University of Hertfordshire UH 25 Research Archive

Citation for published version:
Mohammed Chalouli, Nasr-eddine Berrached, and Mouloud Denai, 'Intelligent Health Monitoring of Machine Bearings Based on Feature Extraction', *Journal of Failure Analysis and Prevention*, Vol. 17 (5): 1053-1066, October 2017.

DOI:
https://doi.org/10.1007/s11668-017-0343-y

Document Version:
This is the Accepted Manuscript version.
The version in the University of Hertfordshire Research Archive may differ from the final published version.

Copyright and Reuse:
© 2017 ASM International. Published by Springer US.
This manuscript version is made available under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Enquiries
If you believe this document infringes copyright, please contact Research & Scholarly Communications at rsc@herts.ac.uk

Intelligent Health Monitoring of Machine Bearings Based On Feature Extraction

Mohammed CHALOULI [1] medchalouli@live.fr

Nasr-eddine BERRACHED [1] laresi.usto.2015@com.com

Mouloud DENAI [2] m.denai@herts.ac.uk

[1] Intelligent Systems Research Laboratory (LARESI), USTO-MB University, Oran – Algeria.
[2] School of Engineering and Technology, University of Hertfordshire, Hatfield, UK

Abstract. Finding reliable condition monitoring solutions for large-scale complex systems is currently a major challenge in industrial research. Since fault diagnosis is directly related to the features of a system, there has been many research studies aimed to develop methods for the selection of the relevant features. Moreover, there are no universal features for a particular application domain such as machine diagnosis. For example, in machine bearing fault diagnosis, these features are often selected by an expert or based on previous experience. Thus, for each bearing machine type, the relevant features must be selected. This paper attempts to solve the problem of relevant features identification by building an automatic fault diagnosis process based on relevant feature selection using a data-driven approach. The proposed approach starts with the extraction of the time-domain features from the input signals. Then, a feature reduction algorithm based on cross-correlation filter is applied to reduce the time and cost of the processing. Unsupervised learning mechanism using k-means++ selects the relevant fault features based on the squared Euclidian distance between different health states. Finally, the selected features are used as inputs to a Self-Organizing Map producing our health indicator. The proposed method is tested on roller bearing benchmark datasets.

Keywords: Failure diagnosis, bearing faults, time-domain features, condition based maintenance, health indicators, relevant features, fault feature extraction.

1 Introduction

Bearing fault are one of the foremost causes of breakdown in rotating machines. It represents over 40% of the motor faults according to the research conducted by Electric Power Research Institute (EPRI) [1], [2]. Most of the existing faults diagnosis methods can identify many bearing faults, but often cannot recognize the fault level accurately [3]. Also, the diversity of symptoms which can develop from the same fault makes diagnostics even harder. This is why, the diagnosis of these critical components has grown strongly in the industrial world as the desire to obtain more efficient and safer production line becomes indispensable.

Replacing a bearing before its end of life leads to unnecessary downtime and parts cost. If, on the other hand, the bearing is used till its end of lifetime leads to unplanned downtime, safety and environmental risks and subsequent damage of other parts.

Thus, the most appropriate maintenance plan in this case is the condition-based maintenance (CBM) which consists of scheduling maintenance activities only when a functional failure is detected. In CBM, the layer responsible of the evaluation of the machine health state is called fault diagnosis and is mainly based on feature extraction.

Condition-Based Maintenance can be described according to the seven functional layers as depicted in Fig. 1 [4][5].

Fig. 1. Open system architecture for condition based maintenance layers

The most important layer is the diagnostics layer or the health assessment because the successor layers such as presentation and decision supports rely strongly and directly on the diagnostics outputs, there will be no decision to make or data to present if the health sate of the machine in unknown. This is why we focused on the diagnostics, noting that the prognostic which is a main layer in CBM can be deducted from diagnostics result.

Achieving the diagnostics task can be done with two approaches data-driven or model-based [6]. In the data-driven approach, which is used in this work, the data collected from experiments are exploited to learn about the system and from that, assesses the system current state of health [7]. Data-driven methods aim at transforming the raw monitoring data into relevant information of the system including the degradation which offers a good diagnostics accuracy; especially when the operating context is variable or in the case of new systems because of a lack of experts. However, the results they offer are less precise than those provided by model-based methods [8].

Degradation of a rolling usually begins with small-sized seizure marks. These localized defects generate periodic shocks of very short duration. Thus, for early diagnosis the raw signals (vibrations) should be exploited for fault detection purpose. Vibration monitoring is carried out using any of the three methods, namely, time domain analysis, Frequency domain analysis, and time frequency analysis. All the three techniques have been described in detail in [9] and compared in [10]. Among the three techniques, time domain analysis is the simplest and easy to implement. It requires the calculation of statistical features, using which the faults can be classified [11]. It is worth noting that the frequency-domain indicators employed in this work are normalized with respect to the rotational speed. Regarding the time-domain indicators, they are independent of the rotational speed when the load is non-rotating [12].

The use of time domain features directly has a number of shortcomings. Noting initially, when a defect appears, the tiny shocks increase the peak level considerably, but have less influence on the RMS. The RMS level may become significantly high in bearings with multiple or spreading defects, resulting in a reduction in the Crest factor. Background noise is also a problem because it increases the RMS level, and consequently decreases the Crest factor.[13].

With the growing techniques of machine learning and In order to solve the choice of relevant features' problem [14][3]. This work were based on many researches in this field, citing the inspiring work done by [15] which is based on extracting time and frequency features for a critical component, followed by an unsupervised feature selection by calculating the pairwise symmetrical uncertainty, for all the input signals. The features are compressed to lower dimension using PCA. And finally, applies the EMD on the projected features, then following the Data's evolution over time until getting a final residual extracted from the compressed features. Noting that the proposed method extracted the trend feature used in prognostics directly without passing by diagnostics task.

Few year earlier, [13] introduced two new parameters where TALAF describes the damage's evolution over time which can be considered as health indicator and the second parameter is THIKAT representing the degree of confidence relative to the use of defective bearing and also confirm the TALAF early diagnosis. This work also spots the light on why time domain features can't be used individually as health indicators and that the crest factor and kurtosis are the most sensitive features for fault diagnosis. [16] used a multi-class Support Vector Machine (MSVM) classifier enhanced with by PCA for pre-processing for bearing fault diagnosis. The PCA is used for feature selection task then the MSVM one against one strategy uses the PCA result as inputs for fault classification. The results are illustrated in 3 outputs: inner race fault, outer race fault and balls fault. Concluding from this work that the Time domain features or frequency domain features can't be used directly in classification. In this work [17], a multi-class SVM using one-against-all strategy with a non-linear kernel is applied to classify bearing health state. This will result in estimated vectors of state transition times. Where, these vectors of transition times will provide the information necessary to carryout prognostics. Noting that, exploiting the diagnostic results for prognostics make the RUL estimation faster and easier for interpretation. This work inspired me for prognostic deduction by calculating the distance between clusters of different health state of bearing. Recently, [18] used the Genetic Algorithm (GA) to determine the initial point of K-means aiming to overcome its drawback where it is probable to get stuck in local optima because an inappropriate starting point. It was concluded from the results that the proposed method is able to identify the rotor condition with more accuracy. This work talks about correcting the K-means drawback for the initialization of the starting point and how it is less accurate to use the raw K-means in bearing diagnosis. In the same year,

[11] described that among the several types of condition monitoring, the vibration monitoring is the most widely used techniques. His work focuses on statistical features extracted from Time domain Analysis of bearing vibration; the features are extracted from the raw vibration signal and also from its time derivative and time integral and used only for diagnosis task using Graphical User Interface (GUI). As results, the features which are useful for fault classification are variance and Root Mean Square (RMS). This has been established by calculating the mean and the standard deviation values of all the features, as these values vary largely. This work is interesting for fault classification once the fault is known but for the fault detection the RMS and the variance aren't that informative as it's described in this work. And currently, [19] introduces the Deep Neural Network as a new kind of machine learning in bearing fault diagnosis which used for feature extraction. The high dimensionality of the input signal is reduced using principal component analysis (PCA). The given results show good and promoting results in this field in case of large amount of data. And finally, [20] intends to review the use of spectral kurtosis in diagnostics prognostics as a potential relevant feature.

In this last decade, more research are focused on fast fault diagnosis and integration of complex machine learning algorithms in diagnostics and prognostics process but there is no study justifying the choice of selected features or there number. Our contribution is illustrated in these aims:

- Automatic relevant feature selection.
- Novel Health indication construction.
- Deducing the Remaining Useful Lifetime from the diagnostics results.

Often researchers overlook the study beyond selecting the adequate number of features [14] which includes the choice of inappropriate features, and this is unsafe because the accuracy of detecting the fault of these machines depends directly on the number and type of features used in diagnostics. Currently, there is no systematic approach to determine the optimum number of features for a given problem and this is still an ongoing field of research [3]. For example, in the case of bearing machine most experts tend to use impulse sensitive features, however, it is not explained how does the machine learning program automatically select the features and how can these features be calculated once the machine is changed or in the case of new machine. This paper will present answers to these questions.

The proposed method is an automatic fault diagnosis based on the extraction of time domain descriptors from the raw vibration signal. Hereafter, unsupervised machine learning algorithm is used to determine only features carrying relevant information related to the diagnostic task. Finally, Health indicator is built for monitoring the health state of the bearing using self-organizing map as dimension reduction approach. This method has the advantage of being simple because it is based on time domain descriptors calculated from vibration signals directly without any frequency response calculation which reduces computation time and costs. The method's simplicity and ease of implementation make it suitable for real-time industrial applications.

The rest of paper is organized as follow: Section 1 presents a brief overview on the maintenance strategy and fault diagnosis of machine bearings. Then, the proposed method methodology is discussed in the next section. Followed by, result and validation in section 3. And finally, conclusion and future work

2 Proposed Fault Diagnosis Methodology

Since most of rotating machines consist of simple components assembled together, any failure of these components may be detected from the sensor signals which monitor the parameters related to these components. The proposed method operates in two phases. First, the fault feature extraction is applied on the vibration signal in order to extract the statistical descriptors then the relevant features are selected using modified k-means clustering algorithm. The relevant features are determined exclusively from the data and without any a priori assumption from the expert or any consideration of the component type. The second phase can be described as the construction of the fault diagnosis health indicator using Self-Organizing Map (SOM) as a nonlinear projection of the relevant features on a two-dimension space which is simple to interpret and coherent to graphical representation. T

The two phases are complementary for the fault feature diagnostics task but can be treated separately providing relevant features identification from the first phase. While, the second phase results a health indicator made by features given as inputs. However, the health indicator in this case is not the optimal choice for diagnostics.

2.1 Experiment setup for data collection

The test rig is composed by twelve bearing Rexnord ZA-2115 double row bearings installed on a shaft. Their vibrations were acquired using PCB 353B33 High Sensitivity Quartz ICP accelerometers installed on the bearing housing with one accelerometer for each bearing. The rotational speed was kept constant at 2000 RPM by an AC motor coupled to the shaft via rub belts. A radial load of 2 721.55 Kg is applied onto the shaft and bearing by a spring mechanism. All bearings are force lubricated. Data acquisition was performed by an NI DAQ Card 6062E and the sampling rate was fixed at 20 kHz.[21]. [22] provides a good description about this test rig.

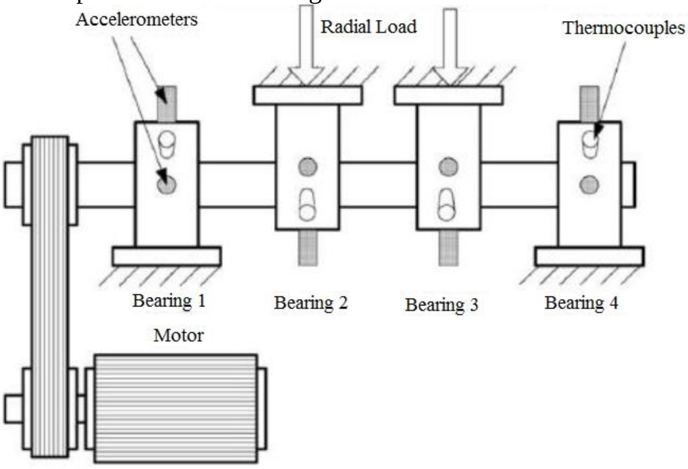


Fig. 2. Bearing test rig and sensor placement illustration [21]

The second data set of Fig. 3 is provided by four bearings. The signal Fig. 3(a) represents the bearing reaching its time to failure, while Fig. 3(b) shows the vibration signal of the bearing before reaching its time to failure.

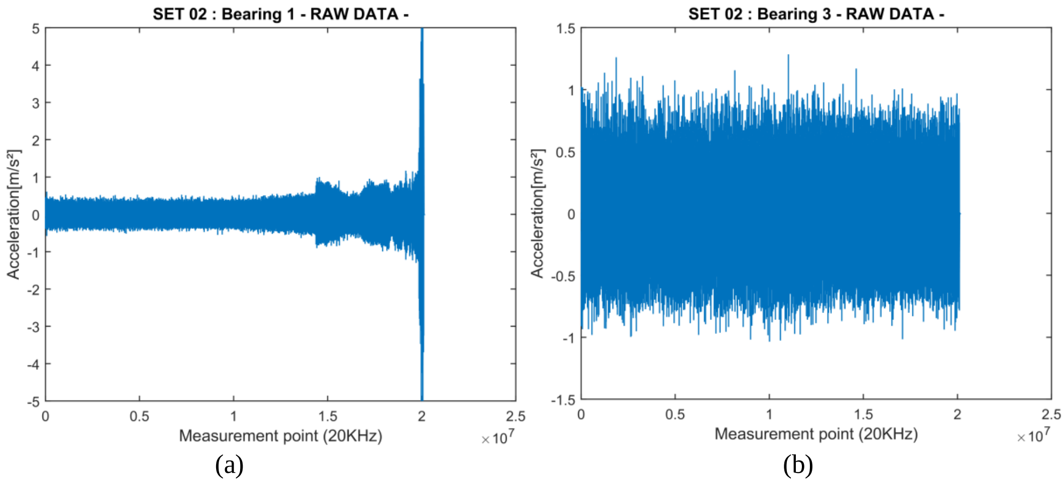


Fig. 3. Vibration signals of the second set of NASA repository database.

It can be observed on the defective signal of Fig. 3(a) that the amplitude has increased significantly reaching a value greater than twice its initial value while working in good condition within its remaining useful lifetime. The kind of features extracted and their domain are discussed in the next section.

2.2 Relevant feature identification

To identify the most relevant features, the following three steps must be performed:

1) **Feature Extraction:** This step determines the signal descriptors according to a specific domain (time, frequency and time-frequency). Time domain descriptors are used in this paper.
2) **Feature Reduction:** A marge dataset of features required more time and cost for computation which is a problem to avoid. Cross-correlation filter is applied in our method.
3) **Feature Selection:** Once the redundant features are ignored, the non-informative feature must be deleted, selecting at the end of process only the relevant features. K-means is used as an unsupervised classifier to perform the classification task.

A. Feature extraction

One of the key problems in diagnostics is the extraction of features from the vibration signals. Vibration monitoring is carried out using time domain analysis, frequency domain analysis or time-frequency analysis. The three techniques have been well described in [9]. The features should be sensitive to machine faults and at the same time robust to background noise. Another important consideration in feature domain selection is that computation complexity for extracting features should be low in order to be suitable for real time diagnosis. The time-domain features are recommended because normal and defective signals differ in their statistical characteristics in the time domain where the calculation is simple and complexity is low [17]. It is worth noting that the time-domain indicators are independent of the rotational speed when the load is non-rotating.[12] and they are calculated from vibration signals directly without any temporal frequency calculations which reduces computation time making it more easily adoptable in industry because the simplicity of its application [23].

There are at least three reasons why feature extraction is an important problem in fault diagnosis process [24]:

- **Dimension Reduction:** When a machine learning program is given too many variables to consider, most of which are redundant or non-informative, it is naturally much harder to make a good decision. It will be better to select only the relevant features in order to gain in computation time and resources.

- **Automatic Exploratory Data Analysis:** Taking the case where the data are so novel that there are no field experts who understand the data well enough to be able to extract the important features prior to the analysis. Under such circumstances, automatic exploratory data analysis becomes an alternative solution.

- **Data Visualization:** The human eye has an amazing ability in recognizing systematic patterns in the data. At the same time, however, humans are usually unable to make good sense of data if it is more than three dimensional which makes data visualization on reduced space more important in feature extraction process.

Some of time-domain features used in literature are listed in Table 1.

Table 1. Time-domain features.

N°	Feature	Symbol	Equation
1	Mean	$\bar{x}$	$\frac{1}{n}\sum_{i=1}^{n}\lvert x_i\rvert$
2	Root mean square	x_{rms}	$\sqrt{\frac{1}{n}\sum_{i=1}^{n} {x_i}^2}$
3	Peak	x_{peak}	$max(\lvert x_i\rvert)$
4	Root amplitude	x_{ram}	$\left(\frac{1}{n}\sum_{i=1}^{n}\sqrt{\lvert x_i\rvert}\right)^2$
5	Standard deviation	x_{std}	$\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$
6	Skewness	x_{skew}	$\frac{\sum_{i=1}^{n}(x_i-\bar{x})^3}{(n-1).{x_{std}}^3}$
7	Kurtosis	x_{kur}	$\frac{\sum_{i=1}^{n}(x_i-\bar{x})^4}{(n-1).{x_{std}}^4}$
8	Hyper Kurtosis	X_{hku}	$\frac{\sum_{i=1}^{n}(x_i-\bar{x})^6}{(n-1).{x_{std}}^6}$
9	Shape factor	x_{shf}	$\frac{x_{rms}}{\bar{x}}$
10	Crest factor (Peak F.)	x_{crf}	$\frac{x_{peak}}{x_{rms}}$
11	Impulse factor	x_{imf}	$\frac{x_{peak}}{\bar{x}}$
12	Clearance factor	x_{clf}	$\frac{x_{peak}}{x_{ram}}$

However, time domain indicators, even if they are well suited for online monitoring, they do not identify the defect responsible of the degradation [25].

B. Feature reduction

Bearing in mind that some features are somehow similar to others and knowing that calculation time and cost are directly proportional to the number of treated features which make feature reduction so important step.

A cross-correlation filter is applied to reduce the feature having almost the same indication as illustrated in Fig. 4

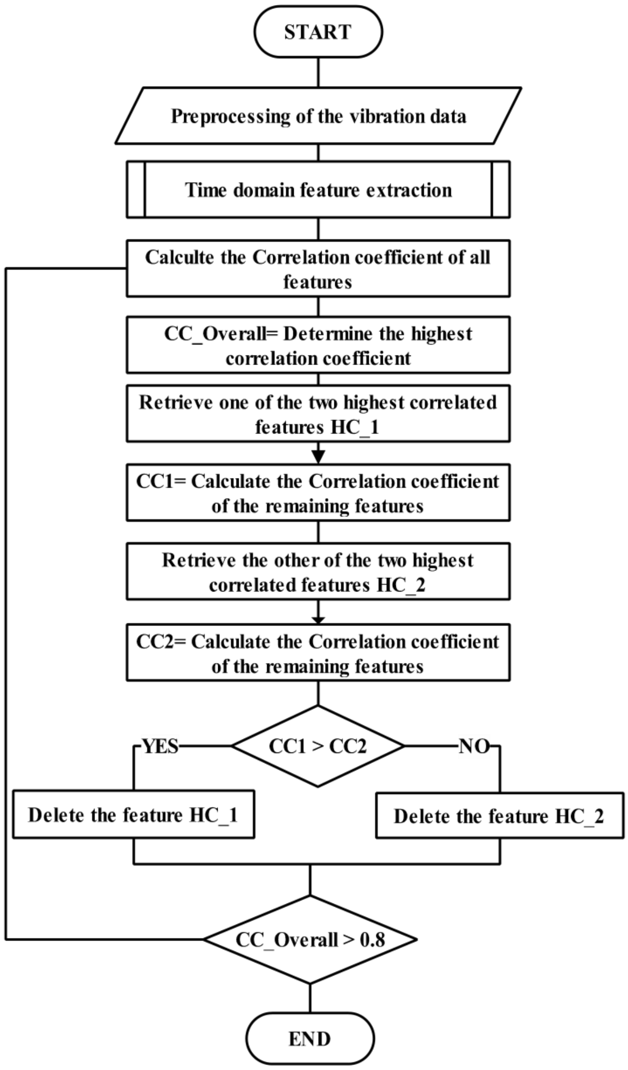


Fig. 4. Feature reduction algorithm.

Where the coefficient of correlation between two features is calculated by equation (1):

$$CC(A,B) = \frac{CV(A,B)}{\sqrt{CV(A,A) * CV(B,B)}} \quad (1)$$

$$CV(A,B) = \frac{1}{\mathrm{N}-1}\sum_{i=1}^{N}(A_i - \mu_A) * (B_i - \mu_B) \quad (2)$$

$$\mu = \frac{1}{\mathrm{N}}\sum_{i=1}^{N} A_i \quad (3)$$

The advantages of this method against others such as Principal Component Analysis (PCA) can be resumed in two main points. The non-alternation of the feature nature where the results are a reduced number of features saving the characteristic of each feature while the results of PCA for example gives a new representation of inputs data in other space. The second point is represented by the direct selection of non-redundant features after the features are reduced which does not require other calculation as compared to the PCA method where the analysis must be done every time the features have to be reduced. This advantage is well illustrated when the relevant features process is invoked more than once like an online and offline phases or repetitive health assessment of the same kind of component. The features left from the reduction phase are used as inputs for the relevant feature selection phase described next.

C. Feature selection

The selection of relevant feature process is based on the features given the best distinction between the different health state of a machine is divided to two phases:

a) Normalisation phase

Normalization is a primordial step before classification of features which have different scalar scales. The normalized data affect the classification results significantly. The normalization proposed in this work in based on projection of data on a unique scale between the lowest value of the overall features and the highest one conserving the ratio between data of each feature; the method is illustrated in equation (6).

$$Min_{all} = \sum_{i=1}^{Nbr\_Feature} \sum_{j=1}^{DATA\_length} \min\big(Feat(i,j)\big) \tag{4}$$

$$Max_{all} = \sum_{i=1}^{Nbr_{Feature}} \sum_{j=1}^{DATA_{length}} \max\big(Feat(i,j)\big) \tag{5}$$

$$\mathrm{VALUE}_{\mathrm{newframe}} = \sum_{i=1}^{Nbr\_Feature} \sum_{j=1}^{DATA\_length} \frac{DATA_{value}(i,j) - Min_{all}}{Max_{all} - Min_{all}} \tag{6}$$

b) Classification phase

Because of the lack of history knowledge about the characteristic of the different health state of a machine, unsupervised classification is the adequate solution to distinguish the faulty machine from the healthy one. K-means is one of the most popular clustering algorithms used in industrial application [26]. It is calculated using the Lloyd's algorithm which begins with k arbitrary centers, chosen randomly among the data points. Once done, each point is assigned to the nearest center. Then, it calculates again each center as the center of all its assigned data points. Finally, the whole process except the random choice is repeated until the process stabilizes. It is the speed and simplicity of the k-means method that make it appealing, not its accuracy [27].
The initial point for the k-means clustering should be chosen properly in order to overcome the problems associated with local optima [18].

K-Means++ algorithm gives more successful results than standard K-Means in terms of accuracy and consistency. Because, the K-Means algorithm works only to find a local optimum and this local optimum often becomes poor by using random initial center points; however, K-Means++ starts with rational initial points, using a proportional probability to the distance between centers given a preference to further points. Thus it approximates the best clustering space. Also, it outperforms the standard k-means in speed, too [28] which is the case of our data. K-Means++ is described in more details in [29][30]

This unsupervised learning method is used to determine the features given the largest distance between bearing health state classes where they represent the different fault level. The applied method calculate the Euclidian distance between clusters for each 2 features and with N number of features we get K probability as it's shown in equation (7)

$$K = \frac{n!}{(n-2)! \times 2!} \tag{7}$$

The combination with the longest distance and the densest cluster is taken as relevant feature to build the health indicator.

2.3 Health indicators

Previous research work has shown that there is no feature suitable for all defect types at all degradation stages. For example, the Kurtosis is more suitable for the detection of incipient defects, whereas the RMS value indicates severe defects. Thus, a reliable performance of fault diagnosis method should take advantage of mutual information from multiple features [31]. Future fusion is a key in this case to get robust health indicator detecting several defect types at different stages. The SOM is an appropriate tool for this task with its unique capability of projecting high dimensional data into a low dimensional space while preserving their inherent topographic relationships [31].

In this paper, the relevant fault features are used to build a health indicator through an automated approach. Empiric thresholds are then used to discriminate between the different types of health state. In our case study, the relevant features selected for fault features extraction are trained using SOM toolbox described by [32] in two steps: First, applying training with large neighborhood radius and learning rate followed by a fine tuning. The next step after training is to calculate the mean quantization which produces the average distance between each data vector and its Best Matching Unit. The theory behind the SOM generally and particularly in bearing can be found in [33], [34].

The SOM algorithm is used as a nonlinear projection of our relevant features on a two-dimension space which is simple to interpret and coherent to graphical representation. SOM approach is best suited for make clusters on the map and correlations between variables [32]. Fusing the k-means and the SOM was already used in this work [28] and also provides good results. The work, presented here, differs in the way the how methods are combined, rather than in the use of K-means++ to initialize the weight values before performing the SOM for clustering. In this work, the K-means++ is used for clustering to select the relevant features than use the SOM for features fusion in order to build a reliable health indicator.

3 Results and Discussion

3.1 Results

The features listed in Table 1 are extracted from the vibration signals of a bearing machine as illustrated in Fig. 5. Then, the redundant features isolation phase was applied to extract the dominant features. The red graphs represent the dominant features while those in blue refer to the redundant features which will not be used for further processing. The remaining features are highly correlated at least with one of the dominant features and practically give the same results.

A cross-correlation filter is applied for the redundant features isolation phase to determine the similar features before eliminating one of each two high correlated features according to the algorithm described in Fig. 4. Once the two features with the highest correlation value are selected, the correlation coefficient is calculated using one of these two with the rest of features. The feature giving the highest correlation coefficient will be removed and the whole process will repeated until the overall coefficient reach a threshold of 0.8.

In bearing machine with roller element, the number of dominant features is around seven, according to experiments performed on 12 bearings in different cases. Noting that, the highest number of features is always taken to avoid the risks of losing significant information. From the results, it can be deduced that even for the same kind of bearing there is a slight difference in the number of dominant features because the conditions under which the measurements were taken.

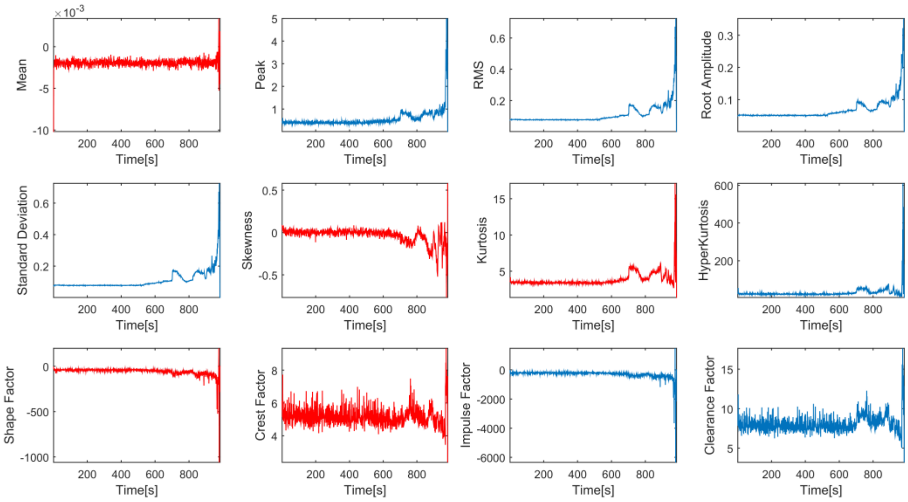
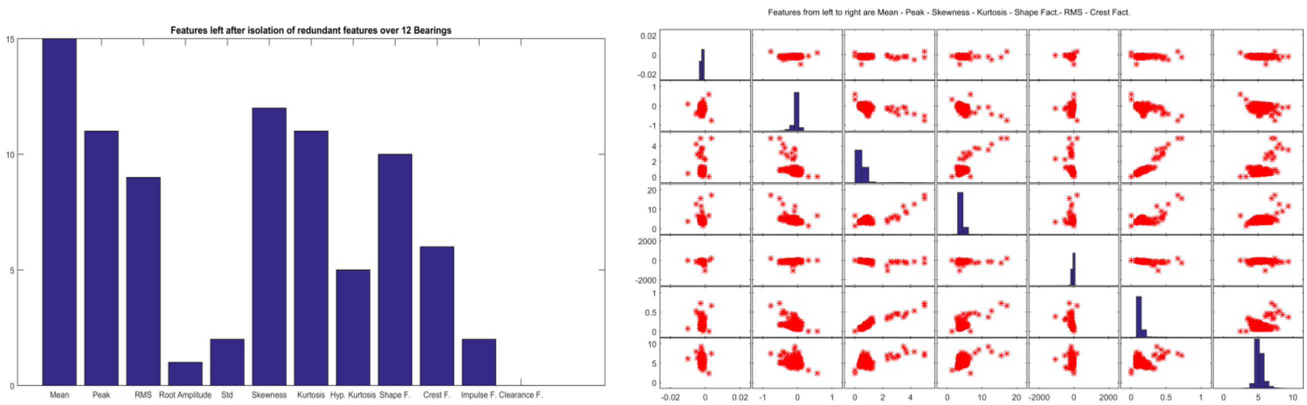


Fig. 5. Feature extracted from the vibration signal of the bearing machine.

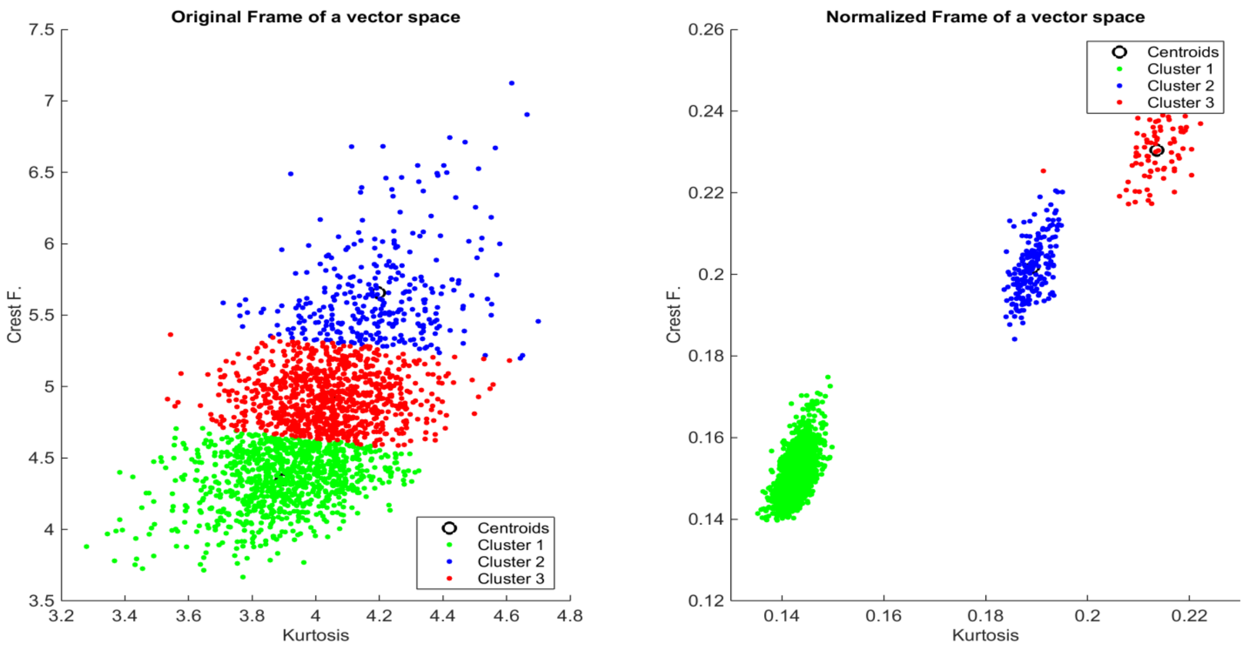
Moreover, to determine accurately the dominant features, many bearing are studied in different cases where the histogram in Fig. 6 (a) reveals how many times the feature was categorized as dominant. Thus, in our case of study the 7 dominant features are: Skewness - Mean – Shape factor - Peak - RMS – Kurtosis and crest factor. Fig. 6(b) shows the relation between these features.



(a) **(b)**

Fig. 6. Dominant features for machine bearings with roller element.

Applying the normalization described by equation (6) improves considerately the classification and the representation of our classes, the axes in Fig. 7 shows how difficult it is to separate the classes, while the normalized data are well separated and easy to classify even with a basic classifier.



(a) **(b)**

Fig. 7. Difference between the normalized and raw features.

The simulation was started with 12 features, and then 7 were eliminated at isolation of redundant features process while the remaining were normalized for the relevant features selection task where the distance of K probability from equation (7) is measured and the two features with the best representation and the highest distance are considered as the relevant fault features. Table 2 show the Squared Euclidean distance between the health centroid and the fault one for the dominant features in diagnostics.

Table 2. Squared Euclidean distance between health and fault state of time-domain features.

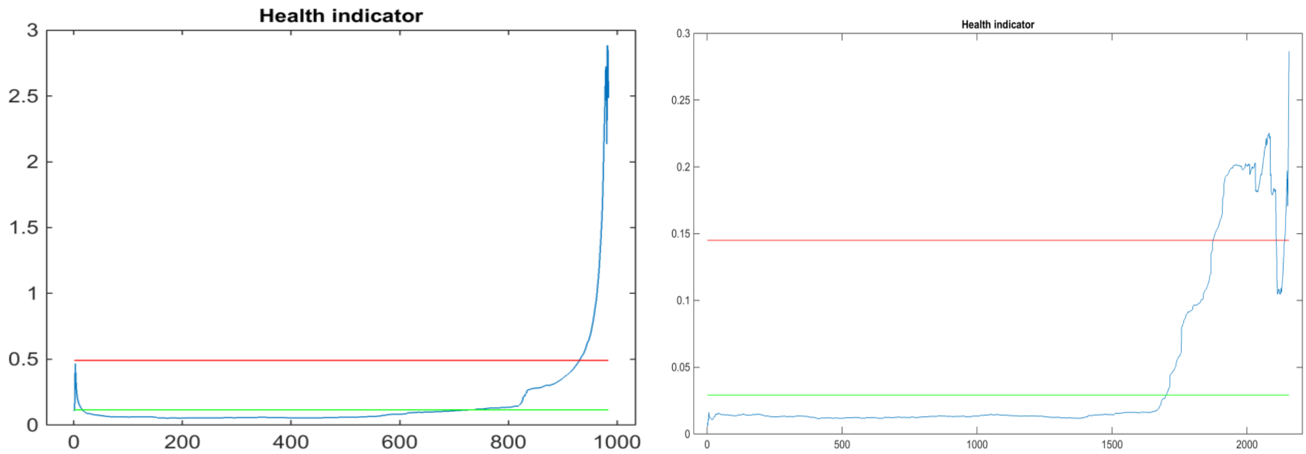
Features	RMS - Skewness	RMS - Kurtosis	RMS - Crest F.	Peak - Kurtosis	Skewness - Crest F.	Skewness - Kurtosis	Kurtosis - Crest F.
Squared Euclidean distance	0.1605	0.1642	0.1648	0.1646	0.1647	0.1641	<u>0.1682</u>

The kurtosis and crest factor give the largest distance. Thus, they are considered as the relevant features and can be described as:

Kurtosis: Defined as the normalized fourth moment and described as the ratio of the fourth moment to the variance. Kurtosis is a measure of peakedness, and hence it measures the degree of peakiness of a distribution compared to a normal distribution. A High value of kurtosis means a longer tail of distribution.

Crest factor: Defined as, the ratio of the peak value over the RMS value. It gives an idea about any impacts present in the signal which detects acceleration bursts even if signal RMS has not changed. A high value of crest factor means the presence of wear or pitting.

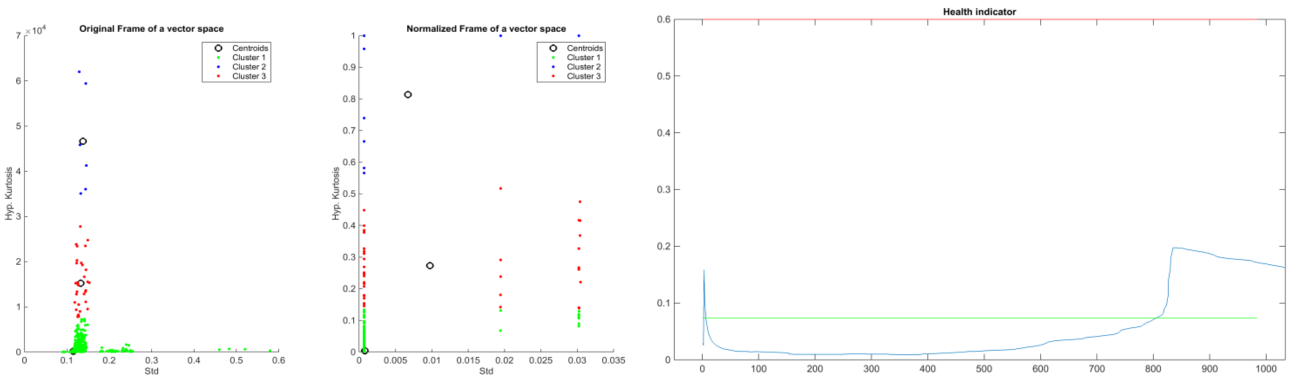
Next, a SOM algorithm is applied to construct the health indicator and the results are shown in the Fig. 8.



(a) (b)

Fig. 8. Health indicator based on Kurtosis-Crest factor.

Fig. 9 illustrates the results of using the Hyper Kurtosis and the standard deviation as non-relevant feature for health assessment. The clustering using these two features provides non-significant result even with normalized data as demonstrated at Fig.9 (a). Thus, the health indicator in Fig. 9 (b) leads to inaccurate and wrong diagnosis.
The results of Fig. 8 and Fig. 9 demonstrate the importance of selecting the adequate relevant features in fault diagnostics tasks.



(a) (b)

Fig. 9. Health indicator based on Hyper Kurtosis-Standard deviation.

The decision then will be made according to the position of the health indicator to the fault threshold (Green) given by the equation (8) and failure threshold (Red) given by the equation (9).

$$\text{Fault} = \frac{\frac{1}{n}\sum_{i=1}^{n}(Hi_i)}{\sqrt{2}} \tag{8}$$

$$\text{Failure } = 2 \times \sqrt{\frac{1}{n}\sum_{i=1}^{n} {Hi_i}^2} \tag{9}$$

A moving average filter is used to smooth our health indicator with experimental span fixed at 300. Smoothing the signal makes it more coherent while graphical representation and easier for decision making.

3.2 Validation

The goodness of the features for diagnostics is basically a measure of separability between data from healthy and faulty equipment. A diagnostics is considered reliable when the health states of machine are far apart from each other and samples from the same health state are close to each other [2].

Many researches confirm that Kurtosis and the Crest factor are the more suitable features for fault diagnosis. These features are non-dimensional magnitudes, so they are immune from weaknesses in the data process due to quality of the sensors or the location where they are mounted[17]. [23] found that the (RMS), crest factor and the kurtosis give a reasonably global defect indication. And according to [22], it can be considered that slight degradation has emerged when these three features values are beyond their threshold. The research done by [35] has shown that Kurtosis value and Crest Factor are sensitive to impulse faults, especially in the incipient stage. Among the most suitable scalar indicators used to characterize the vibrations are the Crest Factor and Kurtosis. [13] deduced that, the Crest factor and Kurtosis are less dependent on the vibration level, but are sensitive to the spikiness of the vibration signals, and they can provide an early indication of significant changes in vibration signals. More details about the effectiveness of these features are discussed in this book [36] and this thesis[37], [38].

To validate the result we took another rolling machine database to apply test the results. Moreover than the NASA Repository [39] the CWRU for Case Western Reserve University data are used [40].

Table 3,4 and 5 demonstrates the efficiency of the normalization and the relevant fault feature choice.

Table 3. Relevant Features classified by measure of separability.

	Occurrence	**Occurrence %**
Kurtosis-Crest.F	**14**	**52%**
Shape.F-Crest.F	11	41%
Mean-Skewness	02	7%
TOTAL	**27**	**100%**

Table 3 shows the results applied on nine cases of classification between healthy bearing and Inner race fault, healthy and Outer race fault and between healthy bearing with Ball bearing fault, from the CWRU database. The 27 cases show that the two features providing the farther distance intra-class are Kurtosis and Crest fact

Table 4 illustrates the importance of normalization applying the same classification method on the original and normalized data.

Table 4. Comparison original data and normalized data for kurtosis and Crest factor.

Kurtosis-Crest Factor			
Health /Inner R_12	(Norm-Org)007	(Norm-Org)014	(Norm-Org)021
RPM 1797	0-0	0-0	0-0
RPM 1750	0-1	0-0	0-0
RPM 1730	0-2	0-2	0-1
Health /Outer R_12			
RPM 1797	0-0	0-12	0-0
RPM 1750	0-0	0-85	0-0
RPM 1730	0-0	0-94	0-1
Health /Ball_12			
RPM 1797	0-12	0-75	0-12
RPM 1750	0-89	0-58	0-36
RPM 1730	0-79	0-67	0-111
		Normalized error:	**0**
		Original errors:	**737**

Table 5 confirms the choice of the relevant features and shows that the normalization data has many error while choosing the wrong relevant features.

Table 5. Comparison of original data and normalized data for Shape Factor and Crest factor.

Shape Factor -Crest Factor			
Health /Inner R_12	(Norm-Org)007	(Norm-Org)014	(Norm-Org)021
RPM 1797	0-0	0-0	0-0
RPM 1750	0-0	0-6	0-0
RPM 1730	2-0	0-0	1-0
Health /Outer R_12			
RPM 1797	0-0	8-94	0-0
RPM 1750	0-0	0-0	0-2
RPM 1730	0-0	0-0	1-2
Health /Ball_12			
RPM 1797	11-1	0-12	0-100
RPM 1750	1-0	0-3	0-0
RPM 1730	13-0	0-0	15-0
		Normalized error:	**43**
		Original errors:	**218**

4 Conclusion & Perspectives

This study discussed a new procedure for fault diagnosis of bearing machines based on fault feature extraction and statistical computer science approach.

Three methods were discussed. In the first one, time-domain descriptors are extracted from vibration signals using statistical algorithms followed by eliminating redundant features using cross-correlation filter. The second method based on improved k-means algorithm to select the features given the best separation between different states of health based on the squared Euclidean distance. The last method used the SOM for health indicator's construction enabling the diagnostics to be faster and more accurate.

The aim from reducing the features and constructing to health indicator is to perform online fault diagnosis. The results demonstrate that the selected features infer to more accurate fault diagnosis method which translates also to reduction in both hardware and computational time from signal processing without compromising the machine health indicator accuracy. The results indicate that the combination of Kurtosis and Crest factor are the most sensitive features for roller element bearings fault diagnosis.

This study also is considered as an introduction to the prognostic process which consists in the prediction of the RUL (Remaining Useful Lifetime). The data gathered are used by a diagnostic module to identify the actual operating mode. This state is then projected in the future in order to predict the system's future state [41]

The working on method relies on calculating the distance between centroid classes. The Remaining Useful Lifetime of the bearing in calculated by matching the actual health state by the data previously stored in databases (Template Maching).

Acknowledgments

The authors wish to thank NSF I/UCR Center for Intelligent Maintenance Systems (IMS) [21] and Case Western Reserve University [40] for providing free access to the bearing vibration experimental data from their Web sites.

References

[1] K. Medjaher, D. Tobon-mejia, and N. Zerhouni, "with application to bearings . To cite this version : Remaining useful life estimation of critical components with application to bearings," vol. 61, no. 2, pp. 292–302, 2012.

[2] K. Medjaher, F. Camci, and N. Zerhouni, "Feature extraction and evaluation for Health Assessment and Failure prognostics.," *Proc. First Eur. Conf. Progn. Heal. Manag. Soc. PHM-E'12.*, pp. 1–6, 2012.

[3] J. Hu, L. Zhang, and W. Liang, "Degradation Assessment of Bearing Fault using SOM Network," pp. 561–565, 2011.

[4] L. Wang and R. X. Gao, *Condition Monitoring and Control for Intelligent Manufacturing*. Springer, 2006.

[5] H. Kenneth, A. Adam, A. Aitor, J. Erkki, M. Julien, and M. Samir, *E-maintenance*. 2010.

[6] R. K. Singleton, E. G. Strangas, and S. Aviyente, "Discovering the hidden health states in bearing vibration signals for fault prognosis," *Proceedings, IECON 2014 - 40th Annu. Conf. IEEE Ind. Electron. Soc.*, pp. 3438–3444, 2014.

[7] C. Nataraj and K. Kappaganthu, "Vibration-based diagnostics of rolling element bearings : state of the art and challenges," *13th World Congr. Mech. Mach. Sci.*, pp. 1–10, 2011.

[8] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot, "A Mixture of Gaussians Hidden Markov Model for failure diagnostic and prognostic," *2010 IEEE Int. Conf. Autom. Sci. Eng. CASE 2010*, pp. 338–343, 2010.

[9] S. Patidar and P. K. Soni, "An Overview on Vibration Analysis Techniques for the Diagnosis of Rolling Element Bearing Faults," *Int. J. Eng.*, vol. 4, no. May, pp. 1804–1809, 2013.

[10] A. Boudiaf, A. Moussaoui, and A. Dahane, "A Comparative Study of Various Methods of Bearing Faults Diagnosis Using the Case Western Reserve University Data," *J. Fail. Anal. Prev.*, 2016.

[11] L. P. P, S. P. V, and V. P. S. Naidu, "Bearing Health Condition Monitoring : Time Domain Analysis," pp. 75–82, 2014.

[12] L. Batista, B. Badri, R. Sabourin, and M. Thomas, "A classifier fusion system for bearing fault diagnosis," *Expert Syst. Appl.*, vol. 40, no. 17, pp. 6788–6797, 2013.

[13] S. Sassi, B. Badri, and M. Thomas, "TALAF and THIKAT as innovative time domain indicators for tracking BALL bearings," in *Proceedings of the 14th Seminar on Machinery Vibration*, 2006, vol. 2, no. 1, pp. 24–27.

[14] V. Sugumaran and K. I. Ramachandran, "Effect of number of features on classification of roller bearing faults using SVM and PSVM," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 4088–4096, Apr. 2011.

[15] A. Mosallam, K. Medjaher, and N. Zerhouni, "Time series trending for condition assessment and prognostics," *J. Manuf. Technol. Manag.*, vol. 25, no. 4, pp. 550–567, 2014.

[16] G. Jia, S. Yuan, and C. Tang, "Fault Diagnosis of Roller Bearing Based on PCA and Multi-class Support Vector Machine," pp. 198–205, 2011.

[17] D. Galar, U. Kumar, and Y. Fuqing, "RUL prediction using moving trajectories between SVM hyper planes," *2012 Proc. Annu. Reliab. Maintainab. Symp.*, pp. 1–6, 2012.

[18] M. M. Ettefagh, "Bearing Fault Diagnosis using Hybrid Genetic Algorithm K-means Clustering," 2014.

[19] W. Lu, X. Wang, C. Yang, and T. Zhang, "A Novel Feature Extraction Method using Deep Neural Network for Rolling Bearing Fault Diagnosis," pp. 2427–2431, 2015.

[20] Y. Wang, J. Xiang, R. Markert, and M. Liang, "Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications," *Mech. Syst. Signal Process.*, vol. 66–67, pp. 679–698, 2016.

[21] H. Qiu, J. Lee, J. Lin, and G. Yu, "IMS Bearing Data The," 2006.

[22] S. Hong and Z. Zhou, "Remaining useful life prognosis of bearing based on Gauss process regression," in *Biomedical Engineering and Informatics (BMEI), 2012 5th International Conference on*, 2012, no. Bmei, pp. 1575–1579.

[23] M. E. Boukhobza, Z. Derouiche, and Z. A. Foitih, "Location and evaluation of bearings defects by vibration analysis and neural networks," *Mechanika*, vol. 19, no. 4, pp. 459–465, 2013.

[24] M. Zhu, "Feature extraction and dimension reduction with applications to classification and the

analysis of co-occurrence data," 2001.

[25] A. Soualhi, K. Medjaher, and N. Zerhouni, "Bearing health monitoring based on hilbert-huang transform, support vector machine, and regression," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 1, pp. 52–62, 2015.

[26] P. Berkhin, "Survey of Clustering Data Mining Techniques," pp. 1–56, 2010.

[27] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," *Proc. eighteenth Annu. ACM-SIAM Symp. Discret. algorithms*, vol. 8, pp. 1027–1035, 2007.

[28] Y. Dogan, D. Birant, and A. Kut, "SOM ++ : Integration of Self-Organizing Map and K-Means ++ Algorithms," *Mldm*, pp. 246–259, 2013.

[29] M. Agarwal, R. Jaiswal, and A. Pal, "k-Means++under approximation stability," *Theor. Comput. Sci.*, vol. 588, pp. 37–51, 2015.

[30] M. M. Öztürk, U. Cavusoglu, and A. Zengin, "a Novel Defect Prediction Method for Web Pages Using K-Means ++," *Expert Syst. Appl.*, vol. 42, pp. 6496–6506, 2015.

[31] H. Qiu, J. Lee, M. Kantardzic, O. Nasraoui, and M. Milanova, "Feature fusion and degradation detection using self-organizing map," *Proc. 2004 Int. Conf. Mach. Learn. Appl.*, pp. 107–114, 2004.

[32] J. Vesanto, J. Himberg, E. Alhoniemi, and J. Parhankangas, "Self-organizing map in Matlab : the SOM Toolbox," *Proc. Matlab DSP Conf.*, pp. 35–40, 1999.

[33] S. Shi, L. Zhang, and W. Liang, "Condition monitoring and fault diagnosis of rolling element bearings based on wavelet energy entropy and SOM," *2012 Int. Conf. Qual. Reliab. Risk, Maintenance, Saf. Eng.*, pp. 651–655, 2012.

[34] P. Jun, L. Jiaquan, and L. Xiangjiang, "Self-organizing map-based fault dictionary application research on rolling bearing faults," *Proc. - 4th Int. Conf. Nat. Comput. ICNC 2008*, vol. 3, pp. 311–315, 2008.

[35] Z. Xia, S. Xia, L. Wan, and S. Cai, "Spectral regression based fault feature extraction for bearing accelerometer sensor signals," *Sensors (Switzerland)*, vol. 12, no. 10, pp. 13694–13719, 2012.

[36] A. Boulenger and C. Pachaud, *Aide-mémoire Surveillance des machines par analyse des vibrations*. 2009.

[37] R. CHAIB, "CONTRIBUTION A L'OPTIMISATION DE LA MAINTENANCE CONDITIONELLE PAR L'ANALYSE VIBRATOIRE," 2007.

[38] O. DJEBILI, "Contribution à la maintenance prédictive par analyse vibratoire des composants mécaniques tournants . Application aux butées à billes soumises à la fatigue de contact de roulement .," 2013.

[39] J. Lee, J. Ni, D. Djurdjanovic, H. Qiu, and H. Liao, "Intelligent prognostics tools and e-maintenance," *Comput. Ind.*, vol. 57, no. 6, pp. 476–489, 2006.

[40] "Bearing Data Center | Personal web site of Bearing Data Center." [Online]. Available: http://csegroups.case.edu/bearingdatacenter/home. [Accessed: 15-Jan-2017].

[41] K. Medjaher, J.-Y. Moya, and N. Zerhouni, "Failure prognostic by using dynamic Bayesian networks," *Dependable Control Discret. Syst.*, vol. 2, no. 1, pp. 257–262, 2009.