Proceedings of ASME Turbo Expo 2011
GT2011
June 6-10, 2011, Vancouver, British Columbia, Canada

GT2011-45101

WIND TURBINE DATA ANALYTICS FOR DRIVE-TRAIN FAILURE EARLY DETECTION AND DIAGNOSTICS

LiJie Yu
General Electric Global Research
Niskayuna, NY, USA

ABSTRACT

In recent years, rapid growth in wind energy as a substantial source of electricity generation has created greater demands on wind turbine system reliability and availability. To reduce service costs and maximize return on investment, wind farm operators have begun to take a more proactive approach to turbine problems by relying on intelligent condition monitoring and automated failure detection systems. The challenge is how to effectively convert large amounts of data into actionable decisions to detect and isolate failures at an early stage. This paper describes a unique data analysis and modeling technique for online turbine health monitoring and automatic root cause assessment. It provides a means to capture failure signatures for specific root causes based on historical events as well as engineering knowledge. Both continuous and discrete turbine condition monitoring data are processed to provide a failure probability assessment. First, statistical trend analysis, feature extraction and classification methods are developed to analyze a continuous sensor data set. Secondly, a pattern recognition method is applied to calculate failure indicators from various discrete control system events, or fault messages. Then failure likelihoods derived from both the continuous and the discrete models are combined in a fusion model to increase predictive accuracy. A demonstration of the method on bearing failure modeling using SCADA data will be provided with promising results.

NOMENCLATURE

SCADA	Supervisory Control and Data Acquisition
ANN	Artificial Neural Network
GMM	Gaussian Mixture Model
RM&D	Remote Monitoring and Diagnostics
TF-IDF	Term frequency-Inverse Document Frequency
ROC	Receiver Operating Characteristic
BN	Bayesian Network

INTRODUCTION

In recent years, rapid growth in wind energy as a substantial source of electricity generation has created greater demands on wind turbine system reliability and availability. To reduce service costs and maximize return on investment, more wind farm operators have begun to take a more proactive approach and rely on intelligent condition monitoring and automated failure detection systems. Effective monitoring and diagnostic systems provide early warning of potential problems, thus helping to prevent damage growth or secondary damage, and change costly unplanned maintenance to less expensive planned and prioritized service work.

Figure 1 illustrates the major components of a variable speed wind turbine from a cross section view of the nacelle. The drive train components, including the main bearing, gearbox, and generator, have often been identified as the leading failure cost items. More advanced sensing systems have been introduced in recent years for modern wind turbine condition monitoring [1] [2], such as vibration monitoring, acoustic emission, oil debris analysis, and generator stator current signature analysis, etc. Oil debris analysis is only suitable for gearbox monitoring, and it is not able to pinpoint the specific failing part. Among the others vibration analysis is the most popular for bearing and gear monitoring. Even though technology has been continuously advanced further over the years, practical reasons are still limiting application of the advanced sensing systems. One factor is the high cost of installation and the difficulty of retrofitting existing systems. Also most advanced sensors require high frequency data sampling and storage, which present both technical and economical challenges. Due to these factors and concerns, the installed base of turbines in service with advanced sensors is still relatively small, especially in older wind farms.

On the other hand, supervisory control and data acquisition (SCADA) system has been widely used in most wind farms. At each recording it collects hundreds of data points relating to turbine operation and condition measurement; and it provides easy data access for both on-site and remote processing. The parameters typically collected by SCADA include various bearing temperatures, cooling oil temperature, ambient conditions (wind speed, temperature), rate of rotation of the rotor and generator, and power produced, etc. The conventional view of the SCADA data is that the data is not sufficiently reliable for turbine condition monitoring and failure detection. A failure mode specific data trend is either non-existent or occurs too late to detect and diagnose failures. We would argue what is really lacking are effective analytical approaches that extract and evaluate suitable data features for failure detection and diagnostics. Advanced data analysis should be carried out to convert raw SCADA data into indicative and noise resistant diagnostic features, which will then be incorporated in failure signature models to provide turbine health condition assessment and service decision support.

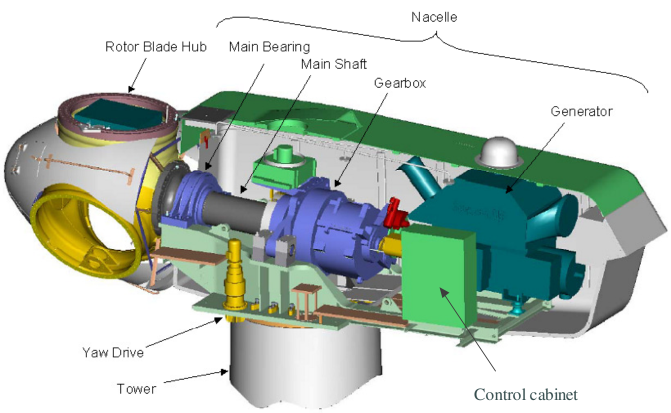


Figure 1 Wind Turbine Major Components

Some recent research have applied statistical and machine learning techniques to analyze wind turbine SCADA data. Kusiak et al. [6] and Yan et al. [7] developed data driven power curve models for turbine performance monitoring. Comparing turbine power output to referenced power curve will help to identify under-performing units. However, power curves alone cannot explain problem root cause. Wigglinkhuizen et al. [5] used a linear correction model to detrend SCADA data, but did not demonstrate capability in failure detection due to the lack of failure events during their research. Sanz-Bobi et al. [3] and Zaher et al. [4] used an artificial neural network to learn the normal behavior of gearbox and generator winding temperature, and derived measurement vs. model deviation for fault detection. They each demonstrated certain failure detection capability with varied failure lead-times.

In this paper, we propose a different technique to analyze SCADA data for failure diagnostics and prediction in a turbine drive train remote monitoring and diagnostic (RM&D) application. Both continuous sensor data and discrete control faults are analyzed statistically for failure signature modeling, and model fusion is used to produce a joint failure likelihood assessment. In the continuous data model, we leverage peer comparison at the wind farm level for data normalization and anomaly detection, then apply statistical trend analysis and feature classification methods for failure assessment. In the discrete fault model, a feature selection and ranking method is developed to identify key faults associated with each failure mode, and then convert fault frequency as well as fault temporal association into a failure likelihood assessment. Historical failure cases are leveraged to both develop and validate the failure signature models. It would be shown that the developed modeling technique is able to effectively isolate failing components from normal units at low false alarm level.

The rest of paper is organized as follows: first, data sources used in the analysis are described. Next, the formulation and demonstration of the continuous data analysis model is presented, followed by the introduction of the discrete data analysis model. Then a diagnostic model fusion method is presented to combine the failure likelihood assessments from the two disparate models. Some concluding remarks will be provided at the end of the paper.

DATA SOURCE

Operational Data

Turbine operation data used in this research is obtained from multiple wind farms in the North America region. All units are 1.5MW GE wind turbines, which is the most widely used wind turbine in the world [8]. Two sources of data are obtained from the plant SCADA system: continuous sensor data and discrete control system fault messages.

The continuous sensor data is collected and averaged every 10 minutes. As our focus is modeling drive-train related failures, engineering knowledge is leveraged to select a subset of subsystem related sensors, which include various bearing temperatures, gearbox cooling oil temperature, ambient temperature, wind speed, rotor and generator speed, and power output. A sample of the continuous time series data is shown in Table 1, where the park name and turbine identifiers have been replaced to protect the operator's identity. 10-minute average is a widely adopted industrial standard in dealing with wind turbine sensor data. However, due to the low frequency data acquisition and averaging, some of the transient information may be lost in the process. Also, some of the sensor input available to turbine control system may not be recorded in the SCADA continuous data stream after all. This information loss and information gap may be addressed by leveraging the second SCADA source, namely the discrete fault.

Turbine control systems constantly examine sensor input in each control loop and apply control logic to assess turbine condition in real time. Once certain pre-determined conditions are met, an event message, or fault will be generated and logged in the SCADA system. A sample of the fault sequence collected by a SCADA is shown in Table 2. Some of these faults messages are benign and only informational, whereas some others reflect more serious control logic violation, which may lead to abnormal turbine behavior or failure. The fault message stream captures transient turbine status based on control logic in real time, and is therefore not affected by the low reporting frequency of the continuous data source. In addition, the fault system provides some additional component information that is not available in the continuous data stream, such as some calculated system state variables based on the turbine control model. In summary, the discrete fault data represents a complementary data set to the continuous sensor data with a certain level of information redundancy, which provides the potential of achieving higher diagnostic accuracy and coverage by fusing information from both sources.

Historical Case Data

Historical failure cases are collected and used for model learning and testing. Failures cases are identified based on service records and field engineers' feedback, which specifies the turbine identity, time of a problem occurrence, symptoms, and the failure's root cause. To learn normal turbine behavior and evaluate model anomaly detection and separation capability, normal operating turbine data are also collected. Normal turbine units are selected from a wide selection of wind parks, and turbines with known failure records are filtered out. Two simulated cases are created for each normal turbine, one in the summer and one in the winter to resemble normal turbine behavior under environmental variation. A total of about 1000 normal cases were created in this study.

Throughout this paper, we will demonstrate the concept of an analysis approach using main bearing failure examples. To this regard, eighteen main bearing failure cases were collected with representative cases for each of several different turbine configurations. Both continuous and discrete SCADA data are extracted for each case. Data series are downloaded for six months prior to the case date. The goal is to identify failure signatures as early as possible, in order to leave enough lead time for service planning, as well as to prevent damage growth causing secondary damage at extra cost.

Another information source that also proved to be instrumental in the modeling process is engineering knowledge. This includes both design engineer knowledge about turbine component variation and operational characteristics, as well as service engineer knowledge about known product defects and failure mode understanding through FMEA analysis. All the engineering support provides valuable guidance in data selection, preparation and interpretation. Compared to an entirely data driven information mining process, the engineering knowledge guided data modeling method is deemed superior since it ensures correct physics of the analysis system and minimizes the influence of artifacts in the data.

Table 1 Continuous SCADA Data Time Series Sample

ParkName	Turbine ID	Timestamp	KW	WIND_SPEED	GEN_RPM	......	TRAN_TEMP	BEARING_A_TEMP	BEARING_B_TEMP
Farm A	T1	11/21/2007 0:10	1550	13.38	1438		51	30.83	28
Farm A	T1	11/21/2007 0:20	1537	13.13	1437		51	30.58	28
	......	......							
Farm A	T2	11/21/2007 0:10	1528	13.38	1437	⋮	51	30	28
Farm A	T2.	11/21/2007 0:20	1525	13.13	1435	⋮	50.01	29.06	27.66
	......	......							

Table 2 Fault Message Sequence Sample

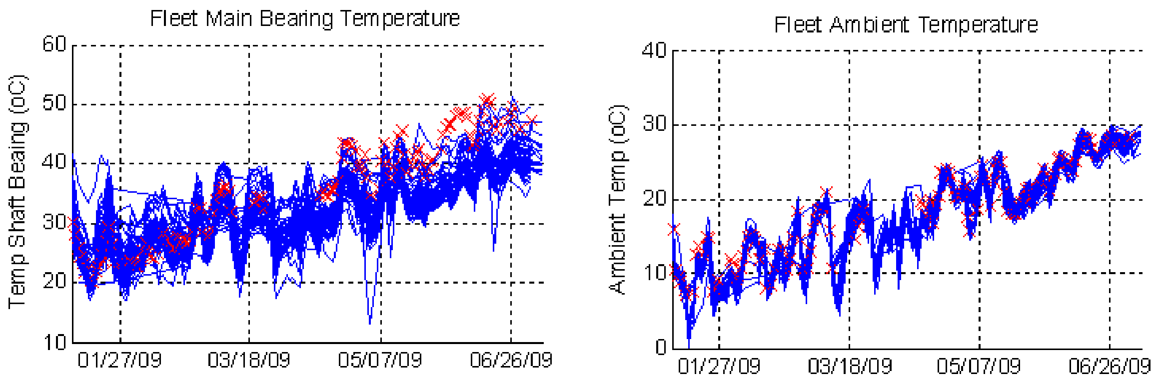
ParkName	Turbine ID	Timestamp	Fault Code	Fault Message
Farm A	T1	11/26/07 11:14	14	Generator Overspeed
Farm A	T1	11/27/07 15:32	144	Blade angle asymmetry
Farm A	T2	11/29/07 19:36	27	Secondary bracking time too long
Farm A	T2.	12/1/08 19:56	3	Manual Stop
	......			

CONTINUOUS DATA MODELING

To monitor the wind turbine bearing health, bearing temperature sensor data are collected. These include temperatures of the main bearing (TEMP_SHAFT_BEARING), high speed gearbox bearing (TRANS_BEARING_TEMP), generator bearings at drive end (BEARING_A_TEMP) and non-drive end (BEARING_B_TEMP). When a bearing defect occurs, such as lack of lubrication, cracks on rollers, pitting or spall on bearing race, the bearing temperature will increase due to excessive friction of the defective component. However, bearing temperature cannot be directly trended for bearing health assessment; this is because it is also affected by environmental conditions, turbine load and speed, as well as component cooling system operation, etc.

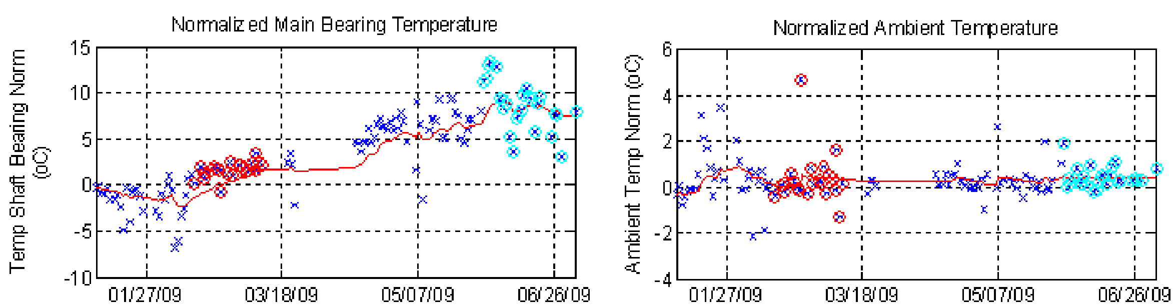
Sanz-Bobi et al. [3] and Zaher et al. [4] separately developed an ANN based non-linear normalizer to detrend turbine parameters for failure detection. In our research, we adopted a simpler and more practical method to detrend the temperature data. First instead of monitoring a particular turbine at a time, turbines in the same wind park are grouped into various fleets, and a fleet may be used as the processing unit for on-line health monitoring. Turbines within the same fleet have the same component configuration (such as gearbox type, bearing type), subjected to similar ambient conditions, therefore they may serve as peers in trend analysis. Secondly, fleet data are separated at different loading levels to remove or reduce the influence of loading. Typically, at high load when a bearing is subjected to maximum stress, the risk of failure is higher, therefore the failure detection at high load carries more weight than those during a low load period. Here load is assessed using power output (KW) and generator speed (GEN_RPM).

Figure 2 shows the corresponding fleet data trend of a turbine with main bearing failure. Note that only two representative variables, main shaft bearing temperature (a) and ambient temperature (b), are selected and shown here for clarity. Data are further aggregated into daily averages. Each blue data trace corresponds to one out of sixty or so turbines within this fleet, and the red-cross shows the trend of the failing turbine. It is obvious both ambient and bearing temperature show seasonal variation. The amount of variation makes direct trending of the temperature variables unusable for anomaly detection. However, except for certain outliers, the variations caused by seasonal conditions are closely correlated among the normal turbines within the fleet. Such correlation indicates that a cohort analysis may provide indicative features to separate abnormal individuals from the rest of the fleet. From Figure 2 we can find that the target turbine initially operated well within the fleet boundary presumably before damage occurred, but bearing temperature started shifting hotter about two months before the final failure date.



(a) **(b)**

Figure 2 Main Bearing Failure Example - Fleet Daily Average Trend



(a) **(b)**

Figure 3 Main Bearing Failure Example - Normalized Trend and Shift Evaluation

Sensor data is then normalized by subtracting the fleet median from each sensor value. Figure 3 shows the normalized data trend (blue cross symbol) for the failed turbine in the above case. It can be found that seasonal or ambient variation has been dramatically removed from the normalized trend, whereas the bearing temperature trend drift caused by the degrading main bearing is retained. The normalized data trends need to be further analyzed to extract statistical features for automatic failure detection and diagnosis using software. Here we adopted a multi-variable shift detection algorithm developed by Yu et al. [9] to automatically extract trend features of shift and current level for each normalized parameters. The algorithm uses piecewise linear regressions to detect the latest and most significant shift within a time series data set. Two sample data sets are selected near the end and start of the identified shifted region, shown as red cycle and cyan cycle respectively in Figure 3, and two-sample t-test is then used to assess the confidence interval of the parameter level and shift.

Other than the two variables shown in the figure, a few other temperature sensors are also used in the main bearing failure analysis, including gearbox bearing temperature, gearbox cooling oil temperature, and nacelle temperature. Some of these variables relate to gearbox health. The rational is that most wind turbine design have the main shaft and gearbox in the same assembly, and failure of the main bearing may cause the gearbox to slide axially causing stress in the gearbox unit. With five sensor variables, and two trend features for each parameter, shift and current level, respectively, a total of ten independent data features are obtained for each case for main bearing health assessment.

To simplify failure signature extraction, PCA (principal component analysis) is used to reduce case features into a lower dimension. Feature clustering is then performed in the principal component (PC) space. A three-component Gaussian mixture model (GMM) is automatically learnt using the first two PCs of normal turbines based on a modified expectation maximization algorithm developed by Figueiredo et al. [10]. As shown in Figure 4, each grey dot shows the PC values of a single normal turbine case, whereas the solid line circles show the bivariate Gaussian density contours for each mixture component.

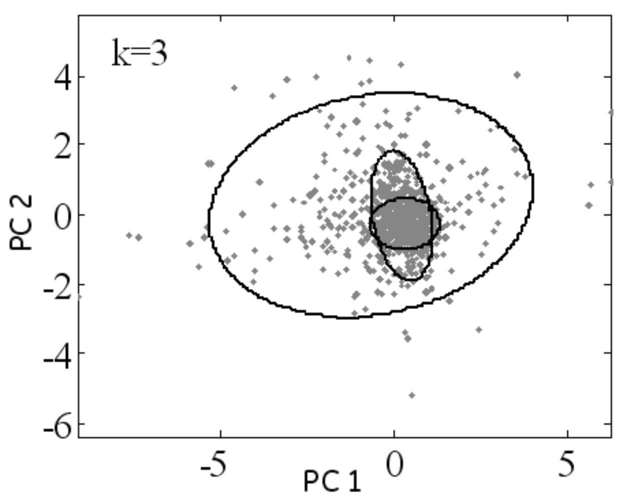


Figure 4 Normal Turbine Three-Component GMM

PC features of each of the eighteen main bearing failure cases are then evaluated using the learnt GMM model to assess probability that the case feature fall in the cluster of each Gaussian component. The PC features of both normal cases and cases with main bearing failure are visualized in Figure 5. As shown in the figure legend, MBi, i=1, 2 or 3, indicates which mixture component a main bearing failure case has the highest probability. A probability threshold is then specified and compared to the outcome of each case. If the case probability to normal is below the threshold, an anomaly flag is raised. In Figure 5, the red dots correspond to normal turbines with an anomaly flag raised; they therefore are considered as false positive alerts, which represent about 2.5% of the normal cases. Eleven blue dots, corresponding to main bearing failure cases detected using the same probability measure, represent true failure detection rate of about 61%.

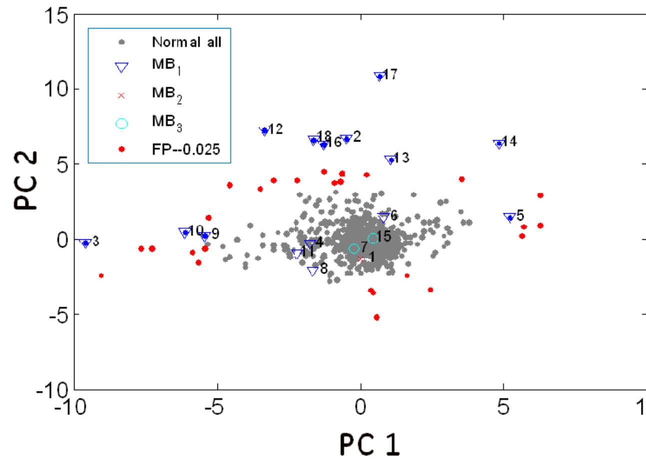


Figure 5 Main Bearing Failures vs. Normal Case Clusters

DISCRETE DATA MODELING

As the example shown in Table 2, various fault messages may be generated during turbine operation. Faults are created by the turbine control system to alert operators of abnormal conditions and protect the turbine from serious damage. Sometimes faults are generated due to random noise or transient operating states, and the turbine will automatically recover to a normal condition so the fault is reset either automatically or manually without need for further investigation. If a fault occurs frequently and persistently, the chance that it is caused by underlying health problems is higher. The challenge is how to process the fault data to filter out nonsense faults, and leveraging fault information to create predictors of incipient turbine failure.

The approach we adopted is similar to the continuous data analysis, in the sense that for each failure mode we analyze features extracted from historical data to maximize separation between failure cases versus normally operating turbines. The fault feature space is the frequency count of each of the selected fault items. It is also hypothesized that besides individual fault occurrences, temporal correlation among faults should also be considered as an indicator for failure identification. To model such temporal correlation between faults, a combo-fault is created if two different fault items occur within a short time interval. Essentially a combo-fault adds extra dimensions in the feature space by taking into account the temporally associated faults. Combo-faults are generated at the preprocessing step, and ranked together with the raw single fault in the next step: feature selection.

Figure 6 shows fault existence comparisons between the normal turbine cases and main bearing failure cases. Two groups of fault existence measures are calculated, where the solid blue bars represent cases with main bearing failure, and the light green bars for normal cases. The horizontal axis is a list of fault items (only those that occurred at least once in main bearing failures), and the vertical axis is the percentage of cases where each fault occurs within the respective group. The fault items have been sorted in order of decreasing percentage difference between the failure and normal group. As shown in the arrow in Figure 6, further left fault items occur more often in main bearing failure cases, and less often in normal turbines, and therefore, have higher failure separation capability than the further right items. The fault codes of the top eight fault items have been listed in the text box of Figure 6, where the fault code with three or less digits are the original fault code, and the six-digit fault code for combo-fault is created by combining fault code of two temporal associated faults.

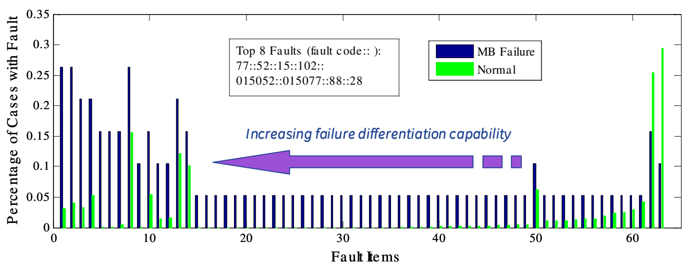


Figure 6 Normalized Fault Frequency Comparison

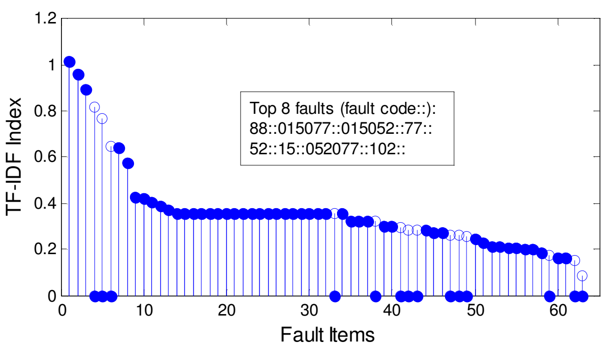


Figure 7 Fault Ranked and Pruned Using TF-IDF Index

Another way of ranking the importance of faults for the failure detection model is term frequency and inverse document frequency (TF-IDF) index, borrowed from text mining methodology. Given a specific document within a large document pool, the task is to rank all the terms in the target document based on their weight or importance for uniquely characterizing the document content. The idea is that if a term occurs evenly in all documents, then its power for identifying the specific document is low, versus a term that occurs more frequently in the target document and rare in other documents, where the identification power is high. By the same token, if a fault occurs frequently in failure cases and less frequently in normal cases, it is a stronger indicator of failure detection. The TF-IDF index is defined as

$$TF\_IDF_i = \frac{n_i}{N} \log \frac{|D|}{1+|\{d : f_i \in d\}|}, \quad i = 1,2,\cdots K$$

where n_i is the number of failure cases with fault i, N is the total failure cases, D is the total normal cases, and $|\{d : f_i \in d\}|$ is the number of normal cases with fault i. Fault TF-IDF ranking for main bearing failure against normal cases is shown in Figure 7. Comparing Figure 7 to Figure 6, it can be seen that the top ranked faults between the two methods are similar, even though the specific ranking order differs. Also shown in Figure 7 is fault pruning, i.e., a single fault is removed from consideration if it has been included in a higher-ranked combo-fault. For example, Fault “77” is removed since Fault “015077” is selected with higher rank. Eliminated faults are those with zero TF-IDF indexes shown as filled-in circles in Figure 7.

Using the ranked and pruned fault items, an iterative process is run to optimize the model performance where different numbers of top ranked faults are selected to form the discrete failure model. For each iteration,, the selected faults are used to fit a logistic regression model that converts fault patterns to case failure probabilities. The fault feature used here is fault occurrence frequency, i.e., the occurrence count of a particular fault within the case duration, for each selected fault items. Models learned with different number of selected faults, K, are then evaluated based on the standard criteria of maximizing the ROC area. For the main bearing analysis, maximum ROC area is obtained when K=3, and the corresponding ROC curve is shown in Figure 8. It is shown that a fairly low false positive rate, about 1%, is obtained from the discrete failure model, whereas the failure detection rate is about 42%.

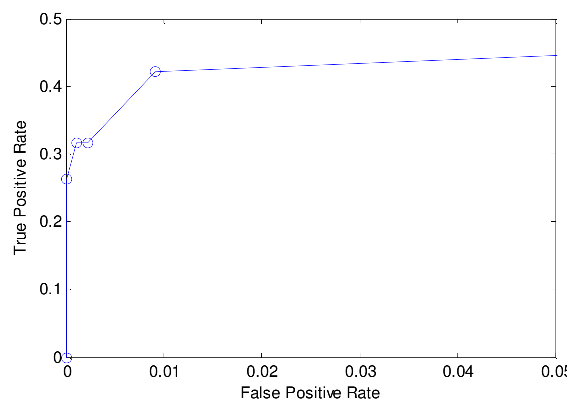


Figure 8 ROC Curve for Main Bearing Fault Model

DIAGNOSTIC MODEL FUSION

The continuous data model and discrete data model independently assess the failure likelihood based on two disparate sets of features. As mentioned earlier, there is both complementary and overlapping information between the two models; therefore, a diagnostic fusion model is needed to fuse the assessment from the two models and to optimize the overall accuracy of failure diagnosis.

Since both models have created a mapping from the feature space to the problem space as failure likelihood scores, there are many options for fusing the two diagnostic results. It could be a straightforward logic operation or averaging. The method we are in favor is a Bayesian Network decision fusion model as described by Yu et al. [11]. Bayesian Network (BN) is a graphical probabilistic reasoning model that contains nodes and edges. A node represents a variable or system state, and an edge connects two nodes and defines the causal relationship between the nodes. A BN fusion model for the main bearing failure is shown in Figure 9, where the node "MB Failure" is the to-be derived fused failure assessment, and the "Continuous Model" and "Discrete Model" represents the failure assessment from the two models described in the previous two sections. Accuracy performance of each individual model is used to specify the conditional probabilities between the decision fusion node and model assessment nodes. An example of probability propagation is shown in Figure 9: when evidence of failure is inserted into the network with probability 1.0 based on assessment of the continuous model, the fused posterior failure probability is obtained as 99%. Depending on the agreement or disagreement from the discrete model, this failure probability will be readjusted by the BN inference engine. Using the BN fusion model, main bearing failure detection accuracy is improved to about 75%, while maintains a fairly low false positive rate at about 5%.

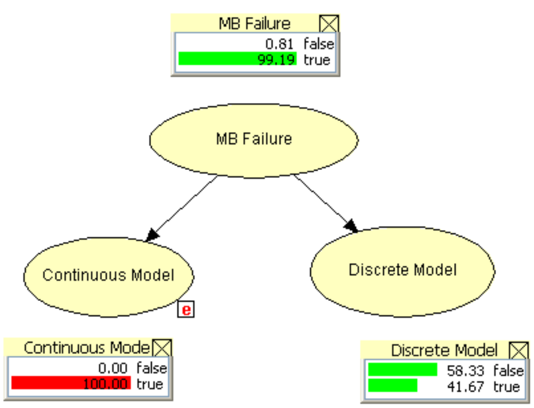


Figure 9 BN Decision Fusion Model for Main Bearing Failure

For each different failure mode, separate pairs of continuous data model and discrete model will be trained. Their assessment of failure likelihood will be independent from the model assessments for other failure modes, since in a real application, multiple failure modes may exist simultaneously. However, the different failure mode assessment is fused in the same fusion model. That is, if there is no interaction between two failure modes, two disjoint groups of fusion nodes will be captured. However, if there is hierarchical interaction between two failure components, they can be modeled easily as connected network nodes, with a connection edge corresponding to the system hierarchy. Evidence derived from one failure component may affect the likelihood of other connected ones, and the probability update is performed automatically through message passing by the Bayesian inference engine.

CONCLUSIONS

This paper has presented a set of analytical methods that systematically process wind farm historical data to derive failure detection and diagnostic signature models. The diagnostic models are derived offline based on patterns extracted from historical RM&D data and turbine failure data. Once validated and deployed, the diagnostic models are then used on-line to analyze remote monitored data and provide real-time turbine health assessment.

It is demonstrated that both continuous data modeling and discrete fault modeling provide indicative and often complementary information regarding likelihood of different failure modes. By leveraging both information sources in a BN-based fusion model, we are able to optimize the diagnostic model performance both in terms of accuracy and failure coverage.

The method developed here is described using SCADA variables due to its wide accessibility. However, the applicability of the method itself is not limited to a specific data source. The same analysis methodology could be carried out for more advanced condition monitoring systems as well, such as vibration indicators or oil debris counts. The advantage gained by adopting this method is that it provides automatic trend analysis to help derive failure signatures, and it also has good potential to reduce false positive alerts of existing systems by leveraging sensor fusion and decision fusion.

ACKNOWLEDGEMENT

The author would like to thank Mark Osborn, John Mihok, and Christopher Dannehy at GE Wind, Jesse Graeter at GE Energy, Robert Horabin and Honor Powrie at GE Aviation for their support for this research.

REFERENCES

[1] Lu, B.;Li, Y.Y. and Yang Z.Z., "A review of Recent Advances in Wind Turbine Condition Monitoring and Fault Diagnosis," *IEEE Power Electronics and Machines in Wind Application*, 2009.

[2] Amirat, Y.; Benbouzid, M.E.H. and Wamkeue, R., "Condition Monitoring and Fault Diagnosis in Wind Energy Conversion Systems: a Review, " *Proceedings IEEE International Electric Machines and Drives Conference*, vol. 2, pp. 1434-1439, 2007.

[3] Sanz-Bobi, M.A.; Garcia; M.C. and Del Pico J., "SIMAP: Intelligent System for Predictive Maintenance Application to the Health Condition Monitoring of a Wind Turbine Gearbox," *Computers in Industry*, vol 57, pp. 552-568, 2006.

[4] Zaher A.S.; McArthur S.D.J. and Infield D.G., "Online Wind Turbine Fault Detection through Automated SCADA Data Analysis," *Wind Energy*, vol. 12, no. 6, pp. 574-593, 2009.

[5] Wiggelinkhuizen, E.; Verbruggen, T.; Braam, H.; Rademakers, L.; Xiang, J.P. and Watson, S, "Assessment of Condition Monitoring Techniques for Offshore Wind Farms," *ASME Journal of Solar Energy Engineering*, vol. 130, August, 2008.

[6] Kusiak, A.; Zheng, H.Y.; and Song Z., "On-Line Monitoring of Power Curves," *Renewable Energy*, no. 34, pp. 1487-1493, 2009.

[7] Yan, Y.J; Osadciw, L.; Benson, G. and White, E., "Inverse Data Transformation for Change Detection in Wind Turbine Diagnostics," *Proceeding of 22nd IEEE Canadian Conference on Electrical and Computer Engineering*, Delta St. John's, Newfoundland and Labrador, Canada, May, 2009.

[8] AWEA U.S. Wind Industry Annual Market Report Year Ending 2009. On line: http://www.awea.org/reports/Annual_Market_Report_Press_Release_Teaser.pdf

[9] Yu, L.J; Cleary, D.J. and Cuddihy, P.E., "A Novel Approach to Aircraft Engine Anomaly Detection and Diagnostics," *Proceedings of IEEE Aerospace Conference*, Big Sky, Montana, 2004.

[10] Figueiredo, M.A.T. and Jain, A.K., "Unsupervised Learning of Finite Mixture Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, " vol. 24, no. 3, March, 2002.

[11] Yu, L.J.; Cleary, D.J.; Osborn, M. and Rajiv, V., "Information Fusion Strategy for Aircraft Engine Health Management," *Proceedings of ASME Turbine Expo,* Montreal, Canada, May, 2007.