Journal of Information Hiding and Multimedia Signal Processing ©2017 ISSN 2073-4212
Ubiquitous International Volume **8**, Number **6**, November **2017**

An Improved Hypersphere Support Vector Machine Method for Vibration Fault Diagnosis of Wind Turbine Gearbox

Chun-Ming Wu and Ji-Hong Yang

College of Information Engineering
Northeast Electric Power University
No.169, changchun Rd., chuanying, 132012, jilin, China
466389144@qq.com; 1944895757@qq.com

Received January, 2017; revised July, 2017

Abstract. *To improve poor performance of SVM which it is difficult for multi-classification, has the high computational complexity and can not reach better accuracy rate of vibration fault diagnosis of gearbox. An improved hypersphere support vector machine method for vibration fault diagnosis of wind turbine gearbox is proposed in this paper. Firstly, a minimum sphere structure classifier with as less sample as possible is constructed for each fault type during training. Then it is directly to input the test samples to the multi-fault classifiers based on hypersphere support vector machine with new classification rule for fault identification in the test. Simulation experimental results have shown that the hypersphere support vector machine has higher diagnostic accuracy than BP neural network in the case of less sample, and the new classification rule improves the accuracy rate of vibration fault diagnosis.*

Keywords: Gearbox, Hypersphere SVM, Multi-classification, Fault diagnosis

---

1. **Introduction.** Wind power as a kind of renewable clean energy is increasingly favored by all countries. However, the wind turbine gear boxes are prone to malfunction because of the strict and poor working conditions, so fault diagnosis of wind turbine gearbox has became a research hotspot[1]. However, the vibration signal of wind turbine gear box is non-linear, non-stationary, and the fault signal characteristics are distributed in different frequency bands, which increase the difficulty of vibration fault diagnosis[2]. At present, artificial intelligence technologies which include expert system method, artificial immune method, Bayesian network method, fault tree diagnosis method, artificial neural network method and fuzzy mathematics method are widely used in the field of fault diagnosis, but the common feature of these methods is that the system has a high ability to acquire knowledge[3,4], and knowledge acquisition is a "bottleneck" problem in constructing artificial intelligence system. It is difficult to obtain a large number of fault samples for the vibration fault diagnosis of wind turbine gearbox. So it has low accuracy rate of fault diagnosis due to the lack of fault samples. And in order to obtain a large number of fault samples which will cause serious economic losses. At the same time for the vibration fault diagnosis of wind turbine gearbox, the above artificial intelligence methods themselves also has some limitations.

Support Vector Machine (SVM) is a new machine learning method which developed on the basis of the limited sample statistical learning theory. SVM solves the practical problems which include less sample, non-linear and high-dimensional pattern recognition, cleverly solves the dimension problem, and makes the algorithm complexity is not related to the dimensions of samples, so SVM has a good generalization performance and becomes a new research hotspot [5,6].

In connection with the multi-classification problem, it is more complex to achieve intelligent classification for support vector machine, and the literature [7] proposed a multi-class identification method based on the ball structure of support vector machine. However hypersphere support vector machine has many advantages, which can effectively deal with unbalanced samples, and can easily extended two types problems to many types problems.

Then, an improved hypersphere support vector method for vibration fault diagnosis of wind turbine gearbox is proposed. The method is to construct the vibration fault diagnosis model based on the less fault sample of the wind turbine gearbox. The new classification rule is adopted when the fault samples are located in the overlapping area of the fault model with multi-hypersphere in the fault classification. And the basic principle of the rule is to calculate the correlation between fault sample and each fault hypersphere of the vibration fault diagnosis model.

2. Hypersphere Support Vector Machine. The hypersphere support vector machine is a classifier that defines each class of data with a hypersphere in the high-dimensional space[8], the basis principle is that the training samples are mapped from the original space to a high-dimension feature space by nonlinear mapping in consideration of the case of linear indivisibility. The hypersphere which is as small as possible and contains all (or nearly all) samples is calculated in the high-dimension feature space[9]. However, it is sensitive to some remote samples for the such definition, so it allows some samples to be outside the hypersphere, therefore, the slack variable ξ_i is introduced. The specific mathematical description of hypersphere support vector machine is as follows: given the training sample set $X = \{x_i | x_i \in R^N, i = 1, 2, ..., m\}$, and the initial optimization problem is:

$$\min_{c_k, R_k} R_k^2 + C\sum_{i=1}^{m} \xi_i$$

$$Subject\ to: ||\varphi(x_i) - c_k||^2 \le R_k^2 + \xi_i \tag{1}$$

$$\xi_i \ge 0, 1, ..., m$$

C_k and R_k is the center and the radius of the smallest enclosing sphere S_k, and C is the penalty factor. The nonlinear datas are treated as training set, and the training datas are mapped to the high-dimensional linear feature space by a non-linear mapping $\phi(x_i)$. But it is not necessary to calculate the non-linear function when solving, just calculate the kernel function $K(x_i, x_j) =< \phi(x_i), \phi(x_j) >$. The Lagrangian multiplier method is used to solve the quadratic programming problem with linear constraint, and getting the dual optimization problem is:

$$\min_{\alpha_{i,k}} \sum_{i,j=1}^{m} \alpha_{i,k}\alpha_{j,k} k(x_i, x_j) - \sum_{i=1}^{m} \alpha_{i,k} k(x_i, x_j)$$

$$Subject\ to: 0 \le \alpha_{i,k} \le C, i = 1, ..., m \tag{2}$$

$$\sum_{i=1}^{m} \alpha_{i,k} = 1$$

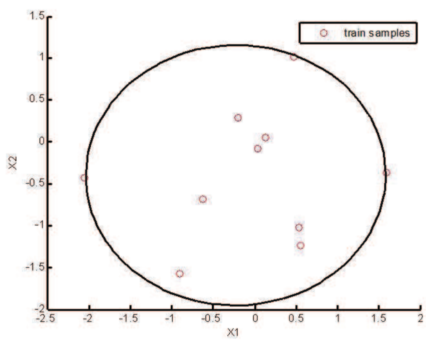
The final decision function is:

$$f_k(x) = \mathrm{sgn}(R_k^2 - \sum_{i,j=1}^{m} \alpha_{i,k}\alpha_{j,k}k(x_i, x_j) + 2\sum_{i=1}^{m} \alpha_{i,k}k(x_i, x) - k(x, x)) \tag{3}$$

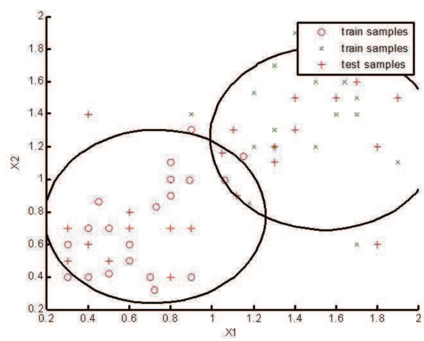
If $f_k(x) = +1$, the sample x is located inside the hypersphere; if $f_k(x) = -1$, the sample x is located outside the hypersphere; if $f_k(x) = 0$, the sample x is located on the hypersphere, the two-dimensional structure diagram is shown in Fig.1(a).

When $\alpha_i > 0$, the corresponding samples x_i are support vectors. The radius R_k can be calculated by making $f_k(x) = 0$ with support vectors. And there are many options for kernel functions of SVM, such as polynomial kernel function $k(x_i, x_j) = (1 + x_i x_j)^p$ and Gaussian kernel function $k(x_i, x_j) = \exp(-||x_i - x_j||^2/(2\sigma^2))$.

Ideally, any two hyperspheres are independent of each other and each new sample can be correctly classified. It will appear the situation which has the overlapping part of two or more hyperspheres in the actual application, and a sample point maybe belongs to several hyperspheres at the same time. How to correctly classify samples of overlapping regions has a great influence for the performance of the result classifier. Fig.1(b) shows the two-dimensional case of two hyperspheres which have overlapping part.



(a)



(b)

Figure 1. (a) The schematic diagram of the minimum enclosing sphere, (b) Two-dimensional case of two hyperspheres have overlapping part.

3. **Improved Hypersphere Support Vector Machine.**

The classification rule for the overlapping parts of the hyperspheres is usually to calculate the distance between the test point and the center of the hyperspheres[10]; the improved classification rule is to calculate the projection of the test sample to the line of two center of the intersecting hyperspheres[11]; and another improved classification rule is to reconstruct the sub-hyperspheres according to certain rules, including the same error samples and the heterogeneous error samples [12].

But there are some disadvantages of these methods, which the methods of spherical center distance and projection have large classification error and the low accuracy rate; and the method of sub-hypersphere is complex of overall thinking structure, which indirectly increases the computational complexity.

For the existing problems, a new classification rule for the overlapping part of the hyperspheres classifier is proposed. The basis principle of the new classification rule is to calculate the connection between the test sample and each hypersphere fault classifiers; and the test sample belongs to the hypersphere classifier with larger connection. The improved formula is as follows, firstly, all samples are initialized and k_0 is a sample of the hypersphere overlapping part, k_i are all samples in any one hypersphere, i is the number of the samples in one hypersphere and each sample is composed of n vectors.

$$k_0 = (k_0(1), k_0(2)......k_0(n)) \tag{4}$$

$$k_i = (k_i(1), k_i(2)......k_i(n)) \tag{5}$$

Calculating the connection between k_0 and k_i :

$$r(k_0(x), k_i(x)) = \frac{1}{1+|(k_0(x+1)-k_0(x))-(k_i(x+1)-k_i(x))|} \tag{6}$$

$$R(k_0, k_i) = \frac{1}{n-1}\sum_{x=1}^{n} r(k_0(x), k_i(x)) \tag{7}$$

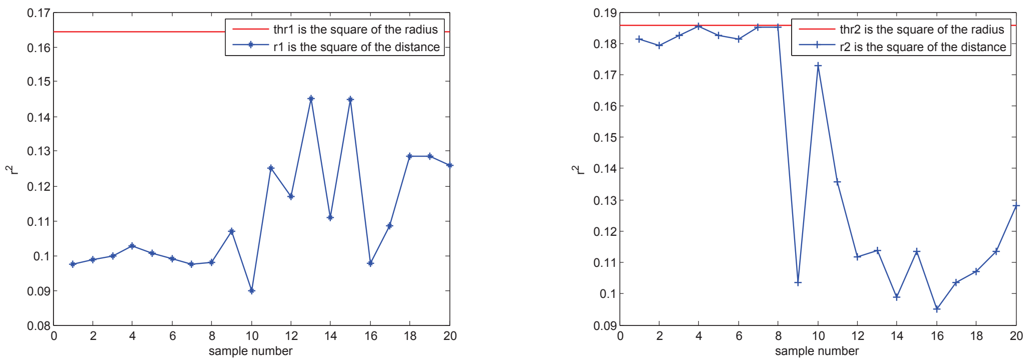
Calculating the connection between k_0 and all samples of any one hypersphere:

$$\varepsilon = \sum_{i=1}^{m} R_i(k_0, k_i) \bullet \frac{1/L_i}{\sum_{i=1}^{m}\frac{1}{L_i}} \tag{8}$$

L_i is the euclidean distance between k_0 and k_i.

Selecting Iris data set in the UCI as the sample data source for the experiments to illustrate the effectiveness and superiority. In order to verity the new classification rule for the overlapping part of the hyperspheres, so the test samples used should be targeted in the verification. Selecting two types samples with four-dimension as training datas in the Iris data set, and 42 samples of each type are selected as training datas to train the hyperspheres; 10 samples of each type are selected as the test datas (20 samples is in overlapping part of the hyperspheres). The radial basis function is chosen as the kernel function in the experimental verification, , which the kernel width is 0.05 and the penalty parameter C is 5.

The line charts of Fig.2(a) and Fig.2(b) respectively represent r1 and r2, which are the square of the distance between the 20 test samples and the center of two hyperspheres. The red lines are thr1 and thr2, which are the square of the radius of two hyperspheres.

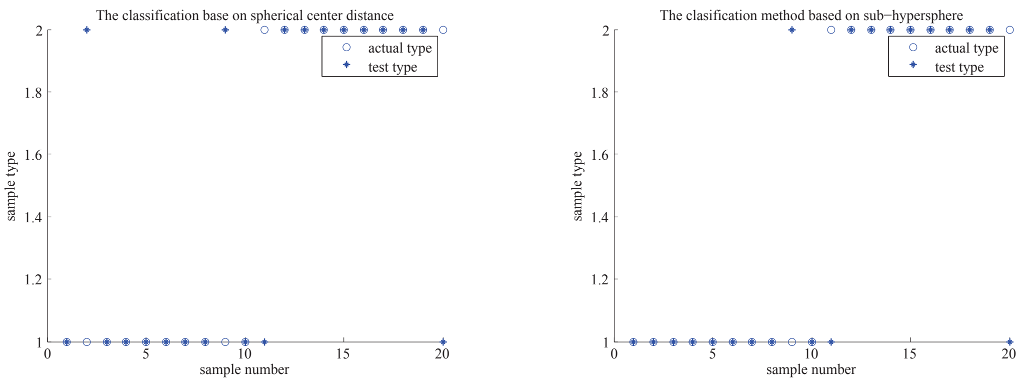


(a) (b)

FIGURE 2. (a) The square of the distance between test samples and center of the first hypersphere, (b) The square of the distance between test samples and center of the second hypersphere.

There are three cases for the classification of the test samples. Case 1 is the test samples in outside the hypersphere; case 2 is the test samples in only one hypersphere; case 3 is the test samples in the overlapping part of two hyperspheres. It can be seen from Fig.2(a) and Fig.2(b) that the 20 test samples belong to Case 3.

So in Case 3(20 test sample points are located in the overlapping part of the two hyperspheres), respectively, the rule of classification based on spherical center distance, the improved classification rule based on the sub-hypersphere, and new classification rule are used for the experimental simulation. Fig.3(a) is the comparison diagram of the test results and actual results based on the classification rule of the spherical center distance, and four test samples are classified incorrectly, so the accuracy rate of the test samples is 80% ((16/20) * 100%). Fig.3(b) is the comparison diagram of the test results and actual results based on the classification rule of the sub-hypersphere, and there test samples are classified incorrectly, so the accuracy rate of the test samples is 85%((17/20) * 100%). Fig.4 is the comparison diagram of the test results and actual results based on the new classification rule, one test sample is classified incorrectly, so the accuracy rate of the test samples is 95%((19/20) * 100%). It can be seen that the new rule of classification is superior to other classification methods, and improves the accuracy rate of classification.



(a) (b)

FIGURE 3. (a) Comparison diagram of test result and actual result based on distance method, (b) Comparison of test result and actual result based on sub-hypersphere method.

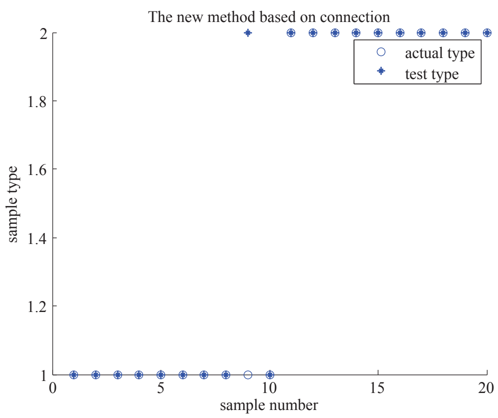


FIGURE 4. Comparison diagram of test result and actual result based on improved method

4. **In the application of the wind turbine gearbox vibration fault diagnosis.**

4.1. **Diagnostic Procedure.** .

(1)Pretreatment for the vibration datas of wind turbine gearbox ;

(2)Using vibration datas of different fault types to train n independent hypersphere SVMs and getting the vibration fault diagnosis model of wind turbine gearbox.

(3)Putting the fault datas into the trained model to identify the fault type, and the rule of classification are:

If the test sample is out of all the hyperspheres, computing the distance from the sample to the centers of all hyperspheres; if the test sample is in the overlapping part of the hyperspheres, the decision is made by the new rule classification.

The flow diagram of vibration fault diagnosis based on the new classification rule is shown in Fig.5.

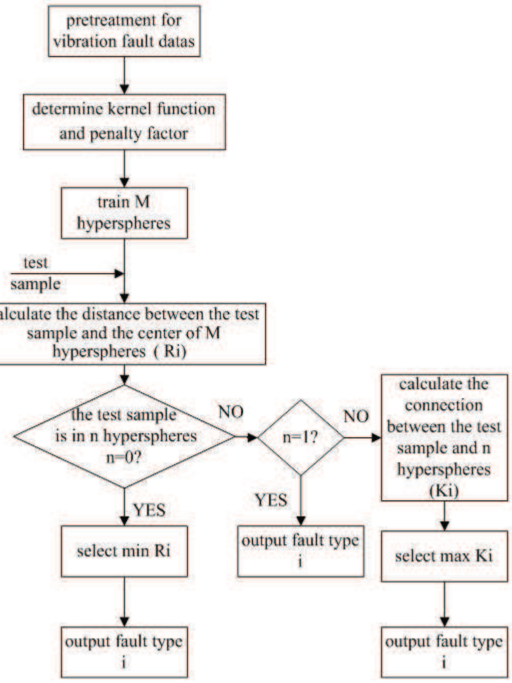


FIGURE 5. The flow diagram of vibration fault diagnosis based on the improved method

4.2. **The selection of kernel function and its parameters.** Commonly, the kernel functions of SVM are linear kernel function, radial basis kernel function, Sigmoid kernel function and polynomial kernel function. Domestic and foreign literature shows that the performance of SVM is not related to the choice of kernel function, so the radial basis kernel function is chosen as kernel function of SVM, which is small calculated quantity and only one parameter which is kernel width σ. And the mathematical formula of radial basis kernel function is $k(x_i, x_j) = \exp(-||x_i - x_j||^2/(2\sigma^2))$, which σ^2 is 1 and the penalty factor C is 10 through several experiments.

4.3. **Experimental verification.** According to the "VDI3834 Wind Power Standard" [13] issued by the German Society of Engineers and the "Guidelines for Vibration Condition Monitoring and Diagnose of Wind Turbine Generators" [14] issued by the National Energy Administration, acceleration sensor should be perferentially chosen to monitor vibration condition in high temperature or strong magnetic field environment. So the condition monitoring of wind turbine gearbox should select the acceleration sensor. In the practical application, the GW50-750 wind turbine which in one region of Burqin wind farm was selected for experiment verification. The vibration datas of the wind turbine gearbox are measured by the standard acceleration sensor placed in the key part of the gearbox; and the collected data for denoising through the EMD; then, the feature vector of the sample is extracted by the wavelet packet. Table 1 gives the characteristic vectors of some training samples.

The four fault states of wind turbine gearbox are bearing inner ring fault, bearing outer ring fault, tooth surface wear and tooth breakage, and 40 samples are used as training samples of each fault state, and 10 samples are used as test samples of each fault state.The fault type is determined by the vibration fault diagnosis model based on the method which are the hypersphere support vector machine, the sub-hypersphere support vector machine and the improved method. Table 2 shows the diagnosis results, it can be seen from Table 2 that the new method has better effectiveness of fault diagnosis than the other two methods.

In order to further verify the accuracy, 40 groups, 80 groups, 120 groups, 200 groups, 500 groups of each fault sample were selected in the harsh environment, 20 fault samples of each fault state as the test samples. Respectively, four methods which are the BP neural network, the hypersphere support vector machine, the sub-hypersphere support vector machine and the improved method are used for fault identification. Fig.6 shows the number of diagnostic errors based on the four methods.

Table 1. The characteristic vectors of some training samples of gear box

operate status	number	characteristic vector of some training samples							
normal	1	0.5680	0.1095	0.0369	0.0689	0.0940	0.0398	0.0351	0.0255
	2	0.5613	0.12	0.0375	0.0734	0.0991	0.0496	0.0391	0.0367
	3	0.5778	0.1312	0.0381	0.0834	0.0562	0.0599	0.0381	0.028
tooth surface wear	4	0.6987	0.0993	0.0271	0.0506	0.0328	0.0163	0.0187	0.0135
	5	0.6855	0.0723	0.0285	0.0542	0.0789	0.0243	0.0289	0.0267
	6	0.6973	0.0833	0.0261	0.0753	0.0613	0.0281	0.0337	0.0345
Tooth breakage	7	0.2439	0.0718	0.1293	0.0702	0.0236	0.1315	0.1687	0.1613
	8	0.2395	0.0459	0.1156	0.0791	0.0275	0.1076	0.1735	0.2356
	9	0.3417	0.0596	0.0937	0.0567	0.0234	0.1097	0.1522	0.2313

It can be seen from Fig.6 that the diagnosis accuracy rate of the hypersphere SVM is superior to that of neural network, at the same time, the improved hypersphere SVM which in this paper improves the accuracy rate of vibration fault diagnosis compared with other methods of the hypersphere SVM.

5. **Conclusion.** The paper uses four fault samples of wind turbine gearbox to train four hyperspheres which structure a vibration fault diagnosis model for fault type identification

Table 2. Fault diagnosis results of gear box

Fault condition	The number of train samples	The number of test samples	The number of diagnostic errors (improved method)	The number of diagnostic errors (sub-hypersphere method)	The number of diagnostic errors (distance method)
Bearing inner ring failure	40	10	0	1	2
bearing outer ring fault	40	10	1	2	3
tooth surface wear	40	10	2	2	4
tooth breakage	40	10	1	2	3

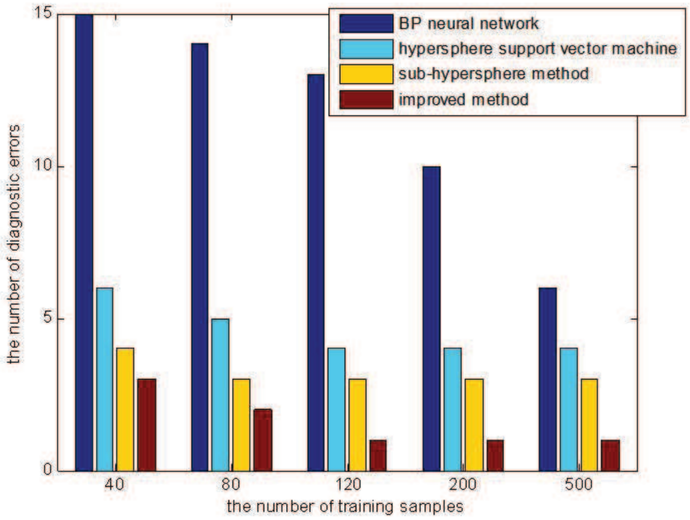


FIGURE 6. The comparison diagram of the number of fault diagnosis errors based on four methods

based on the improved hypersphere support vector machines, and the new classification rule is used in fault identification. Experiments show that the advantage of support vector machine (SVM) is superior to that of neural network in the case of less sample. At the same time, the improved hypersphere support vector machine method has higher effect of classification and higher diagnostic accuracy rate than other classification methods of the hypersphere SVM, and this method is also applied to the fault diagnosis of other engineering parts due to good performance of classification based on the improved method in this paper.

REFERENCES

[1] C. M. Wu, Q. H. Song., Fault Diagnosis of Wind Turbine Vibration Based on Wavelet Transform and Neural Network, *Journal of Information Hiding and Multimedia Signal Processing*, vol.7, no.4, pp. 898–905, 2016.
[2] C. J. Huang, X. W. Wu, W. S. Cao., Based on the full vector envelope technology of LMD and its application in TRT vibration fault diagnosis, *Electric Power Automation Equipment*, vol.35, no.2, pp. 168–174, 2015.
[3] S. Q. Kang, Y. J. Wang, G. X. Yang., *Rolling Bearing Fault Diagnosis Method Using Empirical Mode Decomposition and Hypersphere Multiclass Support Vector Machine, Proceedings of the CSEE*, vol.31, no.14, pp.96–103, 2011.
[4] H. R. Xue, K. Y. Zhang, B. Li., Fault diagnosis of transformer based on the cuckoo search and support vector machine, *Power System Protection and Control*, vol.43, no.5, pp.8–14,2015.
[5] C. H. Zhao., SVM Parameter Optimization in Fault Diagnosis for Wind Power Gear Box, *China Mechanical Engineering*, vol.26, no.16, pp.2222–2226, 2015.
[6] Q. He, D. L. Chu, Z. Zhang., Fault Diagnosis of Rolling Bearing Based on LMD Slice Bispectrum and SVM, *Journal of Northeast Dianli University*, vol.43, no.5, pp.62–67, 2016.
[7] M. L. Zhu., Solving the Problem of Multi-class Pattern Recognition with Sphere-structured Support Vector Machines, *Journal of Nanjing University(Natural Sciences)*, vol.39, no.2, pp.153–159, 2003.
[8] S. Y. Hu., An Approach to Discriminate Overlap Region of Multi-class Classification SVDD for Analog Circuits Fault Diagnosis, *Computer Measurement and Control*, vol.24, no.1, pp.50–54, 2016.
[9] S. Liu, P. Chen, X. Z. Li., A Weighted Hyper-Sphere Support Vector Machine for Multi-class Classification, *Microelectronics and Computer*, vol.32, no.1, pp.19–24, 2015.
[10] S. Q. Kang, Y. J Wang, Y. C. Jang., Fault Classification of Rolling Bearing Based on Hypersphere Spherical Center Distance Multiclass Support Vector Machine, *Proceedings of the CSEE*, vol.34, no.14, pp.2319–2336, 2014.

[11] T. Xu, D. K. He., Theory of Hypersphere Multiclass SVM, *Control Theory and Application*, vol.26, no.11, pp.1293–1298, 2009.

[12] P. Chen, S. Liu L. H. Li., Research on Multiple Sub-Hyper-Sphere Support Vector Machine, *Microelectronics and Computer*, vol.31, no.12, pp.28–33, 2014.

[13] VDI3834 part 1: Measurement and evaluation of the mechanical vibration of wind turbine and their components, 2009.

[14] National Energy Administration. NB/T 31004-2011. *Guidelines for Vibration Condition Monitoring and Diagnose of Wind Turbine Generator*, China Electric Power Press, Beijing, 2011.