclidean distance measure uses the difference between the attributes squared, and thus if a feature is less than normal its distance is still positive and will be considered an increased value with respect to normal; and hence would not be differentiated from a feature that is higher than normal with the same difference. The modified Euclidean distance will consider any feature value below the normal value to be within the normal region and the distance of the specific feature will be zero. Thus the modified distance function D(X,Y) ,presented in Eq. (4), can be considered as the distance from feature vector X to the region bounded by Y rather than the distance from feature vector X to feature vector Y. Giving a weight of zero to features that are below normal or a set threshold in a distance calculation shares some similarities to the non-linear mapping method described by Bechhoefer et al. (2003), used in the vibration based health indicator calculation for the helicopter health and usage monitoring system.

The probability of defect $P_f(d)$ for a given data set f is calculated as shown in Eq. (6).

$$P_f(d) = 1 - \exp^{-D\left(X_f, N_f\right)} \tag{6}$$

Where $P_f(d)$ is the probability of the data set having a defect type d (where d=1,…33); X_f is the normalized feature vector for data set f; N_f is the normalized feature vector for the normal baseline in the same regime of data set f; and $D(X_f,N_f)$ is the modified Euclidean distance function described in Eq. (4). Since in each of the 20 regimes (shown in table 1) there are four normal data sets, then N_f is the mean of the normalized feature vectors of the four normal data sets in that regime. Note that $P_f(d)$ would be a value between zero and one, indicating the probability of each data file of having a specific defect, based on its distance from the normal baseline within its regime. The diagnosis process for each defect type will be based on selecting the appropriate feature vector that is able to isolate the defect from other possible mechanical defects.

6.2 Gear Tooth Breakage Signature

The proposed method was used to diagnose broken tooth problem in the idler shaft gear that was meshing with the output gear. The expected signature of a gearbox with a broken tooth is an overall higher level of vibration energy, impact occurring once per revolution of the broken tooth in the time signal; higher sidebands around the Gear Mesh Frequency (GMF); and natural frequency excitation due to the impact from the broken tooth. Due to the complexity of the given problem and the overlap of different defect signatures, these signatures could also be an indicator of other problems. A subset of features is needed that can isolate this specific defect from other defects that could have similar signatures.

The selected features for this case are provided below:

1. RMS of envelope signal of the input accelerometer.
2. RMS of envelope signal of the output accelerometer.
3. Ratio between RMS of raw signal from output accelerometer to input accelerometer.
4. Ratio between sidebands: (sum of sidebands around gear 3) divided by (sum of sidebands around gear 4) from output accelerometer.

The first two features help detect the overall vibration excitation around the natural frequency; this is quantifying the natural frequency excitation due to the impact defect. Although a bearing defect will also excite high frequency modes of the system; the energy in the envelope spectrum would be at a few peaks and not spread across the entire spectrum for a bearing defect. The third feature indicates that the vibration on the output side is generally higher than the input side (pointing towards either Gear 3 or Gear 4 problems rather than Gear 1 or Gear 2). The fourth feature indicates that the sidebands of Gear 3 are higher than those of Gear 4 and points towards Gear 3 as the probable cause. The combination of these four features indicates the signature of a gear with broken tooth problem; on the output side; and most likely due to a Gear 3 problem.

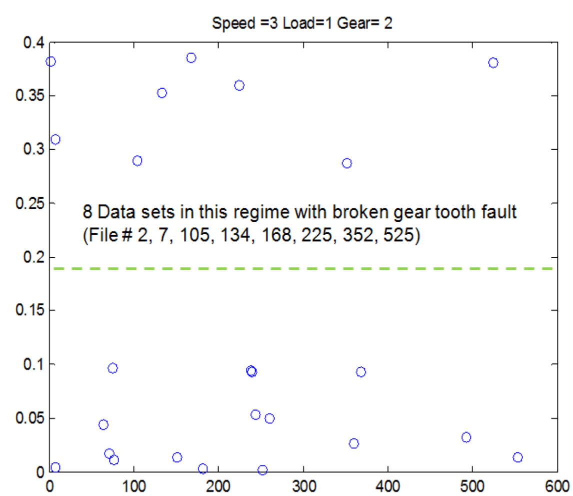

Figure 17: Probability of Gear 3 broken tooth defect for 40Hz speed, light load, and helical gear regime

After calculating the probability of defect for each data set in its regime using the four features described above, it was clear that in each regime 8 data files could be separated from the other data files. Figure 17, shows an example from a regime (speed= 40 Hz, Load=Light, Gear= Helical) where the probability of defect was calculated, and it is clear that eight of these data files can be isolated from the other files because they have a higher probability of defect for a broken tooth in Gear 3. The same method could be used for detecting a broken tooth problem in the other gears or for diagnosing other gear problems, contingent upon selecting the appropriate feature set.

6.3 Shaft Imbalance Fault Detection

The proposed method was also used to diagnose impact shaft imbalance. A mass imbalance for a particular shaft would have a higher peak corresponding to 1X for the particular imbalance shaft and this particular indicator is commonly used to diagnosis shaft imbalance problems for rotating machinery. But for this specific problem, the peak at 1xrpm also overlaps with the peaks at the characteristic bearing fault frequencies; in particular the BPFO of the bearings on the idler shaft (1.0174xrpm of input shaft) and the BPFI of the bearings on the output shaft (0.9895xrpm of input shaft). So, a high amplitude peak at 1xrpm of input shaft could be caused by either of these problems. The selected features are as follows:

1. Peak in time synchronous average spectrum from input accelerometer at input shaft speed.
2. Peak in time synchronous average spectrum from output accelerometer at input shaft speed.
3. Peaks in envelope spectrum at input shaft speed from both input and output accelerometer.

The first two features are from the time synchronous average and should filter out non-synchronous multiples of 1xrpm of input shaft. The third feature indicates amplitude modulation caused by the imbalance. This provides a feature set that provides indication of an input shaft imbalance but also indicators that are not influenced from other potential mechanical defects that have peaks in a similar frequency range.

After calculating the probability of defect for each data file in its regime using the three features described, it was clear that in each regime, 4 data sets could be separated from the other data sets. Figure 18, shows an example from a regime (speed= 50 Hz, Load=Heavy, Gear= Spur) where the probability of defect was calculated, and it is clear that four of these data files can be isolated from the other files because they have a higher probability of defect for input shaft imbalance.

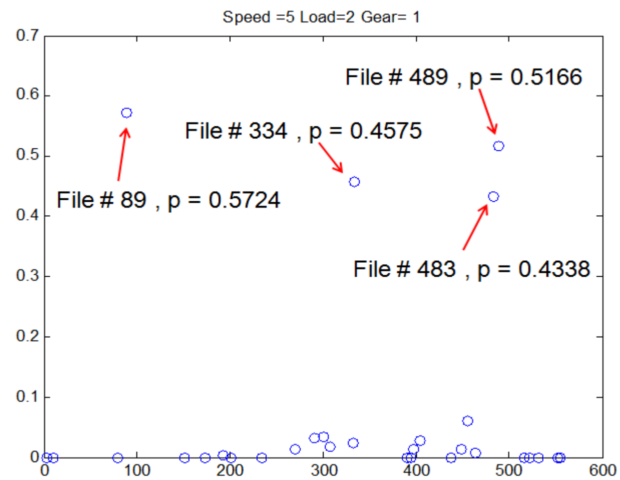

Figure 18: Probability of input shaft imbalance defect for regime of 50Hz speed, heavy load, and spur gears.

6.4 Bent Shaft Diagnosis

A bent shaft signature is dependent on the location of the bent, i.e. bent before the bearings, bent on the bearings, bent on the gears, or bent on the couplings. Each one of the aforementioned cases has a distinct signature. A common signature characteristic of all these potential bent shaft faults is the high amplitude at the 1X harmonic of the shaft speed. Specific additional characteristics would apply to the other cases, such as a bent on the coupling which would have an impact in the time domain signal which was the case for this data.

For detecting a bent shaft, the following features were selected:

1. Kurtosis of the time synchronous average of the output signal vibration.
2. Ratio between peak at 1xrpm of input shaft to next highest peak after wavelet decomposition and FFT of approximation signal.

The wavelet decomposition was used to isolate both the impact and the 1X harmonic component of the input shaft (5X of output shaft); including indicators that quantify the impact and the harmonic component is necessary for the bent shaft diagnosis.

Figure 19, provides an example from a regime (speed= 50 Hz, Load=Light, Gear=Helical) where the probability of defect was calculated, and it is clear that four of these data files can be isolated from the other files because they have a higher probability of defect for a bent input shaft.

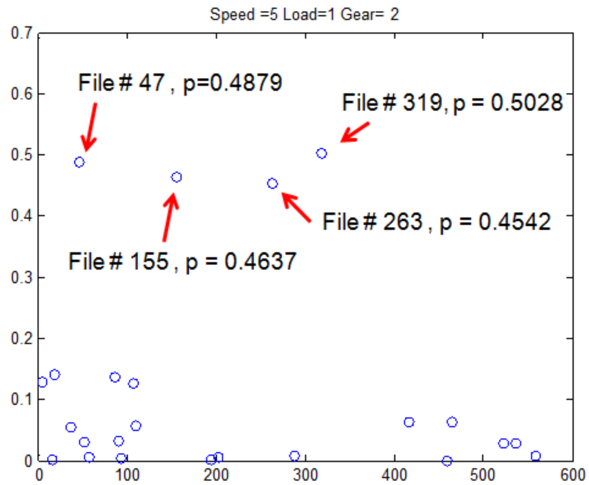

Figure 19: Probability of input bent shaft defect for regime of 50Hz speed, light load, and helical gears.

6.5 Bearing Fault Detection

For detecting the inner race defect in the bearing on the input shaft output side, the feature selected was the ratio of feature 1 below divided by feature two:

1. Peak at Ball Pass Frequency Inner Race (BPFI) from the Envelope spectrum of the output accelerometer, input shaft bearing.
2. Peak in time synchronous average spectrum at input shaft speed from input and output accelerometer.

Figure 20, shows an example from a regime (speed= 50 Hz, Load=Light, Gear=Helical) where the probability of defect was calculated, it is clear that four of these data files can be isolated from the other files because they have a higher probability of inner race defect for the bearing on the input shaft/output side.

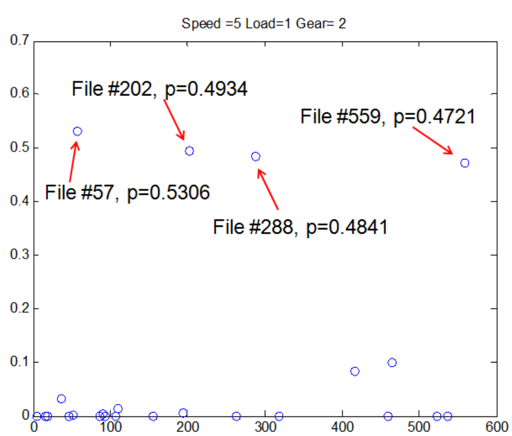

Figure 20: Probability of inner race bearing defect on the input shaft/output side bearing calculated for regime of 50Hz speed, light load, and helical gears.

6.6 Detecting Bad Key Fault

For this gearbox system, the output shaft could be in two different states, either the health state or a defect state; where the defect state is due to a loose coupling between the output shaft and the load because of a faulty key. This faulty shaft key would cause slipping between the output shaft and the load; the vibration signals and the tachometer information can be used to isolate this particular event.

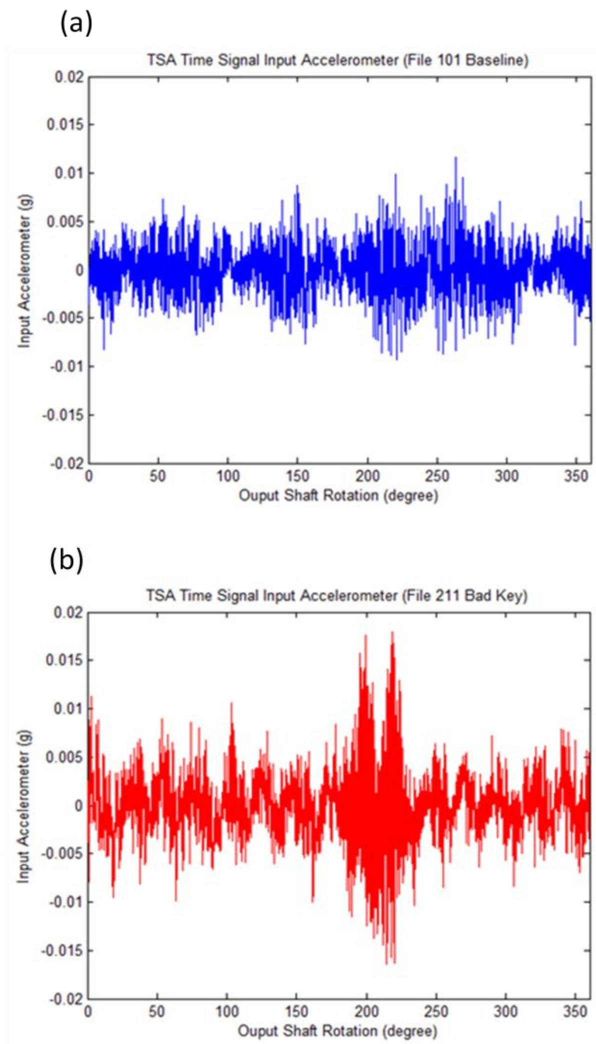

Figure 21: Signature of bad key defect.

Figure 21 (b) shows the time synchronous average signal of a data set with a bad key. It can be noticed from the signal of an impact that is of much longer duration; this is due to the averaging of the time synchronous signal and the slipping of the shaft speed due to the fault key. Notice that this signature of a bad key is in sharp contrast to the time synchronous signal shown in Figure 21 (a) for a gearbox without any faults including a key that is working properly. Also if one would compare the signature of a broken tooth compared to a bad key by examining the time synchronous average time signal; although in both cases there is this transient impact, the bad key has a specific impact that last for a much longer duration due to the slipping of the output shaft. Utilizing a specific set of features that can isolate the bad key from not only the baseline case, but also faults that also have transient impacts is what is needed for providing the specific root cause diagnosis information.

For detecting the bad key defect on the output shaft, the following features were selected:

1. Kurtosis from time synchronous average time signal from output accelerometer.
2. Spectral Kurtosis feature from output accelerometer, for band from 10 KHz to 20 KHz.

Figure 22 shows the probability of defect results for the bad key case; the result show 4 data sets in a particular regime that clearly have this problem.

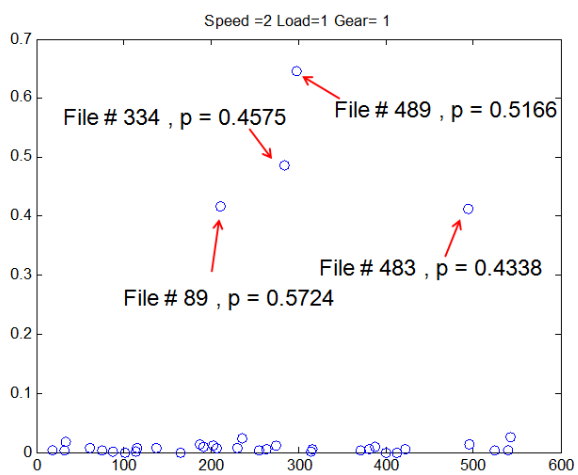

Figure 22: Probability of bad key defect for regime of 35Hz speed, light load, and spur gears.

7. CONCLUSION

This paper introduced a systematic methodology for gearbox health assessment and fault classification. The methodology was validated for 560 data sets of gearbox vibration data provided by the Prognostics and Health Management Society for the 2009 data challenge competition, and won the first place in the student division. The methodology involves the utilization of a comprehensive set of signal processing and feature extraction methods; in that the use of a single signal processing method would not be applicable for a mechanical system that consisted of a multitude of different faults. A regime segmentation approach was necessary to provide a fair comparison between data sets and grouped the data by load, speed, and gear type. A health assessment algorithm was used to classify and find the 80 baseline healthy data sets. Using the baseline data sets provided by the health assessment method, a fault diagnosis method based on a modified Euclidean distance calculation from normal along with specific features correlated to different fault signatures is used to diagnose specific faults. The fault diagnosis method is evaluated for the diagnosis of five different gearbox fault types, and could be further extended for other faults as long as a set of features can be correlated with a known fault signature. The methodology can be further applied to other rotating machine applications involving gear, shaft, or bearing components

8. SUGGESTIONS FOR FUTURE WORK

Some of the future work looks to further refine some of the techniques and methods employed as follows:

1. Refine the distance calculation algorithm used for fault diagnosis. The current modified Euclidean distance function used in this paper has proven to be very useful for vibration-based diagnosis applications and can be further enhanced to take baseline variation into consideration whenever such data is available

2. Additional signal processing methods such as the empirical mode decomposition could be evaluated to see if a more robust feature set can be provided for gearbox mechanical defects.

3. A “regime-independent” fault signature discovery method would be evaluated. Such a method would be very useful for diagnosis; in such that whenever a specific fault is identified in one regime, the signature could be captured and used to find similar faults in other regimes.

REFERENCES

J. Antoni, and R.B. Randall (2006). The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines, *Mechanical Systems and Signal Processing*, vol. 20, pp. 308-331.

J.Antoni. (2006). The spectral kurtosis: a useful tool for characterizing nonstationary signals, *Mechanical Systems and Signal Processing*, vol. 20, pp. 282-307.

E. Bechhoefer, and A. Bearnhard (2003). Setting HUMS Condition Indicator Thresholds by

Modeling Aircraft and Torque Band Variance, *IEEE Aerospace Conference Proceedings*.

F.K. Choy, S. Huang, J.J. Zakrajsek, R.F. Handschuh, and D.P. Townsend (2004). Vibration signature analysis of a faulted gear transmission system, *NASA Technical Memorandum*, NASA TM-106623/ARL-TR 475/AIAA-94-2937.

H.J. Decker, and D.G. Lewicki. (2003). Spiral Bevel Pinion Crack Detection in a Helicopter Gearbox, *ARL-TR-2958*, U.S. Army Research laboratory, NASA Lewis.

P. Grabill, J. Berry, L. Grant, and J. Porter. (2001). Automated Helicopter Vibration Diagnosistcs for the US Army and National Guard, *57th Annual Forum of the American Helicopter Society*, Washington, DC, pp.1831-1842.

J. Keller and P. Grabill. (2003). Vibration Monitoring of UH-60A Main Transmission Planetary Carrier Fault, *The American Helicopter Society 59th Annual Forum*, Phoenix, Arizona.

B. Li, M. Chow, Y. Tipsuwan, and J. Hung (2000). Neural-Network Based Motor Rolling Bearing Fault Diagnosis, *IEEE Transactions on Industrial Electronics*, vol. 47, pp. 1060-1069.

J. Ma. (1995). Energy Operator and Other Demodulation Approaches to Gear Defect Detection, *Proceedings of the 49th Meeting of the Society for Machinery Failure Prevention Technology*, Vibration Institute, Willobrook, Illinois, pp.127-140.

Z.K. Peng, and F.L. Chu. (2003). Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography, *Mechanical Systems and Signal Processing*, vol. 18, pp. 199-221.

PHM 2009 Data challenge Competition, (2009). [http://www.phmsociety.org/competition/09]

PHM Society Public Datasets, (2009).PHM 2009 Data challenge Competition, Labeled Dataset. [https://www.phmsociety.org/references/datasets]

P.D. Samuel and D.J. Pines. (2005). A review of vibration-based techniques for helicopter transmission diagnostics, *Journal of Sound and Vibration*, vol. 282, pp. 475-508.

J. Sheeley, R. Xu, Z. Ren, B. Ayhan, W. Lee, M.S. Sahni, Hu Qiaohui, and T. McClerran (2009). Initial Operational Evaluation of a Novel Corona Monitoring System, *Proceedings of the Society for Machinery Failure Prevention Technology (MFPT)*.

P.W. Tse, Y.H. Peng, and R. Yam (2001). Wavelet Analysis and Envelope Detection for Rolling Element Bearing Fault Diagnosis-Their Effectiveness and Flexibilities, *Journal of Vibration and Acoustics*, vol.123, pp.303-310.

P. Vecer, M. Kreidl, and R. Smid (2005). Condition Indicators for Gearbox Condition Monitoring Systems, *Acata Polytechnica*, vol. 45, No. 6

H. Zheng, Z. Li, and X. Chen. (2002). Gear fault diagnosis based on the continuous wavelet transform, *Mechanical Systems and Signal Processing*, vol.16, pp. 447-457.