

# Research on Performance Degradation Assessment Method of Train Rolling Bearings under Incomplete Data

Yong Qin, Xuejun Zhao, Dandan Wang, Zhipeng Wang, Limin Jia  
State Key Laboratory of Rail Traffic Control and Safety  
Beijing Jiaotong University  
Beijing, China

Xuejun Zhao  
School of Traffic and Transportation  
Beijing Jiaotong University  
Beijing, China  
15120930@bjtu.edu.cn

**Abstract**— This paper mainly discusses the performance degradation assessment of train rolling bearings under incomplete data, by using the support vector data description (SVDD) and dynamic particle swarm optimization (DPSO). The proposed method is based on the similarity weight for the assessment of the train rolling bearings under incomplete data. Firstly, to obtain effective features of bearing performance degradation from collected vibration data, the local mean decomposition (LMD) is employed to decompose the vibration data. Secondly, the high-dimensionality of features is reduced by the principal component analysis (PCA). And then, on the basis of choosing the kernel parameter and penalty weight, a degradation method based on SVDD is proposed. Finally, the experimental results verified that the proposed method has a better optimization performance than the traditional method and can assess the performance degradation of train rolling bearings under incomplete data.

**Keywords**- rolling bearings; performance degradation assessment; support vector; particle swarm optimization.

## I. INTRODUCTION

Rolling bearing contributes to a significant portion of train operation. It is a kind of rotating machinery part which is huge in quantity and has complex coupling relationship with other parts. Its working status is directly related to the operation safety of the whole vehicle and has a very important influence on the stable and efficient operation of the train<sup>[1]</sup>. However, under the harsh circumstance of high speed, heavy load, the working face of the train bearing is very easy to wear, crack or emerge indentation even rupture due to the repeated action of contact stress. According to statistics, only 10%~20% of train rolling bearings can reach the designed life<sup>[2]</sup>. Therefore, timely and accurate assessment of rolling bearing performance degradation can provide theoretical basis for health management and prediction key components of the train so as to reduce the rate of train failure and promote the active protection of the safety of the train operation.

At present, researches of the train rolling bearing are mainly about the condition monitoring and fault diagnosis.

Most of the research results aim to complete the identification and state identification of fault types and the identification results mostly aim at a single time point while ignoring the slight and long-term characteristics of the performance degradation of bearings<sup>[3-5]</sup>. Therefore, this paper aims at operation data under normal state of the train bearings and discuss performance degradation assessment.

The concept of equipment performance degradation was proposed by Jay Lee<sup>[6]</sup> in 1970s, mainly including two aspects: feature extraction, performance degradation assessment. At present, in the research of rolling bearings, the above two aspects have been widely studied by many scholars both at home and abroad. In aspects of feature extraction, Fourier analysis, Wigner-Ville distribution, wavelet transform, Hilbert-Huang Transform (HHT) and other time-frequency analysis methods have been proposed<sup>[7-10]</sup>. In the evaluation of performance degradation there are three main methods based on statistical theory, artificial intelligence and models<sup>[11-14]</sup>. On the basis of above research results, this paper summarizes the characteristics of the performance degradation of the train bearing. Based on the feature extraction, the performance degradation evaluation method is proposed based on the support vector data description (SVDD) and dynamic particle swarm optimization (DPSO).

## II. EXTRACTION BASED ON THE LMD AND PCA

In order to provide basic data for the performance degradation assessment of the third section of this paper, the whole feature extraction process of the train rolling bearing is required. Performance degradation assessment focuses on the judgment of the whole process of equipment performance degradation and the description of the dynamic performance of the equipment. Therefore, the feature extraction method needs to satisfy: Keep relatively complete vibration information to facilitate monitoring of the comprehensive state of the whole aspect of the device. The extracted features have high sensitivity, which can reflect the early abnormal state of the equipment. The extracted features can distinguish the degree

of degradation to identify the degraded state. For these reasons, based on the original data of vibration, the vibration data are segmented and local mean decomposition(LMD) is performed firstly. Secondly, based on the decomposition of obtained data, calculating the real-time, statistical and multi-dimensional characteristics. Then use principal component analysis (PCA) to reduce high dimensional feature.

The specific algorithms of LMD and PCA can be referenced in [15] and [16] respectively. When calculating the real time features, to obtain the information of features as much as possible, 12 indexes including RMS, peak, peak factor, kurtosis, RMS amplitude, the average absolute value, skewness, form factor, pulse factor, energy, energy moment, Shanon entropy are calculated. During the calculation of statistical features, data are preprocessed and 4 kinds of common statistical distribution model including the Weibull distribution, gamma distribution, normal distribution and logistic regression are used for distribution fitting and feature vectors are constructed by using the parameters of the distribution model.

### III. DEGRADATION ASSESSMENT BASED ON SVDD AND DPSO

This sector mainly discusses the incomplete data of rolling bearings and proposes the performance degradation assessment method based on the SVDD. The data of normal operation are applied for training SVDD hypersphere to find a minimum-volume hypersphere. And the degradation degree of the rolling bearing is obtained by calculating the distance between the feature value of rolling bearing and the hypersphere center. However, owing to the invisibility of multidimensional vector space, the DPSO is applied for the selection of parameters. The flow chart is shown as Fig.1.

#### A. Principle of support vector data description

Support vector data description was originally proposed by Tax and Duin<sup>[17]</sup>. Given a target object set  $x_i \in R^d$   $i=1, \dots, N$ , the basic idea of SVDD is to find a minimum-volume hypersphere in high dimensional space with center  $a_F$  and

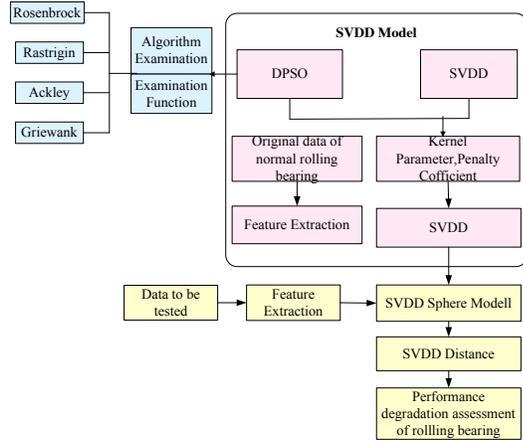


Figure 1 Main steps of bearing performance degradation assessment

radius R to enclose most of the objects, as shown in (1)

$$\text{Minimize } O_p(R, a_F, \xi) = R^2 + c \sum_{i=1}^N \xi_i \quad (1)$$

$$\text{Subject to } \|\phi(x_i - a_F)\|^2 \leq R^2 + \zeta_i$$

$$\zeta_i \geq 0 \quad \forall i=1, \dots, N$$

where c is the penalty weight which gives the trade-off between the volume of the hypersphere and the number of errors.  $\zeta_i$  are slack variables which allows a probability that some of the training samples can be wrongly classified.  $\phi$  is a nonlinear mapping which maps the input object into a high dimensional feature space F.

The dual problem of (1) is as(2). Where,  $K(x_i, x_j)$  is the kernel function.

$$\text{Maximize } O_d(\alpha) = 1 - \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K(x_i, x_j) \quad (2)$$

$$\text{Subject to } \sum_{i=1}^N \alpha_i \quad 0 \leq \alpha_i \leq C \quad i=1, \dots, N \quad C \in [1/N, 1]$$

In this study, the Gaussian kernel,  $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2)$  is selected. It is because Gaussian kernel has only one free parameter to be turned and is shown to yield tighter boundaries than other kernel choices<sup>[18]</sup>. Where,  $\alpha$  is Lagrange multiplier. According to the Kuhn-Tucker conditions, the objects can be classified into three categories: the object with  $\alpha_i = 0$  are inside of the hypersphere; the objects whose  $0 < \alpha_i < C$  are on the hypersphere boundary; and the objects whose  $\alpha_i = C$  fall outside the hypersphere and have nonzero  $\xi_i$ . The objects with  $\alpha_i > 0$  are the support vectors. Objects lying on the hypersphere boundary ( $0 < \alpha_i < C$ ) are also called unbounded support vectors. Objects lying outside the hypersphere ( $\alpha_i = C$ ) are also called bounded support vectors. The center can be expressed as (3). And its radius R can be determined by utilizing the distance between  $a_F$  and any support vector x on the ball boundary (unbounded support vectors), as(4). Finally, for the test object x, the output can be obtained by comparing its distance to the center  $a_F$  with radius D in space F. The SVDD decision function is as (5):

$$a_F = \sum_{i=1}^{N_s} \alpha_i \phi(x_i) \quad (3)$$

$$R = \left( 1 - 2 \sum_{x_i \in SV_s} \alpha_i K(x_i, x_k) + \sum_{x_i \in SV} \sum_{x_j \in SV_s} \alpha_i \alpha_j K(x_i, x_k) \right)^{\frac{1}{2}} \quad (4)$$

$$D(x) = \|\phi(x_i - a_f)\|^2 - R^2 = c - 2 \sum_{i=1}^{N_s} \alpha_i K(x, x_i) \quad (5)$$

where  $c = (1 - R^2) + \sum_{i=1}^{N_s} \alpha_i K(x_i, x_j)$  is a constant. For the rolling bearing fault detection, the real-time monitoring data  $x$  are accepted as target objects if  $D(x) \leq D$ , which indicates the rolling bearing is normal. Otherwise, it is rejected as an outlier, which indicates the rolling bearing is abnormal. There are two parameters needed to be tuned,  $C$  and  $q$ .  $C$  controls the tradeoff between the volume of the hypersphere and the classification error of the model. It can be tuned to achieve the determined confidence level of the fault detection process. By changing the value of the width parameter  $q = 1/2\sigma^2$  in the Gaussian kernel, the description transforms from a solid hypersphere to a Parzen density estimator.

#### B. Kernel parameter and penalty weight selection method based on DPSO

In 1995, James Kennedy and Russell Eberhart proposed the particle swarm optimization (PSO)<sup>[19]</sup>. PSO encourages constructive cooperation and information sharing between particles, which enhance the search for a global optimal solution. However, a potentially dangerous property in PSO still exists: stagnation due to the lack of momentum, which makes it impossible to arrive at the global optimum<sup>[20]</sup>. To avoid these drawbacks of the basic PSO, some improvements such as the time-varying parameters and random perturbation have been proposed. These improvements can enhance convergence of PSO toward the global optimum, to find the optimum solution efficiently. Therefore, DPSO is proposed to avoid the potential drawbacks. The principal of the DPSO is shown as follows.

Definition: The similarity degrees of two particles have to satisfy the following rules:

- (1)  $s(i, i) = 1$
- (2) When  $d(i, j) \rightarrow \infty$ ,  $s(i, j) \rightarrow 0$
- (3) For any particle such as  $i$  and  $j$ ,  $s(i, j) \in [0, 1]$

Based on the rule, the similarity between  $i$  and  $j$  is calculated through (6)

$$s(i, j) = \begin{cases} 1, & d(i, j) \leq d_{\min} \\ 1 - \left[ \frac{d(i, j)}{s_{\max}} \right]^\alpha, & d_{\min} \leq d(i, j) \leq d_{\max} \\ 0, & d(i, j) \geq d_{\max} \end{cases} \quad (6)$$

Where  $d(i, j)$  is the Euclidian Distance between particle  $i$  and particle  $j$ .  $d_{\max}$  and  $d_{\min}$  are constants.

Set iteration times  $t$  of  $d_{\max}$  and  $d_{\min}$  respectively and calculate similarity degree  $s(i, j)$ . When the similarity is zero, the inertia weight of the particle is the largest one representing by  $\omega_{\max}$ . When the similarity is one, the inertia weight of the particle is the lowest one representing by  $\omega_{\min}$ . When the similarity is between zero and one, similarity decreases accordingly. The calculated equations of inertia weight are show as follows:

$$\omega_i = \omega_{\max} - s(i, g)(\omega_{\max} - \omega_{\min}) \quad (7)$$

$$\omega_i = \omega_{\min} - (\omega_i - \omega_{\min}) \cdot \frac{t_{\max} - t}{t_{\max}} \quad (8)$$

Finally, the kernel parameter and penalty weight selection method based on DPSO is made up of 6 steps:

Step1: Generate  $n$  locations and initial speeds of kernel parameter or penalty weight

Step2: Evaluate the adaptation of every kernel parameter or penalty weight.

Step3: Confirm the best location  $p_i$  of every particle and the global best location  $p_g$

Step4: Calculate particle and global similarity of every penalty or penalty weight according to (6) and calculate particle and global weight according to (7) and (8).

Step5: Update the location and position of kernel parameter or penalty weight.

Step6: If the result satisfies the stopping condition, output the result. Otherwise, turn to Step2.

## IV. EXPERIMENTAL RESULT AND ANALYSIS

### A. Original data acquisition

In this paper, the rolling bearing life vibration data are from the University of Cincinnati intelligent maintenance system (IMS) center. The test apparatus is shown in Fig.2. Four ZA-2115 Rexnord rolling bearings are mounted on the same output shaft with different positions. The speed of the output shaft is 2000rpm. In the shaft and bearing, the radial load of 6000lb is exerted by the spring mechanism. Collect the data by data collection card NI DAQ6062EJ and the sample rate is 20KHz. Collect the data every ten minutes and the collection time is 1 second. The data length is 20480.

This paper takes the original data collecting during every turn of the bearing as the division basis which means that 600 points for a segment. So the data collected in 1 second will be divided into 34 segments.

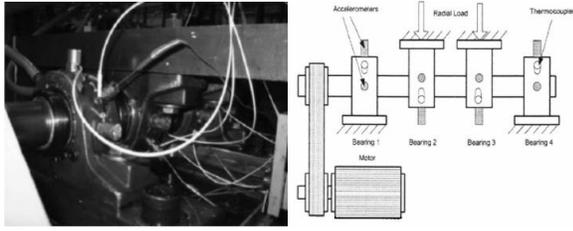


Figure 2. Bearing test equipment and sensors mounting position

## B. Features Extraction

1) *Real Time Feature Extraction*: The data of each segment are calculated by the product function (PF) component of LMD and its 12 real time characteristic values are calculated respectively.

Then feature dimension which the 1s original data correspond to is calculated by (9). However, the PF component derived from each segment is different. The minimum decomposition value of the PF component is recorded as the final PF dimension numbers of the (the first 5 PF of each piece of data decomposition result is obtained, the original data per second corresponding to is  $34 \times 5 \times 12$ ). To strengthen the difference of the feature value of each stage of degradation, this paper takes the 34 segments of data feature value as the final feature value of the data collected per second (Take the root-mean-square (RMS) value as an example, every group consists of 34 data segment, each segment of data for five PF, each piece of data will get and 5 RMS value corresponding to 5 PF. 34 pieces of data will get 34 RMS corresponding to the first PF, the 34 RMS values are added to obtain the value of RMS the first PF corresponding to. Similarly, other feature value which PF corresponds to can also be obtained. feature dimension in 1s can be figured out.

$$D = N \times M \times L \quad (9)$$

Where D: feature dimensions

N: number of data segments

M: PF numbers of each segment

L: numbers of features

2) *Statistical Feature Extraction*: Due to the limitation of several statistical distribution for fitting data value is greater than zero, so PF components does not satisfy the requirements obviously, data preprocessing is needed. Because of the good effect of Hilbert transform on early bearing fault information extraction, do the Hilbert transformation after the decomposition of the PF. Carry out 4 kinds of distributions and fitness tests using the envelope value after the transformation. This paper makes the distribution fitting test with Minitab software and the value of fit test is calculated. The results show that the Logistic distribution and normal distribution are not good. Therefore, the extracted features are four including the Weibull shape parameter, Weibull scale parameter, gamma shape parameter gamma scale parameter.

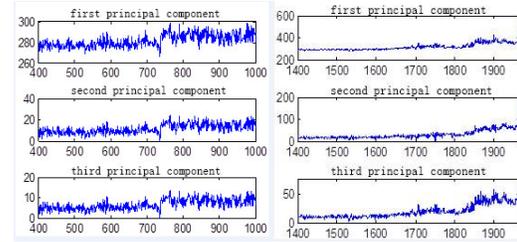
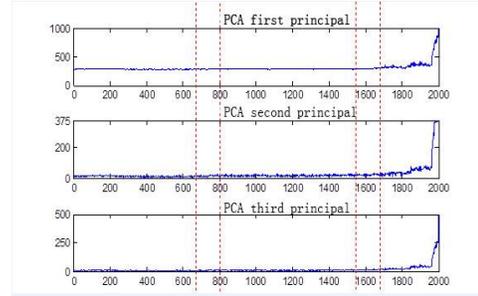


Figure 3. dimension reduction results of PCA (the 1st ~ 3rd principal component)

3) *PCA Dimension Reduction*: As mentioned before, to avoid the trouble of dimensionality and improve the efficiency of the algorithm. On the basis of the extracted multiple real-time and statistical feature, the PCA is used to reduce the dimension.

This paper utilizes the cumulative variance contribution rate method (95%) to choose first 16 feature vectors after dimension reduction. Fig.3 shows the reduced dimension after the first 1~3 main element. It can be seen that the performance degradation trend are well preserved.

## C. The Examination of DPSO

To research the optimization ability of the DPSO, several commonly used functions are applied to the performance examination, including Rosenbrock, Rastrigin, Griewank and Ackley. The optimization ability between PSO and DPSO is simulated via four examination functions. The optimization process between the proposed algorithm and the PSO is shown as Fig. 4. Concluding from the Fig.4, both DPSO and PSO converge quickly at the earlier stage. However, Fig.4d) shows that PSO stops declining from the beginning of the iteration, indicating PSO is easily affected by the partial optimum solution. Therefore, the DPSO has a better simulation result.

## D. Degradation Assessment

1) *Parameter selection*: Different kernel functions will lead to different high dimensions. Therefore, the performance of SVDD will be affected by the kernel function. As far as the development of the kernel functions, Gauss radial basis function is applied widely in the field of fault diagnosis and its kernel function has characteristic  $K(x_i, x_i) = 1$ . Therefore, the use of Gauss radial basis function can simplify the objection function and improve the calculation speed. This paper

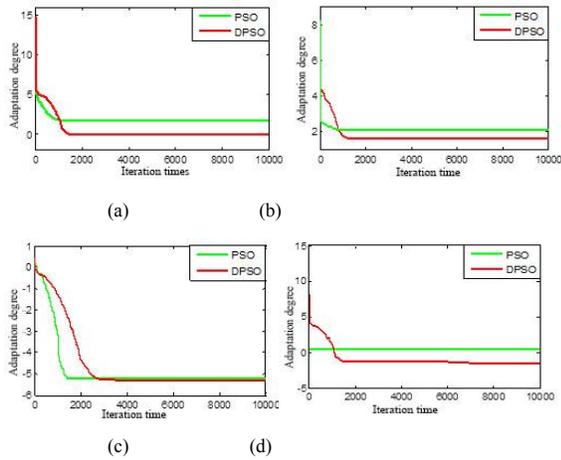


Figure 4. Convergence performance comparison of two methods

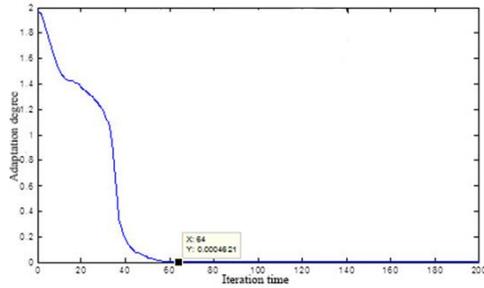
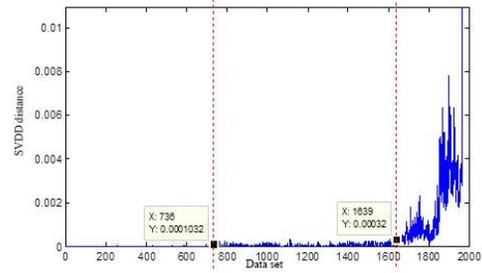


Figure 5. The curve of adaptation

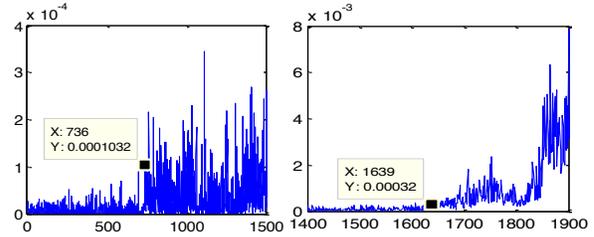
employs DPSO for the selection of kernel parameter and penalty weight and chooses mean square error (MSE) for adaptation and the number of the particle is 20. Besides, the inertia weight is decreasing within the range between 0.95 to 0.4 and the iteration times are 200. The adaptation curve of the SVDD is shown as Fig.5 and the result converges at the sixty-fourth times. The calculated kernel is 15.6832 and the penalty weight is 0.5136.

2) *Degradation assessment based on SVDD*: The degradation assessment of the rolling bearing is carried out with the feature extracted via PCA. This paper mainly discusses the incomplete data and take the 50 groups normal data out of 2000 groups for training. After standardizing process, the data are input into the SVDD model for finding the center and radius of the SVDD hypersphere. Then data to be tested are input into the existing SVDD model and the degradation degree is obtained by calculating the distance between the data and the hypersphere center.

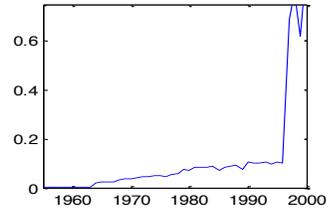
The SVDD distance is shown as Fig.6 and it is divided into three stages. The first stage covers the data between 0 and 735, which means the rolling bearing is normal stage and the SVDD distance is very small and stable. The second stage covers the data between 736 and 1638, which means SVDD distance increases and fluctuates in a larger scale indicating that the rolling bearing enters into an initial stage. The third



(a) SVDD distance



(b) 0-1500 amplification (c) 1400-1900 amplification



(d) 1955-2000 amplification

Figure 6. The curve of performance degradation

stage covers the data between 1639 and 2000 and the SVDD distance increases rapidly up to another scale indicating that the rolling bearing enters into a deep degradation stage.

## V. CONCLUSION

This paper proposes a performance degradation assessment method of train rolling bearings based on SVDD and DPSO. After the decomposition and extraction from the raw vibration signal, with the configuration of the kernel parameter and penalty weight, the SVDD is employed for performance degradation assessment. The experimental results showed that the proposed method is capable to identify the degradation and reflect the health information accurately during the operation of rolling bearings.

## ACKNOWLEDGMENT

This research is supported by the National Key Research and Development Program of China (No.2016YFB1200200), the National Key Technology R&D Program of China (NO.2015BAG12B01) as well as the State Key Laboratory of

REFERENCES

- [1] Choe H C, Wan Y, Chan A K, Neural pattern identification of railroad wheel-bearing faults from audible acoustic signals: comparison of FFT, CWT, and DWT features[J]. Proceedings of the SPIE - The International Society for Optical Engineering, 1997, 3078:480-96.
- [2] Qin T L, Yang Y, Cheng H, Rolling bearing fault diagnosis based on intrinsic mode function energy moment and BP neural network. Journal of Vibration[J], Measurement & Diagnosis, 2008, 28:229-232.
- [3] Nabiyev N K, Diagnostics of axle boxes bearings based on identification measuring method[J]. Transactions of University, Karaganda State Technical University, 2010,(1):77-79.
- [4] He W, Zhou X, Application of the wavelet-SOFM network in roll bearing defect diagnosis[C]. 2009 WRI Global Congress on Intelligent Systems, 2009:8-12.
- [5] He Guangjian, Xing Zongyi, Zuo Cheng, Zhang Yong, Fault diagnosis method for rolling bearing of metro vehicle based on EMD and SVM [J]. Railway Computer Application,2015,(8):1-4,15.
- [6] Lee J, Intelligent Maintenance System (IMS) Technologies[OL],[2013-5-30].
- [7] Yin Q, Qian S, Feng A, A fast refinement for adaptive Gaussian chirplet decomposition[J]. IEEE transactions on signal processing,2002,50:1298-1306.
- [8] Liang Yu, Jia Limin, Cai Guoqiang, Liu Jinzhao, Fault diagnosis method of rolling bearing based on AFD algorithm [J]. China Railway Science,2013,(1):95-100.
- [9] Shi Linsuo, Mi Wenpeng, Application of Wigner-Ville-Distribution-based spectral Kurtosis algorithm to fault diagnosis of rolling bearing [J]. Journal of vibration measurement & diagnosis,2011,31(1):27-31.
- [10] Shen Lu, Li Junsheng, Wang Hongjun, Application of improved Hilbert-Huang transform method in gear fault diagnosis[J]. Journal of aerospace power, 2009, 24(8):1899-1903.
- [11] Hu T G, Pan Y Z, Zhang J S, Integration of soft and hard classifications using linear spectral mixture model and support vector machines[J]. Spectroscopy and Spectral Analysis, 2011, 31(2):508-511.
- [12] Chen Chang, Tang Baoping, L Zhongliang, Degradation trend prediction of rolling bearings based on Weibull distribution and least squares support vector machine [J]. Journal of Vibration and Shock,2014,(20):52-56.
- [13] Gao Hongli, Li Dengwan, Xu Mingheng, Intelligent monitoring system for screw life evaluation [J]. Journal of southeast jiaotong university,2010,45(5):685-691
- [14] Cong Feiyun, Chen JinDong Guang-ming, Performance degradation assessment by Kolmogorov-Smirnov test and prognosis based on AR model [J]. Journal of Vibration and Shock,2012,31(10):79-82.
- [15] Wang Mingda, Zhang Laibin, Liang Wei, Duan Li-xiang. Local mean decomposition method based on B-spline interpolation [J]. Journal of vibration and shock, 2010,29(11): 73-77,254.
- [16] Richard A J,Dean W W, Applied Multivariate Statistical Analysis (Sixth Edition) [M]. New Jersey: Prentice Hall, 2007.
- [17] Tax D.M.J, Duin R.P.W, Support vector domain description[J].Pattern recognition letters,1999,20(11):1191-1199.
- [18] Yang Zhao,Shengwei Wang,Fu Xiao. Pattern recognition-based chillers fault detection method using SupportVector Data Description (SVDD)[J].Applied Energy,2013,112:1041-1048.
- [19] J. Kennedy, R. Eberhart, Particle swarm optimization[C]:Neural Networks, 1995. Proceedings., IEEE International Conference on. IEEE, 1995:1942-1948.
- [20] Shi Y, Eberhart R. A, Modified particle swarm optimizer[C]. In:IEEE World Congress on Computational Intelligence, Ahchorage, AK, USA, 1998:69-73.