Device Health Estimation by Combining Contextual Control Information with Sensor Data and Device Health Prognostics Utilizing Restricted Boltzmann Machine

Linxia Liao[1], Tomonori Honda[2], and Radu Pavel[3]

[1,2] *Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA, 94304, USA*
linxia.liao@parc.com
tomo.honda@parc.com

[3] *TechSolve, Inc., 6705 Steger Dr, Cincinnati, OH, 45237, USA*
pavel@techsolve.org

Abstract

The goal of this work is to bridge the gap between business decision-making and real-time factory data. Beyond real-time data collection, we aim to provide analysis capability to obtain insights from the data and converting the learnings into actionable recommendations.

For device health estimation, we focus on analyzing device health conditions and propose a data fusion method that combines sensor data with limited diagnostic signals with the device's operating context. We propose a segmentation algorithm that provides a temporal representation of the device's operation context, which is combined with sensor data to facilitate device health estimation. Sensor data is decomposed into features by time-domain and frequency-domain analysis. Principal component analysis is used to project the high-dimensional feature space into a low-dimensional space followed by a linear discriminant analysis to search the optimal separation among different device health conditions. Our industrial experimental results show that by combining device operating context with sensor data, our proposed segmentation and linear transformation approach can accurately identify various device imbalance conditions even for limited sensor data which could not be used to diagnose imbalance on its own.

For device health prediction, we propose a restricted Boltzmann machine based method to automatically generate features that can be used for remaining useful life prediction, which is performed by a random forest regression algorithm. The proposed method was validated through run-to-failure dataset of a machine tool spindle test-bed.

1. Introduction

The growing Internet of Things is predicted to connect 30 billion devices by 2020 (MacGillivray, Turner, & Lund, 2013). This will bring in tremendous amounts of data and drive the innovations needed to realize the vision of Industry 4.0—Cyber-Physical systems monitoring physical processes, and communicating and cooperating with each other and with humans in real time. One of the key challenges to be addressed is how to analyze large amounts of data to provide useful and actionable information for businesses intelligence and decision making. In particular, to prevent unexpected downtime and its significant impact on overall equipment effectiveness (OEE) and total cost of ownership (TCO) in many industries. Continuous monitoring of equipment, early detection of incipient faults, and prediction of failure before it happens can support optimal maintenance strategies, prevent downtime, increase productivity, and reduce costs.

A significant number of anomaly detection and diagnosis methods have been proposed for device fault detection and health condition estimation. Chandola et al. (Chandola, Banerjee, & Kumar, 2009) discusses various categories of anomaly detection technologies and their assumptions as well as their computational complexity. Several approaches such as statistical methods (Markou & Singh, 2003), neural network methods (Markos & Singh, 2003) and reliability methods (Guo, Watson, Tavner, & Xiang, 2009), have been applied to detect anomalies for various types of equipment. The philosophies and techniques of monitoring and predicting machine health with the goal of improving reliability and reducing unscheduled downtime of rotary machines are presented by Lee et al. (Lee et al., 2014).

Many of these methods focus on analyzing, combining, and modeling sensor data (e.g. vibration, current, acoustics signal) to detect machine faults. One issue that remains mostly unaddressed in these methods is that they rarely consider the varying operating context of the machine. In many cases, false alarms are generated due to a change in machine operation (e.g. rotational speed) rather than a change in machine condition. A major challenge in addressing this issue is that most machine controllers are built with proprietary communication protocols, which leads to a barrier in obtaining control parameters to understand the context under which the machine is operating. Recently, the MTConnect open protocol (Standard, 2009) was developed to connect various legacy machines independent of the controller providers. MTConnect provides an unprecedented opportunity to monitor machine operating context in real-time. In this paper, we leverage MTConnect to diagnose machine health condition by combining sensor data with operating context information. Additionally, we investigate whether it is possible to diagnose machine health condition using less sensor data when it is combined with context information.

Many methods have been proposed in the literature for device remaining useful life (RUL) prediction. These methods can be generally classified as data-driven method, physics-based method, and hybrid method (Liao & Kottig, 2014). Since detailed information of the assembled components is not available in our case, physics-based modeling is unfeasible. Hence, data-driven method becomes the primary approach in our work for prediction. To enable an accurate prediction using data-driven method, feature extraction is a critical step. If an extracted feature is well correlated with the fault propagation process (e.g. vibration root mean square increases as the machine degrades), a good prediction can be expected by extrapolating the historically observations to the future. Related work can be seen in (Coble & Hines, 2009), which used genetic algorithm to find the optimal feature subset, and in (Liao, 2014), which used genetic programming to discover novel features for prediction. In most of the cases, engineering expertise is need to a certain extent to guide the feature extraction, which might not be directly available for complex systems. We would like to explore automatic feature generation method for remaining useful life prediction when engineering expertise is unavailable. Deep learning has recently gained popularity in machine learning based on learning layers of network structure based on restricted Boltzmann machines (RBM). RBM has been widely used as a generative model in many applications such as image classification, speech recognition, and word representation. It has recently been applied in prognostics health management area for health state classification (Tamilselvan & Wang, 2013). Instead of using RBM in a classification scenario, we explore RBM as a feature extraction tool in a RUL prediction scenario.

Prior work (Pavel, Snyder, Frankle, Key, & Miller, 2010) has demonstrated that vibration data could be used for diagnosing machine imbalance fault conditions. Our study focuses on extending prior work by exploring various types of sensor and control data for diagnosing the imbalance of the machine tools. Prior work (Pavel & Iverson, 2012) proposed self organizing maps and polynomial curve fitting for RUL prediction based on domain specific features such bearing signature frequencies. Our study focuses on automatic feature generation assuming domain specific expertise is unavailable.

Our contribution includes the following:

- Combining control and sensor signals for machine health condition estimation, while utilizing a different set of sensor data such as temperature, power, flow, and lubricant/coolant pH instead of vibration.
- A novel method of using Restricted Boltzmann Machine as a feature generation model and coupling with a random forest algorithm in remaining useful life prediction applications.

Our hypothesis is that these advancements to prior work will aid in improving the diagnosis and prognostics capability, as well as reducing the cost of machine diagnostics by utilizing cheaper sensors and saving engineering effort in feature engineering for predictive maintenance tasks.

2. TECHNICAL APPROACH

This section contains two subsections to describe the technical approaches for: (1) device health estimation by combining contextual control information with sensor data; and (2) remaining useful life prediction using Restricted Boltzmann Machine and random forest.

2.1. Device Health Estimation

For each extension to prior work listed in Section 1, we performed two main steps for diagnostics:

- Feature Extraction & Synthesis
- Model Selection

2.1.1. Feature Extraction & Synthesis

There are various approaches for condensing time series information into data mining features. Prior work has utilized transfer functions to map control signals to vibration sensor data (Pavel et al., 2010). The diagnosis step is then reduced to comparing the features of transfer function-predicted vibration data and the sensor-derived vibration data. This approach makes sense when the control signal directly impacts the output variables of the machine. For motion control of machine tools, the estimated transfer function should be similar to the transfer function of the implemented control (like PI or PID). Typical vibration data features would include average, standard deviation, and maximum FFT values (Deng, Runger, Tuv, & Vladimir, 2013).

However, we would like to diagnose the state of machine using not only accelerometers, but also other sensors, such as temperature sensors. Since temperatures at various locations are not part of active control loops, there may not exist well defined transfer functions that can map control signals to temperature sensor data very accurately. In such cases where conventional features extracted from temperature signals are not correlated with the fault (imbalance) to a sufficient degree. Additionally, if the associated sensors are too expensive to install, then data fusion may be applied.

There are three data fusion approaches typically used in machinery diagnostics (Liu & Wang, 2001; Jardine, Lin, & Banjevic, 2006)—data-level fusion, feature-level fusion, and decision-level fusion. Data-level fusion involves combining sensor data before feature extraction, such that features contain information gathered from multiple sensors. Feature-level fusion involves generating features from each sensor separately, then fusing this set of features generated from all of the sensors coherently for diagnostics. Finally, decision-level fusion creates diagnostics from each sensor separately, then aggregates these diagnostics into a single diagnostic output.

The choice of the three types of data fusion methods is often application specific. In our application, we found that temperature sensor data cannot resolve imbalance conditions by itself and control signal data is too coarse-grained to aid in classifying imbalance conditions using the standard data-fusion techniques. Note that we did not focus on spindle acceleration data, which could diagnose imbalance on its own (see Subsection 3.1.1) since that would require retrofitting existing machine tools with new expensive sensors and data acquisition hardware. Ideally we would like to use the readily accessible control signals and data from inexpensive temperature sensors to diagnose imbalance. To achieve this goal, we proposed a different type of data fusion approach. We used the control signal to provide the contextual information for temperature sensor data. The control signal is used for the segmentation of sensor data, but does not directly map into feature vectors (see Subsection 3.1.2).

2.1.2. Model Selection

Since the data sets are statistically small and dimensionality of the data is increased by feature synthesis, the models to be used for imbalance classification need to be carefully chosen to avoid over-fitting. The high-dimensional data needs to be projected to a much smaller sub-space to prevent over-fitting[1] To accomplish this, the main techniques used in this study are Principal Component Analysis (PCA) (Wold, Esbensen, & Geladi, 1987) and Linear Discriminant Analysis (LDA) (Koehler & Erenguc, 1990). These techniques are based on linear coordinate transformation, which makes them more likely to under-fit and less likely to over-fit (Yang, Chen, & Wu, 2011).

[1]Note that complexity of model is positively correlated with likelihood of over-fitting. Thus, creating a classifier that takes high-dimensional input will have higher degree of fredoom (i.e. higher complexity) compare to low-dimensional inputs, which results in higher likelihood of over-fitting.

2.2. Device Remaining Useful Life Prediction

The remaining useful life (RUL) prediction algorithm can be summarized in Figure 1. The pre-processed data is input to the Restricted Boltzmann Machine to automatically generate features. The preprocessed data can actually be the raw signals, e.g. vibration signals, or time/frequency domain features of vibration, or features extracted by signal processing techniques e.g. discrete wavelet transform. The generated features are then input to a predictor, which is random forest in this case, to predict RUL.

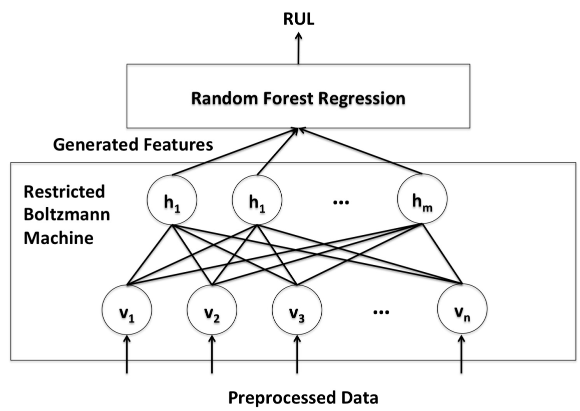


Figure 1. RUL prediction method.

2.2.1. Feature Generation

Restricted Boltzmann Machine (RBM) can be considered as a two-layer network which consists a visible layer and a hidden layer. The visible layer corresponds to the observed input units (v), and the hidden layer corresponds to the feature detectors which are hidden units (h). Since we consider Gaussian input for both the input and hidden units, the energy function of the RBM is more complex than the common binary case. We defined the energy function as:

$$E(v,h) = \sum_{i \in vis} \frac{(v_i - a_i)^2}{2\delta_i^2} + \sum_{j \in hid} \frac{(h_j - b_j)^2}{2\delta_j^2} - \sum_{i,j} \frac{v_i}{\delta_i} \frac{h_j}{\delta_j} \omega_{ij}, \tag{1}$$

where v_i, h_j are the states of the visible unit i and hidden unit j, a_i, b_j are their levels, δ_i, δ_j are the standard deviations, and ω_{ij} is the weight between them. The probability that the RBM network assigns to a visible vector is given by summing over all hidden vectors:

$$P(v) = \frac{1}{Z} \sum_h \exp(-E(v,h)), \tag{2}$$

where $Z = \sum_{v,h} \exp(-E(v,h))$. Now we can define:

$$P(v,h) = \frac{\exp(-E(v,h))}{Z} \tag{3}$$

$$P(h|v) = \frac{\exp(-E(v,h))}{\sum_h \exp(-E(v,h))} \tag{4}$$

Then we can use the negative log likelihood gradient to update the parameters ($a_i, b_j, \delta_i, \delta_j, \omega_{ij} \in \theta$) using:

$$\begin{aligned}
\frac{d}{d\theta}(-logP(v)) &= \frac{d}{d\theta}(-log\sum_h P(v,h)) \\
&= \frac{d}{d\theta}(-log\sum_h \frac{\exp(-E(v,h))}{Z}) \\
&= -\frac{Z}{\sum_h \exp(-E(v,h))}(\sum_h \frac{1}{Z}\frac{d\exp(-E(v,h))}{d\theta} \\
&\quad -\sum_h \frac{\exp(-E(v,h))}{Z^2}\frac{dZ}{d\theta}) \\
&= \sum_h (\frac{\exp(-E(v,h))}{\sum_h(-E(v,h)}\frac{dE(v,h)}{d\theta}) + \frac{1}{Z}\frac{dZ}{d\theta} \\
&= \sum_h P(h|v)\frac{dE(v,h)}{d\theta} - \frac{1}{Z}\sum_{v,h}\exp(-E(v,h))\frac{dE(v,h)}{d\theta} \\
&= \sum_h P(h|v)\frac{dE(v,h)}{d\theta} - \sum_{v,h} P(v,h)\frac{dE(v,h)}{d\theta}
\end{aligned} \tag{5}$$

The positive part in the last line of Eq. 5 is the so called positive phase contribution and the negative part is the so called negative phase contribution. The algorithm updates the parameters through iterations coupling with a learning rate and/or a momentum parameter until a stop criterion is met. The hidden unit states are used as the extracted features for RUL prediction.

2.2.2. RUL prediction

We treat the RUL prediction as a regression problem, in which we will train a supervised learner to match the extracted features and the expected RUL. In our case, we picked random forest algorithm as our prediction algorithm to demonstrate how to make predictions based on the features extracted from RBM. Random forest (Breiman, 2001) is an ensemble algorithm for classification or regression by aggregating the decision result from multiple decision trees. A simple pseudo algorithm of random forest training is described in Algorithm 1. After training, the algorithm outputs a RUL given a feature vector extracted from the RBM model described in Section 2.2.1.

Algorithm 1 Random forest training algorithm

- Draw N bootstrap data samples from original dataset D;
- For each of the bootstrap data samples, build a decision tree. For each node of the decision tree, randomly sample M of the predictors (observations in our case), and choose the best split among the selected predictors;
- Make a prediction by aggregating the predictions of the N trees (e.g. majority votes);
- At each bootstrap iteration, predict data not in the bootstrap sample (called out-of-bag data) using the tree built with the bootstap sample. Aggregating the out-of-bag error rate, and repeat the process until a preset threshold is met (i.e. error rate or maximum number of iteration)

3. RESULTS

This section contains two subsections to demonstrate: (1) device health estimation using data collected from a machine tool including sensor data and MTConnect data; (2) remaining useful life prediction using run-to-failure dataset collected from a machine tool spindel testbed.

3.1. Device imbalance condition estimation

We have explored three imbalance scenarios to investigate our hypothesis of diagnostics using:

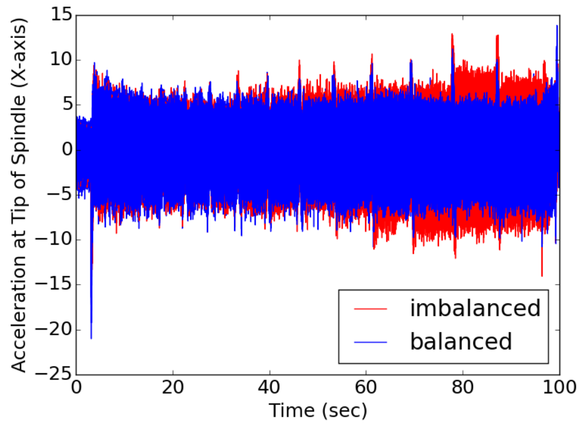
- Sensor based diagnostics
- Control based temporal segmentation followed by sensor based diagnostics

3.1.1. Sensor based Diagnostics

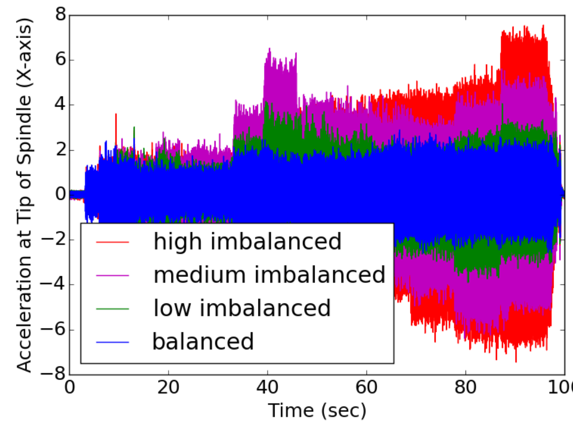
In this case, each sensor signal was analyzed separately to determine if any of the sensor signals contains enough diagnostic information to detect imbalance on its own. By plotting the time series data we find that spindle acceleration sensors (which captures vibration) show higher oscillation amplitudes (see Figure 2) with increasing imbalance. Since imbalance actually impacts moment of inertia of the spindle, this change in acceleration is expected.

We also considered measuring imbalance through temperature. From the energy flow perspective, additional acceleration caused by imbalance should result in higher energy consumption from the power source and higher energy dissipation to thermal inertias due to friction, which should result in temperature increase in parts of the machine tool. However, the time series data, from each of the temperature sensors, did not show distinguishing features similar to the acceleration sensors. An example of temperature sensor time series data is shown in Figure 3.

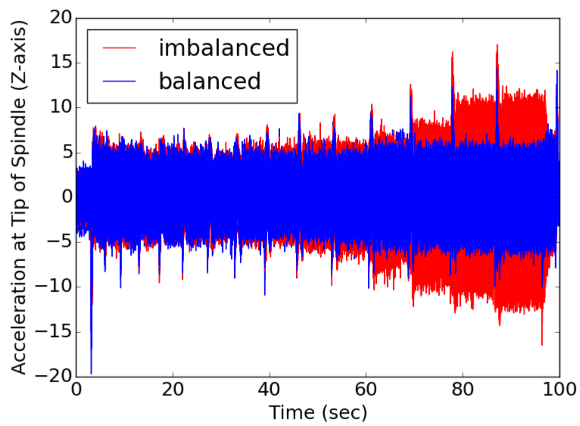
For this sensor data analysis, the features extracted are (i) average, (ii) standard deviation, (iii) maximum amplitude of FFT, and (iv) frequency for maximum amplitude of FFT. These four features are inspected visually to determine if im-



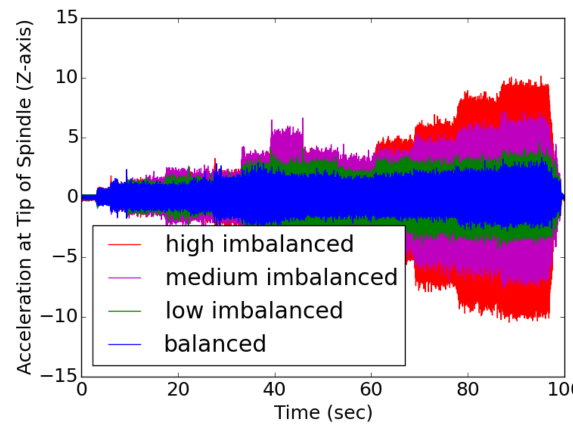
(a) Spindle x-axis acceleration: 2009 data.



(b) Spindle a-axis acceleration: 2010 data.



(c) Spindle z-axis acceleration: 2009 data.



(d) Spindle z-axis acceleration: 2010 data.

Figure 2. Spindle acceleration data for different imbalance level.

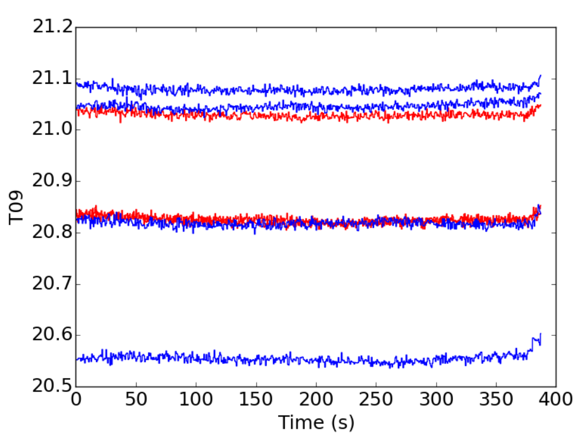
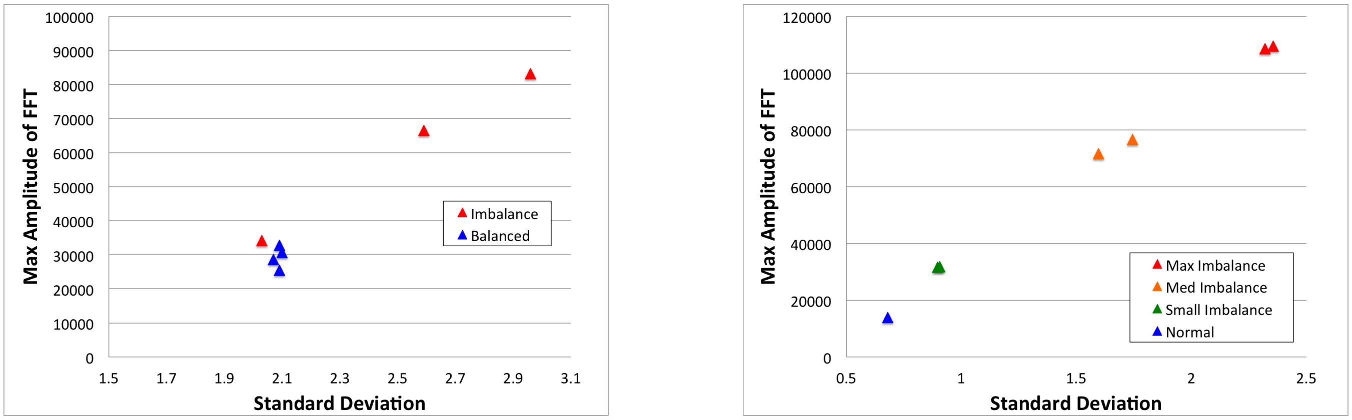


Figure 3. Sample temperature sensor Data (fluid temperature): blue and red traces indicate nominal and faulty conditions respectively.

balance could be classified by a simple linear classifier. The spindle acceleration (X, Y, and Z) feature (maximum amplitude of FFT) showed easily visible characteristics that can distinguish between degrees of imbalance. See Figure 4 for an example of visual classification based on X-axis acceleration data. Other sensor signals like power, pH, flow, and temperature did not exhibit such classification capability.

3.1.2. Control-based Segmentation followed by Sensor-based Diagnostics

The second diagnostic approach that we explored combines both sensor and control data in a coherent manner. The first step in this approach is to utilize the control signal to provide temporal segmentation, i.e., assuming quasi-steady state, the goal is to find the time intervals in which the following conditions are satisfied: (i) all experiments display same values for the primary control signal (actual spindle speed) , and (ii) all the control signals are constant over the same period. Note that, to investigate the dynamic response, rather than quasi



(a) Spindle a-axis acceleration for 2009 data.

(b) Spindle x-axis acceleration for 2010 data.

Figure 4. Visual classification using spindle x-axis acceleration sensor.

steady state response, the control signals should be consistent across the experiments so that responses are compared under the same set of control inputs. Figure 5 (a) shows the result of this temporal segmentation scheme. For each of the control signals, we have computed the standard deviation at the each time step and identified the periods with standard deviation below a set threshold to find the consistent time intervals (shown as colored segments along the time axis in Figure 5 (b)). Then we find the intersection of the sets of consistent time intervals over all the control signals to determine the aggregate time intervals over which the control signals are statistically consistent (shown as black segments along the time axis in Figure 5 (c)).

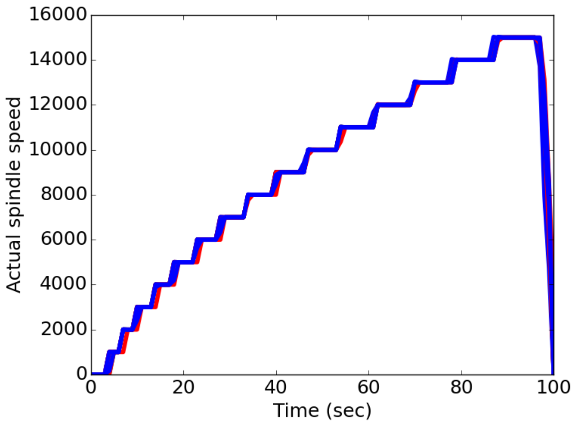
These temporal segments are then mapped to sensor data to facilitate diagnostics. For each of 16 temporal segments, we computed features including (i) average, (ii) standard deviation, (iii) maximum FFT amplitude value, and (iv) FFT frequency at maximum amplitude. This step produces a 64 dimensional feature space to diagnose machine imbalance. As mentioned before, to avoid the overfitting we focus on linear transformation based approaches. We implemented Principal Component Analysis (PCA) to reduce the dimensionality from 64 to 4 (postulating that there should be 4 unique dimensions given the 4 uncorrelated features that we have selected). The PCA step is followed by Linear Discriminant Analysis to find the optimal coordinate transformation that provides maximum separation between classes. Result of this PCA-LDA analysis is shown in Figure 6 for Fluid Temperature sensor data. Another temperature sensor located at Spindle Motor also exhibits similar diagnostic capability after application of control based temporal segmentation. This demonstrates that control data can be used to provide context to sensor data in a way that helps diagnose machine imbalance. Thus, temperature sensor which had inferior diagnostic performance without context data, could classify imbalance perfectly when it is combined with additional context from control signal.

3.2. Spindle Remaining Useful Life Prediction

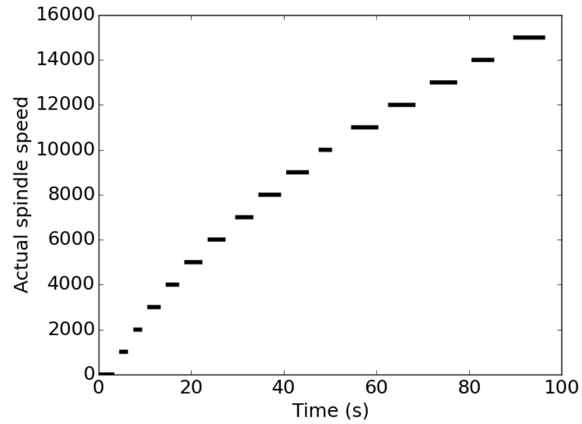
3.2.1. Experiment Setup and Data

The spindle test-bed was built at TechSolve using a frequency drive, a motor, a poly-V belt transmission, and a simplified spindle using two bearings identical to the ones used in the horizontal machining center. Figure 7 presents a picture of the spindle test-bed showing the motor, the belt transmission, and the actual spindle. A loading mechanism pulling on the nose of the spindle was added to accelerate the degradation. The force was kept constant as 35 lbs. The spindle was rotating at a constant speed of 9120 resolution per minute. The spindle motor was shut down automatically by the frequency drive when the bearing was locked at the end of life. A current sensor was installed on one phase of the power cable in the control box which controls the speed of the motor. An accelerometer was installed on the housing of the spindle to collect vibration data. Four thermal couples were installed to collect temperature data of the motor, spindle bearing, loading deck and ambient temperature, respectively. The sampling rate was 25600 Hz and 765440 data points were collected every hour.

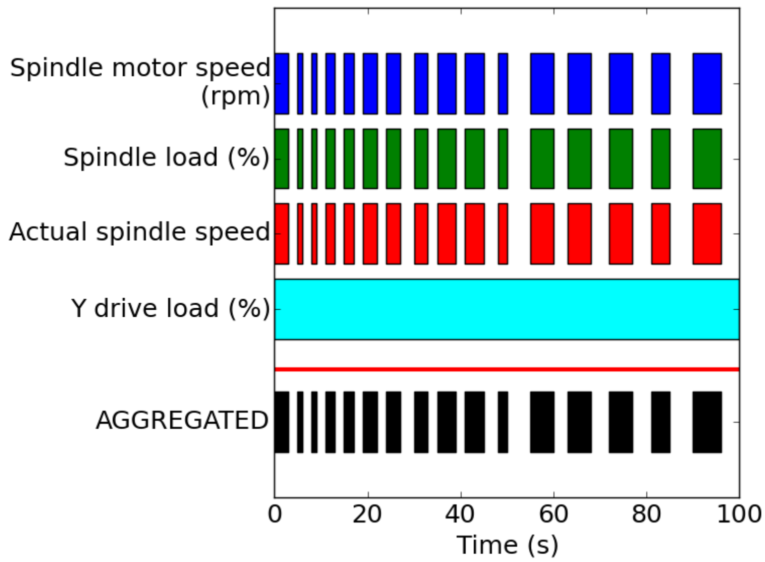
We first examined the time series statistics of the vibration signal and energy based on the rotational speed which was 152 Hz. The purpose was trying to find the feature(s) that has(have) trending though the life time so that the feature(s) can be used for prediction based on the historical trend. The result revealed that there was no obvious trending in none of the features at least from visual inspection. Figure 8 showed some of the features that we examined. The solid red lines were the smoothed features using a moving average window of length 3.



(a) Raw spindle speed control.



(b) Spindle speed control with consistent time segment.



(c) Aggregating control signals.

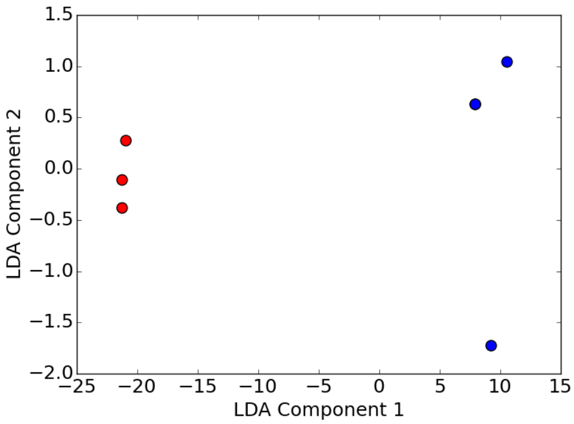
Figure 5. Time series segmentation.

3.2.2. RUL Prediction Using Features Generated by RBM

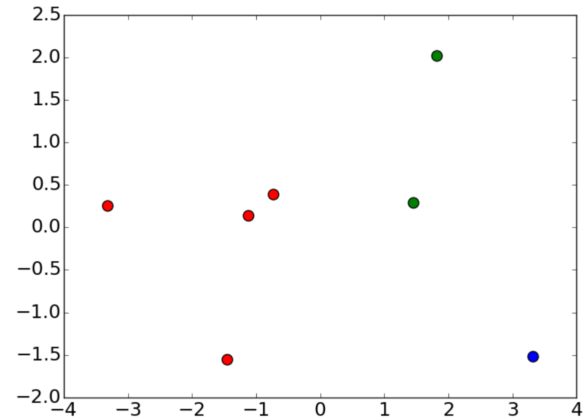
Since there was no obvious trending in the features that we examined, the RUL prediction was supposed to be not straight forward. We would like to use the proposed RBM method to generate features automatically. In order to test the generative capability of the RBM and assuming there is no engineering guidance on feature extraction, we arbitrarily selected the frequency amplitude values ranging within 76 Hz and 532 Hz. It covered the frequency range upto 3.5 times of the rotating frequency. The frequency amplitude values were used as the input to the RBM. The RBM learning parameters were chosen by trial and error because there was no general guidance available to a practical application. The number of hidden nodes was set to 1000. The number of maximum of epochs was set to 100. The learning rate was chosen as 0.1 and it was fixed through the iterations.The momentum was set to be 0.1. The number of iteration for Gibbs sample of Contrastive Divergence algorithm was set to be 1.

After training, we input the training data itself to the trained RBM, and selected the hidden units whose standard deviations are greater than zero. The purpose was to avoid those hidden units that did not contain any variance information. As a result, 88 hidden nodes were selected.

The ground truth of RUL was calculated by the hours from the time stamp when the data was collected to the failure time stamp. All RUL hours were normalized to be in the range between 0 and 1. The selected 88-dimensional features and

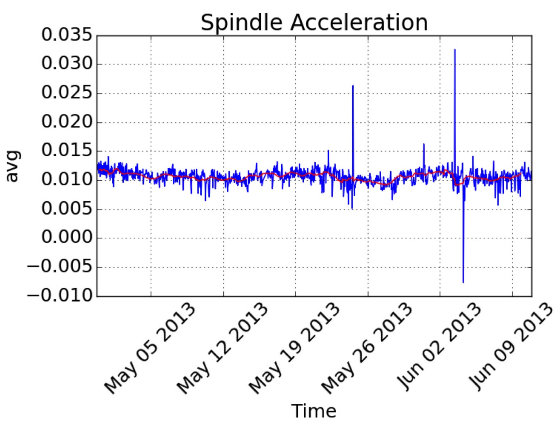


(a) Group 1.

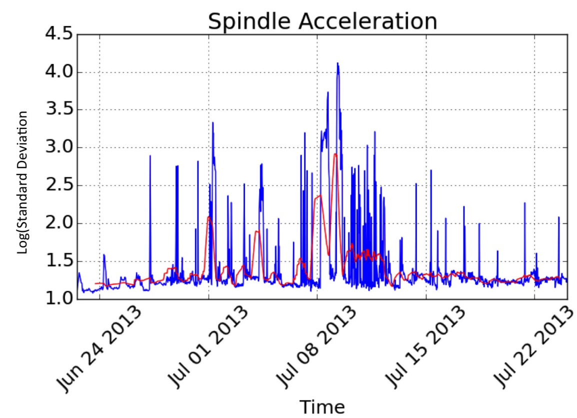


(b) Group 2.

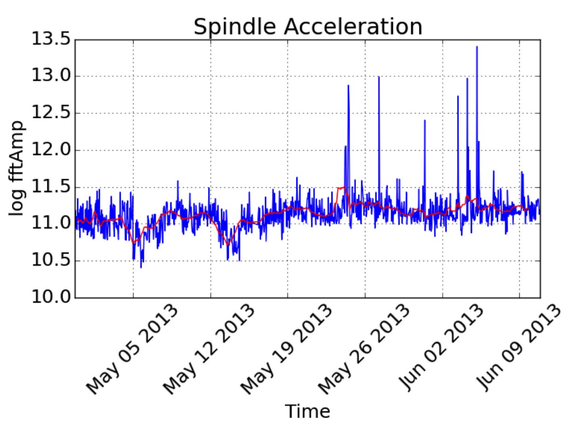
Figure 6. PCA-LDA result using fluid temperature.



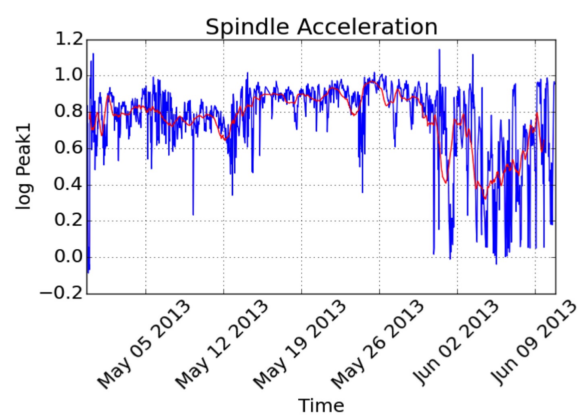
(a) Average acceleration.



(b) Log of acceleration standard deviation.



(c) Log of maximum FFT amplitude.



(d) Log of maximum acceleration peak.

Figure 8. Vibration features.

the corresponding normalized RUL were trained by the random forest regression algorithm. A five-fold cross validation was used to validate the prediction. For comparison purposes, we also directly applied the random forest regression to the

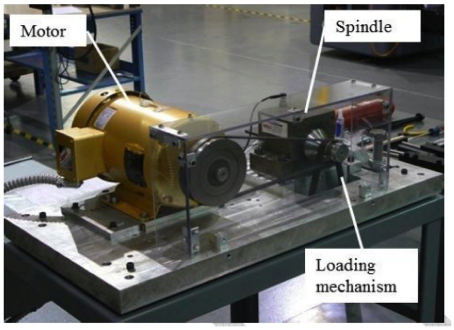


Figure 7. Spindle testbed.

Table 1. Comparison of prediction methods.

Method	Avg. RMSE	Avg. MAE
Random Forest	0.077	0.065
Random Forest + RBM	0.047	0.029

pre-processed data without using RBM feature generation. A five-fold cross validation was also applied for this case.

The result was shown in Table 1, which showed that the performance of the proposed RBM and random forest RUL prediction method was superior to the random forest method without using RBM to generate features in terms of both RMSE (Root Mean Square Deviation) and MAE (Mean Absolute Error) criterion. To show the RUL prediction result, 90% of the data was randomly selected for training and the rest was used for testing. The testing result was plotted in Figure 9. The solid line was the ground truth of RUL and the circles were the predicted RUL values. The predicted RUL was very close to the ground truth and the prediction RMSE was 0.043.

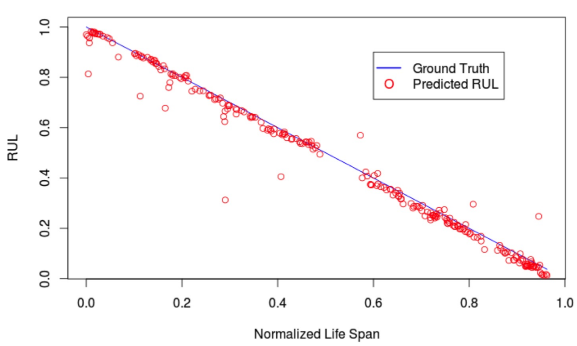


Figure 9. RUL Prediction

4. CONCLUSION

We proposed a method based on principal component analysis and linear discriminant analysis to combine control and sensor signals for machine health condition estimation. This work explored various types of sensory and control data for diagnosing the imbalance conditions of the machine tools. Our finding was that by combining context information gained from the control signal, certain sensors can have better sensitivity in diagnosing faulty conditions. In our case, the temperature sensor was able to classify machine imbalance conditions with much higher sensitivity than using itself alone. For practical implementation, using thermal couples can reduce the cost of both sensors and data storage comparing to using accelerometers.

For future work on device health estimation, we will explore diagnostics based on control signal alone. Given that relying on sensor data typically requires adding sensors to existing machine tools, it would be ideal if we could diagnose imbalance of the machine from control signals that are usually recorded (i.e. no additional hardware required). The expectation is that if a machine tool uses feedback controls, then the control signal should be impacted by any change in the operational characteristics (in this case the imbalance of the machine tools).

We also proposed a novel method of using Restricted Boltzmann Machine as a feature generation model and coupling with a random forest algorithm in remaining useful life prediction applications. RBM has been explored in many classification scenarios,but it hasn't been explored in the RUL prediction scenario to our best knowledge. The run-to-failure test showed RBM can potentially generate useful features for RUL prediction with high accuracy.

For future work on RUL prediction, considering using a discriminative restricted Boltzmann machine (Larochelle & Bengio, 2008) model to integrate feature extraction and prediction as a unified task is of interest. One of the advantages is that as a unified task, model selection, parameter tuning, and initialization can be done only once comparing to using two learning phases (feature generation followed by prediction). Currently, the features generated from RBM barely have physical meaning or are hard to explain. If an extra objective term can be added in addition to the energy function 1 of RBM, it may generate features that are easier to understand e.g. a feature which has a better trending over the life span.

REFERENCES

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, *41*(3), 15.

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

Deng, H., Runger, G., Tuv, E., & Vladimir, M. (2013). A time series forest for classification and feature extraction. *Information Sciences*, *239*, 142–153.

Guo, H., Watson, S., Tavner, P., & Xiang, J. (2009). Reliability analysis for wind turbines with incomplete failure data collected from after the date of initial installation. *Reliability Engineering & System Safety*, *94*(6), 1057–1063.

Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, *20*(7), 1483–1510.

Koehler, G. J., & Erenguc, S. S. (1990). Minimizing misclassifications in linear discriminant analysis*. *Decision sciences*, *21*(1), 63–85.

Larochelle, H., & Bengio, Y. (2008). Classification using discriminative restricted boltzmann machines. In *Proceedings of the 25th international conference on machine learning* (pp. 536–543).

Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., & Siegel, D. (2014). Prognostics and health management design for rotary machinery systems-reviews, methodology and applications. *Mechanical Systems and Signal Processing*, *42*(1), 314–334.

Liao, L. (2014). Discovering prognostic features using genetic programming in remaining useful life prediction. *Industrial Electronics, IEEE Transactions on*, *61*(5), 2464–2472.

Liao, L., & Kottig, F. (2014). Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction. *Reliability, IEEE Transactions on*, *63*(1), 191–207.

Liu, Q. C., & Wang, H.-P. B. (2001). A case study on multisensor data fusion for imbalance diagnosis of rotating machinery. *AI EDAM*, *15*(03), 203–210.

MacGillivray, C., Turner, V., & Lund, D. (2013). Worldwide internet of things (iot) 2013–2020 forecast: Billions of things, trillions of dollars. *IDC. Doc*, *243661*(3).

Markos, M., & Singh, S. (2003). Novelty detection: a review-part 2: neural network based approaches. *Signal Processing*, *83*(12), 2499–2521.

Markou, M., & Singh, S. (2003). Novelty detection: a review-part 1: statistical approaches. *Signal processing*, *83*(12), 2481–2497.

Pavel, R., & Iverson, J. C. (2012, April). Machine tool health characterization using condition monitoring and prognostic based technologies. In *MFPT 2012 conference.* Dayton, OH.

Pavel, R., Snyder, J., Frankle, N., Key, G., & Miller, L. (2010, April). Machine tool health monitoring using prognostic health monitoring software. In *MFPT 2010 conference.* Huntsville, AL.

Standard, M. (2009). Part 1-overview and protocol, version 1.01. *MTConnect Institute*.

Tamilselvan, P., & Wang, P. (2013). Failure diagnosis using deep belief learning based health state classification. *Reliability Engineering & System Safety*, *115*, 124–135.

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, *2*(1), 37–52.

Yang, B., Chen, S., & Wu, X. (2011). A structurally motivated framework for discriminant analysis. *Pattern Analysis and Applications*, *14*(4), 349–367.

BIOGRAPHIES

Linxia Liao Linxia Liaos research interests include predictive and prescriptive analytics; system and components fault diagnostics and prognostics; signal processing; and machine learning algorithms as well as their integration on embedded systems. Currently, he is developing solutions for device health prognostics and patient disease progression modeling. Prior to joining PARC, Linxia worked as a research scientist with Siemens Corporate, Corporate Technology (previous Siemens Corporate Research) located in Princeton, NJ. He conducted research and implemented various prognostics and health management-related applications in the fields of manufacturing, energy, and transportation. He also previously worked at Siemens Technology-To-Business (TTB) Center in Berkeley, CA to transfer the patented Methods for prognosing mechanical systems technology from the university to industry applications. Dr. Liao received his Ph.D. degree in Industrial Engineering from the University of Cincinnati, where he conducted research at the NSF I/UCR Center for Intelligent Maintenance Systems (IMS). He earned his M.S. and B.S. degrees in Mechanical Science and Engineering at Huazhong University of Science and Technology (HUST) in China. Linxia has one issued patent and seven pending patents, and he has published one book chapter and 25+ papers in leading journals and conferences.

Tomonori Honda Tomonori Honda received his Ph.D. and M.S. in Mechanical Engineering from California Institute of Technology and B.S. in Mechanical Engineering and Nuclear Engineering from University of California, Berkeley. He had continued his Ph.D. research at Massachusetts Institute of Technology as a Postdoctroal Scholar and Research Scientist for several years. He has published over 30 papers in journals and peer-reviewed conferences. He has also won the 2013 JALA Ten Award from Society of Laboratory Automation and Screening (SLAS) and the 2013 ASME Design Theory and Methodology Best Paper Award. Prior to joining PARC, Tomonori served as a principal data scientist at Edmunds.com, where he led the modeling and analytics effort including reinforcement learning framework for realtime ads targeting using big data. He also worked as a junior analytics manager at Opera Solutions, where he explored different techniques to improve model performances and won an in-house modeling competition. He is a member of ASME, AIAA, and AAAI, and has been a reviewer for several engineering design journals including ASME Journal of Mechanical Design, Research in Engineering Design, Journal of Engineering Design, and IEEE System Journal. His main research areas include design for system reliability, prognosis, and maintenance scheduling, design synthesis for complex multidisciplinary systems, and behavioral design theory.

Radu Pavel Dr. Radu Pavel has over 20 years of research and teaching experience in mechanical and manufacturing engineering. Dr. Pavel is knowledgeable in the areas of modeling, design of experiments, test-bed instrumentation, data analysis, sensor technology, and training. His expertise include advanced grinding and finishing techniques for new materials such as superabrasive grinding of ceramics and composites, ELID grinding, and hard turning. Dr. Pavel has published multiple papers in refereed conference proceedings and journals, and organized symposia focused on advances in material processing and inspection for the ASME International Manufacturing Science and Engineering Conferences.

Dr. Pavel obtained his bachelor, master and doctoral degrees in mechanical engineering from University "Politehnica" of Bucharest, Romania. Dr. Pavel also obtained a Ph.D. in Manufacturing Engineering from Univerity of Toledo, Ohio, U.S.A.

Dr. Pavel was hired by TechSolve in February 2004 on the position of machining/grinding research engineer. For the first three years he has been a part of the research team of the "Intelligent Optimization and Control of Grinding Processes" project sponsored by the NIST-ATP division. Next, Dr. Pavel was selected to be part of the team working for the Smart Machine Platform Initiative (SMPI) program. His main focus areas were Tool Condition Monitoring and Condition-Based Maintenance for the Smart Machine. Under the SMPI program, Dr. Pavel evaluated innovative technologies, and contributed to the development of new technologies including those oriented towards the diagnostics and prognostics of health and maintenance of the machine tool.

Dr. Pavel was also the technical lead for the Adaptive Machining Program (AMPI). In this program, vision and laser sensing technologies were used to accurately measure the position and shape of a workpiece in the three dimensions (3D), before machining. The goal is to adapt the cutter path to the actual geometry of the workpiece and avoid the inaccuracies resulting from a cutter path derived from the ideal virtual model of the part. The 3D scanning and point-cloud processing techniques were used to determine the actual geometry of the part before and after machining.

Dr. Pavel is currently TechSolve' s Chief Technology Officer, leading a team of engineers from the Advanced Machining R&D department.