A Semi-supervised Support Vector Data Description- based Fault Detection Method for Rolling Element Bearings based on Cyclic Spectral Coherence

Chenyu LIU [1,2], Konstantinos GRYLLIAS [1,2]

[1]Department of Mechanical Engineering,
Faculty of Engineering Science, KU Leuven
[2]Dynamics of Mechanical and Mechatronic Systems, Flanders Make
Celestijnenlaan 300, BOX 2420, 3001 Leuven, Belgium
chenyu.liu@kuleuven.be, konstantinos.gryllias@kuleuven.be

Abstract

Modern Internet-of-Things (IoT)-driven condition monitoring exploits data from various surveillance tools to reflect the health status of machinery. However, many diagnostic tasks are still hardly achievable, considering the complex operating environment of mechanical components. Detecting rolling element bearing faults, for timely repair and maintenance before a total breakdown, is one of the challenges which has a significant impact on rotating machinery. Nowadays advanced signal processing techniques are combined with high level machine learning approaches, focusing towards automatic fault diagnosis. A plethora of diagnostic indicators have been proposed to track the system degradation. Cyclic Spectral Correlation (CSC) and Cyclic Spectral Coherence (CSCoh) have been proved as powerful tools in signal processing community. Due to the periodic mechanism of the bearing faults' impacts, the diagnostic indicators extracted from CSCoh domain can better detect bearing faults. On the other hand, supervised machine learning approaches with labeled training and testing datasets cannot be realistically obtained under industrial conditions. In order to overcome this limitation, a novel fault detection approach based on semi-supervised learning and Support Vector Data Description (SVDD) is proposed in this paper. The proposed SVDD model utilizes CSCoh domain indicators to build the feature space and fits a hyper-sphere to calculate the Euclidean distances in order to isolate the healthy and faulty data. Meanwhile a systematic fault detection decision strategy is proposed to estimate the bearing status simultaneously with the detection of fault initiation. Two run-to-failure bearing datasets are used to implement the proposed fault detection framework and the results show that the SVDD model with CSCoh indicators can precisely detect the bearing faults. Moreover, the CSCoh based indicators show superior characteristics in the detection process comparing to time and Fast Kurtogram (FK) based Squared Envelope Spectrum (SES) domain indicators.

1 Introduction

Rolling element bearings are important but fragile components and are highly sensitive to external operation environment. Premature bearing failure caused by mounting errors, poor lubrication, debris contamination etc. may trigger a breakdown of machinery or even a fatal accident [1]. Therefore, the early bearing fault detection arises as a critical mission in the frame of predictive maintenance, receiving extensive attention in recent years.

Data-driven based fault detection methods have been widely applied to monitor rolling element bearings in rotation machinery. Its framework consists of four major steps: 1) data acquisition, 2) construction of Health Indicators (HI), 3) determination of decision threshold, 4) detection of anomaly [2]. Vibration signals are widely used in the data acquisition step to detect machine faults. HIs, derived from different domains, are used to track the bearing fault progression. Time domain HIs, such as root mean square (RMS), kurtosis, skewness etc., are widely employed [3]. They represent the statistical features of the time sequence but cannot characterize nonlinearities of bearing vibrations. On the other hand, a plethora of signal processing methods have been developed to enhance the embedded defect information, including Envelope Analysis, Spectral Kurtosis,

Short-Time Fourier Transform, Cyclic Spectral Analysis etc. [4]. These approaches allow the extraction of sophisticated HIs which correlate with bearing fault characteristic frequencies with more expressiveness and robustness in practical applications.

Among the above mentioned signal processing methods, cyclic spectral analysis recently gained much attention due to its capability of revealing the second-order cyclostationary periodicities of rolling element bearing signals. It has been proved that the vibration signals of faulty bearings, especially with localized fault, will exhibit cyclostationary behaviour [6]. Cyclic Spectral Correlation (CSC) and Cyclic Spectral Coherence (CSCoh) have been developed as efficient tools in cyclostationary analysis [7]. They represent the potential fault modulation information into frequency-frequency domain bivariable maps. The integration over the spectral frequency, leads to the estimation of the Enhanced Envelope Spectrum (EES), which demonstrates the modulation frequencies and their harmonics. Due to the periodic mechanism of the bearing faults' impacts, the EES can provide a clearer representation of bearing faults compared to other methods such as the Squared Envelope Spectrum (SES) [8]. Some attempts have been proposed to construct HIs from EES. In [9], a self-running bearing diagnosis framework with scalar indicators derived from EES has been set up . Moreover the summation of the amplitudes of three harmonics, which appear at the characteristic fault frequencies of EES, has been used initially as diagnostic indicators [10], and further as bearing prognostic indicators.

Another key problem is the definition of a threshold, which can trigger the detection based on the HIs, estimated on signals emitted by faulty bearings. On one hand, traditional statistic decision threshold strongly depends on the expert knowledge of the indicators' statistical distribution as well as on the empirical understanding of bearing faults. On the other hand, the detection threshold should be able to precisely isolate the faults meanwhile tolerating outliers to maintain a robust performance of the false alarm and misdetection rate under different external environments. Nowadays machine learning approaches show great potential to achieve this comprehensive target. In the research of automatic fault diagnosis, machine learning algorithms are considered as black-box models which can exhibit system behaviours from data. Therefore, it can provide continuous information for the rapid deterioration of bearings and detect anomalies. Support vector machine (SVM) is one of the initial machine learning attempts in real-world fault detection application [11]. The major principle of SVM is to seek an optimal hyper-plane as the decision surface that can maximize the margin of separation between two classes. In the case of bearing fault detection, the hyper-plane of SVM has been used as the detection threshold to separate indicators extracted from bearing signals into healthy and faulty class [12, 13]. Artificial neural network (ANN) is also widely applied to fulfill the task of fault detection. A supervised non-linear autoregressive neural network with exogenous input, was trained in [14], to model the healthy condition of bearings meanwhile setting up the detection threshold using the Mahalanobis distance to detect the bearing faults.

In real industrial conditions, the bearings are operating for most of the time in healthy conditions and the faulty datasets are rather limited and usually not labelled. Therefore standard supervised machine learning approaches, which are based on known labels of the classes, cannot be used for fault detection. In order to overcome this limitation, semi-supervised fault detection techniques have been proposed, based on the training of the methodology exclusively on healthy data. Support Vector Data Description (SVDD) has been proposed by Tax and Duin [15] and has been successfully used in one-class classification problems for various applications. SVDD is inspired by SVM but uses a hyper-sphere, instead of a hyper-plane, to solve the non-linear separable problem. SVDD provides a geometric description of the observed data by calculating the Euclidean distances between the data and the center of the hyper-sphere. It is trained by minimizing the volume of the sphere with most training data. When testing with the trained SVDD, any data found to have larger distance than the boundary of the sphere, are considered to be anomalies. In the bearing fault detection, the hyper-sphere of the SVDD can be used as the detection threshold.

In this paper, the CSCoh based diagnostic indicators and the SVDD classifier are combined in order to detect bearing faults. A comprehensive bearing status decision strategy is proposed within this framework. The methodology is tested and evaluated on experimental data from run-to-failure bearing datasets. The results demonstrate the efficacy of the method, presenting high detection rate with low false alarm and misdetection rate. The rest of the paper is organised as following: the background of Cyclic Spectral Coherence is discussed in Section 2 and the theory of SVDD is introduced in Section 3. The proposed methodology and the experimental setup are illustrated in Section 4 and 5, respectively. The results are presented and analysed in Section 6. Finally, the paper closes with some conclusions.

2 Cyclic Spectral Coherence

The concept of cyclostationarity was first proposed in the field of telecommunications and later was introduced to the mechanical engineering community. As a subcategory of non-stationary processes, cyclostationarity is a stochastic process that exhibits hidden periodicities embedded in systems. When a fault is generated in rolling element bearings in different locations (inner race, outer race or rolling element), a series of shocks in repetition form will be generated simultaneously, modulated by other frequencies like the shaft speed. This phenomenon can be described as cyclostationary and can be exploited to detect bearing damage. Different orders of cyclostationarity are defined based on the order of moments. The first-order cyclostationarity (CS1) is described as:

$$R_{1x}(t) = R_{1x}(t+T) = \mathbb{E}\{x(t)\} \tag{1}$$

where the first-order moment, called the statistical mean, $R_{1x}(t)$ is periodic with period T. $x(t)$ represents the signal function at time t. $\mathbb{E}\{x(t)\}$ is the ensemble average, which represents the average of the same stochastic process of repeated experiments. CS1 exhibits as periodic waveforms in vibration signals of rotating machinery and can be generated by imbalances, misalignments or flexible coupling. Moreover, second-order cyclostationary (CS2) depicts the periodicity of second-order moments with the autocorrelation function:

$$R_{2x}(t, \tau) = R_{2x}(t+T, \tau) = \mathbb{E}\{x(t)x(t-\tau)^*\} \tag{2}$$

where τ is the time-lag. CS2 provides a distinction of the stochastic process with amplitude or frequency modulation. CS2 has been proved extremely effective to achieve diagnosis on rotating components that are not completely phase-locked with shaft speeds, such as rolling element bearings.

The Cyclic Spectral Correlation (CSC) is designed as a tool to describe CS1 and CS2 signals in frequency-frequency domain. It is defined as the autocorrelation function of two frequency variables using the two-dimensional Fourier transform, as shown in Eq. 3:

$$CSC_x(\alpha, f) = \lim_{W \to \infty} \frac{1}{W} \mathbb{E}\{\mathscr{F}_W[x(t)]\mathscr{F}_W[x(t+\tau)]^*\} \tag{3}$$

where α is the cyclic frequency related with the modulation and f is the spectral frequency representing the carrier. $\mathscr{F}$ stands for the Fourier transform and W is the time duration. Furthermore, a normalization procedure can be added to the CSC, to minimize uneven distributions, which is known as Cyclic Spectral Coherence (CSCoh). The CSCoh is estimated as:

$$CSCoh_x(\alpha, f) = \frac{CSC_x(\alpha, f)}{\sqrt{CSC_x(0, f)CSC_x(0, f-\alpha)}} \tag{4}$$

The integration of the CSCoh over the spectral frequency f from zero to the Nyquist frequency can lead to the Enhanced Envelope Spectrum (EES) which is an improved version of the Squared Envelope Spectrum, as described in Eq. 5:

$$IES(\alpha) = \frac{1}{F_2 - F_1} \int_{F_1}^{F_2} |CSCoh_x(\alpha, f)| df \tag{5}$$

3 Semi-supervised SVDD

3.1 Theory

Semi-supervised learning approaches focus on partially labelled datasets. The objective of a semi-supervised model is to classify some of the unlabelled data, leveraging information from the labelled part. SVDD has been developed as one-class classification method [15] in the way of semi-supervised learning. It can be seen as an extension of SVM, which uses a hyper-sphere instead of a hyper-plane as the classification decision surface. The hyper-sphere is characterized by the radius R and the center a, which are the decision variables. The labelled data "targets" are regarded as the training set of SVDD. The data, being exactly on the boundary, are the support vectors and the data outside are "outliers". The primary goal of SVDD is to construct a hyper-sphere with a minimum radius, which simultaneously contains the maximum number of targets. The objective function can be described by Eq. 6:

$$F(R,a) = \min R^2 + C\sum_{i=1}^{n} \xi_i \qquad (6)$$

This function subjects to two constraint conditions: $\|x_i - a\|^2 \leq R^2 + \xi_i, \forall i = 1,\ldots,n$ and $\xi_i \geq 0, \forall i = 1,\ldots,n$. Here $x_i \in \mathbb{R}$ is the training dataset with n samples. ξ_i is the slack variable which represents the tolerance of targets being outside the boundary. The slack variable is used to avoid an extremely large radius hyper-sphere that reduces the descriptive ability. C is the penalty factor that controls the trade-off between the radius and the rejected number of data points. It can also be described as $C = \frac{1}{nf}$, where f is the outlier fraction. Eq. 6 is a convex quadratic programming problem which cannot be solved for unknown R. However, it can be transformed to an equivalent dual problem with Lagrange duality. By applying Lagrange multipliers, the constraint conditions can be fused into the objective function, which turns it to a dual form as Eq. 7:

$$L = \max \sum_{i=1}^{n} \alpha_i(x_i \cdot x_i) - \sum_{i,j} \alpha_i \alpha_j(x_i \cdot x_j) \qquad (7)$$

where $(x_j \cdot x_i)$ stands for the inner product of x_i and x_j, $\alpha_i \in \mathbb{R}$ are the Lagrange multipliers. New conditions are set for the dual form: $\sum_{i=1}^{n} \alpha_i = 1$ and $0 \leq \alpha_i \leq C, \forall i = 1,\ldots,n$. The dual form is composed of the data itself which makes the problem solvable. The training data x_i and its corresponding α_i are related with the radius R and the center a of the hyper-sphere according to Eq. 8:

$$\begin{cases} x_i - a^2 < R^2 \iff \alpha_i = 0 \\ x_i - a^2 = R^2 \iff 0 \leq \alpha_i \leq C \\ x_i - a^2 > R^2 \iff \alpha_i = C \end{cases} \qquad (8)$$

3.2 Flexible data description

The inner product $(x_i \cdot x_j)$ in Eq. 7 can be replaced with kernel functions $K(x_i, x_j)$ to make a more flexible description for non-linear datasets. The kernel function can map the data to a higher dimension feature space which makes the non-linear data separable. The kernel functions have various choices such as Gaussian, linear, polynomial etc. Among all the functions, Gaussian kernel is the most common choice and is also adopted in this paper as described in Eq. 9:

$$K(x_i, x_j) = \exp\frac{-\|x_i - x_j\|^2}{2s^2}; \qquad (9)$$

where s represent the kernel width parameter. By using the kernel function, the distance of any observation z to the center of the sphere can be described as follows:

$$\text{dist}^2(z) = K(z,z) - 2\sum_{i} \alpha_i K(x_i,z) + \sum_{i,j} \alpha_i \alpha_j K(x_i,x_j) \qquad (10)$$

Considering the previous spherical data boundary conditions in Eq. 8, $\text{dist}^2(z) < R^2$ represents the position inside the sphere and $\text{dist}^2(z) = R^2$ is on the boundary of the sphere. When $\text{dist}^2(z) > R^2$, then z is recognized as an outlier. It should be noticed that the width of the kernel function K will influence on the calculation of the distance and is optimized in the proposed methodology, as explained in the following section.

4 Proposed methodology

4.1 Health Indicator construction

In order to realise the fault detection, HIs need to be constructed to evaluate the current condition as well as the degradation level of the bearing. The goodness of HIs can be characterized by monotonicity, trendability and prognosability. In real applications, the HIs should be able to detect bearing faults in early stage meanwhile minimizing the false alarm and misdetection rate. Time domain statistic features, such as Root Mean Square (RMS), kurtosis, skewness etc. are the most used HIs since they are effective to capture instant faults, such as using kurtosis in the detection of bearing spall initiation. However, the effectiveness of time domain HIs is restricted when the faults are distributed across the bearing surface. During the developing of damage, the

vibration signals exhibit more cyclostationarity with masked fault signatures. To overcome this limitation, the HIs, in this paper, are constructed using EES features from the CSCoh domain, as shown in Tab. 1. The CSCoh domain HIs are defined as the amplitude of the harmonics of the bearing characteristic frequencies extracted from the EES, i.e. the Shaft Rotation speed (SR), the Ball Pass Frequency Outer race (BPFO), the Ball Pass Frequency Inner race (BPFI) with sidebands and the Ball Spin Frequency (BSF) with sidebands. Four (4) harmonics and their sum are chosen to keep the robustness of the HIs and to facilitate the early start of detection.

S1R	EES peak at SR	S2I	EES peak at 2nd harmonic BPFI and sidebands
S2R	EES peak at 2nd harmonic of SR	S3I	EES peak at 3rd harmonic BPFI and sidebands
S3R	EES peak at 3rd harmonic of SR	S4I	EES peak at 4th harmonic BPFI and sidebands
S4R	EES peak at 4th harmonic of SR	SIR	Sum of 4 EES peaks of BPFI harmonics and sidebands
SSR	Sum of 4 EES peaks of SR harmonics	SIOR	Sum of 4 EES peaks of BPFO and BPFI harmonics
S1O	EES peak at BPFO	S2B	EES peak at two times BSF and sidebands at FTF
S2O	EES peak at 2nd harmonic of BPFO	S4B	EES peak at four times BSF and sidebands at FTF
S3O	EES peak at 3rd harmonic of BPFO	S6B	EES peak at six times BSF and sidebands at FTF
S4O	EES peak at 4th harmonic of BPFO	S8B	EES peak at eight times BSF and sidebands at FTF
SOR	Sum of 4 EES peaks of BPFO harmonics	SBF	Sum of 4 EES peaks of BSF and sidebands
S1I	EES peak at BPFI and sidebands		

Table 1: Health indicators from CSCoh domain

To compare the classification performance by using different health indicators, 9 statistic features from the time domain are also extracted during the test, as shown in Tab.2. Meanwhile, the same 21 frequency domain features from the Fast Kurtogram (FK) based SES are also extracted as another comparison group. The results from these indicators are illustrated in the following section.

RMS	$RMS = \sqrt{\frac{\frac{1}{n}\sum_{i=1}^{n} x_i^2}{n}}$	Variance	$VR = \frac{\sum_{i=1}^{n}(x_i - m)^2}{(n-1)\sigma^2}$	Shape Factor	$SF = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}}{\frac{1}{n}\sum_{i=1}^{n}|x_i|}$
Kurtosis	$KU = \frac{\sum_{i=1}^{n}(x_i-m)^4}{(n-1)\sigma^4}$	Mean	$m = \frac{\sum_{i=1}^{n} x_i}{n}$	Impulse Factor	$IF = \frac{\max|x_i|}{\frac{1}{n}\sum_{i=1}^{n}|x_i|}$
Skewness	$SK = \frac{\sum_{i=1}^{n}(x_i-m)^3}{(n-1)\sigma^3}$	Crest Factor	$CF = \frac{\max|x_i|}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}}$	Clearance Factor	$CLF = \frac{\max|x_i|}{(\frac{1}{n}\sum_{i=1}^{n}\sqrt{|x_i|})^2}$

Table 2: Health indicators from time domain

4.2 Kernel width optimization using cross validation with grid search
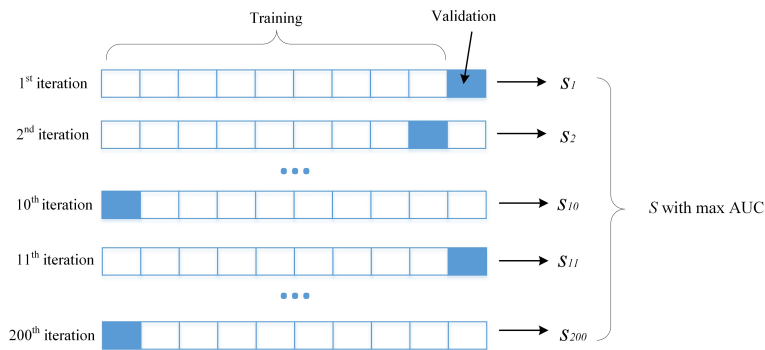

Figure 1: Optimization of kernel width using 10-fold Cross validation with grid search

The boundary of the hyper-sphere can be described with a kernel function for the SVDD classifier. In this paper, the Gaussian kernel is used since it has been proved effective in former research [15]. On the other hand, the kernel width acts as a hyperparameter in the SVDD model and influences the classification performance. The Receiver Operating Characteristic (ROC) curve is commonly used to evaluate the performance of a binary

classifier. The corresponding area under the ROC curve (AUC) is the key metric. The classifier with a larger AUC value performs better than the others. In order to tune the kernel width meanwhile proceeding cross validation during the training process, cross validation with grid search is adopted to search for the kernel width, that can maximize the AUC of the classifier. With 10-fold cross validation method, 200 iterations are proceeded to search for the kernel width s, which gives the maximum AUC values as depicted in Fig. 1. In this way, the model could achieve an automatic parameter optimization to avoid a stochastic selection of the hyperparameter. The results from the ROC curve will be discussed in Section 6.

4.3 Bearing fault detection procedure

The practical need for the bearing fault detection approach is a systematic decision strategy as a complement of the SVDD classifier based on CSCoh indicators. The flowchart of the proposed bearing fault detection process is shown in Fig. 2. The entire procedure consists of three major parts: feature extraction, classification and fault detection. The CSCoh based HIs are extracted from the EES to construct the feature space in the feature extraction stage. The features are then normalized by subtracting the mean value and dividing by the standard deviation to avoid the calculated Euclidean distances be governed by particular features in the SVDD classifier. Semi-supervised methods are able to learn from fewer labelled data points with the help of a large number of unlabelled data points, therefore the training data are selected from the health conditions which normally can be gathered at the beginning period of the bearing operation. In the classification stage, the entire dataset is split into training and testing sets with 25% and 75% of the whole dataset, respectively. The training set includes the validation part to realise the parameter optimization and meanwhile is labelled as healthy. A banana dataset is generated as artificial outliers around these real data to train the hyper-sphere as the classifier.
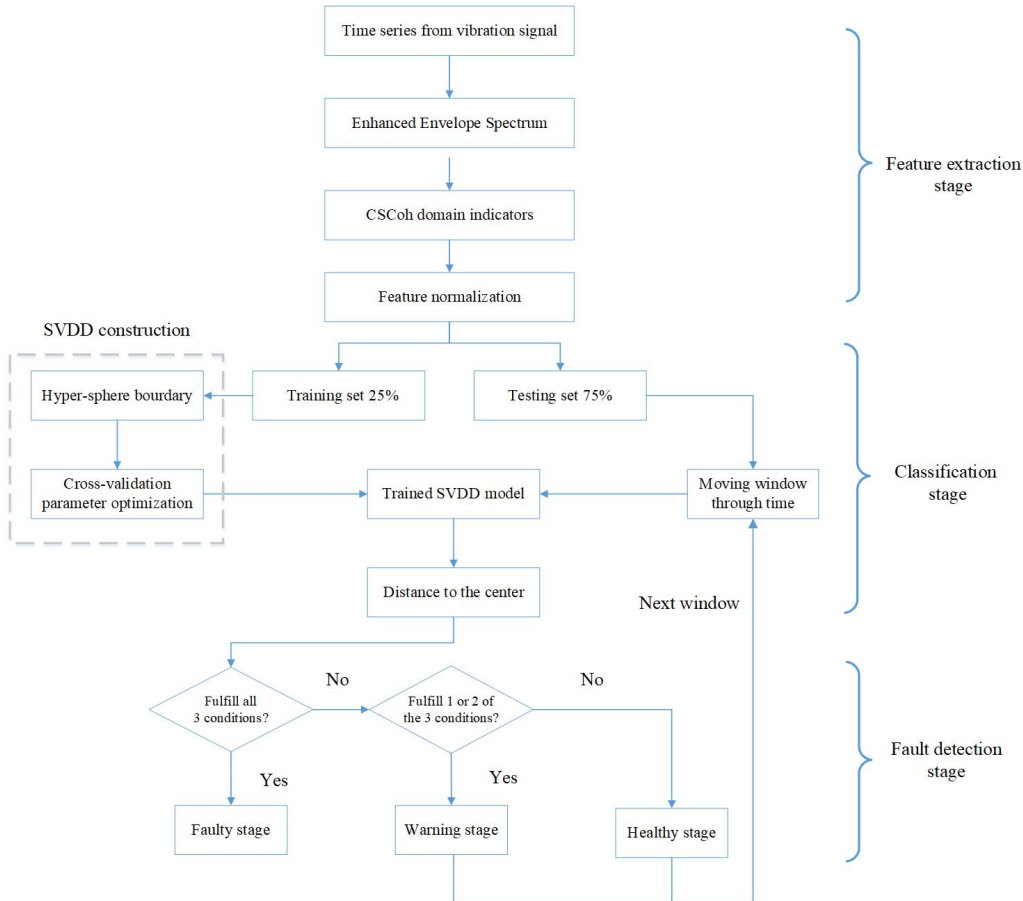

Figure 2: Procedure for the SVDD-based fault detection

Once the SVDD classifier is trained, a moving window is applied on the testing set. The data in the window are sent to the SVDD model to calculate the Euclidean distances to the center of the hyper-sphere. Considering the distances and the threshold, i.e. the boundary of the hyper-sphere, a fault detection decision strategy is built following 3 conditions: 1) 50% or more of the distances in the window are above the threshold. 2) 50% or

more of the distances in the window continuously pass the threshold. 3) The average distance of the data in the window is equal or greater than the threshold. These conditions can keep the detection threshold robust from the influence of random dominant outliers meanwhile reducing the misdetection rate. Three bearing statuses are then defined, according to the conditions. When all the three conditions are fulfilled, the status of the bearing is considered as faulty. In contrast, if less than 3 conditions are satisfied, then the status is considered as warning. If the distances from the bearing data obey none of the three conditions, it is considered as in a healthy status. Additionally, the starting point of the warning status represents the existence of a premature anomaly and the starting point of faulty status is the detection point of a mature bearing defect.

5 Experimental setup and dataset

The proposed SVDD based semi-supervised methodology is tested and validated using the NASA Intelligent Maintenance Systems (IMS) dataset [16]. The IMS dataset was collected during the rolling element bearing endurance experiment by using a dedicated test rig in the University of Cincinnati. The layout of the test rig is shown in Fig. 3, which consists of an electric motor coupled with a rub belt with 2000 rpm stationary speed, 4 double row Rexnord ZA-2115 bearings mounted on a common shaft lubricated by a circulation system, a 6000 lbs radial load applied on the bearing 2 and 3, a shaft and PCB 353B33 High Sensitivity Quart ICP accelerometers. The collecting of data was stopped when the accumulated debris exceeded a certain level on a magnetic plug and was considered as the end of bearing life. The physical parameters and the characteristic frequencies can be found in Fig. 3.
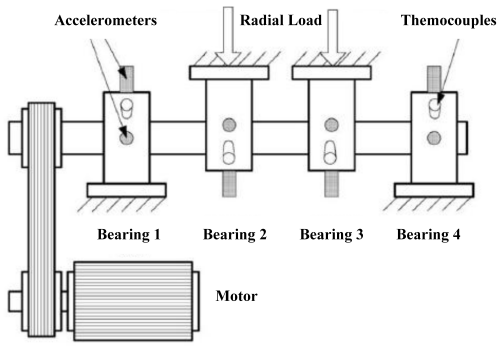
Physical Parameters	
Pitch diameter	71.5 mm
Rolling element diameter	8.4 mm
Number of rolling element per row	16
Contact angle	15.7
Static load	26690 N
Characteristic Frequencies	
Shaft rotation speed	33 Hz
Ball pass frequency outer race (BPFO)	236 Hz
Ball pass frequency inner race (BPFI)	297 Hz
Ball spin frequency (BSF)	139 Hz
Fundamental train frequency (FTF)	15 Hz

Figure 3: IMS test rig layout (left) and characteristics of bearings (right)

The run-to-failure experimental results were published in three datasets with numerous of files. Some details should be noticed when processing these datasets. According to the document describing the experiment, the sampling frequency should be 20 kHz but the file for one second includes 20,480 points so the sampling frequency probably was 20.48 kHz and this one is used in this paper. Moreover, the measurement for Dataset 1 is not continuous but with several interruptions during recording. These interruptions influences the behaviour of some HIs and this issue will be discussed in the following section. The signals from the dataset will be marked with time (Day) and number of signal (# signal number).

6 Results and discussion

6.1 Case 1: Dataset 1 Bearing 3

The 9 time domain HIs, the 21 FK based SES domain HIs and the 21 CSCoh domain HIs estimated processing the Dataset 1 Bearing 3, after normalization, are presented in Fig. 4. It is noticeable that the time axis is not in equal intervals for this map and other figures involving Dataset 1 in this section, due to the recording interruptions. The time axis represents the real time points in days from beginning (Day 0) to the end of the experiment (Day 34.5) but not the recording time. Analysing the time and the FK based SES features in the map, it can be visually observed that there is a significant changing step before Day 7.1 (#157), caused by a long period interruption. Therefore the data before Day 7.1 are abandoned to avoid contamination of the
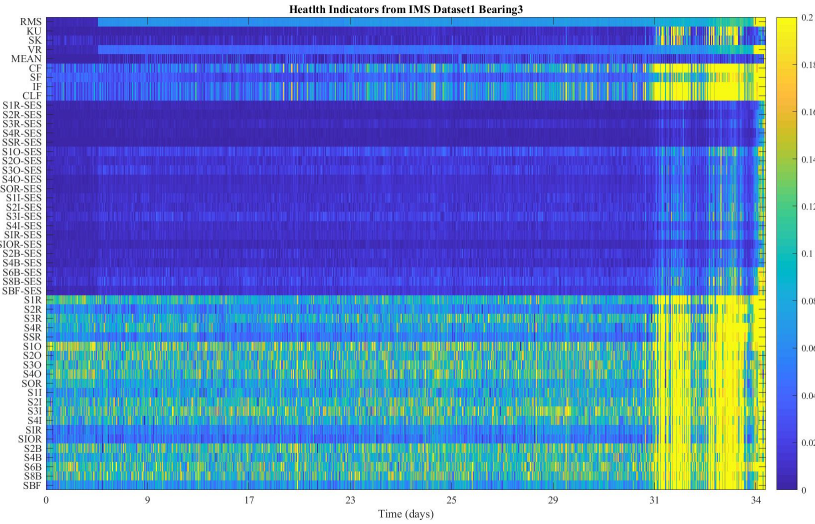
Figure 4: Health indicator map for Dataset 1 Bearing 3

healthy training set, which is selected from Day 7.4 (#200) to Day 17 (#600). The remaining data are used for testing.

The distances of the CSCoh indicators to the SVDD center are shown on the left of Fig. 5. Based on the proposed decision framework, the fault detection result is shown on the right of Fig. 5. The warning stage starts from Day 29.2 (#1570) till Day 30.9 (#1795) and the fault is detected also at Day 30.9 (#1796).
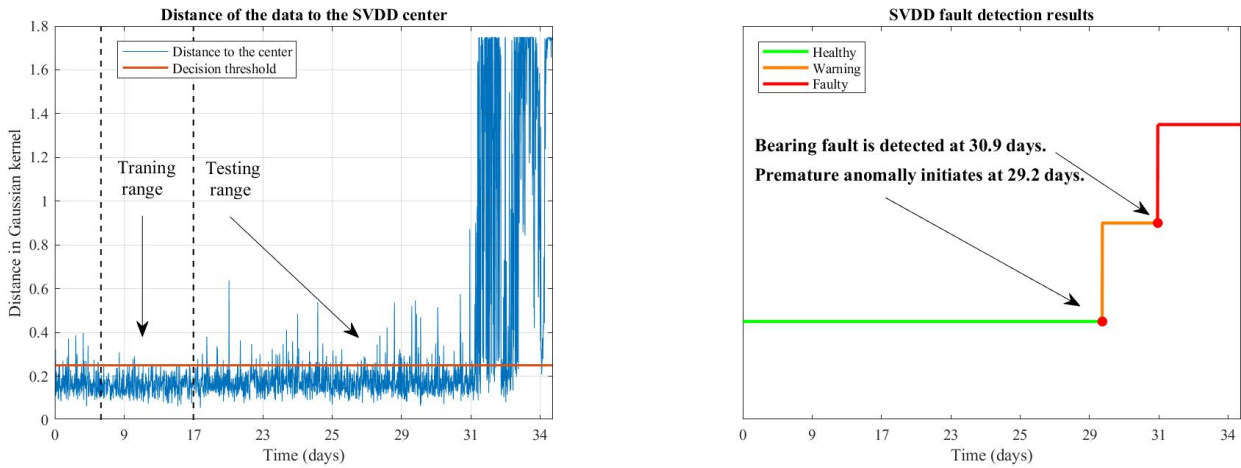

Figure 5: Fault detection restuls (left) and testing ROC curve (right)
of CSCoh indicators of Dataset 1 Bearing 3

The evaluation of the detection performance uses results from the study [17] as reference. The testing ROC curve of the SVDD classifier is plotted in Fig. 6. As discussed in Section 4, the ROC curve is employed to evaluate the classification ability of a binary classifier. The horizontal and vertical axis of a ROC curve represent the False Positive rate (FPr) and the True Positive rate (TPr). The classifier behaves better when the curve is closer to the left up point where FPr=0, TPr=1. Additionally, the area under ROC curve (AUC) can be used as a quantitative index in the evaluation. A larger AUC represents a better classification performance. Furthermore, more indexes are used to evaluate the classifier in the later part.

The HIs from the time and the FK based SES domain are also used to train the SVDD model as shown in Fig. 7 and 8. Compared to the results of CSCoh, it is obvious that the distances of the time and the FK based SES domain indicators exhibit wide fluctuations around the threshold. The outliers above the threshold spread in the whole testing range for both the two HI groups. Hence, the fault detection results show longer warning stages for the two groups with 7.9 and 8 days respectively. The time indicators detected the fault at Day 31 (#1820) and the FK based SES indicators give the detection at the same time as the CSCoh indicators at Day
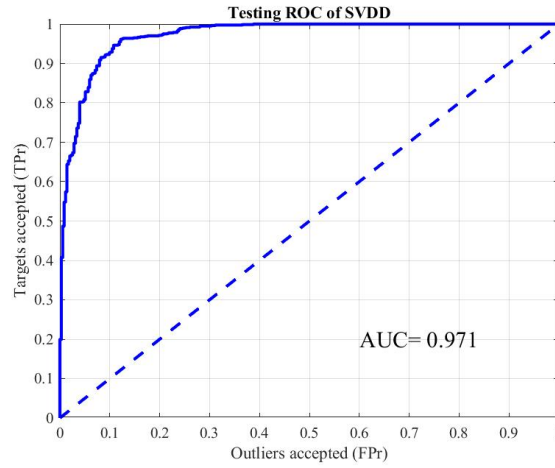
Figure 6: ROC curve for SVDD with CSCoh indicators of Dataset 1 Bearing 3

30.9 (#1796). The detection results show that time and FK based SES indicators are sensitive to the changing of external environment due to the recording interruptions. On the contrary, the distances of SVDD results from CSCoh indicators keep steady during most of the testing period which indicates that the CSCoh HIs are robust in detecting the bearing faults.
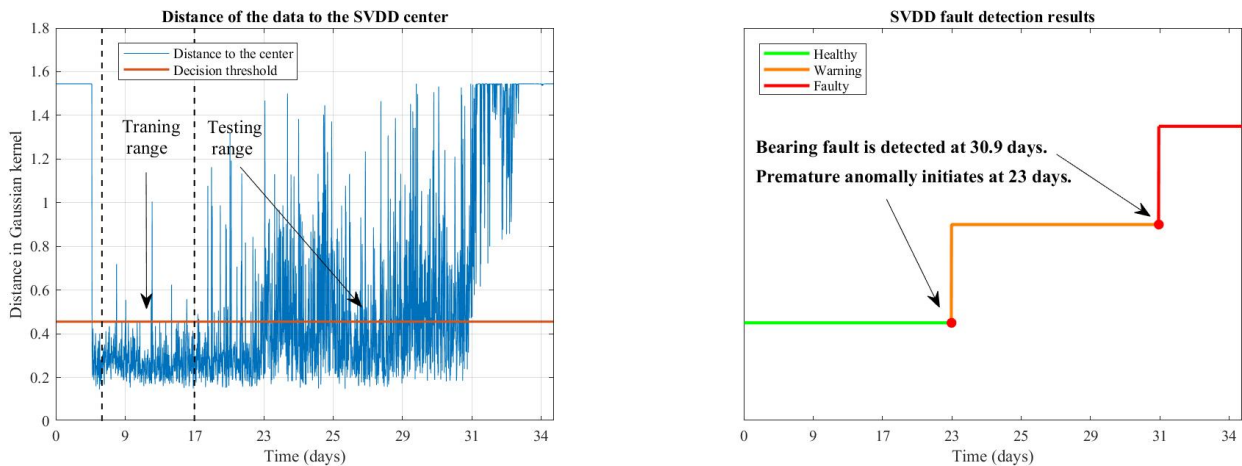

Figure 7: Distances to SVDD center (left) and fault detection results (right) for time domain indicators of Dataset 1 Bearing 3

The metrics derived from the classification confusion matrix are adopted to compare the performance of SVDD with different HIs, including the False Positive Fraction (FPF), the False Negative Fraction (FNF), the False Positive Rate (FPR) and the False Negative Rate (FNR). The calculation of the four values is shown in Fig. 9. The metrics represent the classification performance, measured from the estimated labels and the true labels, based on [17]. More specifically, FPR can be seen as the evaluation of false alarms and FNR represents the misdetection of the classifier. Lower values of these metrics indicate a better classification performance. The metrics together with the AUC value of all three group of HIs are listed in Tab. 3. The SVDD classifier apparently performs better with indicators from CSCoh than the other two groups, presenting superiority in the accurately detection of bearing faults.

6.2 Case 2: Dataset 2 Bearing 1

The HI map of Dataset 2 Bearing 1 is shown in Fig. 10. Since the measurement of this dataset is continuous without any interruption, all the indicators behave more monotonic than in Dataset 1. A clear separation can be seen at Day 4.9 for all the indicators which represent a dramatic increasing of the HIs. On the other hand, the HIs from the CSCoh show a forehead increase at Day 3.8 for the indicators constructed with the harmonics and
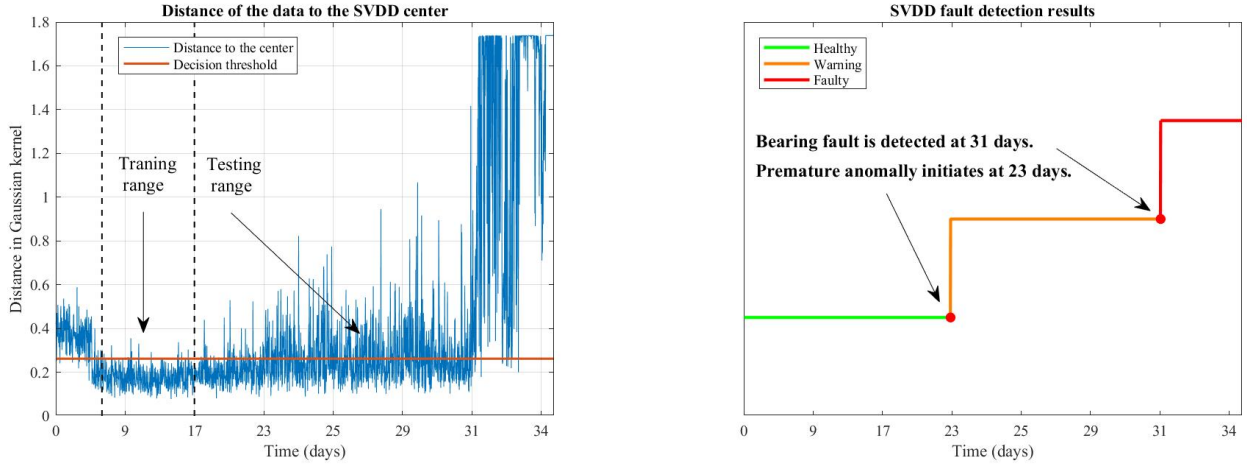
Figure 8: Distances to SVDD center (left) and fault detection results (right) for
FK based SES indicators of Dataset 1 Bearing 3

Figure 9: Calculation of the evaluation metrics.

	FPF	FNF	FPR	FNR	AUC
Time	0.474	0.195	0.370	0.085	0.905
FK based SES	0.248	0.040	0.525	0.023	0.963
CSCoh	0.076	0.023	0.177	0.019	0.971

Table 3: Performance of SVDD classifiers with different indicators for Dataset 1 Bearing 3

the sum of BPFO, which is the fault type in this case.

The distances from the SVDD and the fault detection results of the CSCoh indicators are shown in Fig. 11. The training set is selected from Day 0.3 (#50) to Day 2.1 (#300) and the remaining data are used for testing. Less distances pass the threshold which represent the monotonicity of the indicators. The warning stage is 0.1 days from Day 3.6 (#521) to Day 3.7 (#533) which is relatively short due to the indicators with more tendency. Then the bearing fault is detected at Day 3.7 (#534).

The indicators from the time and the FK based SES are also sent to the SVDD model to realise fault detection and the results are shown in Fig. 12 and 13. The CSCoh indicators provide the earliest detection compared to the time indicators (Day 3.8, #550) and the FK based SES indicators (Day 3.7, #541). The amplitudes of the characteristic frequencies and their harmonics from EES present more monotonicity and trendability which provide evidence in the construction of feature space during the training of SVDD model. As a result, the detection performs better using the CSCoh indicators.

The metrics of the classifiers' performance among different groups of HIs show less variance, compared to the first case. The CSCoh indicators gain lower FPF, FPR and FNR. The time domain indicators get lower FNF but the AUC is still lower than for the CSCoh.
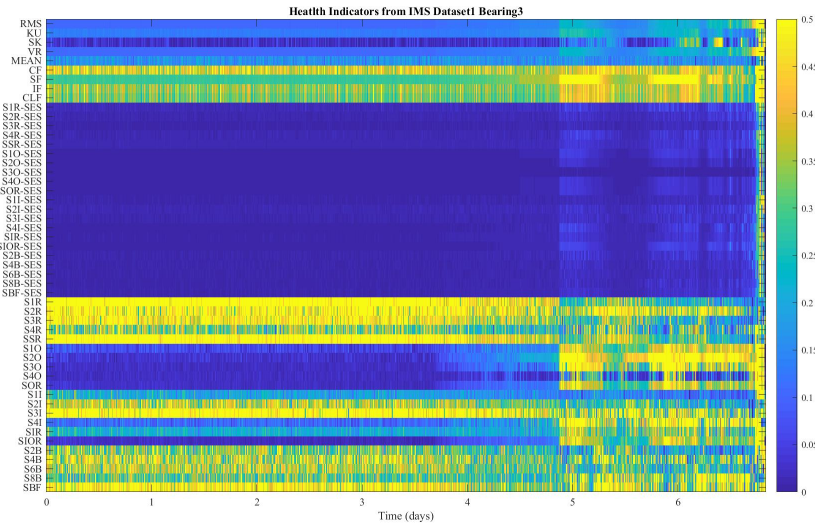
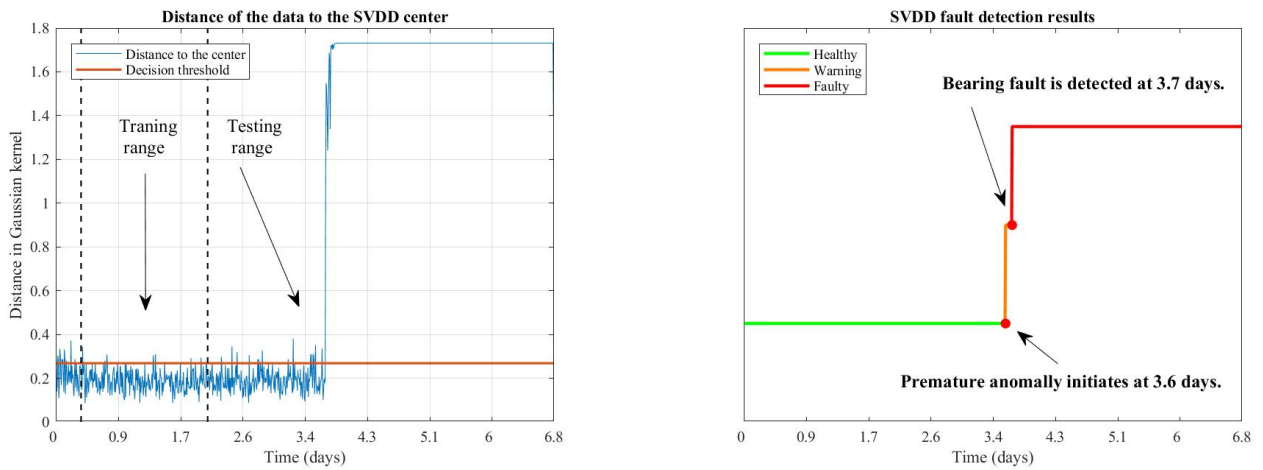Figure 10: Health indicator map for Dataset 2 Bearing 1

Figure 11: Distances to SVDD center (left) and fault detection results (right) for
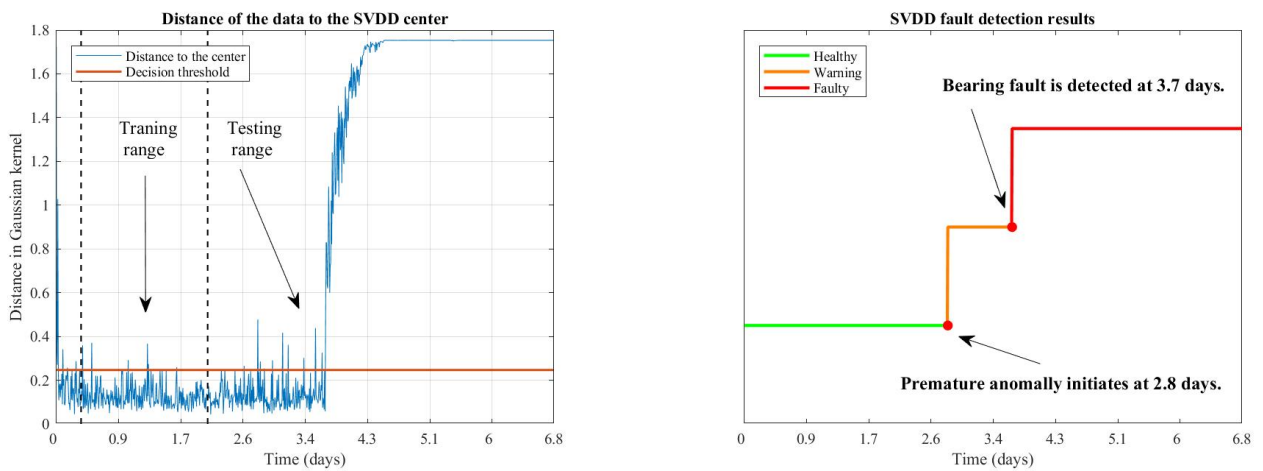CSCoh indicators of Dataset 2 Bearing 1

Figure 12: Distances to SVDD center (left) and fault detection results (right) for
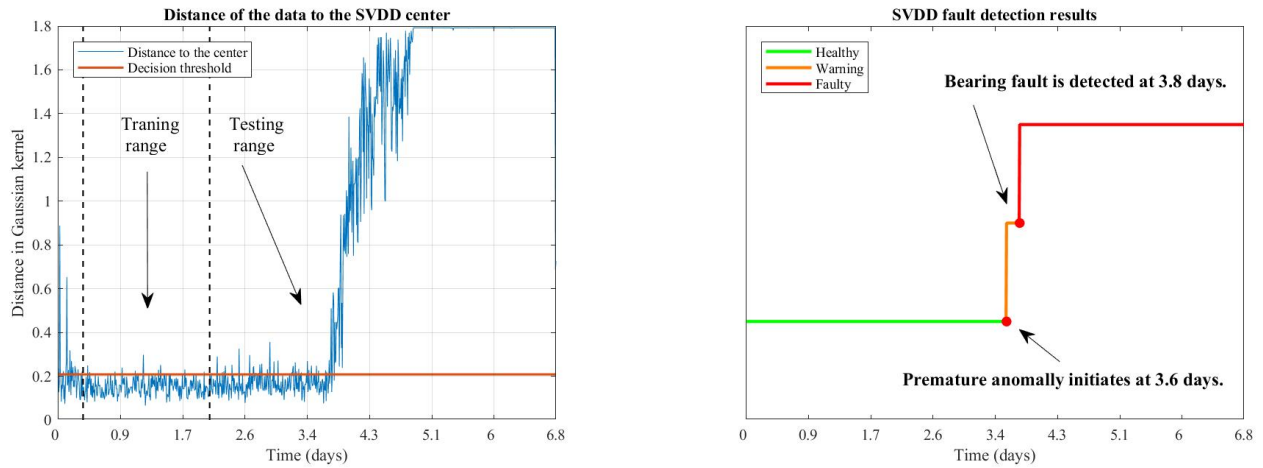time domain indicators of Dataset 2 Bearing 1

Figure 13: Distances to SVDD center (left) and fault detection results (right) for
FK based SES indicators of Dataset 2 Bearing 1

	FPF	FNF	FPR	FNR	AUC
Time	0.078	0.018	0.028	0.041	0.991
FK based SES	0.090	0.020	0.038	0.005	0.992
CSCoh	0.056	0.020	0.028	0.005	0.992

Table 4: Performance of SVDD classifiers with different indicators for Dataset 2 Bearing 1

7 Conclusion

In this paper, a novel rolling element bearing automatic fault detection approach is proposed by combining CSCoh domain indicators and a semi-supervised SVDD technique. Experimental results from run-to-failure bearing datasets prove that the bearing faults can be accurately detected with the proposed methodology. Compared to the indicators from time domain and the FK based SES domain, CSCoh diagnostic indicators are more robust to the changing of external environment and improve the performance of the SVDD classifier. The proposed semi-supervised methodology therefore has a strong practical significance in industrial applications since it can be effective in both bearing failure warning and damage detection.

Acknowledgements

References

[1] R. Randall , J. Antoni, *Rolling element bearing diagnostics-A tutorial*, Mechanical Systems and Signal Processing, 2011 February, 1, vol. 25, no. 2, pp.485-520.

[2] B. Zhang, G. Georgoulas, M. Orchard, A. Saxena, D. Brown, G. Vachtsevanos, S. Liang. *Rolling element bearing feature extraction and anomaly detection based on vibration monitoring*, 16th Mediterranean Conference on Control and Automation, 2008 June 25, pp.1792-1797.

[3] K. Gryllias, S. Moschini, J. Antoni, *Application of cyclo-nonstationary indicators for bearing monitoring under varying operating conditions*, Journal of Engineering for Gas Turbines and Power, 2018 Jan 1, vol 140, no. 1.

[4] C. Yi, D. Wang, W. Fan, KL. Tsui, J. Lin, *EEMD-based steady-state indexes and their applications to condition monitoring and fault diagnosis of railway axle bearings*, Sensors, 2018 Feb 27, vol. 18, no. 3, pp.704.

[5] Y. Lei, F. Jia, J. Lin, S. Xing, SX. Ding, *An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data*, IEEE Transactions on Industrial Electronics, 2016 May, vol. 63, no. 5, pp.3137-3147.

[6] J. Antoni, F. Bonnardot, A. Raad, M. El Badaoui, *Cyclostationary modelling of rotating machine vibration signals*, Mechanical Systems and Signal Processing, 2004 Nov 1, vol. 18, no. 6, pp.1285-1314.

[7] A. Mauricio, A. Miguel, W. Smith, R. Randall, J. Antoni, K. Gryllias, *Cyclostationary-based tools for bearing diagnostics*, Online Proceedings of ISMA 2018 including USD 2018, 2013, pp.905-918.

[8] A. Mauricio, J. Qi, K. Gryllias. *Vibration-Based Condition Monitoring of Wind Turbine Gearboxes Based on Cyclostationary Analysis*, Journal of Engineering for Gas Turbines and Power, 2019 Mar 1, vol. 141, no. 3.

[9] S. Kass, A. Raad, J. Antoni, *Self-running bearing diagnosis based on scalar indicator using fast order frequency spectral coherence*, Measurement, 2019 May 1, vol. 138, pp.467-484.

[10] J. Qi, A. Mauricio, K. Gryllias, *Enhanced Particle Filter and Cyclic Spectral Coherence based Prognostics of Rolling Element Bearings*, In PHM Society European Conference, 2018 Jun 28, vol. 4, no. 1.

[11] K. Gryllias, I. Antoniadis, *A Support Vector Machine approach based on physical model training for rolling element bearing fault detection in industrial environments*, Engineering Applications of Artificial Intelligence, 2012 Mar 1, vol. 25, no. 2, pp.326-44.

[12] T. Shon, Y. Kim, C. Lee, J. Moon, *A machine learning framework for network anomaly detection using SVM and GA*, In Proceedings from the sixth annual IEEE SMC information assurance workshop, 2005 Jun 15, pp.176-183.

[13] R. Perdisci, G. Gu, W. Lee, *Using an Ensemble of One-Class SVM Classifiers to Harden Payload-based Anomaly Detection Systems*, In ICDM, 2006 Dec 18, vol. 6, pp.488-498.

[14] P. Bangalore, LB Tjernberg, *An artificial neural network approach for early fault detection of gearbox bearings*, IEEE Transactions on Smart Grid, 2015 Mar, vol. 6, no. 2, pp.980-987.

[15] D. Tax, R. Duin, *Support vector data description*, Machine learning, 2004 Jan 1, vol. 54, no. 1, pp: 45-66.

[16] H. Qiu, J. Lee, J. Lin, G. Yu. *Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. Journal of sound and vibration*, 2006 Feb 7, vol. 289(4-5), pp.1066-1090.

[17] K. Gryllias, I. Antoniadis, C. Yiakopoulos, *A novel semi-supervised mathematical morphology-based fault detection & classification method for rolling element bearings*, In Proceedings of the 22nd International Congress on Sound and Vibration, 2015.