# Bearing Performance Degradation Assessment based on A Combination of Multi-Scale Entropy and K-medoids Clustering

Long Zhang, Chengyang Song, Chaobing Wang, Tianpeng Xu, Jianyu Yi
School of Mechatronics & Vehicle Engineering
East China Jiaotong University
Nanchang, China
Songcy94@163.com

***Abstract*—Bearing performance degradation assessment (PDA) underlies the residual useful life prediction and maintenance decision-making. In bearing fault diagnosis and PDA, a commonly used method is clustering analysis, among which K-medoids clustering is not susceptible to extreme data. The amalgamation of K-medoids clustering and Degree of Membership (DOM) is expected to give a health indicator with a determined value range for PDA. In this study a diagnostic model for bearing PDA based on Multi-scale entropy (MSE) and K-medoids clustering is proposed. Extracting the multi-scale entropy of the rolling bearing vibration signal to construct the K-medoids clustering model. The tested data are input into the model to obtain the Degree of Membership as the index to evaluate the current bearing operating state. Resultant health indicator is used to depict bearing health condition with the help of an adaptive threshold. Results on artificially induced faults and bearing run-to-failure data demonstrate the proposed method is able to track the progress of bearing faults and detect them at incipient stage.**

***Keywords-bearings; multi-scale entropy; k-medoids clustering; Performance degradation assessment***

## I. Introduction

Rolling bearings are one of the most widely used parts in mechanical equipment, and their operating conditions are related to the overall performance and life of the equipment. The vibration signal of the rolling bearing can reflect the information of fault such as the location and the degree. Analyzing it can help us to understand the operating status of the equipment and predict the remaining life, which provides a basis for formulating the maintenance strategy [1].

There are two types of degradation assessment models: probability models and distance models. Gaussian mixture model (GMM) and hidden Markov model (HMM) belong to the probability model. Support vector machine (SVM) and support vector description (SVDD) are the distance models. Hong [2] used GMM to calculate the density distribution of the feature. After the vibration signal was decomposed into EEMD, the difference between the basic feature space and the test feature space was calculated to evaluate the degradation of the bearing. Guo [3] combined the wavelet packet analysis with the support vector machine and used the geometric distance as the fault degree index. Rai [4] used the EMD and the K-medoids clustering model to achieve the degradation assessment. However, only the normal samples are used to construct the model in the above literature, and the use of the failed samples is lacking, the vibration data are not fully utilized.

This paper proposes a degradation assessment method based on Multi-scale entropy and K-medoids clustering. The K-medoids clustering model is established by extracting the multi-scale entropy of the bearing normal data and the failure data. Compared with the K-means clustering, K-medoids has better robustness to deal with abnormal data and noise data [5], and it is not easily affected by extreme data [6]. The membership function can express the similarity in a specific interval [0,1]. The degradation evaluation model can be established by combining K-medoids and membership function to evaluate the degree of failure and realize the quantitative evaluation of the degree of bearing failure.

The membership function can express the similarity in a specific interval [0,1]. The degradation evaluation model can be established by combining K-medoids and membership function to evaluate the degree of failure and realize the quantitative evaluation of the degree of bearing failure.

## II. MULTI-SCALE ENTROPY

### A. *Sample Entropy*

Entropy can be used to describe the uncertainty of information. Sample entropy is an improved algorithm of approximate entropy. Sample entropy algorithm is as follows [7]:

1) Given a raw data set of length n: $X=[x_1,x_2,x_3,\cdots x_n]$, It can obtain an m-dimensional vector $X(i)=[x_i,x_{i+1},\cdots,x_{i+m-1}]$, $1\leq i\leq n-m$, with given similar tolerance r and embedding dimension m.

2). Define the distance $d_{ij}$ Between x(i) and x(j) as the maximum absolute value of the difference.

3) For each i, the distance $d_{ij}$ ($1 \le i \le n-m$), $j \ne i$ between x(i) and the other vectors x(j) is calculated, and the number of $d_{ij}$ Less than r is counted and recorded as num($d_{ij}$<r). Defined $B_{im}(r)$ as follows:

$$B_{jm}(r) = num(d_{ij} < r) / (N - m - 1) \tag{1}$$

4) Caculate the average value $B^m(r)$ of $B_{im}(r)$.

5) For dimension m+1, repeat the steps (1)-(4) to compute the $B^{m+1}(r)$.

6) Finally, calculate the Sample entropy (SE) of the order using (2).

$$SE(M, r, n) = \ln B^m(r) - \ln B^{m+1}(r) \tag{2}$$

The value of Sample entropy is related to embedding dimension m, similarity tolerance r and data length n. In general, m = 2; r=0.15×SD (SD is the standard deviation of the original data).

### *B. Multi-scale Entropy*

Multi-scale entropy is to calculate the Sample entropy on multiple scales of the original signal. Different scales are obtained by coarse-graining procedure [8]. The calculation steps of MSE are as follows:

1) For a given original data $x=[x_1, x_2, x_3, \cdots, x_n]$, the length n, similar tolerance r and embedding dimension m , a new coarse-graining vector $P^{(s)}$ Can be computed by a (3).

$$p_j^{(s)} = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} x_i \tag{3}$$

$i \le j \le n/s$, scale factor $s=[1,2,\cdots,s_{max}]$ is positive integer. In the actual test, the calculation efficiency is low when the value of scale factor is too large, and the information of vibration signal can't be extracted completely when the value is too small. In this paper, $s_{max}$=20.

2) Calculate the Sample entropy of each coarse-grained sequence, then the Sample entropy of s coarse-grained sequences is obtained. The Sample entropy is plotted as a function of scale factor s, which is called multi-scale entropy analysis.

## III. K-MEDOIDS CLUSTERING

K-medoids is a clustering method based on partition and an improvement of K-means algorithm. Among various K-medoids algorithms, Partitioning Around Medoid (PAM) proposed by Kaufman and Rousseeuw (1990) [9] is the most effective one. Given the data object $x = \{x_1, x_2, \cdots, x_n\}, x \in R^s$, the aim of the algorithm is to divide the data set into K clusters, $C=\{c_1, c_2, \cdots, c_k\}$ is each cluster, so that the object with the most central position in the cluster is the representative object, the central point, while the other objects are non-representative objects, and the sum of the differences between all non-representative objects and the central point of the cluster is the smallest. The clustering effect is evaluated by the following (4).

$$J = \sum_{j=1}^{k} \sum_{x_i \in k_j} D(x_i, o_j) \tag{4}$$

$x_i$ Is data object, $O_j$ Is the cluster center of cluster $c_j$, $D(x_i, o_j)$ is the spatial euclidean distance of $x_i$ And $o_j$.

$$D(x_i, m_j) = \sqrt{\left\| x_i - o_j \right\|^2} \tag{5}$$

The steps involved in K-medoids algorithm are as follows:

*1) Initialization.* Random selection of K objects $\{O_1, O_2, \cdots, O_K\}$ as initial centers in data set X.

*2) Cluster Establishment.* The remaining objects are divided into clusters nearest to the center point and computed j.

*3) Exchange.* In cluster $C_j$ , a non-representative object $O_{rand}$ Is randomly selected to exchange with the central point $O_i$, replacing $O_i$ And recalculating the clustering effect, which is denoted as $J'$.

*4) Selection.* If $J'$<J, then $O_{rand}$ Replaces $O_i$, otherwise it will not change.

*5) Circulation.* Repeat steps 3 and 4 until the central point does not change.

## IV. MEMBERSHIP FUNCTION

The membership function can be used to express the relationship of membership between elements and sets. The membership function is given by (6).

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{ik}}\right)^{z/(m-1)}} \tag{6}$$

Where $\mu_{ij}$ Is the membership of the i th sample to the j th cluster center, $0 \le \mu_{ij} \le 1$. When $\mu_{ij}$=1, representing i belongs to j, when $\mu_{ij}$=0, representing i does not belong to j; c is the number of clustering centers; $d_{ij}$ Is the Euclidean distance from the i sample to the j clustering center; $d_{ik}$ Is the distance from the i sample to the K clustering center; q is the fuzzy weighting index; m is the fuzzy weighting index.

As for m, Bezdek [10] considers that it is most appropriate to consider m = 2 from the physical interpretation of FCM, Pal [11] considers the value range of m from the perspective of cluster validity is [1.5, 2.5]. In this paper, m=2 is selected according to experience.

Calculating the distance between the measured bearing life data and the two clustering centers, and inputting it into the membership function to obtain the DOM of the normal clustering center, which is used as the evaluation index DI (Degradation Index) of the performance degradation evaluation of bearings.

## V. K-MEDOIDS METHOD BASED ON MULTI-SCALE ENTROPY

Multi-scale entropy feature extraction of rolling bearing signal is presented. Two clustering centers can be obtained by training K-medoids model with the early normal samples of bearings to be detected and the same type of training bearings healthy samples and failure samples. It includes healthy sample clustering center $O_{normal}$ And fault sample clustering center $O_{failure}$. Performance evaluation of tested bearings based on $O_{normal}$. Fuzzy evaluation is introduced to calculate the DOM of cluster centers with normal samples as evaluation indexes. The specific steps in this paper are as follows:

*1) Extracting Fault Features*. The values of MSE are related to three parameters, similarity tolerance r, embedding dimension m and scale factor s. In this paper, 20 coarse-grained vector sequences are obtained by taking scale factor s=20. Then, by calculating the Sample entropy of each sequence, a 20-dimensional feature vector can be obtained, and the obtained vector can be regarded as input.

*2) Establishing Evaluation Model*. K-medoids clustering model is trained by feature vectors, and the clustering centers $\{O_{normal}, O_{failure}\}$ of fault and normal states are obtained.

*3) Performance Degradation Assessment*. For test samples to be evaluated, the DOM of clustering centers of normal state signals are calculated by equation (11) as evaluation index to evaluate bearing status. The closer the test sample is to the healthy state, the the DOM is closer to 1, otherwise, the the DOM is closer to 0.

The flow is shown in Fig. 1.

## VI. DATA ANALYSIS OF DIFFERENT FAULT LEVELS

The rolling bearing data come from the Case Western Reserve University. As shown in Fig. 2, the experimental equipment consists of a three-phase induction motor with a power of 1.49kw on the left, a torsion sensor in the middle, and a dynamometer on the right. This paper chose the bearing mounted on the drive end and the fault is implanted into the experimental bearing. In this experiment, a total of four bearing fault types were simulated, each of which was divided into different fault levels.

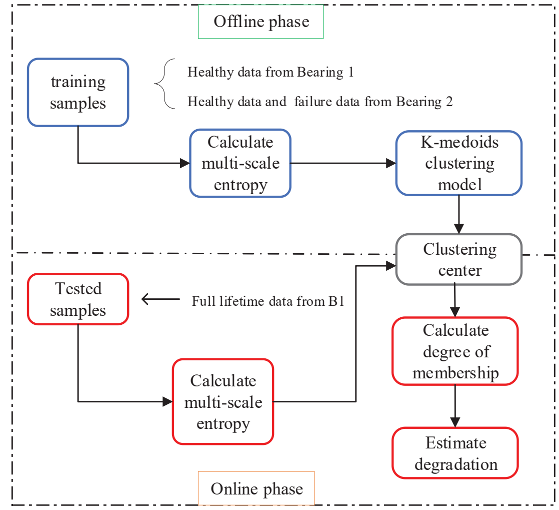



Figure 1. Flow chart of performance degradation assessment of rolling bearings

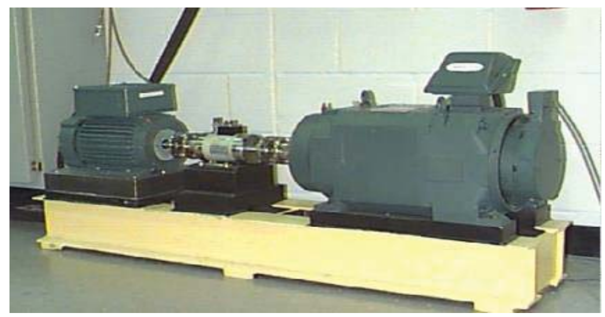


Figure 2. Rolling bearing fault experimental equipment

In this paper, we get an outer ring fault with three different fault levels: normal, 14 mm, 21 mm. During the experiment, the data sampling frequency was 12000 Hz and the motor speed was 1797 r/min. And three sets of signals were collected for each of three different outer ring fault levels. Fig. 3 shows the RMS values of the nine signals and the RMS mean of each fault degree. It can be found that the 14 mm RMS is smaller than the normal RMS, which is inconsistent with the actual situation, indicating that the RMS does not distinguish the different fault levels very well. Using the method of this paper to identify different levels of faults, the data set is divided into two parts: training set and testing set. The training set is used to train the K-medoids model, and the 9 sets of signals are input into the model to calculate the DOM of the healthy center, a DOM as shown in Fig. 4 is obtained. In this figure, it can be seen that as the degree of failure deepens, the DOM becomes smaller and smaller. It is verified that the combination of multi-scale entropy and K-medoids clustering model has a good effect on the identification of different states of rolling bearings.

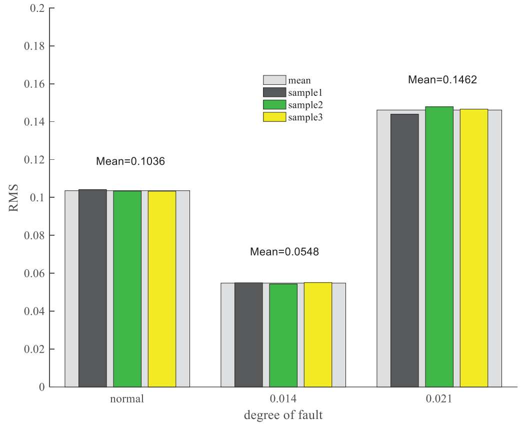



Figure 3. RMS values of different outer race fault levels

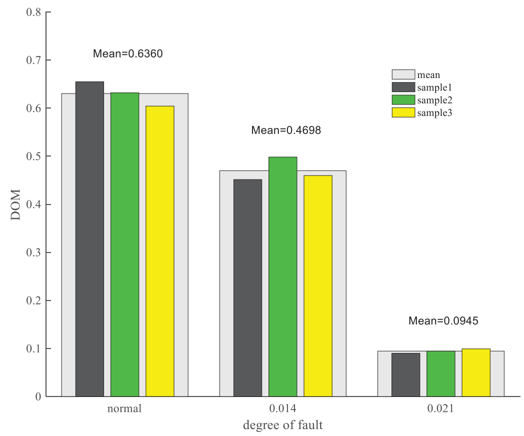



Figure 4. DOM of different fault degrees based on K-medoids

## VII. Fatigue Test Data Analysis

### *A. Experimental equipment*

In order to verify the early diagnosis and tracking ability of the method in the actual environment, this paper analyzes the actual fatigue test data of the rolling shaft. The test data were taken from the Intelligent Maintenance System Center of the University of Cincinnati, USA. Fig. 5 is the sensor layout diagram of the rolling bearing fatigue experimental equipment [12].

The machine rotates at 2000 r/min and drives the spindle to rotate through the belt drive. Four double row roller bearings are mounted on the spindle, model Rexnord ZA-2115. The bearings on both ends of the experimental equipment are fixed to the body, and the middle bearing applies a radial load to the main shaft. In the test, the accelerometer is mounted on the bearing housing, and the sampling frequency is 20000 Hz. The sampling interval is 10 min. A total of 984 data files are Collected. Each file contains 4 columns, each column has 20480 points of data. The first column is the bearing 1 data, and the second column is the bearing 2 data. In this paper, the sample data of bearing 1 and bearing 2 are taken as testing sample and training sample respectively, and the normal data and failure data of bearing 2

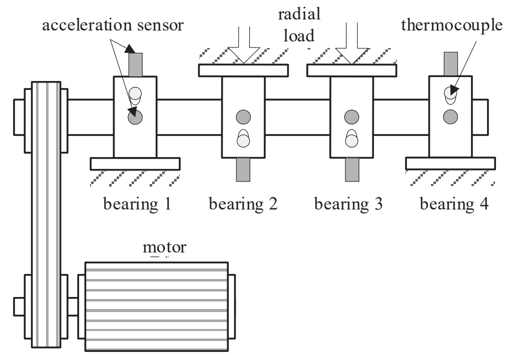



Figure 5. Rolling bearing fatigue experimental equipment

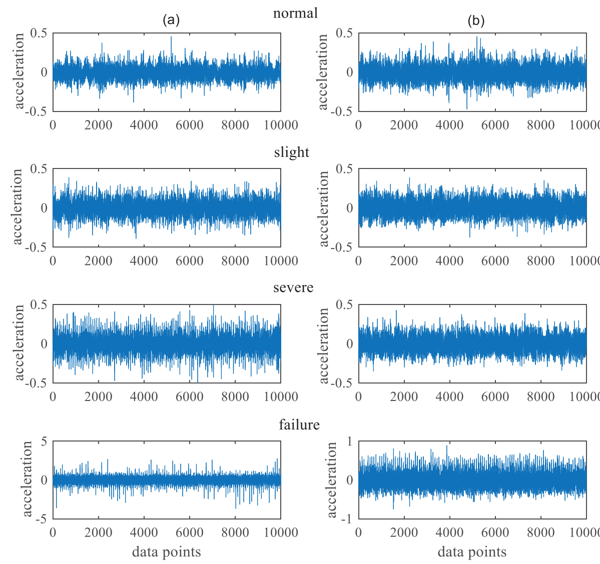



Figure 6. Signal time domain diagram of different states:(a)bearing1,(2)bearing2

and the normal data training model of bearing 1 are used to build model.Analyzing the degradation state of bearing 1. The time domain diagrams for the different fault states of bearing 1 and bearing 2 are shown in Fig. 6: normal, slight fault, severe fault and failure.

### *B. K-medoids analysis*

Performing K-medoids cluster analysis. The multi-scale entropy feature extraction is performed on the vibration data of the bearing of bearing 1 and bearing 2 to obtain 20 vector sequences. The test sample is the bearing 1 data, the training sample is the bearing 1 normal data and the bearing 2 normal data and the failure data, and the K-medoids clustering model is established to obtain two cluster centers. As shown in Fig. 7, the clustering effect is made by selecting three frequency bands with feature numbers 1, 2, and 3. It can be seen that the two states are well distinguished, and each state has a cluster center. The distance d1 between the test sample and the cluster center

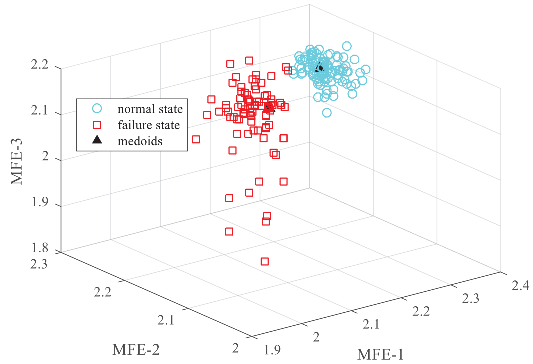


Figure 7. MSE clustering result

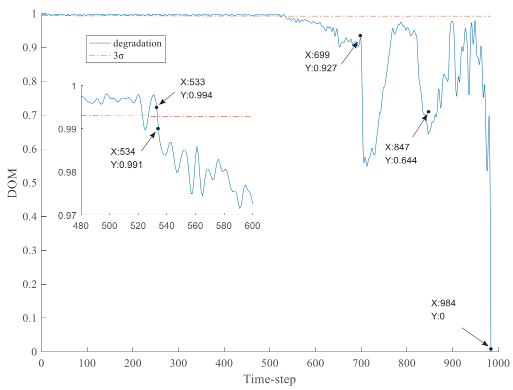


Figure 8. Performance degradation curve of K-medoids

of the normal sample is calculated by the Eq (10), and the distance d2 between the test sample and the fault center can be calculated by the same reason. Inputting d1 and d2 into the membership function (11), and calculating the DOM of the cluster center with the normal sample, and obtain the DI value of the bearing in the whole life. The performance degradation curve is shown in Fig. 8.

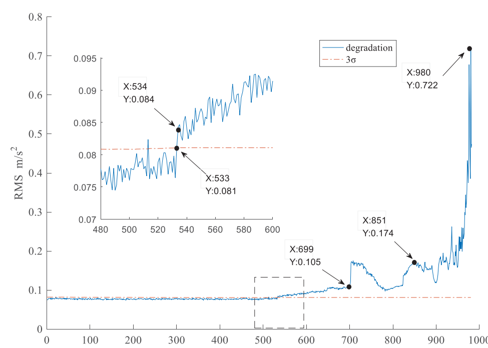
In this paper, the 3σ criterion is used as the adaptive alarm threshold. For a normal distribution with a mean of μ and a standard deviation of σ, or a random variable with a nearly normal distribution, the probability that the sample value falls at (μ-3σ, μ+3σ) is 99.73%. When a value exceeds the range, it can be considered as not belonging to current status. It can also be assumed that the performance degradation index DI value also conforms to the normal distribution. When multiple consecutive DI values exceed the range determined by the previous DI value, the degree of failure of the device can be considered to be greatly changed.

As can be seen in the figure, when the threshold is exceeded at the 534th sample, the fault begins to appear. At the 699th point, the sharp drop occurred, and the degree of bearing failure was deepened. The bearing failure between 699 and 847 points appeared as a period of deepening after the failure occurred and then smoothed, and then deteriorated sharply at the 984th sample. At the highest point, the bearing is completely ineffective. The four phases of the fault state can be divided into: normal state (1~533), slight state (534~699), severe state (700~847), failure state (848~984).

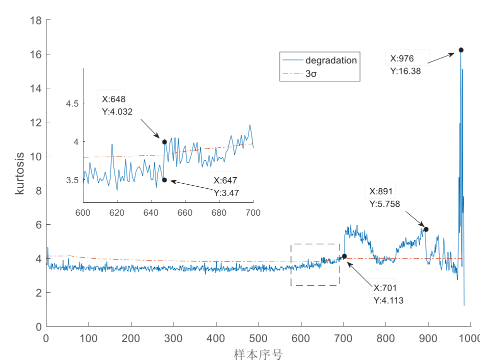
### *C. RMS and kurtosis analysis*

Dimensional time domain parameters and dimensionless time domain parameters are common fault detection parameters, and the RMS value and Kurtosis can be used as evaluation indexes to evaluate the reliability of bearings. The RMS value of the bearing signal will increase with the increase of the fault degree, and the kurtosis value is sensitive to the impact of the signal, which can be used as the basis for early judgment of the fault. The performance degradation curve of the rms value and the kurtosis value of the whole life data of the test bench bearing is shown in Fig. 9.

The figure shows four states normal state, early fault failure, medium failure state, and severe failure state. The curve rises slowly in the early failure state. In the moderate failure state, the curve rises and falls back significantly. In the severe failure state, the curve rises sharply and the bearing fails completely. Using the K-medoids method in this paper, the performance degradation curve changes significantly at the threshold and is sensitive to early faults. The RMS does not change significantly at the threshold, the Kurtosis performance degradation curve only exceeds the threshold at 648 points, which is not conducive to timely maintenance strategy. In addition, neither RMS nor Kurtosis has a clear upper limit. Compared with the degree of membership [0,1] interval, there is no quantitative determination of fault degree.



RMS



Kurtosis

Figure 9. PDA of bearing 1 based on RMS and Kurtosis

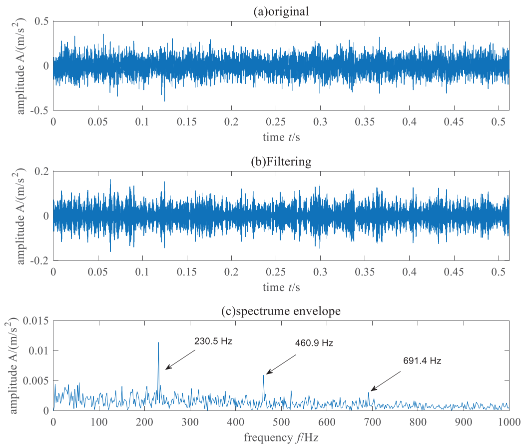



Figure 10. Envelope demodulation diagram for the 534th sample

### D. Envelope spectrum analysis

The membership index in this paper considers that an early fault occurs at 534 moments. According to the structural parameters and rotational speed of the bearing, the fault frequency of the bearing is calculated to be 235.4 Hz. The optimal Morlet wavelet resonance demodulation analysis was performed on 534 samples to verify its correctness. Calculating the peak factor of the envelope spectrum of the filtered signal, the ratio of the maximum envelope spectrum to the root mean square, the optimal center frequency is 4730.8 Hz and the effective bandwidth is 900 Hz. The 534th sample was subjected to resonance demodulation using the optimal center frequency and effective bandwidth, and the result is shown in Fig. 10. The 230.5 Hz and its octave components, which are close to the frequency of the fault signature, appear in the envelope spectrum, indicating that an early fault occurred at this time. The samples before the analysis of 534 times are all frequency components close to the fault frequency, so it is inferred that the 534 samples are the moments when the fault occurs, which is consistent with the evaluation results in this paper.

## VIII. CONCLUSIONS

This paper proposes a bearing performance degradation assessment method based on multi-scale entropy and K-medoids clustering. The multi-scale entropy is used to extract features of the signal on multiple scales, and the membership function can display the similarity on a specific interval [0,1]. The K-medoids clustering model is used to find the center of the normal and failure signal respectively. The membership function is used to calculate the degree of membership as an evaluation index, which can achieve the quantitative evaluation of the fault degree. The method in this paper is verified by the bearing failure test, shows that the index is more effective and sensitive to the degree of fault than other classical methods such as RMS and Kurtosis. This method can detect early faults in time and quantitatively assess the degree of faults.

## REFERENCES


[1] C. Mariela, S. René-Vinicio, L. Chuan, P. Fannia, C. Diego, V. José. et al, “A review on data-driven fault severity assessment in rolling bearings,” J. Mechanical Systems and Signal Processing. vol. 99, pp. 169-196, 2018.

[2] H. Sheng, B. Q .Wang, G. Q. Li, and H, Qian, “Performance degradation assessment for bearing based on ensemble empirical mode decomposition and Gaussian mixture model,” J. Journal of Vibration & Acoustics. vol. 136, 61006, 2014.

[3] L. Guo, X. L. Li, C. Wu, and C. Z. Liu, “Assessment method for performance degradation of rolling bearings based on SVM,” J. Bearing. 2012(8), pp. 46-50.

[4] R. Akhand, and U. S.H, “Bearing performance degradation assessment based on a combination of empirical mode decomposition and k-medoids clustering,” J. Mechanical Systems and Signal Processing. vol. 93, pp.16-29, 2017.

[5] P. Chu, and L. Ke, “Improved k-medoids clustering algorithm based on improved granular computing,” J. Joumal of Computer Applications, vol. 34(7), pp. 1997-2000, 2014.

[6] X. M. Gao, Y. Feng, and X. J. Feng, “Incremental clustering algorithm based on k-medoids,” J. Computer Engineering, 2005(S1), pp. 181-183.

[7] S. R.J, and J. R.M, “Physiological time-series analysis, using approximate entropy and sample entropy,” J. American Journal of Physiology-Heart and Circulatory Physiology, vol. 278(6), H2039-H2049, 2000.

[8] A. G B, K. P C, and M C, “Multiscale entropy analysis of biological signals,” J. Physical Review E Statistical Nonlinear & Soft Matter Physics. vol. 71(2), 2005

[9] K. Leonard, and R. Peter J, “Partitioning around medoids (program PAM)” M. 2008.

[10] B. James. C. A, “Physical Interpretation of fuzzy isodata,” J. IEEE Transactions on Systems Man & Cybernetics. 2007, vol. 6(5), pp. 387-389.

[11] N. R. Pal, J. C. Bezdek, “On cluster validity for the fuzzy c-means model,” J. IEEE Transactions on Fuzzy Systems, vol. 3(3), pp. 370-379, 1995.

[12] H. Q, Jay. Lee, J. Lin, and G. Yu, “Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics,” J. Journal of Sound and Vibration. vol.289(4-5), pp. 1066-1090, 2006.