

Anomaly Detection Using Self-Organizing Maps-Based K -Nearest Neighbor Algorithm

Jing Tian¹, Michael H. Azarian², and Michael Pecht³

^{1, 2, 3} *Center for Advanced Life Cycle Engineering (CALCE), University of Maryland, College Park, MD, 20742, U.S.A.*

jingtian@calce.umd.edu

mazarian@calce.umd.edu

pecht@calce.umd.edu

ABSTRACT

Self-organizing maps have been used extensively for condition-based maintenance, where quantization errors of test data referring to the self-organizing maps of healthy training data have been used as features. Researchers have used minimum quantization error as a health indicator, which is sensitive to noise in the training data. Some other researchers have used the average of the quantization errors as a health indicator, where the best matching units of the trained self-organizing maps are required to be convex. These requirements are not always satisfied. This paper introduces a method that improves self-organizing maps for anomaly detection by addressing these issues. Noise dominated best matching units extracted from the map trained by the healthy training data are removed, and the rest are used as healthy references. For a given test data observation, the k -nearest neighbor algorithm is applied to identify neighbors of the observation that occur in the references. Then the Euclidean distance between the test data observation and the centroid of the neighbors is calculated as a health indicator. Compared with the minimum quantization error, the health indicator extracted by this method is less sensitive to noise, and compared with the average of quantization errors, it does not put limitations on the convexity or distribution of the best matching units. The result was validated using data from experiments on cooling fan bearings.

1. INTRODUCTION

Anomalies are patterns in data that do not conform to a defined notion of normal behavior (Chandola, Banerjee, & Kumar, 2009). Anomaly detection is used in the prognostics and health management (PHM) of mechanical and electronic systems to detect the existence of a fault before failure happens. The performance of currently available

anomaly detection methods leaves room for improvement because some systems are still failing without warning. For example, even though maintained regularly, bearings remain the top contributor to failures of systems like computer cooling fans (Tian, 2006) and induction motors (Bianchini, Immovilli, Cocconcelli, Rubini, & Bellini, 2011).

The data used in anomaly detection for mechanical and electronic systems are signals that are sensitive to faults. For example, in rotating machinery, time series like vibration signals and motor current signals have been used because they are sensitive to faults, widely available, and non-intrusive. Some other signals like acoustic emission signals were found to be sensitivity to a fault at an early stage (Oh, Azarian, & Pecht, 2011), and they have been used as precursor parameters in the health monitoring of cooling fan bearings (Oh & Shibutani, 2012).

Sensor signals may not be adequate for users to identify an anomaly of the system so fault features have been extracted from the sensor signals to increase separation of the normal and abnormal behavior of the system. For time series signals, commonly used features include peak-to-peak, rms, and kurtosis of the signal's amplitude in the time domain, characteristic frequency components, wavelet coefficients, and empirical mode decomposition energy in frequency and time-frequency domains. Some researchers have introduced more sophisticated features (Tian, Morillo, & Pecht, 2013).

The extracted features need to be transformed to understandable information to determine if a test observation is an anomaly or not. There are two approaches to perform this task. One is the physics-of-failure (PoF) approach. Variables of PoF models are monitored and compared to the calculated value from the model. When deviation of the monitored value from the model value exceeds a predetermined threshold, an anomaly is identified. Another is data-driven approach, where data mining techniques are applied to explore the structure of the data of the extracted features. Based on the structure, deviation of the system from a normal state is estimated. The PoF

approach requires physical models of the system failure mechanisms, which are not available in many applications. The data-driven approach does not have this requirement, but it needs more data than the PoF approach. With the rapid development of data acquisition techniques, the obstacle to obtain data is weakened, and therefore data-driven approaches are preferred in many applications.

The data-driven approach is usually realized by machine learning techniques. Based on the use of the data, machine learning techniques can be classified as supervised machine learning techniques and unsupervised machine learning techniques (Pecht, 2008). To detect a fault, supervised machine learning techniques require healthy training data and faulty training data to construct regions of healthy conditions and faulty conditions, and then a test data observation is classified to be healthy or faulty depending on which region it falls into. In anomaly detection, representative supervised machine learning techniques include support vector machine (SVM) (Sotiris, Tse, & Pecht, 2010) and k -nearest neighbor (KNN) algorithms (He, Li, & Zhu, 2013). Application of these techniques is limited by the availability of training data of anomalies.

Unsupervised machine learning techniques do not need training data. They group observations into different clusters according to their mutual similarity. For example, during clustering, normal data and anomalies have different performance. Normal data may form large and dense clusters, and anomalies may form small and sparse clusters. Popular unsupervised machine learning techniques for anomaly detection in mechanical and electronic systems include self-organizing maps (SOM) (Huang, Xi, Li, Liu, Qiu & Lee, 2007) and k -means clustering (Wang, Liu, & Cui, 2012). Success of these techniques depends on the assumed relationship between the characteristics of the clusters and the anomalies.

In many cases, normal data are abundant and the anomalies that can be used for training are scarce. Semi-supervised learning techniques are preferred in these cases. Some researchers identify the class for normal data and use these data as references to calculate the Mahalanobis distance (MD) of the test data (Jin, Ma, Cheng, & Pecht, 2012). The test data are classified as anomalies if their MD values are above a certain threshold. When the normal data are distributed in several clusters, current applications of MD cannot reflect the degree of deviation of the test data from being normal. Some researchers have used self-organizing maps (SOM) to cluster the data in terms of best matching units (BMUs) (Huang et al. 2007). The smallest distance of a test data observation to the BMUs, which is called the minimum quantization error (MQE) is used as an indicator for anomaly detection. In the presence of noise, which is introduced into the signals via sources like other interfering signals and errors of measurements, MQE can be the

distance of the test data observation to a noise-dominated BMU, resulting in false detection.

In this study, the semi-supervised application of SOM in anomaly detection is improved. After the maps are trained by normal training data, some BMUs are removed to reduce the influence of noise, and the neighbors in the BMUs of a given test data observation are identified by the k -nearest neighbor algorithm. Then the Euclidean distance between the test data observation and the centroid of the neighbors is calculated as an anomaly indicator.

The rest of the paper is organized as follows: in section 2, the theoretical background of SOM and its application in system health monitoring are introduced. The SOM-based KNN algorithm developed in this study is introduced in section 3, and the algorithm is validated with an experimental study in section 4. In section 5, conclusions from this study are presented.

2. SELF-ORGANIZING MAPS

Self-Organizing Maps (SOM), also called Kohonen neural network, is a type of unsupervised machine learning technique based on competitive learning (Kohonen, 1990). It creates a network that maintains information on the topological relationships within the training data.

2.1. Theoretical Background of SOM

An SOM consists of a number of neurons. Each neuron is represented by a weight vector that has the same dimension of the training data. The neurons are organized according to their similarity where the neurons with the similar weight vectors are grouped as neighbors. This neighborhood relationship describes the structure of the map, which reflects the relationship in the training data.

To create an SOM, at first the input data is normalized per variable by calculating the z-score of each observation. The size of the map is then determined by calculating the number of neurons from the number of observations in the training data using Eq. (1).

$$M \approx 5\sqrt{N} \quad (1)$$

where M is the number of neurons, which is an integer close to the result of the right hand side of the equation, and N is the number of observations.

The neurons are organized in a 2-dimensional map. The ratio of the side lengths of the map is approximately the ratio of the two largest eigenvalues of the training data's covariance matrix.

Elements of the weight vector of each neuron are initialized randomly. A training data observation is then picked as an input vector to calculate its Euclidean distance between all the neurons. For each input observation, the neuron that has

the minimum distance is found. This neuron is called the best matching unit (BMU) of that input observation. Neighbors of the BMU are selected, and their weight vectors are updated using a neighborhood function in described Eq. (2).

$$h_{ci}(t) = a(t)e^{(-\frac{\|r_c - r_i\|^2}{2\sigma^2(t)})} \quad (2)$$

where h_{ci} is the neighborhood function between the BMU c and a neuron i . t is the index of iterations of training. a is the learning rate. r_c is the vector of the BMU c , and r_i is the vector of neuron i . σ is the radius around c .

The neighborhood function is a non-increasing function of t and the distance between neuron i and the BMU c so that the neurons close to the BMU c are moving closer to c and the rate of moving is decreasing over the iterations of training.

The neurons are updated according to Eq. (3).

$$W_i(t+1) = W_i(t) + h_{ci}(t)[x(t) - W_i(t)] \quad (3)$$

where $W_i(t)$ is the weight vector of neuron i at t^{th} iteration of training. h_{ci} is the neighborhood function, and $x(t)$ is the input observation of the BMU c .

The SOM is trained iteratively until all the weight vectors of the map are grouped into clusters according to their distance. When the learning process finished, the SOM is created. The procedure is summarized in Figure1. Details of SOM can be found in (Kohonen, 1990).

2.2. Application of SOM in Mechanical and Electronic System Health Monitoring

Researchers have explored the performance of SOM in health monitoring of mechanical and electronic systems where minimum quantization error (MQE) of a test data observation to the SOM has been used as a indicator to evaluate the health of the system (Qiu, Lee, Jin, & Yu, 2003).

Quantization error describes the distance between the input data observation and the BMU of the SOM. MQE is calculated as in Eq. (4):

$$Q = \min_k \|D - B_k\| \quad (4)$$

where Q is the MQE, D is a test data observation, and B_k is the weight vector of the k^{th} BMU.

To monitor health conditions, at first the SOM is trained by the healthy training data, and then the MQE of a test data observation to the SOM is obtained. Large MQE indicates that the test data observation belongs to a space which is not covered by the training data. Based on the assumption that any deviation from the space covered by the normal training

data is regarded as a deviation of the system from being normal, MQE can be used to indicate the severity of the system's deviation from normal. This assumption is evaluated in the studies of (Kang & Birtwhistle, 2003). When MQE values are calculated for different stages in the life cycle of the system, the trend of the system's health condition is obtained.

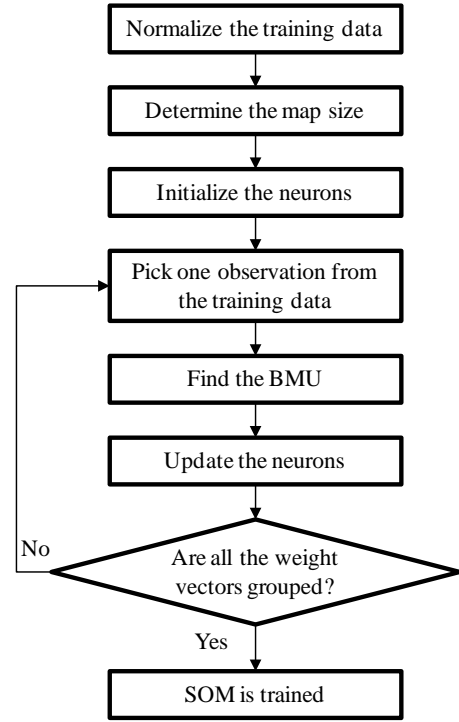


Figure 1. Flow chart of the training process of SOM

In practical situations, the normal training data are inevitably contaminated by noise. It is likely that during the training process, noise may have dominant influence on some BMUs in the map. During the testing process, when a test data observation is close to one of the noise dominated BMUs, its value of MQE is small, and it would be classified as normal. As a result, false detection may occur.

One method to reduce the influence of noise dominated BMUs is to use the average of all the quantization errors as an indicator. This is equal to the Euclidean distance between the test observation and the centroid of all the BMUs. However, if the BMUs are distributed in different clusters, or if they are non-convex, the centroid of the BMUs may fail to represent the collective information of the BMUs. Therefore, a method is needed to improve the application of SOM in anomaly detection under noisy conditions.

3. SELF-ORGANIZING MAPS-BASED K -NEAREST NEIGHBOR ALGORITHM

As discussed in section 2, MQE is subject to the influence of noise in the training data, and the average of quantization errors fails to work for the training data that are non-convex or have isolated clusters. These shortcomings can be overcome by selecting a subset of the BMUs and calculating their average quantization errors as an anomaly indicator.

At first, a threshold is applied to the BMUs to remove the noise dominated BMUs. The normal training data contains information of both the dynamics of the system and noise. The dynamics of the system are stable and the data should concentrate on certain neurons in the SOM. As a result, some neurons become BMUs multiple times. The noise is random and the data from the noise do not concentrate on any neuron. As a result, even if some neurons become BMUs because of the noise, these neurons do not become BMUs very often. By removing the BMUs with relatively few hits, the influence of noise can be reduced.

A subset of the BMUs is then selected. The average quantization error of a test observation to the BMUs in the subset is calculated as an anomaly indicator. Using the subset of BMUs has two benefits. First, by calculating the average of the quantization errors of the subset, the influence of noise is further reduced. Second, for a certain size of the subset in a local region, the data of the subset can be confined to the same cluster and be approximately convex, and therefore, the centroid of the subset is representative of the health condition of this subset.

A main task is to select the BMUs that form the subset as the normal reference. In this study, the BMUs in the subset are selected as the nearest neighbors of the test data observation. If one nearest neighbor is selected, the health indicator is the MQE. If k nearest neighbors are selected, the health indicator is calculated as the average of the MQE, the second minimum quantization error, and up to the k^{th} minimum quantization error. By including multiple nearest neighbors, the influence of noise is reduced.

Identification of the nearest neighbors is performed by the k -nearest neighbor (KNN) algorithm. In most cases, KNN is used as a classification technique, where a test data observation is classified to a class if it is closer to the nearest neighbors in that class. In this study, KNN is used as a semi-supervised learning technique, where KNN is only used to identify the nearest neighbors in the reference BMUs. The distance of the test data observation to the centroid of the identified neighbors is calculated. In this study this distance is called KNN distance. It is used as the health indicator. The use of KNN in this study is illustrated in Figure 2.

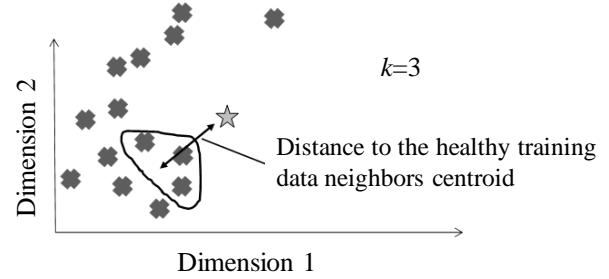


Figure 2. KNN distance of a test data observation when $k=3$

A flow chart of the method developed in this study is shown in Figure 3.

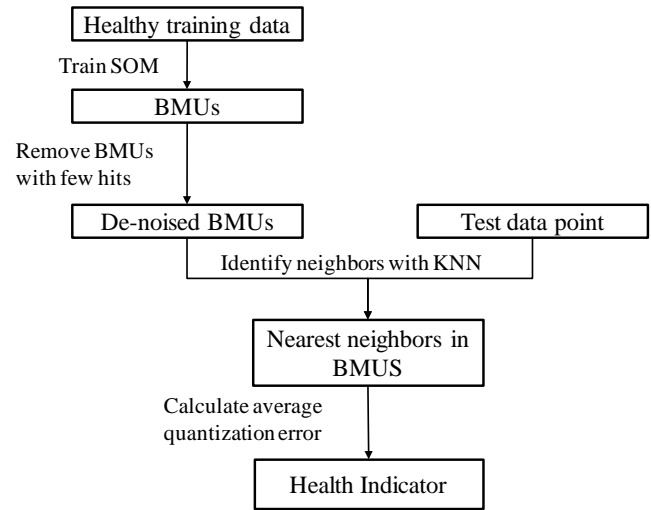


Figure 3. Flow chart of the SOM-based KNN Algorithm

In anomaly detection, the method is first applied to the healthy training data to get the sample of the value from the health indicator of the healthy system. A percentile of the sample is then selected as the anomaly threshold.

4. EXPERIMENTAL STUDY

The data from a cooling fan accelerated life test was used to validate the method developed in this study. The data have a tendency to form several clusters and they contain noise, which can present difficulties when used with conventional methods. The SOM approaches which involve either directly using the MQE or taking the distance to the centroid of multiple BMUs as the healthy reference produce erroneous detection results when used with this type of data or with non-convex data. The method developed in this paper was designed to address these two issues, which is demonstrated using data collected from the cooling fan bearing in the experiment.

4.1. Setup of the Experiment

A new cooling fan with ball bearings was tested. The ball bearings were designed to be lubricated by grease and oil. To accelerate the degradation, the bearings were only lubricated by oil. After an initial measurement, the cooling fan was run at its rated speed of 4,800 rpm in a chamber at the fan's maximum rated temperature of 70 °C. The cooling fan under test is shown in Figure 4.



Figure 4. The cooling fan under test

4.2. Data Acquisition

The vibration acceleration signal and the motor current signal have been identified as sensitive to bearing faults (Immovilli, Bellini, Rubini, & Tassoni, 2010). The two signals were monitored in this study. The measurements were collected while the cooling fan was run at room temperature for a brief time between stressing periods. Signals collected at each measurement have a time span of 10 seconds and consist of 1,024,000 observations, where the sampling rate is 102,400 Hz. Before the accelerated life test, three measurements of signals were collected as training data, which form a 3,072,000 by 2 matrix. Each row is an observation, and each column is a signal. Test data were collected after 0 hours, 8 hours, 16 hours, 24 hours, 48 hours, and 72 hours of accelerated life testing, which form 6 stages of the test. At each stage there was one measurement of the signals, which form a 1,024,000 by 2 matrix. The 0 hour signal was one of the three measurements from the training data.

The data were cut into segments sequentially. Each segment has 0.2 seconds of data, each containing 20,480 observations. For one measurement, there are 50 segments. Features were extracted from these segments. The structure of a measurement is shown in Table 1.

Table 1. Structure of a measurement

Segment index	Observation index	Vibration	Current
1	1	0.06	-6.84
	2	0.12	-6.66

	20480	0.11	5.66
2	20481	0.15	5.75

	40960	0.26	6.34
...

50	921601	-0.23	5.84

	1024000	0.29	-6.98

4.3. Feature Extraction

Some commonly used fault features were extracted from the segments of the data for both the vibration signal and the current signal. These features include peak-to-peak, rms, standard deviation, skewness, and kurtosis of the amplitude. For each signal, there are 5 features, and for both the vibration and current signals, together there are 10 features. After feature extraction, the data of each measurement is a 50 by 10 matrix, where the row is an observation of the features, and the column is a feature.

4.4. Analysis Result

All the data were normalized by calculating z-scores referring to the mean and standard deviation of the training data. The size of the SOM was determined according to Eq. (1). The training data have three measurements, each of which has 50 observations, so there are 150 observations for training. According to Eq. (1), the map size was determined as 9 by 7 with 63 neurons. Each neuron is a vector with 10 elements, corresponding to the number of features.

After training, BMUs were identified in the map, as shown in Figure 5. Each lattice cell represents a neuron, and the number in a cell is the number of times the neuron has become a BMU, or the number of hits. The map shows the data tend to form several clusters.

To determine the threshold to remove noise dominated neurons, hits of the neurons were sorted in a descending manner. The percentage of the cumulative sum of hits was plotted in Figure 6.

The x axis is the index denoting the neurons with the same number of hits. For example, 3 denotes the neurons that have 3 hits. There are 9 such neurons, and altogether they account for 45 hits in the map. The y axis is the cumulative fraction of hits for the neurons referring to the total sum of hits.

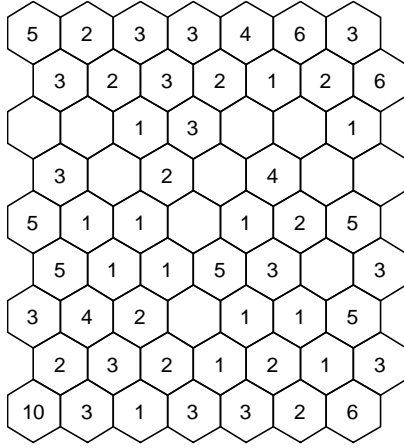


Figure 5. Hits of BMUs in the trained SOM

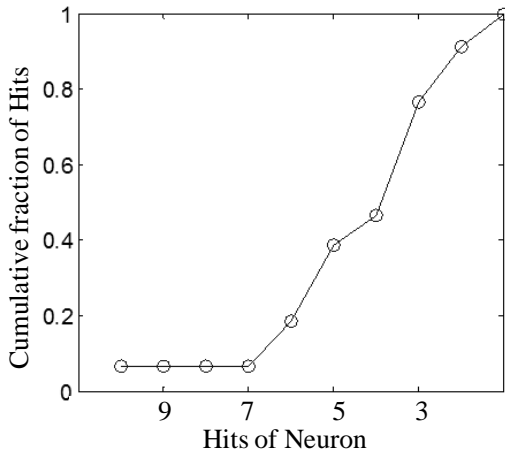


Figure 6. Cumulative fraction of hits

Among the 63 neurons, 52 neurons have become BMUs at least once. The sum of hits from neurons with more than 1 hit account for 91.3% of the total hits. If we accept that 90% of the BMU neurons are not dominated by the noise, the neurons that have 1 hit should be removed. The remaining BMU neurons were used as reference data for KNN analysis.

For each observation from the features of the test data, KNN found k nearest neighbors in the reference data, which are the BMUs. A larger k reduces influence of noise better, but it makes the algorithm more sensitive to the convexity of the data. Also, an odd value of k can help the algorithm to avoid tied votes. In this study, k was set to 3, considering that one neighbor is too sensitive to noise, and the next odd number is 3. The Euclidean distance of the test observation to the centroid of the BMUs neighbors was calculated as a health indicator. Values of the health indicator for the training data were calculated to establish a baseline for healthy condition. Distribution fitting of the health indicator value for the training data is shown in Figure 7. The Kolmogorov-

Smirnov goodness-of-fit test verified that the data could be fitted with a lognormal distribution. Using the 99.7 percentile as the threshold to separate healthy data and anomalies, the anomaly threshold on the health indicator was found to be 3.6. If the value of the health indicator of a test observation is higher than this value, the observation is classified as an anomaly.

The algorithm was applied to the data at all six stages of the test. Results are shown in Figure 8.

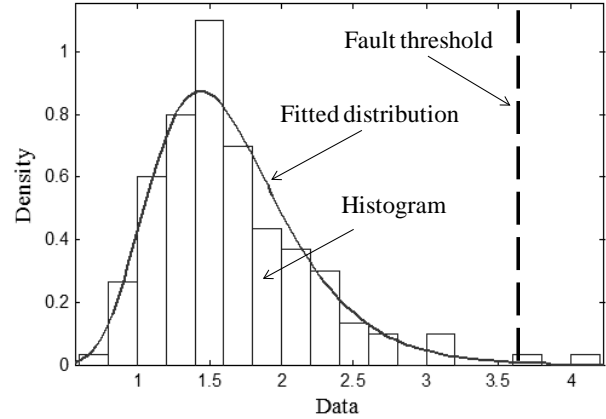


Figure 7. Distribution fitting of the health indicator of the healthy training data

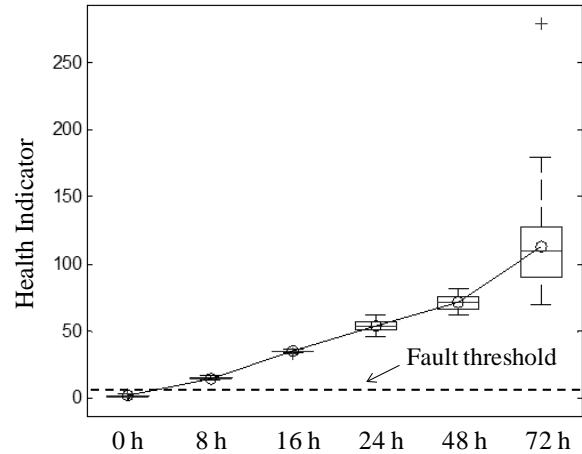


Figure 8. Health indicator for the test data

The health indicator values of the test data at each stage are shown in box plots. For each box, the central mark is the median, the edges are the 25th and 75th percentiles, and the whiskers extend to the most extreme data observations not considered outliers. Outliers are observations which are outside 2.7 standard deviations from the mean value of the data and are marked as crosses. The circles are the means. Means at different time intervals are linked by straight lines.

According to the health indicator value at each stage, the cooling fan bearings began to have anomalies after 8 hours

of test. The health indicator indicates that the bearing degraded monotonically until the end of the test after 72 hours of test, where the bearing failed with audible sound emitted. The results are consistent with the observations in the experiment. The increase of the health indicator, which is the distance between the test data to their nearest neighbors in the reference BMUs, occurred because the reference BMUs established a region representing healthy conditions of the bearings. Larger distances to this region indicate a larger deviation from the healthy conditions of the bearings. As the bearings degraded, their condition deviates from being healthy, so the distance to the healthy region, which is the health indicator, increased.

Besides the mean value, the standard deviation of the health indicator value is also increasing with the degradation of the bearings. This observation can be directly seen in Figure 8. The standard deviation of the health indicator value at each stage of the test is shown in Figure 9.

The increase of the standard deviation can be explained as occurring because, as bearings degrade, random fluctuations become more frequent and intense in the vibration signal and the current signal. Values of the fault features extracted from the signals are distributed in a wider range due to these fluctuations, and as a result, the health indicator has a larger standard deviation as it combines the information of the features.

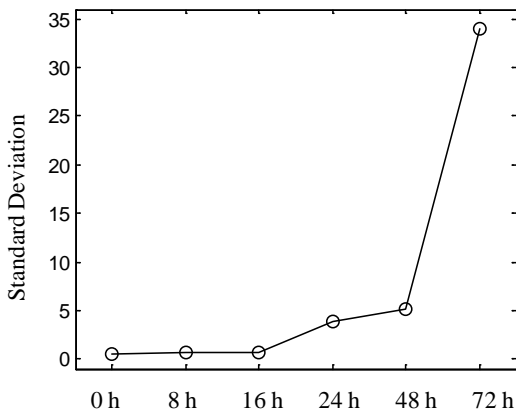


Figure 9. Standard deviation of the health indicator

In summary, although the data tended to form clusters, and contained noise, the method monitored the degradation of the bearing, and successfully detected the anomalies. The unsupervised learning method employed in this study has the benefit of reduced sensitivity to noise in the data and the ability to accommodate data non-convex distributions including data with multiple clusters. The requirements of this method are that training data are needed that sample the full range of healthy behavior (i.e., represent all the possible healthy clusters and the complexity of their distribution). This can impose practical limitations on the use of this method, since it can be costly or time consuming to collect this type of data for some systems. Furthermore, this

method needs to be combined with other algorithms for diagnostic or prognostic functions, since it is limited to anomaly detection.

5. CONCLUSIONS

This paper presents a self-organizing maps-based k-nearest neighbor algorithm for anomaly detection, which is applied in the health monitoring of mechanical and electronic systems. BMUs of the SOM trained by the healthy training data are extracted as healthy references. BMUs with small hits are removed from the references to reduce the influence of noise. For a test data observation, its Euclidean distance to the nearest neighbors in the reference BMUs is calculated as its health indicator value.

The algorithm provides a measure of health monitoring and anomaly detection of bearings where the influence of the noise from the monitoring signals is reduced by removing noise dominated BMUs and by averaging neighboring reference BMUs. The influence of the distribution of the healthy training data is reduced by using KNN to take a subset of BMUs in a local region as references. Outputs of the algorithm include a health indicator that monotonically increases with the degradation of the system, and an anomaly detection threshold on the value of the health indicator. Moreover, the standard deviation of the health indicator can also be used as a measure of degradation for the system.

The algorithm can be implemented in applications where the healthy training data are non-convex, for example, the data have several clusters. The algorithm can also reduce the influence of noise.

ACKNOWLEDGEMENT

The authors would like to thank the more than 100 companies and organizations that support research activities at the Prognostics and Health Management Group within the Center for Advanced Life Cycle Engineering at the University of Maryland annually.

REFERENCES

- Bianchini, C., Immovilli, F., Cocconcelli, M., Rubini, R., & Bellini, A. (2011). Fault detection of linear bearings in brushless AC linear motors by vibration analysis. *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 1684-1694.
- Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, pp. 15:1-15:58.
- He, D., Li, R., & Zhu, J. (2013). Plastic bearing fault diagnosis based on a two-step data mining approach. *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3429-3440.
- Huang, R., Xi, L., Li, X., Richard Liu, C., Qiu, H., & Lee, J. (2007). Residual life predictions for ball bearings based

- on self-organizing map and back propagation neural network methods. *Mechanical Systems and Signal Processing*, vol. 21, no.1, pp. 193-207.
- Immovilli, F., Bellini, A., Rubini, R., & Tassoni, C. (2010). Diagnosis of bearing faults in induction machines by vibration or current signals: A critical comparison. *IEEE Transactions on Industry Applications*, vol. 46, no. 4, pp. 1350-1359.
- Jin, X., Ma, E. W., Cheng, L. L., & Pecht, M. (2012). Health monitoring of cooling fans based on Mahalanobis distance with mRMR feature selection. *Instrumentation and Measurement, IEEE Transactions on*, vol. 61, no.8, pp. 2222-2229.
- Kang, P., & Birtwhistle, D. (2003). Condition assessment of power transformer onload tap changers using wavelet analysis and self-organizing map: field evaluation. *IEEE Transactions on Power Delivery*, vol. 18, no.1, pp. 78-84.
- Kohonen, T. (1990). The self organizing map. *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464-1480.
- Lau, E. C., & Ngan, H. W. (2010). Detection of motor bearing outer raceway defect by wavelet packet transformed motor current signature analysis. *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 10, pp. 2683-2690.
- Oh, H., Azarian, M. H., & Pecht, M. (2011). Estimation of fan bearing degradation using acoustic emission analysis and Mahalanobis distance. *Proceedings of the Applied Systems Health Management Conference* (pp. 1-12), May 10-12, Virginia Beach, VA, U.S.A.
- Oh, H., Shibutani, T., & Pecht, M. (2012). Precursor monitoring approach for reliability assessment of cooling fans. *Journal of Intelligent Manufacturing*, Vol. 23, no. 2, pp. 173-178.
- Pecht, M. (2008). *Prognostics and Health Management of Electronics*. New York: Wiley-Interscience, New York.
- Qiu, H., Lee, J., Lin, J., & Yu, G. (2003). Robust performance degradation assessment methods for enhanced rolling element bearing prognostics. *Advanced Engineering Informatics*, vol. 17, no. 3, pp. 127-140.
- Sotiris, V. A., Tse, P. W., & Pecht, M. G. (2010). Anomaly detection through a Bayesian support vector machine. , *IEEE Transactions on Reliability*, vol. 59, no. 2, pp. 277-286.
- Tian, J., Morillo, C., & Pecht, M. G. (2013). Rolling element bearing fault diagnosis using simulated annealing optimized spectral kurtosis. *IEEE Conference on Prognostics and Health Management (PHM)*, (pp. 1-5), June 24-27, Gaithersburg, MD, U.S.A.
- Tian, X. (2006). Cooling fan reliability: failure criteria, accelerated life testing, modeling and qualification. *Proceedings of 2006 Reliability and Maintainability Symposium* (pp. 380-384), Jan 23-26, Newport Beach, CA, U.S.A.
- Wang, G., Liu, C., & Cui, Y. (2012). Clustering diagnosis of rolling element bearing fault based on integrated Autoregressive/Autoregressive Conditional Heteroscedasticity model. *Journal of Sound and Vibration*, vol. 331, no. 19, pp. 4379-4387.

BIOGRAPHIES

Jing Tian received the B. Eng degree in Machinery Design and Manufacturing and Automation from the University of Electronic Science and Technology of China. He is a Research Assistant doing Ph.D. research at the Center for Advanced Life Cycle Engineering (CALCE), University of Maryland, College Park. His sponsored research projects include helicopter drive train data analysis for condition-based maintenance, prognostics and health management (PHM) software design, and bearing-gear union data analysis for fault diagnosis. Prior to joining CALCE in 2010, he had worked at the No.30 Institute of China Electronics Technology Group Corporation and Lenovo Corporate Research as an engineer and researcher for 7 years, where he performed data analysis and design work on several rugged computer products, which have been well accepted by the market. His research focuses on machine learning and its application in PHM.

Michael H. Azarian received the B.S.E. degree in chemical engineering from Princeton University and the M.E. and Ph.D. degrees in materials science and engineering from Carnegie Mellon University.

He is a Research Scientist with the Center for Advanced Life Cycle Engineering (CALCE), University of Maryland, College Park. Prior to joining CALCE in 2004, he spent over 13 years in the data storage, advanced materials, and fiber optics industries. His research focuses on the analysis, detection, prediction, and prevention of failures in electronic products. He has advised many companies on reliability of electronic products, and is the author or co-author of numerous publications on solder joint degradation, electrochemical migration, capacitor reliability, electronic packaging, and tribology. He is the holder of five U.S. patents for inventions in data storage and contamination control.

Dr. Azarian is co-chair of the Miscellaneous Techniques subcommittee of the SAE G-19A standards committee on detection of counterfeit parts. He was previously the vice-chairman of the working group for IEEE Standard 1332-2012, "Reliability Program for the Development and Production of Electronic Products," and the Technical Editor of IEEE Standard 1624 on organizational reliability capability, for assessing suppliers of electronic products. He also co-chaired iNEMI's Technology Working Group on Sensor Technology Roadmapping. He is on the Editorial Advisory Board of Soldering & Surface Mount Technology.

Michael Pecht received the M.S. degree in electrical engineering and the M.S. and Ph.D. degrees in engineering mechanics from the University of Wisconsin, Madison.

He is the Founder of the Center for Advanced Life Cycle Engineering, University of Maryland, College Park, where he is also a George Dieter Chair Professor in mechanical engineering and a Professor in applied mathematics. He has been leading a research team in the area of prognostics for the past ten years and has formed a prognostics and health management consortium at the University of Maryland. He has consulted for over 100 major international electronics companies, providing expertise in strategic planning, design, test, prognostics, IP, and risk assessment of electronic products and systems. He has written more than 20 books on electronic-product development and use and supply chain management and over 400 technical articles.

Dr. Pecht is a Professional Engineer and a fellow of ASME and IMAPS. He is the editor-in-chief of IEEE Access, and served as chief editor of the IEEE Transactions on Reliability for nine years, chief editor for Microelectronics Reliability for sixteen years, an associate editor for the IEEE Transactions on Components and Packaging Technology, and on the advisory board of IEEE Spectrum. He was the recipient of the highest reliability honor, the IEEE Reliability Society's Lifetime Achievement Award in 2008. He was also the recipient of the European Micro and Nano-Reliability Award for outstanding contributions to reliability research, the 3M Research Award for electronics packaging, and the IMAPS William D. Ashman Memorial Achievement Award for his contributions in electronics reliability analysis.