Xu *et al.*
Chinese Journal of Mechanical Engineering (2025) 38:106
https://doi.org/10.1186/s10033-025-01263-1

Chinese Journal of Mechanical Engineering

ORIGINAL ARTICLE **Open Access**

Rolling Bearing Early Fault Detection Method Based on Feature Clustering Fusion Degradation Index

Xiangyang Xu[1*], Haotian Wang[1], Xihui Liang[2], Chuan Zhao[3] and Ziyuan Ren[1]

Abstract

The research on rolling bearing early fault detection is mainly focused on degradation index extraction and adaptive setting of alarm threshold. The mainstream methods are to extract degradation indicators based on adaptive features and set adaptive alarm thresholds based on the Shewhart control chart. However, the adaptive feature extraction method does not consider the correlation between features, and the Shewhart control chart is not sensitive to small fluctuations caused by early faults. In this study, a rolling bearing early fault detection method based on a feature clustering fusion degradation index is proposed. The multidomain statistical features are extracted to form the initial feature set, and the improved hierarchical clustering algorithm is combined with the feature evaluation index to select features to form a preferred feature subset, to ensure the richness of index information and reduce redundancy. After the construction of the degradation index, to suppress the interference caused by nonstationary and abnormal shocks in early fault detection, the accurate evaluation method and anomaly determination strategy of control chart parameters are studied, and an improved exponential weighted move average control chart is designed to monitor the degradation index. The effectiveness and superiority of the proposed method are verified by public data sets. This research provides a rolling bearing early fault detection method, which can provide comprehensive degradation indicators, eliminate interference caused by random anomalies and running in periods, and achieve an accurate detection of early bearing failures.

Keywords Rolling bearing, Feature selection, Degradation index, Modified exponential weighted move average control chart, Early failure detection

1 Introduction

As a key component of large rotating machinery equipment, it is necessary to monitor the performance degradation of rolling bearings effectively and accurately. However, in specific equipment such as aero-engines and large submersible motors, limited by the structural layout and working environment, only one vibration sensor is usually installed near the bearing [1]. Owing to the long transmission path, serious attenuation, and complex frequency components of the collected unidirectional vibration signals, it is very difficult to mine the bearing performance degradation information and identify the weak characteristic signals of early faults, which poses higher requirements for the analysis and fault detection methods of vibration signals. Therefore, it is of significance to carry out accurate evaluation and fault detection of online performance degradation of rolling bearings based on unidirectional vibration signals [2].

*Correspondence:
Xiangyang Xu
xxiangyang@hotmail.com
[1] School of Mechatronics and Vehicle Engineering, Chongqing Jiaotong University, Chongqing 400074, China
[2] School of Mechanical Engineering, University of Manitoba, Winnipeg, MB R3T 2N2, Canada
[3] School of Mechanical and Electrical Engineering, North China Institute of Aerospace, Hebei 065000, China

In the bearing early fault detection, it is crucial to construct degradation indicators that accurately characterize the bearing condition and quantify its degradation trend. Physical degradation indicators typically consist of time- and frequency-domain statistical features, which are of significance. However, these features are limited in their ability to produce optimal evaluation results, as they are strongly dependent on specific failure and operating conditions. Virtual degradation metrics are obtained by fusing multiple features. Scholars mostly use machine learning and deep learning methods to construct degradation models by adaptively extracted features [3–7]. However, adaptively extracted features are poorly interpretable and lack explicit expressions, and thus it is challenging to understand the features in depth. Compared to adaptively extracted features, statistical features can be used to analyze the signal in multiple domains, have physical meaning, and are easy to understand and use. Meddour et al. [8] used correlation, monotonicity, average increase rate, and robustness to select features, and adaptive-network-based fuzzy inference system (ANFIS) to construct degradation indicators to achieve bearing remaining useful life prediction. Zhou et al. [9] extracted time-domain features and energy entropy, used three features of correlation, monotonicity, and robustness to evaluate and select features, and combined them with an optimized support vector machine to construct degradation indicators for early fault detection and recognition. Niu et al. [10] selected statistical features using rank mutual information to form a subset of features and constructed degradation indicators by similarity-based modeling. The above studies provide useful references for construction of degradation indicators. However, there has not been extensive discussion on selection of features to form feature subsets. The feature evaluation index can be used to select features with a high correlation with the bearing life and good trend and robustness to form a feature subset. However, the correlation between the features has not been considered, and there are some redundant features with a high correlation in the subset of features, which affects the construction of subsequent degradation indices. Therefore, in this study, feature selection is performed based on two aspects: interfeature correlation and feature performance. By analyzing the correlation between features, the initial feature set is clustered into multiple feature clusters. By combining feature evaluation indices, the best feature is selected from each feature cluster to form a preferred feature subset, achieving the goal of performance maintenance of the feature subset while reducing the internal correlation.

After the construction of degradation indicators, Shewhart control charts based on the Lajda criterion (6 σ principle) are often used to set alarm thresholds adaptively to detect early failures in the bearing degradation stage. Chen et al. [11] described the health state of the bearing by the change in the abnormality score of the adjacent graph model and used Shewhart control charts to make a decision on whether to generate early failures. Lu et al. [12] reconstructed the one-dimensional signal segmentation into two-dimensional grayscale images, used the similarity of adjacent cycle images as a degradation indicator, and considered that an early failure occurred when the control limit of the Shewhart control chart was exceeded. Dong et al. [13] used the ratio of fault feature energy to noise energy as a performance degradation indicator of rolling bearings and set the health threshold with the help of the Shewhart control chart. Although the Shewhart control chart has been increasingly applied in early bearing failure detection, there are still problems, as follows. (1) The data of the Shewhart control chart should be normally distributed, while the vibration data of the rolling bearing operation have noise and other interference, and their specific distribution is often difficult to determine. (2) The rolling bearing performance degradation is a continuous change process, while the Shewhart control chart uses only the data collected at the current moment for detection, which results in insensitivity to the small fluctuations caused by early failures that continue to rise or fall. Compared to the Shewhart control chart, the exponential weighted move average (EWMA) control chart [14] has the following advantages. The EWMA control chart not only can be applied to the case of unknown data distribution, but also can use historical data information and is more sensitive to the detection of small drifts in the data. Therefore, EWMA control charts are considered for application in early fault detection of rolling bearings. However, in the actual application, the inaccurate estimation of control chart parameters that affect the threshold setting and abnormal shocks in the normal operation stage of rolling bearings that cause false alarms hinder the detection of early faults by EWMA control charts.

In this study, we propose a rolling bearing degradation index construction and early fault detection method. The absolute value of the Pearson correlation coefficient is used to characterize the correlation between features, which is then employed to divide feature clusters using an improved hierarchical clustering algorithm. The algorithm is integrated with feature evaluation indices to select a preferred feature subset, and degradation indicators are derived through autoencoder fusion. Based on the degradation indicators, a few degradation indicator values are selected for parameter estimation. An improved EWMA control chart is designed to operate in conjunction with an anomaly determination strategy for early detection of faults.

The paper is organized as follows. Section 2 describes the degradation indicator construction method by feature clustering and fusion. Section 3 presents the modified EWMA control chart anomaly detection method. Section 4 describes the whole process of the method. Section 5 provides an experimental validation of the proposed method. Section 6 presents the conclusion.

2 Degradation Index Based on Feature Clustering Fusion

2.1 Degradation Feature Extraction

The traditional statistical feature is a powerful tool to characterize the change in rolling bearing vibration signal when a fault occurs. In this study, 66 features are selected to form the original feature set, including 46 time-domain features, 16 frequency-domain features, and 4 entropy features.

The defining equation of time-domain features is shown in Appendix Table 5, where $x(n)$ is the original vibration signal, $n = 1, 2, 3, \ldots, N$, N is the number of samples in the sample, α is the truncation factor, $r(n)$, $d(n)$, and $b(n)$ denote the residual signal, differential signal, and first-order sideband signal, respectively, $r_m(n)$, $d_m(n)$, and $b_m(n)$ denote the mth time series of $r(n)$, $d(n)$, and $b(n)$, respectively, $e(n)$ denotes the current time envelope, $e(n) = |b(n) + j \times H(b(n))|$, where $H(b(n))$ is the Hilbert transform of $b(n)$, $e_m(n)$denotes the mth time series envelope of $b(n)$, M denotes the total number of time records to date, M' denotes the total number of time records of normal bearing, s_i is the amplitude of the ith harmonic, H is the total number of harmonics in the frequency range, and $\Delta x(n) = x(n)^2 - x(n-1)x(n+1)$.

A total of 16 features are chosen in the frequency domain [15], as shown in Appendix Table 6, where $y(k)$ is the spectral amplitude of $x(n)$, $k = 1, 2, 3, \ldots, K$, K is the number of spectral lines, and f_k is the frequency value of the kth spectral line.

In addition, entropy features have been applied in fault feature extraction of rotating machinery. Entropy as a statistical metric can be used to quantify complexity and detect dynamic changes. In this study, we add commonly used primitive entropy methods, approximate entropy, sample entropy, permutation entropy, and fuzzy entropy, to the initial feature set for testing, as defined and applied in Ref. [16].

Table 1 Variable range setting

Variable	Range
l	10–50
d	5–30
a	3–5
b	3–20

Table 2 Intraclass feature and interclass feature correlations

Feature cluster	Correlation	Feature cluster	Correlation
I	0.812	I-V	0.431
II	0.815	II-III	0.595
III	0.794	II-IV	0.428
IV	0.814	II-V	0.384
V	0.689	III–IV	0.526
I-II	0.556	III-V	0.476
I-III	0.462	IV–V	0.601
I-IV	0.418		

2.2 Modified Hierarchical Clustering Algorithm

The commonly used clustering algorithms are based on hierarchical clustering, K-means [17]. The K-means algorithm requires the number of clustering centers to be set in advance and is prone to sticking in a local optimum. In contrast, the hierarchical clustering enables to either set the number of feature clusters in advance or determine the number of clusters after an analysis of the clustering results. This approach is more aligned with the goal of reducing redundancy in feature subsets and leveraging the distinct information provided by various features.

The traditional hierarchical clustering algorithm uses Euclidean distance in the calculation of the distance between features, and classifies the features with close distances into the same category. However, this is not suitable for the classification of degraded features. Pearson correlation coefficient is a widely accepted metric across various disciplines for quantification of the degree of linear correlation between two variables. The redundancy or similarity of the feature subset selected from the initial feature set should be sufficiently small, so that the absolute value of the Pearson correlation coefficient can be used to directly measure the similarity between the two features (F_i, F_j).

$$|PCC(F_i, F_j)| = \left| \frac{Cov(F_i, F_j)}{\sigma_{F_i}\sigma_{F_j}} \right|, \tag{1}$$

where $Cov(F_i, F_j)$ is the covariance and $\sigma_{Fi}\sigma_{Fj}$ is the standard deviation.

We propose a modified hierarchical clustering algorithm, which uses the absolute value of the Pearson correlation coefficient in the calculation of the adjacent matrix to measure the similarity between features. The algorithm implementation steps are as follows.

(1) Each feature is regarded as a cluster, and the adjacency matrix is calculated.
(2) The two nearest clusters are combined into one cluster.
(3) The mean connection method is used to calculate the proximity between the new cluster and other clusters, and the proximity matrix is updated.
(4) Steps (2) and (3) are repeated until the defined number of termination clusters is reached.

In the mean connection method, the mean of the absolute values of the Pearson correlation coefficients of any feature in two clusters is used as proximity between the two clusters. A flow chart is shown in Figure 1.

2.3 Feature Evaluation Index

To quantify the feature performance, three feature evaluation indices of correlation, monotonicity, and robustness proposed in Ref. [18] are introduced to evaluate the feature performance.

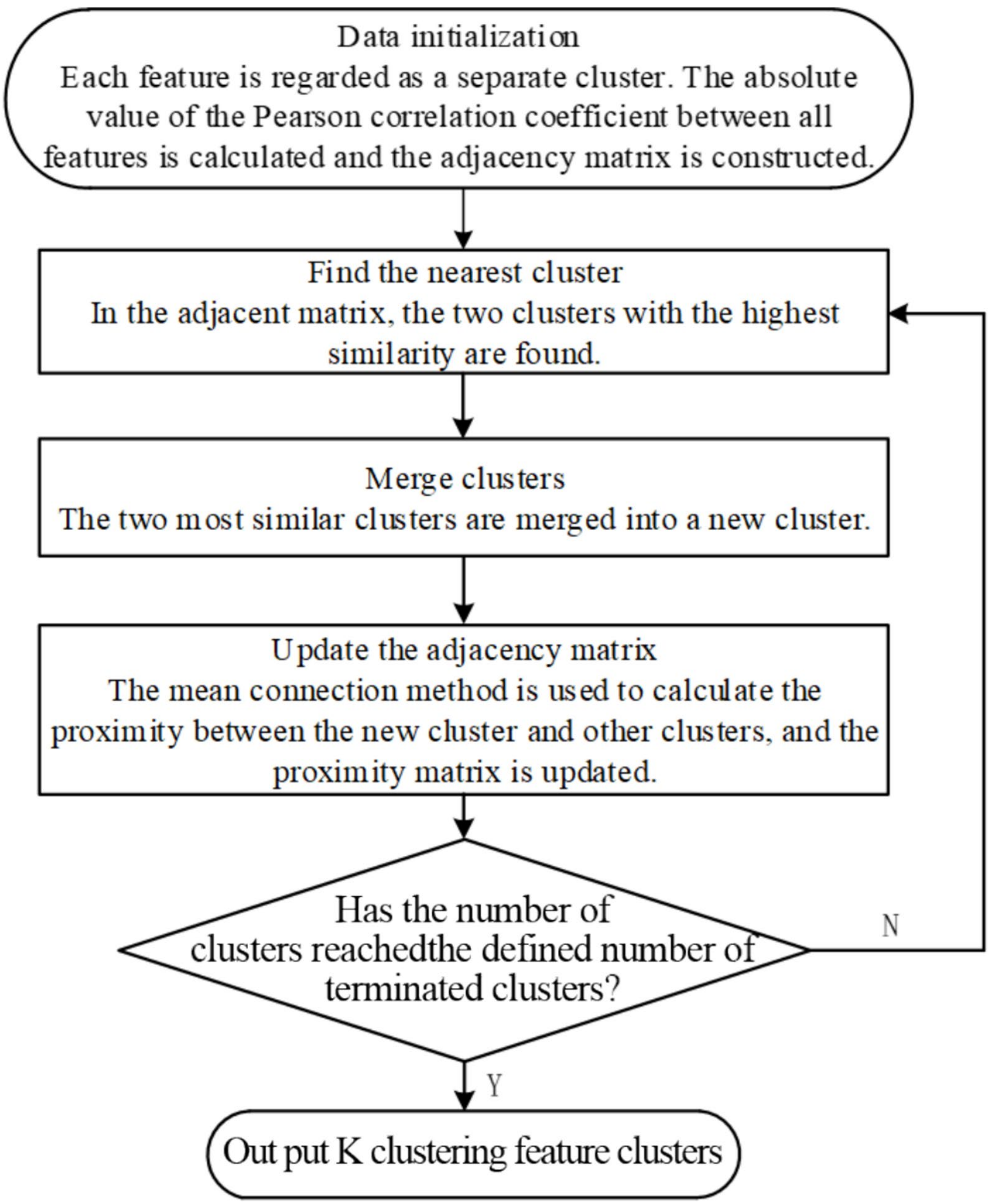

Figure 1 Modified hierarchical clustering algorithm flowchart

The moving average method is used to decompose the features. This approach makes the trend in the trend component clearer while allowing the random component to reflect the degree of change in the feature trend. The calculation formula is

$$F(m) = F_T(m) + F_R(m), \tag{2}$$

where $F(m)$ is the feature of the mth sample, $F_T(m)$ is the trend item, and $F_R(m)$ is the random term.

The correlation evaluation index is used to describe the correlation between the features and bearing life. The calculation formula is

$$Corr = \frac{\left|M\sum_{m=1}^{M}F_T(m)m - \sum_{i=1}^{M}F_T(m)\sum_{m=1}^{M}m\right|}{\sqrt{\left[M\sum_{m=1}^{M}F_T(m)^2 - \left(\sum_{m=1}^{M}F_T(m)\right)^2\right]}} \times \frac{1}{\sqrt{\left[M\sum_{m=1}^{M}m^2 - \left(\sum_{m=1}^{M}m\right)^2\right]}}. \tag{3}$$

The monotonicity evaluation index is used to evaluate the monotonicity of the features with the deepening of bearing degradation. The calculation equation is

$$Mon = \frac{1}{M-1}\left|\sum_{m=1}^{M-1}\varepsilon(F_T(m+1) - F_T(m)) - \sum_{m=1}^{M-1}\varepsilon(F_T(m) - F_T(m+1))\right|. \tag{4}$$

The robustness evaluation index [19] is used to describe the interference resistance performance of the feature and is calculated by

$$Rob = \frac{1}{M}\sum_{m=1}^{M}\exp\left(-\left|\frac{F_R(m)}{F(m)}\right|\right). \tag{5}$$

In Eqs. (3), (4), and (5), M is the total number of samples and $\varepsilon(\,)$ is the step function.

A single evaluation index cannot comprehensively evaluate the features, and thus a constrained weighted combination of three indicators is constructed to select the most suitable features.

$$O = \omega_1 Corr + \omega_2 Mon + \omega_3 Rob, \quad \begin{cases} \omega_i > 0, \\ \omega_1 + \omega_2 + \omega_3 = 1, \end{cases} i = 1, 2, 3, \tag{6}$$

where O is the comprehensive feature evaluation index and ω_i is the weight of each index.

The comprehensive feature evaluation index is positively correlated with the feature performance, and $O\in[0, 1]$. When O is closer to 1, the feature performance is better, and vice versa. As the preferred features are mainly oriented to the early fault identification of rolling bearing state, the correlation and monotonicity should be given larger weights, $\omega_1 = 0.4$, $\omega_2 = 0.4$, $\omega_3 = 0.2$.

2.4 Autoencoder

The autoencoder is a self-supervised learning algorithm, as shown in Figure 2, which includes an encoder and decoder [20]. The encoder compresses the high-dimensional input into an efficient abstract representation, while the decoder reconstructs this abstract representation into the desired output. The input and output of the autoencoder are unlabeled data, aimed to learn the internal structure of a high-dimensional input and extract nonlinear features [21].

The objective function of autoencoder learning is

$$S_{W,W\prime,b_1,b_2}(X) \approx X. \tag{7}$$

Assuming that the input is X, the weight matrix of the input layer and hidden layer is W, the coding layer node offset is b_1, and the decoding layer node offset is b_2. $S(\cdot)$ is divided into two stages. Eq. (8) shows the encoding stage from the input layer to the hidden layer, while Eq. (9) shows the decoding stage from the hidden layer to the input layer.

$$h = f(W \times X + b_1), \tag{8}$$

$$Y = g\left(W' \times h + b_2\right). \tag{9}$$

To reconstruct the input X, the square error loss function and cross entropy loss function are defined as

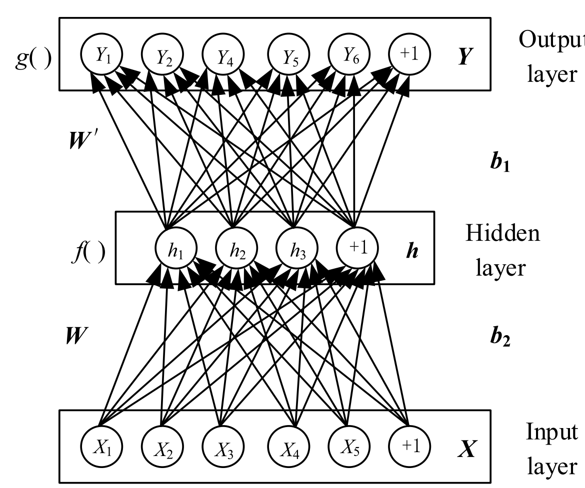

Figure 2 Autoencoder structure

$$L_S\left(W,W',b_1,b_2;X\right)=\frac{Y}{2}-X^2, \tag{10}$$

$$\begin{aligned}&L_C\left(W,W',b_1,b_2;X\right)\\&=-[X\log Y+(1-X)\log(1-Y)].\end{aligned} \tag{11}$$

In this study, the middle layer of the autoencoder has only one neuron h_1, and the output of the middle layer h_1 is the degradation index (DI).

3 Bearing Early Fault Detection Method Based on a Modified EWMA Control Chart

3.1 EWMA Control Chart

The EWMA control chart can set the alarm threshold according to the DI to achieve early fault detection. It reflects more information during the detection process near the current moment. Based on this concept, different weights are assigned to the data, and the EWMA statistics are obtained through an iterative summation [22], leading to the design of the EWMA control chart. As shown in Figure 3, the EWMA control chart contains the mean line, upper control limit (UCL), and lower control limit (LCL). The control limit is the set threshold.

The calculation formula of the EWMA statistics is

$$Z_i=\lambda x_i+(1-\lambda)Z_{i-1}, \tag{12}$$

where Z_i is the EWMA value at point i, x_i is the eigenvalue at point i, and λ is a value in (0, 1). A smaller λ leads to a higher sensitivity to a small drift. In this study, λ = 0.2. The initial value of the EWMA statistics Z_0 is $E(x)=\mu$.

According to the iterative method,

$$Z_i=\lambda\sum_{j=0}^{i-1}(1-\lambda)^i x_{i-j}+(1-\lambda)^i Z_0. \tag{13}$$

In this regard, for older historical data, the weight is smaller, which decreases exponentially. The calculation equation of standard deviation in the EWMA control chart is

$$\sigma_{Z_i}=\sigma\sqrt{\left(\frac{\lambda}{1-\lambda}\right)\left[1-(1-\lambda)^{2i}\right]}, \tag{14}$$

where σ is the standard deviation of the observed value.

According to Eq. (14), the control limit of the EWMA control chart is

$$\begin{cases}UCL|=\mu+3\sigma\sqrt{\left(\frac{\lambda}{1-\lambda}\right)\left[1-(1-\lambda)^{2i}\right]},\\CL|=\mu,\\LCL|=\mu-3\sigma\sqrt{\left(\frac{\lambda}{1-\lambda}\right)\left[1-(1-\lambda)^{2i}\right]}.\end{cases} \tag{15}$$

3.2 Modified Parameter Estimation Methods

The design of the EWMA control chart is divided into two stages. Stage I is used to collect the data in the controlled state, eliminate the factors affecting the stability of the process, and estimate the parameters of the control chart in the controlled state. Stage II is used for construction of the corresponding control chart, collection of online data, and detection.

In the actual operation of rolling bearings, the EWMA control chart parameters for monitoring of the DI are unknown. It is necessary to select the controlled data to estimate the mean and standard deviation to set the alarm threshold. At present, the whole health stage data are usually selected for parameter estimation, but this method cannot provide accurate parameters. Figure 4 shows that the amplitude of the vibration signal in the healthy stage of the rolling bearing is not stable, and that the amplitude of the signal in the healthy stage A is significantly larger. The use of data estimation parameters from the entire healthy stage can lead to an increase in standard deviation, and thus the upper and lower control limits become more lenient, which results in delayed test results. The data used for parameter estimation should be as stable as possible and close to the degradation stage, so

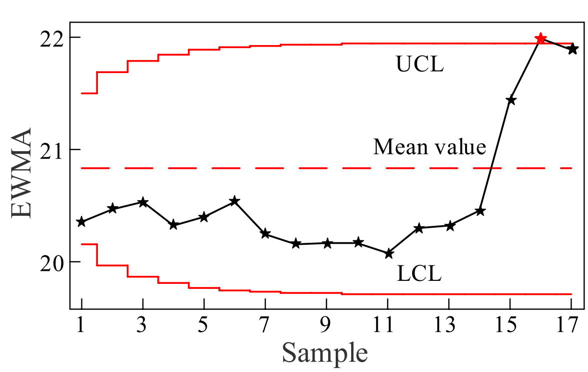

Figure 3 Example of EWMA control diagram

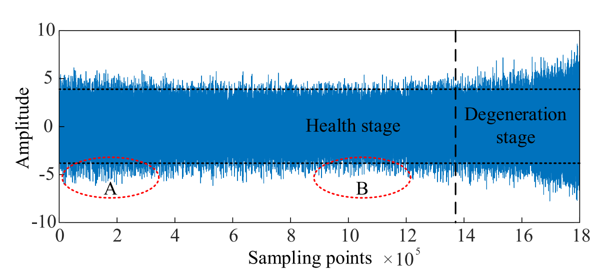

Figure 4 Time-domain vibration signal

that the EWMA control chart can be designed to monitor the subsequent operation process, timely detect early faults, and alarm.

To obtain accurate parameters, we propose a phase I parameter estimation method. The main implementation steps are as follows.

(1) DI decomposition. DI $F = \{f_1, f_2, \cdots, f_M\}$, M is the number of samples, and f is the DI value. The moving average method is used to decompose F, $F = Q + S$, where $Q = \{q_1, q_2, \cdots, q_M\}$ is the trend term and $S = \{s_1, s_2, \cdots, s_M\}$ is the random term.

(2) Normalization and grouping of trend items. The trend term Q is normalized, and then grouped by a sliding window. The number of data points contained in each group is l, while the sliding distance is d. The first group is $Q_1 = \{q_1, q_2, \cdots, q_l\}$, while the h group is $Q_h = \{q_{1+(h-1)d}, q_{2+(h-1)d}, \cdots, q_{l+(h-1)d}\}$.

(3) Dispersion calculation and analysis of each group of trend items. The variance σ^2_h of Q_h is calculated as the dispersion to describe the bearing state, where $h = 1, 2, \cdots, H$. H is the total number of groups. The trend term variance sequence $\{\sigma^2_1, \sigma^2_2, \ldots, \sigma^2_H\}$ can be obtained. In the normal operation stage, the abnormal impact is mostly random and discontinuous, and thus we consider the continuous multiple sets of dispersion changes to avoid false alarms. It is assumed that, from the hth group to the $(h + a)$ th group, the value of σ^2_h to σ^2_{h+a} increases monotonically, and $\sigma^2_{h+k}/\sigma^2_h \geq b$ indicates that the sudden increase in the dispersion of the DI has been out of control and produces a warning. In the determination of k, it is required that the trend data points corresponding to σ^2_h and σ^2_{h+k} do not have intersection, to avoid the interference of overlapping data. Therefore, it is necessary to satisfy $l + (h - 1)d \leq 1+(h + k - 1)d$, and thus $k = \lceil (l - 1)/d \rceil$, where $\lceil * \rceil$ denotes the upward rounding operation.

(4) Control chart parameter estimation. In warning, the corresponding trend term is grouped as Q_h, and the corresponding trend term sequence in the normal operation stage is $Q_i^{health} = \{q_i^1, q_i^2, \cdots, q_i^{(h-1)d}\}$. Considering the interference of unstable data on the calculation of control limits, some stable data are selected for the calculation. It is considered that Q_i^{health} corresponds to the normal working stage of the life cycle, and thus Q_i^{health} is divided into two sections. The second section is taken for the analysis to meet the demand. However, to avoid accidental phenomena including extreme special cases, we divide Q_i^{health} into three sections and calculate the variance of the last two sections. Finally, a set of corresponding trend data points with a small variance is selected as controlled data to estimate the parameters μ and σ.

The variables involved in the above method are l, d, a, and b. In practical applications, the detection time should be reduced as soon as possible to find early faults, and thus l, d, and a should not be too long. Simultaneously, there may be impacts in the features due to sensor loosening and other reasons, making l, d, and a too short and yielding a false alarm. Additionally, the parameter b affects the sensitivity of detection. If it is too small, it may cause false alarms, while, if it is too large, it may fail to accurately detect early faults. Therefore, the range of variables in this section is shown in Table 1.

To avoid the interference of subjective factors and improve the accuracy of the method, a genetic algorithm is used to optimize the relevant variables [23]. The fitness function in the genetic algorithm is defined by minimizing the absolute error between the manually labeled early fault sample number and early fault sample number detected by the proposed method. The absolute error is

$$ae = |T_{ar} - T_{ma}|, \tag{16}$$

where T_{ar} is the sample serial number of the artificial label and T_{ma} is the sample serial number detected by the proposed method.

3.3 Modified Exception Determination Strategy

The normal operation of rolling bearings is often accompanied by noise and other interferences, leading to random abnormal shocks [24] in the collected data, as illustrated in Figure 5(a). This figure shows the time-domain vibration signal, where these random shocks are evident even in the early stages of operation. The DI in Figure 5(b) shows that these random shocks yield abnormal points during the health stage. However, these isolated abnormal points are not sufficient indicators of early failure. As shown in Figure 5(c), the improved EWMA control chart reflects these dynamics more effectively. The DI remains stable, and, only when it consistently exceeds the alarm threshold, it signifies the onset of failure, marking the transition into the degradation stage. If the number of consecutive data points is too large, the method becomes less sensitive to changes in the performance degradation stage, which reduces its ability to detect these transitions. Conversely, if the number of consecutive data points is too small, the likelihood of false positives increases, as minor anomalies may be incorrectly identified as a shift in the performance degradation stage [25]. Therefore, we enhance the anomaly detection strategy by selecting five consecutive data points for monitoring. When

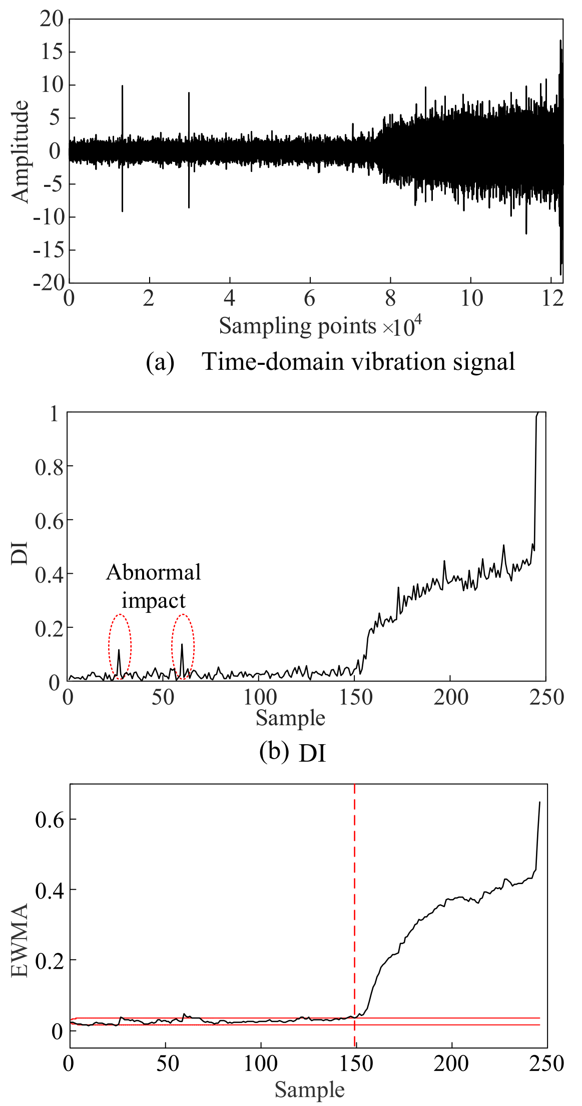

Figure 5 Improved EWMA control chart detection results when there are abnormal values in the normal stage

the five consecutive data points are outside the control limit, the anomaly is determined and alarmed.

4 Proposed Bearing DI Construction and Early Fault Detection Process

We propose an early fault detection method for rolling bearings based on a degradation index derived from feature clustering fusion and improved control chart. The specific steps are as follows.

Training phase:

(1) 66 features are extracted from the original vibration signal as an initial feature set.
(2) With the modified hierarchical clustering algorithm, the initial feature set is divided into K feature clusters according to the correlation between features.
(3) The comprehensive feature evaluation index is used to evaluate each feature, calculate the average feature ranking, and select the features with the highest average ranking in each cluster to form the preferred feature subset. The preferred feature sequence number is retained as the criterion for extraction of the preferred feature subset in the test set.
(4) The K-dimensional preferred feature is input into the autoencoder to reduce to one dimension, and the DI is obtained.
(5) The modified EWMA control chart is constructed to detect the early fault of the DI. The optimized optimal variables in the modified parameter estimation method are retained as the criteria for design of EWMA control charts in the test set.

Test phase:

(1) The preferred feature subset is extracted from the original vibration signal and input in the autoencoder to obtain the DI.
(2) The modified EWMA control chart is constructed according to the optimal variables to monitor the DI.
(3) Early fault detection is carried out. The early fault of the proposed method is determined as follows. When five consecutive data points are outside the alarm threshold after the DI trend term $q_i^{(h-1)d}$, the early fault is considered to have occurred and is alarmed.

The early fault detection method of rolling bearing is shown in Figure 6.

5 Experimental Verification

5.1 Data Sets and Experimental Settings

To prove the effectiveness of the proposed method, the XJTU-SY data set [26] and IEEE PHM 2012 data set [27] are used for verification. The XJTU-SY data set is provided by Xi'an Jiaotong University and Changxing Shengyang Technology Company. The LDK-UER204 rolling bearing is used, including three working conditions and five rolling bearings under each working condition. In the experiment, the sampling frequency is set to 25.6 kHz, the sampling interval is 1 min, and the sampling time is 1.28 s. The verification is divided into three tests

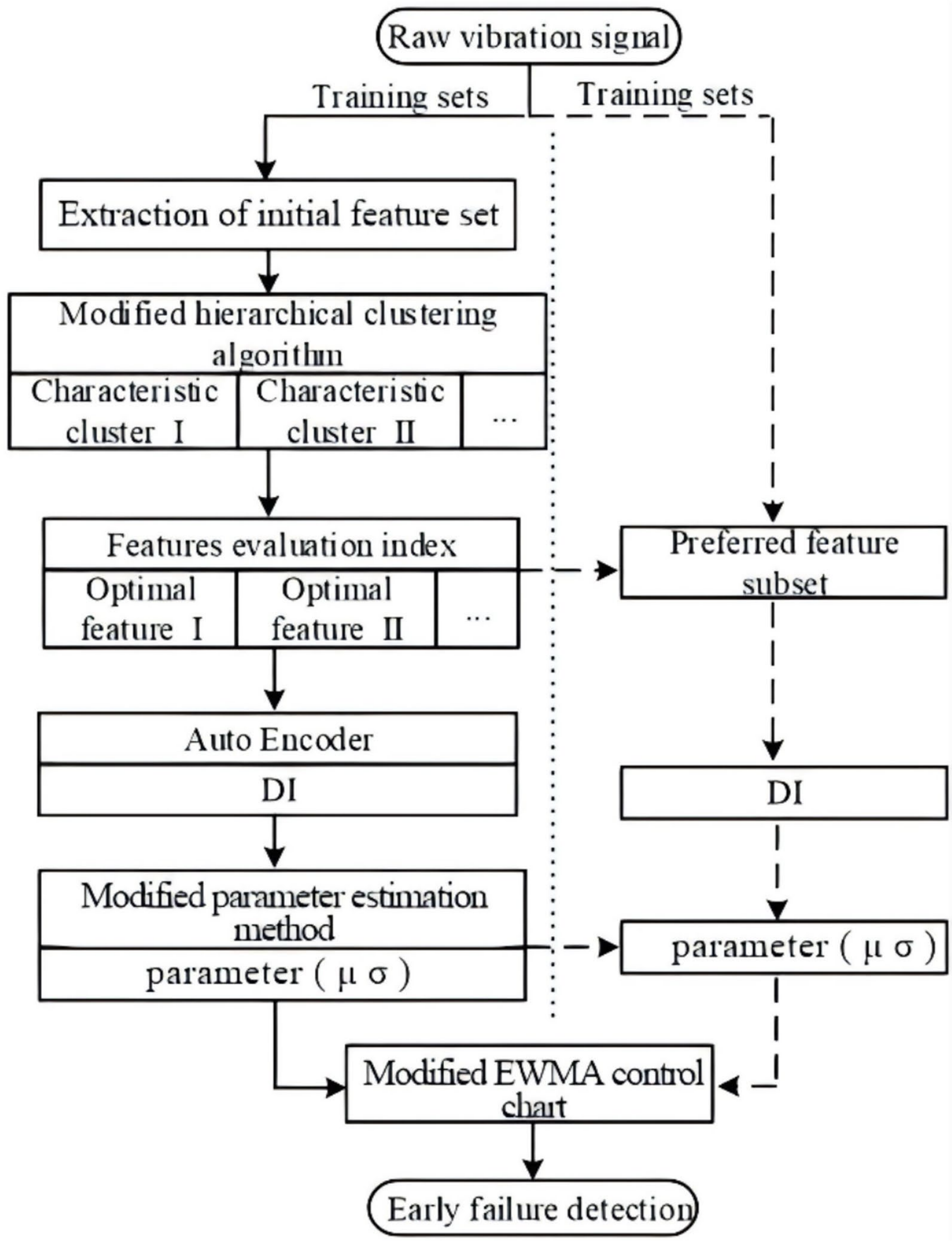

Figure 6 Method flow chart

according to the working conditions. One bearing is randomly selected from each working condition for the test set, while the rest is used for the training set to verify the effectiveness of the proposed method. The test bearings are Bearing 1-1, Bearing 2-4, and Bearing 3-2.

The IEEE PHM 2012 data set is collected on the PRONOSTIA platform, which contains a total of 17 life cycle vibration signals of rolling bearings under three working conditions, including 7 signals under the first working condition, 7 signals under the second working condition, and 3 signals under the third working condition. The sampling frequency is 25.6 kHz, the sampling interval is 10 s, and the sampling time is 0.1 s. To verify the effectiveness of the proposed method more comprehensively, one bearing is selected from the third working condition for the test, while the rest is used for the training set to simulate the situation of few historical data of the current working condition and many historical data of similar working conditions. The number of sampling points in each cycle can be calculated by the formula $N = f_s \times 60/n_s$ (f_s is the frequency, n_s is the revolution speed). To fully ensure integrity and credibility of the fault information of each data sample, the length of each sample is set to 10 cycles.

5.2 Genetic Algorithm Setup and Tagged Data

The individual coding form is based on binary coding in this study. Compared to the commonly used floating-point coding method, the binary coding has the advantages of an easy implementation of the gene crossover algorithm and high efficiency of small-scale search. The initial parameters of the genetic algorithm are set as follows: population size of 40, number of generations of 50, crossover probability of 0.7, mutation probability of 0.01, and generation gap coefficient of 0.9. When the genetic algorithm is optimized, it needs to contain label data, i.e., the sample number of the early fault. The envelope spectrum analysis is used to determine the sample of the first early fault (as much as possible).

For XJTU-SY Bearing 1-1 as an example, the whole life vibration signal is shown in Figure 7. The vibration signal waveform of area A increases slowly, and there is a significant fault in area B. The early fault occurs in area A. Therefore, the envelope spectrum analysis of the samples in area A is carried out to determine the samples where the early fault starts. According to the structural parameters, speed, and failure position of the rolling bearing, the theoretical value of the fault feature frequency of the outer ring of the LDK-UER204 bearing is 107.9 Hz, calculated by the fault feature frequency formula. The vibration signal of each sample is subjected to Hilbert transform and fast Fourier transform to obtain the corresponding envelope spectrum. Finally, the spectral peak frequency and theoretical fault feature frequency in the envelope spectrum are analyzed. As shown in Figure 8, the sample 343 mainly contains frequency components such as those at 34.3, 108.6, 217.2, and 291.4 Hz. The spectral peak at 108.6 Hz is close to the theoretical value of 107.9 Hz. There is no corresponding peak in the sample before this, as shown in Figure 9. The difference

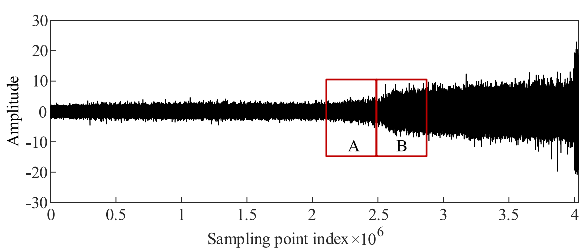

Figure 7 Life cycle vibration signal of rolling bearing 1-1

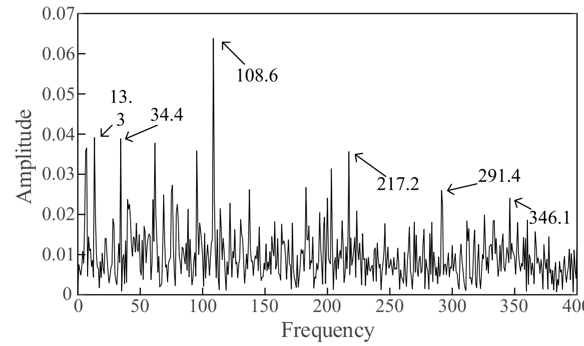

Figure 8 Spectrum of sample 343

between the frequency of the spectral peak and theoretical value is due to the slip effect [28]. Therefore, it is considered that the early fault of Bearing 1-1 occurs in sample 343. When there are multiple failure positions or unknown failure positions, the known failure positions or all possible failure positions are analyzed separately to determine the earliest fault type and sample number.

5.3 Feature Cluster Correlation Analysis

To verify the effectiveness of the modified hierarchical clustering algorithm, the correlation between a single feature and other features in the same feature cluster is calculated, and the correlation triangular matrix shown in pattern 17 is obtained. The mean value of the matrix represents the correlation of the feature cluster. Similarly, the correlation between a feature in the feature cluster and other feature clusters is calculated, and the correlation triangular matrix is obtained. The mean value of the matrix represents the correlation between feature clusters. For Bearing 1-2 in the XJTU-SY data set as an example, 66 statistical features are extracted and normalized as an initial feature set. After the modified hierarchical clustering algorithm, it is divided into multiple feature clusters according to the correlation. As shown in Figure 10, it can be divided into five feature clusters.

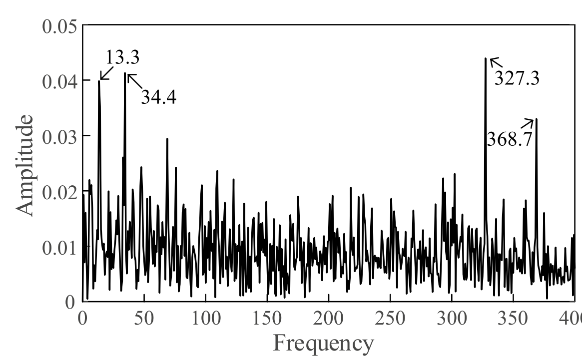

Figure 9 Spectrum of sample 342

In the diagram, the vertical axis represents dissimilarity, reflecting the measure of similarity between different data points. Table 2 shows the correlations within and between feature clusters of the XJTU-SY dataset for bearing 1-2, where I is the correlation within the feature cluster I, I–II is the correlation between feature cluster I and feature cluster II, and so on.

$$C = \begin{pmatrix} |c_{11}| & \cdots & |c_{1n}| \\ & & |c_{nn}| \end{pmatrix}, \tag{17}$$

where n is the number of features of a cluster and $|c_{1\mathrm{n}}|$ is the correlation between the first feature of a cluster and nth feature.

Table 2 shows that the correlation between feature clusters is lower than that within feature clusters, which indicates that the modified hierarchical clustering algorithm is effective in dividing feature clusters with a high intraclass correlation and low interclass correlation.

To further verify the effectiveness of the modified hierarchical clustering algorithm, the feature subset selected by the initial feature set only through the feature evaluation index is compared to the preferred feature subset obtained by the modified hierarchical clustering combined with the feature evaluation index. With the test set as an example, the correlation within the feature subset is calculated; the results are shown in Table 3. The preferred feature subset selected by the feature selection method proposed in this study retains the differences between different types of features.

5.4 Comparative Analysis of Different Methods

The DIs of six rolling bearings in the test set are shown in Figure 11, where (a)–(c) show results for the XJTU-SY data set, while (d)–(f) show results for the IEEE PHM 2012 data set. The constructed DI has a good monotonicity, and the long-term trend is increasing, which can better characterize the degradation process of rolling bearings. In the case of an early fault, the DI increases suddenly, which indicates that the DI extracted by this method is sensitive to an early fault and is suitable for early fault detection of rolling bearings [29].

To verify the advantages of the proposed method, six early fault detection methods are compared including the method proposed in this study, root mean square (RMS) + modified EWMA control chart detection method, DI + Shewhart control chart detection method constructed in this study, RMS + Shewhart control chart detection method, envelope spectrum analysis method, and isolated forest anomaly detection method. Among them, RMS is a classic physical DI method. The Shewhart control chart is a classic anomaly detection method, while the envelope spectrum analysis is a typical method for early bearing fault detection. Isolated forest [30] is a classical unsupervised machine learning anomaly detection algorithm. The number of isolated trees is 100, while the maximum number of required data points is 256. The vibration signal is subjected to spectral transformation. The first 1280 frequency points are used. The deep autoencoder is used to adaptively extract features for detection. The encoder dimension changes to 1280-700-400-200. The abnormal alarm detected first by each method is used as a result of the early fault detection of rolling bearings, as shown in Table 4.

The XJTU-SY dataset of Bearing 1-1 and IEEE PHM 2012 dataset of Bearing 3-2 are used as examples for the analysis. The detection results of the first four methods are shown in Figure 12 and Figure 13. The abscissa

Table 3 Intraclass feature and interclass feature correlations

Data set	Bearing	Method value	Features evaluation index
XJTU-SY	Bearing 1-1	0.512	0.784
	Bearing 2-4	0.383	0.659
	Bearing 3-5	0.425	0.821
PHM 2012	Bearing 3-1	0.483	0.776
	Bearing 3-2	0.462	0.795
	Bearing 3-3	0.531	0.715

Figure 10 Improved hierarchical clustering results

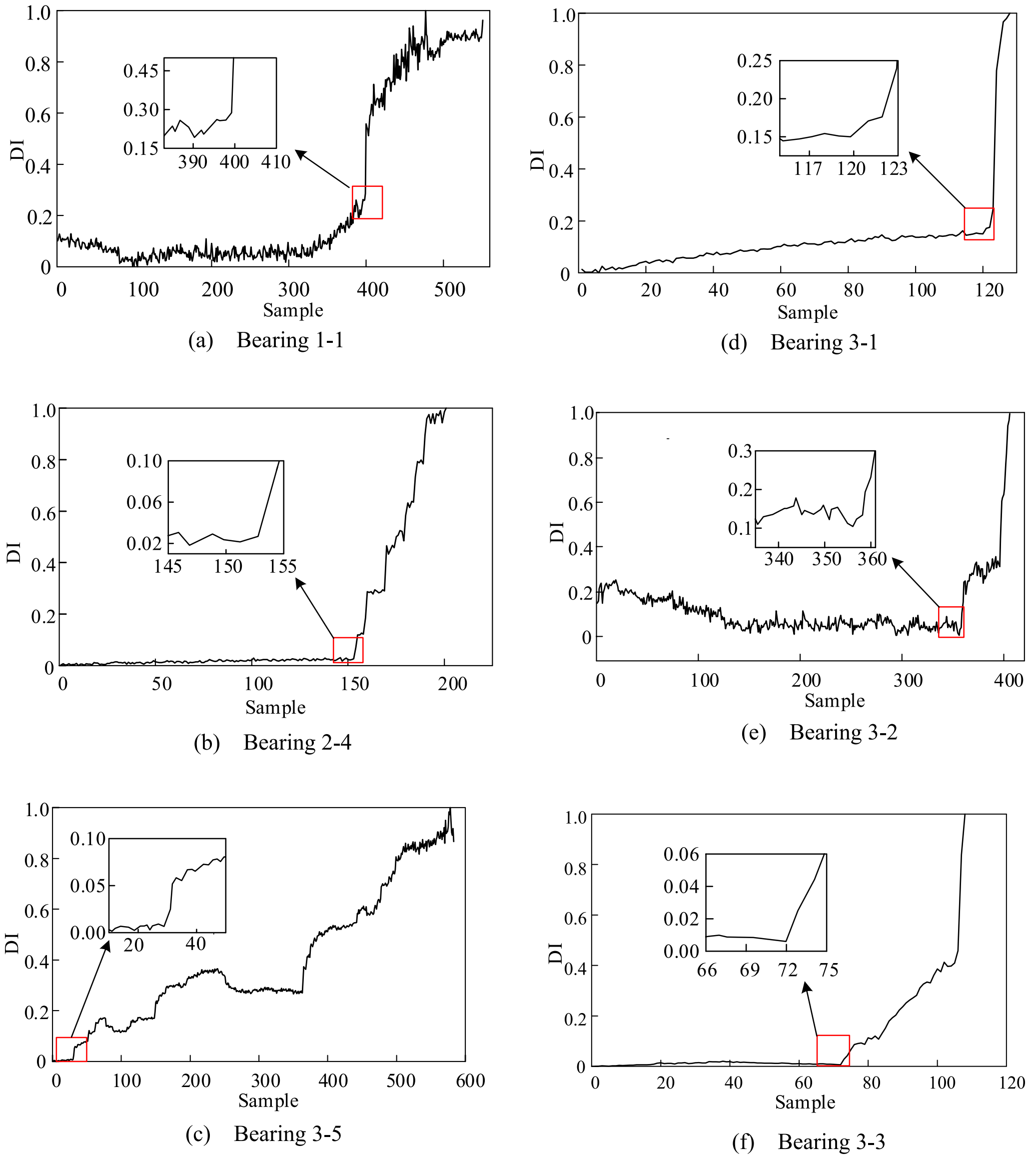

Figure 11 DIs of test set bearings

represents the sample serial number, which, in turn, represents the working time. The red dotted line represents the early fault alarm, while the red solid lines represent the upper and lower control limits.

The envelope spectrum analysis method is employed as a data label production method in this study, which can be used as a significant reference for the accuracy of early fault detection results. As shown in Table 4, the proposed method, RMS + modified EWMA, DI + Shewhart method constructed in this study, and envelope spectrum analysis method do not produce false alarms or detection failures. The detection results of the proposed method and envelope spectrum analysis method are consistent, which indicates the effectiveness of the proposed method. The detection results of the RMS + Shewhart method for the IEEE PHM 2012 data set of Bearing 3-2 are quite different from those of the envelope spectrum analysis method. The detection results of the RMS + Shewhart method contain false alarms and are not accurate. There is a failure of the isolated forest method and lag in the detection results of the remaining methods. According to Figures 12(c), (d) and 13(c), (d), compared to the classical DI, the DI constructed by this method not only can better characterize the degradation process of bearings, but also is more sensitive to early faults. Shewhart control charts

Table 4 Alarm samples of different methods

Data set	Bearing	Proposed method	RMS + modified EWMA	DI + Shewhart	RMS + Shewhart	Envelope spectrum analysis	Isolated forest
XJTU-SY	1-1	343	361	360	367	343	349
	2-4	151	157	155	162	151	168
	3-5	28	31	31	33	28	35
IEEE PHM 2012	3-1	120	124	121	125	120	Failure
	3-2	361	362	369	36	361	375
	3-3	72	73	73	76	72	72

are unable to accurately detect early faults due to the presence of outliers and inaccurate parameter estimation during the normal operation phase. This results in either false alarms or delayed alarms caused by overly broad threshold ranges. The modified EWMA control chart proposed in this study can effectively address the abnormal interference during the normal operation of the bearing, and adaptively set the threshold to detect early faults. The envelope spectrum analysis method requires manual experience. There are redundant features in the features adaptively extracted by the deep autoencoder in the isolated forest method, which leads to the inaccuracy of the subsequent anomaly detection model and affects the detection results. In summary, the proposed method has certain advantages.

6 Conclusions

(1) We propose an improved hierarchical clustering algorithm enabling feature cluster division based on the correlation between features. By combining it with feature evaluation metrics, a high-performance low-correlation optimal feature subset can be selected. The use of an autoencoder for fusion results in a comprehensive degradation indicator, preserving the advantages of certain features, which is more conducive to early fault detection in bearings.

(2) The experimental verification was conducted using the XJTU-SY and IEEE PHM 2012 data sets. Compared to the other five commonly used detection methods, the comprehensive degradation index combined with the improved EWMA control chart detection method proposed in this study is more sensitive to early faults, with detection results obtained earlier than for other methods, and absence of false alarm problems.

(3) The improved hierarchical clustering algorithm can reduce interference and redundancy within the feature set, and integrate the advantages of each feature into a comprehensive degradation index, which can effectively characterize the degradation of bearings. Combined with the EWMA control chart, interference information can be suppressed, which can effectively achieve early fault detection of bearings.

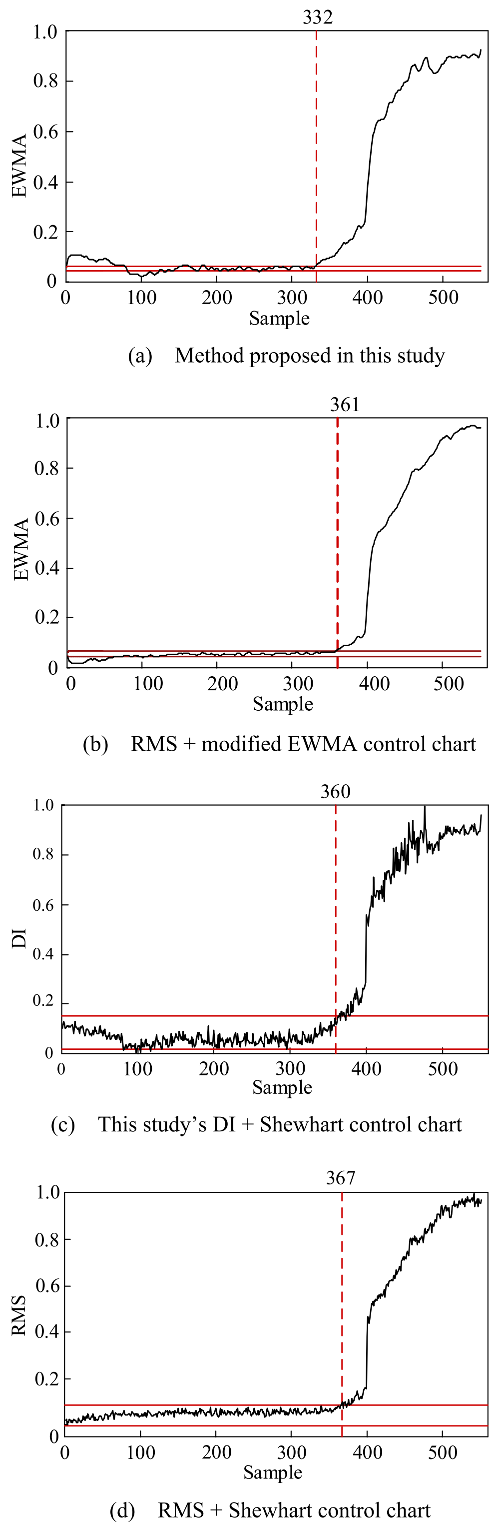

Figure 12 Bearing 1-1 test results

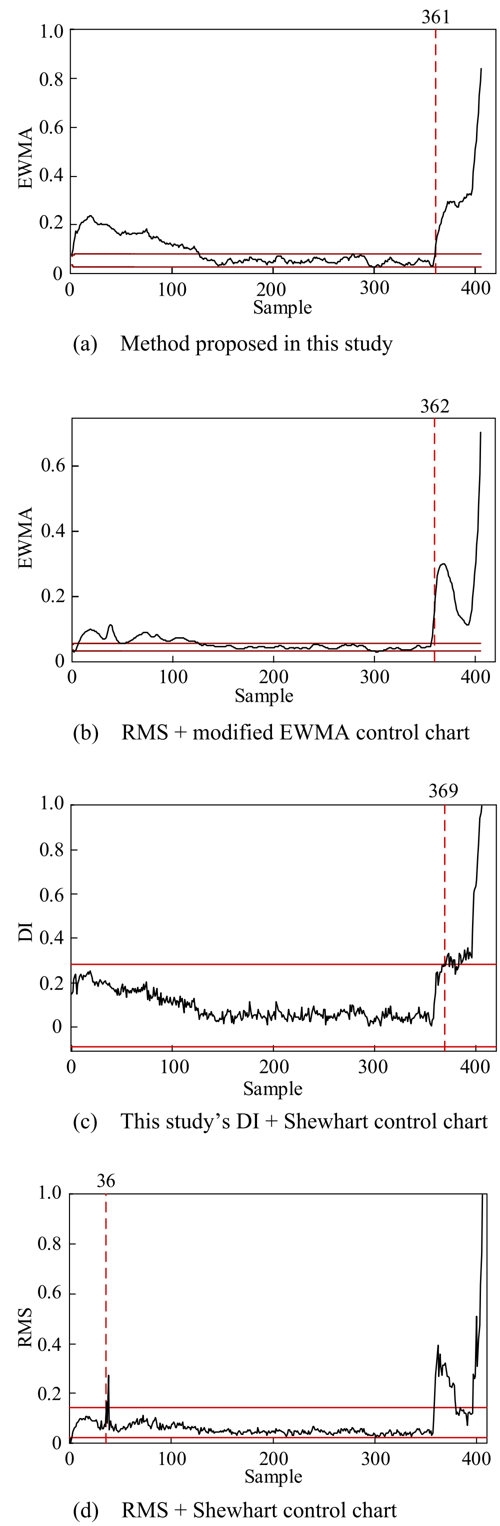

Figure 13 Bearing 3-2 test results

Appendix

See Tables 5, 6

Table 5 Time domain feature definition

Numbering	Feature	Definition
1	Mean value	$F_1 = \frac{1}{N}\sum_{n=1}^{N} x(n)$
2	Maximum value	$F_2 = \max(x(n))$
3	Minimum value	$F_3 = \min(x(n))$
4	Peak value	$F_4 = \max\lvert x(n)\rvert$
5	Peak-to-peak value	$F_5 = F_2 - F_3$
6	Variance	$F_6 = \frac{1}{N-1}\sum_{n=1}^{N}(x(n) - F_1)^2$
7	Standard deviation	$F_7 = \sqrt{F_6}$
8	Root amplitude	$F_8 = \left(\frac{1}{N}\sum_{n=1}^{N}\sqrt{\lvert x(n)\rvert}\right)^2$
9	Absolute mean	$F_9 = \frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert$
10	kurtosis	$F_{10} = \frac{\frac{1}{N}\sum_{n=1}^{N}(x(n)-F_1)^4}{F_6^2}$
11	Skewness	$F_{11} = \frac{\frac{1}{N}\sum_{n=1}^{N}(x(n)-F_1)^3}{F_7^3}$
12	Mean square value	$F_{12} = \frac{1}{N}\sum_{n=1}^{N} x(n)^2$
13	Root mean square	$F_{13} = \sqrt{F_{12}}$
14	F14	$F_{14} = \frac{\sum_{n=1}^{N}(x(n)-F_1)^3}{(N-1)F_7^3}$
15	F15	$F_{15} = \frac{\sum_{n=1}^{N}(x(n)-F_1)^4}{(N-1)F_7^3}$
16	Skewness Factor	$F_{16} = \frac{F_{11}}{F_{13}^3}$
17	Peak factor	$F_{17} = \frac{F_4}{F_{13}}$
18	Shape factor	$F_{18} = \frac{F_{13}}{F_9}$
19	Clearance Factor	$F_{19} = \frac{F_4}{F_8}$
20	Impuls Factor	$F_{20} = \frac{F_4}{F_9}$
21	kurtosis index	$F_{21} = \frac{F_{10}}{F_{13}^4}$
22	Harmonic mean	$F_{22} = \frac{N}{\sum_{n=1}^{N}\frac{1}{x(n)}}$
23	Trimmed mean	$F_{23} = \frac{x([N\alpha]+1)+\cdots+x(n-[N\alpha])}{N-2[N\alpha]}$
24	zero crossing rate	$F_{24} = number\ of\ zero\ crossings/N$
25	Three-order central moment	$F_{25} = \frac{1}{N}\sum_{n=1}^{N}(x(n) - F_1)^3$
26	Fourth-order central moment	$F_{26} = \frac{1}{N}\sum_{n=1}^{N}(x(n) - F_1)^4$
27	Energy operator	$F_{27} = \frac{\frac{1}{N}\sum_{n=1}^{N}(\Delta x(n)-\Delta F_1)^4}{\left(\frac{1}{N}\sum_{n=1}^{N}(\Delta x(n)-\Delta F_1)^2\right)^2}$
28	Quartile distance	$F_{28} =$ the 1st quartile subtracted from the 3rd quartile
29	NA4	$F_{29} = \frac{\frac{1}{N}\sum_{n=1}^{N}(r(n) - \bar{r})^4}{\left(\frac{1}{M}\sum_{m=1}^{M}\left(\frac{1}{N}\sum_{n=1}^{N}(r_m(n) - \bar{r}_m)^2\right)\right)^2}$
30	NA4*	$F_{30} = \frac{\frac{1}{N}\sum_{n=1}^{N}(r(n) - \bar{r})^4}{\left(\frac{1}{M'}\sum_{m=1}^{M'}\frac{1}{N}\sum_{n=1}^{N}(r_m(n) - \bar{r}_m)^2\right)^2}$
31	FM4	$F_{31} = \frac{\frac{1}{N}\sum_{n=1}^{N}(d(n)-\bar{d})^4}{\left(\frac{1}{N}\sum_{n=1}^{N}(d(n)-\bar{d})^2\right)^2}$

Table 5 (continued)

Numbering	Feature	Definition
32	FM4*	$F_{32} = \frac{\frac{1}{N}\sum_{n=1}^{N}(d(n)-\overline{d})^4}{\left(\frac{1}{M'}\sum_{m=1}^{M'}\left(\frac{1}{N}\sum_{n=1}^{N}\left(d_m(n)-\overline{d}_m\right)^2\right)\right)^2}$
33	M6A	$F_{32} = \frac{\frac{1}{N}\sum_{n=1}^{N}(d(n)-\overline{d})^6}{\left(\frac{1}{N}\sum_{n=1}^{N}(d(n)-\overline{d})^2\right)^3}$
34	M6A*	$F_{34} = \frac{\frac{1}{N}\sum_{n=1}^{N}(d(n)-\overline{d})^6}{\left(\frac{1}{M'}\sum_{m=1}^{M'}\left(\frac{1}{N}\sum_{n=1}^{N}\left(d_m(n)-\overline{d}_m\right)^2\right)\right)^3}$
35	M8A	$F_{35} = \frac{\frac{1}{N}\sum_{n=1}^{N}(d(n)-\overline{d})^8}{\left(\frac{1}{N}\sum_{n=1}^{N}(d(n)-\overline{d})^2\right)^4}$
36	M8A*	$F_{36} = \frac{\frac{1}{N}\sum_{n=1}^{N}(d(n)-\overline{d})^8}{\left(\frac{1}{M'}\sum_{m=1}^{M'}\left(\frac{1}{N}\sum_{n=1}^{N}\left(d_m(n)-\overline{d}_m\right)^2\right)\right)^4}$
37	NB4	$F_{37} = \frac{\frac{1}{N}\sum_{n=1}^{N}(e(n)-\overline{e})^4}{\left(\frac{1}{M}\sum_{m=1}^{M}\left(\frac{1}{N}\sum_{n=1}^{N}(e_m(n)-\overline{e}_m)^2\right)\right)^2}$
38	NB4*	$F_{38} = \frac{\frac{1}{N}\sum_{n=1}^{N}(e(n)-\overline{e})^4}{\left(\frac{1}{M'}\sum_{m=1}^{M'}\left(\frac{1}{N}\sum_{n=1}^{N}(e_m(n)-\overline{e}_m)^2\right)\right)^2}$
39	Median	$F_{39} = \left(\frac{N+1}{2}\right)\text{th } \mathit{observation}$
40	Root sum squares	$F_{40} = \sqrt{\sum_{n=1}^{N} x(n)^2}$
41	Mean absolute deviation	$F_{41} = \frac{1}{N}\sum_{n=1}^{N}\lvert x(n)-F_1\rvert$
42	Median absolute deviation	$F_{42} = \frac{1}{N}\sum_{n=1}^{N}\lvert x(n)-F_{39}\rvert$
43	Geometric mean	$F_{43} = \sqrt[n]{\prod_{n=1}^{N} x(n)}$
44	Mobility	$F_{44} = \frac{stdx'(n)}{F_7}$
45	Complexity	$F_{45} = \mathit{mobility\ of}\ x'(n)/\mathit{mobility\ of}\ x(n)$
46	F46	$F_{46} = -\sum_{n=1}^{N}\log\left[\frac{1}{F_7\sqrt{2\pi}}\exp\frac{-(\mathrm{x(n)}-F_1)^2}{2F_2^7}\right]$

Table 6 Frequency domain feature definition

Numbering	Feature	Definition
47	Center gravity frequency	$F_{47}=\frac{\sum_{k=1}^{K}(f_k y(k))}{\sum_{k=1}^{K}y(k)}$
48	Mean square frequency	$F_{48}=\frac{\sum_{k=1}^{K}\left(f_k^2 y(k)\right)}{\sum_{k=1}^{K}y(k)}$
49	Root mean square frequency	$F_{49}=\sqrt{F_{48}}$
50	Frequency variance	$F_{50}=\frac{\sum_{k=1}^{k}\left[(f_k-F_{47})^2 y(k)\right]}{\sum_{k=1}^{k}y(k)}$
51	Frequency standard deviation	$F_{51}=\sqrt{F_{50}}$
52	Mean frequency	$F_{52}=\frac{\sum_{k=1}^{K}y(k)}{K}$
53	F53	$F_{53}=\frac{\sum_{k=1}^{K}(y(k)-F_{47})^2}{K-1}$
54	F54	$F_{54}=\frac{\sum_{k=1}^{K}(y(k)-F_{52})^3}{K\left(\sqrt{F_{53}}\right)^3}$
55	F55	$F_{55}=\frac{\sum_{k=1}^{K}(y(k)-F_{52})^4}{F_{53}^2 K}$
56	F56	$F_{56}=\sqrt{\frac{\sum_{k=1}^{K}\left[(f_k-F_{47})^2 y(k)\right]}{K}}$
57	F57	$F_{57}=\sqrt{\frac{\sum_{k=1}^{K}f_k^4 y(k)}{\sum_{k=1}^{K}f_k^2 y(k)}}$
58	F58	$F_{58}=\frac{\sum_{k=1}^{K}f_k^2 y(k)}{\sqrt{\sum_{k=1}^{K}y(k)\sum_{k=1}^{K}f_k^4 y(k)}}$
59	F59	$F_{59}=\frac{F_{56}}{F_{47}}$
60	Frequency skewness	$F_{60}=\frac{\sum_{k=1}^{K}\left[(f_k-F_{47})^3 y(k)\right]}{F_{56}^3 K}$
61	F61	$F_{61}=\frac{\sum_{k=1}^{K}\left[(f_k-F_{47})^4 y(k)\right]}{F_{56}^4 K}$
62	F62	$F_{62}=\frac{\sum_{k=1}^{K}\left[(f_k-F_{47})^{1/2} y(k)\right]}{K\sqrt{F_{56}}}$

Acknowledgements
Not applicable

Authors' Contributions
Xiangyang Xu is responsible for the overall structure and direction of the paper; Haotian Wang and Ziyuan Ren wrote the manuscript and processed the data; Chuan Zhao and Xihui Liang modified and polished the details of the paper. All authors read and approved the final manuscript.

Funding
Supported by National Key Research and Development Program (Grant No. 2023YFB4203402), National Natural Science Foundation of China (Grant No. 52375042), Chongqing Technology Innovation and Application Development Project (Grant No. CSTB2022TIAD-KPX0078), Chongqing Transportation Technology Project (Grant No. CQJT-CZKJ2024-10).

Data availability
Data will be made available on request.

Declarations

Competing Interests
The authors declare no competing financial interests

Received: 10 July 2023 Revised: 21 March 2025 Accepted: 9 May 2025
Published online: 23 June 2025

References

[1] P Liu, L Q Wang, C W Zhang, et al. Research status and development trend of condition monitoring on main-shaft bearings used in aircraft engines. *Journal of Aerospace Power*, 2022, 37(2): 330-343.
[2] X H Jin, Y Wang, B Zhang. Failure prediction and health management driven by industrial big data. *Computer Integrated Manufacturing System*, 2022, 28(5): 1314-1336.
[3] L T Chen, G H Xu, S C Zhang, et al. Health indicator construction of machinery based on end-to-end trainable convolution recurrent neural networks. *Journal of Manufacturing Systems*, 2020, 54: 1-11.
[4] A Al-Dulaimi, S Zabihi, A Asif, et al. A multimodal and hybrid deep neural network model for remaining useful life estimation. *Computers in Industry*, 2019, 108: 186-196.
[5] A Z Hinchi, M Tkiouat. Rolling element bearing remaining useful life estimation based on a convolutional long-short-term memory network. *Procedia Computer Science*, 2018, 127: 123-132.
[6] B B Li, L Peng. Bearing vibration anomaly detection based on improved self coding network. *Computer Science and Exploration*, 2022, 16(1): 163-175.
[7] W T Mao, L Ding, S Y Tian, et al. Online detection for bearing incipient fault based on deep transfer learning. *Measurement*, 2020, 152: 107278.
[8] I Meddour, S E Messekher, R Younes, et al. Selection of bearing health indicator by GRA for ANFIS-based forecasting of remaining useful life. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 2021, 43: 1-14.
[9] J M Zhou, F L Wang, C C Zhang, et al. An intelligent method for rolling bearing evaluation using feature optimization and GA-SVM. *Journal of Vibration and Shock*, 2021, 40(4): 227-234.
[10] G Niu, F Qian, B K Choi, et al. Bearing life prognosis based on monotonic feature selection and similarity modeling. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2016, 230(18): 3183-3193.
[11] Z X Chen, Z J Zhu, G L Lu, et al. Novel early fault detection and diagnosis for rolling element bearings in graph spectrum domain. *Journal of Vibration and Shock*, 2022, 41(6): 51-59.
[12] G L Lu, G Y Chen, S J Zeng, et al. Method of anomaly detection of mechanical operating state based on vibration image features. *Journal of Vibration, Measurement and Diagnosis*, 2022, 42(1): 148-153+201.
[13] S J Dong , X W Pei, B P Tang, et al. Recognition of rolling bearing life status based on FNER performance degradation indicator and IDRSN. *Journal of Mechanical Engineering*, 2021, 57(15): 105-115. (in Chinese)
[14] A Haq, J Brown, E Moltchanova, et al. New exponentially weighted moving average control charts for monitoring process mean and process dispersion. *Quality and Reliability Engineering International*, 2015, 31(5): 877-901.
[15] H Q Wang, Z Y Liu, W Lu, et al. Frequency domain group sparse rolling bearing feature extraction method. *Journal of Vibration Engineering*, 2022, 35(5): 1242-1249.
[16] Y B Li, A Z Wang, Z B Liu, et al. The entropy algorithm and its variants in the fault diagnosis of rotating machinery: A review. *IEEE Access*, 2018, 6: 66723-66741.
[17] X G Guo, J Wang, C Cheng. An image segmentation algorithm which combines hierarchical clustering algorithm and graph-based segmentation algorithm. *Journal of National University of Defense Technology*, 2019, 44(3): 194-200.
[18] B Zhang, L Zhang, J Xu. Degradation feature selection for remaining useful life prediction of rolling element bearings. *Quality and Reliability Engineering International*, 2016, 32(2): 547-554.
[19] X L Wang, S T Zhang. Identification and robustness analysis of influence nodes in aviation network. *Science Technology and Engineering*, 2019, 19(8): 297-302.
[20] H W Zhang, W B Tang, P F Li, et al. Detection of yarn-dyed shirt cut defects based on denoised convolution autoencoder. *Journal of Basic Science of Textile University*, 2019, 32(2): 119-125+132.
[21] J Lai, X D Wang, Q Xiang, et al. Review of autoencoders and their applications. *Journal of Communications*, 2021, 42(9): 218-230.

[22] F F Shen, H Z Yang. Fault detection of chemical process with outlier value based on self-association kernel regression. *Control Theory and Applications*, 2023, 40(3): 583-592.
[23] H J Su, Y C Luo, F Liu. Overview of the efficiency evaluation and support technology of the equipment system. *Space and Space Defense*, 2023, 6(3): 29-38.
[24] X Q Yang, R M Liu, J Liu. Bearing acoustic signal enhancement method based on Kalman filter and subband selection. *Mechanical and Electrical Engineering*, 2023, 40(11): 1673-1681.
[25] P Shakya, M S Kulkarni, A K Darpe. A novel methodology for online detection of bearing health status for naturally progressing defect. *Journal of Sound and Vibration*, 2014, 333(21): 5614-5629.
[26] B Wang, Y G Lei, L P Li, et al. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. *IEEE Transactions on Reliability*, 2018, 69(1): 401-412.
[27] P Nectoux, R Gouriveau, K Medjaher, et al. PRONOSTIA: An experimental platform for bearings accelerated degradation tests. I*n IEEE International Conference on Prognostics and Health Management*, 2012: 1-8.
[28] L Niu, H Cao, Z He, et al. A systematic study of ball passing frequencies based on dynamic modeling of rolling ball bearings with localized surface defects. *Journal of Sound and Vibration*, 2015, 357: 207-232.
[29] Z H Zhao, L H Li, S P Yang, et al. An unsupervised bearing health index and early fault detection method. *China Mechanical Engineering*, 2022, 33(10): 1234-1243.
[30] G K Hu, Q T Zhang. Biological cluster coefficient of similarity analysis the applicability of the study. *Marine Limnetic Bulletin*, 2019, (4): 140-145.

Xiangyang Xu , born in 1981, is currently a professor at *Chongqing Jiaotong University, China*. His research interests include vibration, shock high performance transmission and intelligence.

Haotian Wang , born in 1998, received his master degree in mechanical engineering from *Chongqing Jiaotong University, China*, in 2023.

Xihui Liang , born in 1984, is currently a professor at *University of Manitoba, Canada*. His research interests include mechanical equipment fault diagnosis forecasting and health management.

Chuan Zhao , born in 1988, is a currently a lecturer at *School of Mechanical and Electrical Engineering, Beijing Institute of Aerospace Technology, China*. His research interests include signal processing, machine learning and mechanical intelligent fault diagnosis.

Ziyuan Ren , born in 1998, received his master degree in mechanical engineering from *Chongqing Jiaotong University, China*, in 2023.