Machinery Fault Diagnosis Based on Improved Algorithm of Support Vector Domain Description and SVMs

Qiang Wu[1]
[1] Marine Engineering College, Dalian Maritime University, Dalian,Liaoning, 116026,China
wuqiangdl@sohu.com

Chuanying Jia[2]
[2] Navigation College, Dalian Maritime University, Dalian,Liaoning, 116026,China

Wenying Chen[3]
[3] Department of Precision Instruments and Mechanology, Tsinghua University,Beijing 100084, China

Xiaoshuai Ding[4]
[4]Department of Education, Tibet Nationalities Institute, Xianyang, Shaanxi, 712082,China

***Abstract*—In order to improve accuracy of fault diagnosis based on SVMs, an improved algorithm of support vector domain description (ISVDD) is proposed, used to pretreat the fault data. ISVDD constructs the recognizer of fault data by introducing an optimal sphere instead of the minimum sphere. The recognizer can sift out the fault data belonging to new unknown fault types and avoid erroneous diagnosis. A new method of fault diagnosis is given based on ISVDD and hierarchy structure SVMs for the multi-fault problem. Numerical experiments are performed on a real dataset. The results show that ISVDD can be used to pretreat the fault data effectively and that the new method of fault diagnosis has higher precision and can be used in practice.**

***Keywords*—fault diagnosis; support vector domain description; pretreating process of fault data; SVMs**

I. INTRODUCTION

Rotating machines are widely used in many industrious areas. There are many kinds of rotating machinery faults and the reasons of these faults are complicated. Monitoring and diagnosing for rotating machinery faults are important to keep machines work well [1]. The essence of fault diagnosis is pattern recognition of machine states [2].

The algorithm of support vector machines (SVMs) proposed by Vapnik is one of the important tools in pattern recognition [3], [4]. Originally, it was developed for two-class classification. Its main idea was to find an optimal hyperplane that separates data into two classes. This hyperplane was decided by the support vectors on it.

SVMs have exhibited great advantages in machinery fault diagnosis field [5],[6],[7]. A classifier of SVMs is built according to the data of known fault types. A fault datum in field test will be classified as one of those types by the classifier of SVMs. When the new fault types appear, they will be classified as those known fault types always. Because known fault types cannot cover up all fault types, sometimes, error decision will be made. Thus, it is necessary to construct the recognizer which is used to sift out the data of some new unknown fault types.

Tax and Duin presented the support vector domain description (SVDD). SVDD constructed a minimum sphere for a set of data in some high-dimensional space and the sphere separated these data from outliers [8].

This paper proposes an improved algorithm of support vector domain description (ISVDD). The fault recognizer based on ISVDD can be used to separate the data of known fault types from those of new unknown fault types. ISVDD constructs an optimal sphere to encircle the data of known fault types, while the data outside the optimal sphere are believed to belong to new unknown fault types. And then a fault classifier is constructed based on the algorithm of hierarchy structure SVMs for classifying the data of known fault types [9]. The data of known fault types inside the optimal sphere are classified by the fault classifier.

The paper is organized as follows. Section II gives a brief introduction to SVDD and constructs the recognizer based of SVDD. Section III proposes ISVDD and constructs the recognizer based of ISVDD. Section Ⅳ introduces hierarchy structure SVMs for multi-class classification. In section V, a new method of fault diagnosis is given and the experiments are performed to verify the performance of the recognizer based on ISVDD and the method of fault diagnosis proposed in this paper. Finally, section Ⅵ concludes the paper with some remarks.

II. SVDD AND THE FAULT RECOGNIZER BASED ON IT

A. SVDD

Given a dataset of known faults S, $S=\{\boldsymbol{x}_i\,|\boldsymbol{x}_i\in R^n, i=1,2,\ldots,l\}$, we try to find a sphere with minimum volume for S. This sphere is named the minimum sphere and is represented by U. The minimum sphere U contains all (or most of) the data points of the set S with center a and radius R (where R is as small as possible). This minimum sphere is very sensitive to the most outlying points, so some points are allowed outside the sphere. The following optimization problem is to find the minimum sphere U:

$$\begin{aligned} \min \quad & R^2 + C\sum_{i=1}^{l}\xi_i \\ s.t. \quad & \begin{cases} \|\boldsymbol{x}_i - \boldsymbol{a}\|^2 \le R^2 + \xi_i \\ \xi_i \ge 0, \quad i=1,\cdots,l. \end{cases} \end{aligned} \tag{1}$$

where the ξ_i are slack variables and the variable $C > 0$ gives the trade-off between volume of the sphere and the number of known fault points rejected.

By the Lagrange multiplier method, (1) is transformed into a simpler dual problem:

$$\max \quad \sum_{i=1}^{l}\alpha_i(\boldsymbol{x}_i\boldsymbol{\cdot}\boldsymbol{x}_i)-\sum_{i,j=1}^{l}\alpha_i\alpha_j(\boldsymbol{x}_i\boldsymbol{\cdot}\boldsymbol{x}_j)$$

$$s.t. \quad \begin{cases}\sum_{i=1}^{l}\alpha_i=1\\ 0\le\alpha_i\le C\\ i=1,\cdots,l.\end{cases} \tag{2}$$

where α_i are Lagrange multipliers. The minimal sphere can be obtained by solving the quadratic programming problem. The support vectors of the sphere are the data points whose α_i are non-zero.

The method stated above constructs the sphere which includes all data in the input space. In most cases, these data are not spherically distributed, even when the most outlying points are ignored. The points in the original input space are mapped into a feature space by a kernel function where the kernel function $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ satisfies Mercer's theorem. The feature space is high-dimensional. The inner product of two vectors $(\boldsymbol{x}_i\boldsymbol{\cdot}\boldsymbol{x}_j)$ in (2) can be replaced by the kernel function. In the feature space, the quadratic programming problem (2) for the minimal sphere can be expressed as

$$\max \quad \sum_{i=1}^{l}\alpha_i K(\boldsymbol{x}_i,\boldsymbol{x}_i)-\sum_{i,j=1}^{l}\alpha_i\alpha_j K(\boldsymbol{x}_i,\boldsymbol{x}_j)$$

$$s.t. \quad \begin{cases}\sum_{i=1}^{l}\alpha_i=1\\ 0\le\alpha_i\le C\\ i=1,\cdots,l.\end{cases} \tag{3}$$

To judge a test point $\boldsymbol{x}$, the distance r_x to the center of the sphere has to be calculated. The distance r_x can be calculated by

$$r_x^2 = K(\boldsymbol{x},\boldsymbol{x}) - 2K(\boldsymbol{x},\boldsymbol{a}) + K(\boldsymbol{a},\boldsymbol{a}). \tag{4}$$

Typical kernel functions used in pattern recognition include:

Polynomial kernel function:

$$K_1(\boldsymbol{x}_i,\boldsymbol{x}_j) = ((\boldsymbol{x}_i\boldsymbol{\cdot}\boldsymbol{x}_j)+1)^q$$

Gaussian kernel function:

$$K_2(\boldsymbol{x}_i,\boldsymbol{x}_j) = exp(-\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2/2\sigma^2)$$

Sigmoid kernel function:

$$K_3(\boldsymbol{x}_i,\boldsymbol{x}_j) = tanh(v(\boldsymbol{x}_i\boldsymbol{\cdot}\boldsymbol{x}_j)+\theta).$$

A Gaussian kernel function is used in this paper.

B. the Fault Recognizer Based on SVDD

For expressing easily, two terms are defined. The *known faults* mean that the types of faults have been known. The *new faults* mean that the types of faults are unknown and there are not training data for these types of faults. For example, the types of known faults are A, B and C. One datum is one of the known fault data if it belongs to A or B or C, otherwise it is one of the new fault data.

When support vector machines (SVMs) are used in fault diagnosis, a classifier is obtained by training known fault data. The fault data in field test will be classified as one of those types by the classifier of SVMs. However, when the new fault types appear, they will be classified as those known fault types wrongly.

A fault recognizer is constructed based on SVDD here. The function of the fault recognizer is to sift the fault data before the classification of them.

The fault recognizer can pick out the new fault data which do not belong to known fault types in case these new fault data enter the classifier to be classified.

When SVDD is used to construct the fault recognizer, the boundary of the minimum sphere is used to sift data. The data points inside the sphere or on the surface of the sphere belong to the known faults, while the data outside it belong to new faults, so they should be sifted out. That is, if R^2-$r_x^2 \ge 0$, $\boldsymbol{x}$ is inside the sphere or on the surface of the sphere and should be accepted as a known fault datum. On the contrary, if R^2-$r_x^2 < 0$, $\boldsymbol{x}$ is outside the sphere and should be rejected as a new fault datum.

III. ISVDD AND THE FAULT RECOGNIZER BASED ON IT

A. ISVDD

The minimum sphere of SVDD is formulated by the mathematical theory. But from the viewpoint of pattern recognition, the complexity of classification problem lies in that there is not pure objective criterion [10]. A classification problem is not only a pure mathematical problem but also a biological problem. When a man judges fault data, he considers that the data near but outside the minimum sphere belong to the known fault data. Therefore a better sphere used here should be a sphere whose radius is longer than minimum sphere's and it is concentric with the minimum sphere. It is called an optimal sphere. The improved algorithm of support vector domain description (ISVDD) is used to construct the optimal sphere to sift out new fault data which lead to make incorrect decision.

According to the discussion above, the radius R_{os} of the optimal sphere is obtained from

$$R_{os} = \sqrt{R^2 + \Delta R^2}. \tag{5}$$

where ΔR is defined as the increment parameter and can be determined by experiments.

B. the Fault Recognizer Based on ISVDD

When ISVDD is used to construct the fault recognizer, the boundary of the optimal sphere is used to sift data. The data points inside the sphere or on the surface of the sphere belong to the known fault data, while the data outside it belong to new unknown fault data, so they should be sift out.

To judge whether a datum point $\boldsymbol{x}$ is a new fault data points, the distance r_x to the center of the optimal sphere have to be calculated. It can also be obtained from (4).

If $R_{os}^2 - r_x^2 \geq 0$, $\boldsymbol{x}$ is a known fault datum and can be classified by classifier later.

If $R_{os}^2 - r_x^2 < 0$, $\boldsymbol{x}$ should be picked out in order to avoid classification wrongly.

Suppose the input space is two dimensional. Figure 1 shows the difference between the recognizer based of SVDD and the recognizer based on ISVDD. The known fault data points and their sphere in feature space are mapped back on the input space for a Gaussian kernel. The dark points and the dark crosses indicate the known fault data points and the dark line indicates the surface of the sphere. Figure 1a is for the minimum sphere of SVDD and Figure 1b is for the optimal sphere of ISVDD. The dark circles indicate new fault data points in Figure 1.

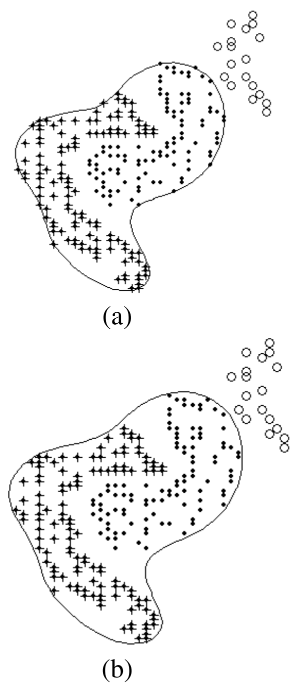

Figure 1. Comparison of two recognizers mapped back in the 2-dimensional input space. (a) for SVDD; (b) for ISVDD.

It is easy to find that both recognizers can pick out the new fault successfully. However, some known fault points outside the boundary are picker out by the minimum sphere in Figure 1a. Comparatively, the closed region in Figure 1b is larger than that in Figure 1a and the known fault points outside the boundary in Figure 1a is inside the closed curve in Figure 1b. Therefore, the known fault points could not be picked out when they are near but outside the boundary of the minimum sphere.

IV. HIERARCHY STRUCTURE SVMS FOR MULTI-CLASS CLASSIFICATION

SVMs were originally developed for two-class classification.

Suppose the training data $(\boldsymbol{x}_i, y_j)$, $\boldsymbol{x}_i \in R^n$ $y_i \in \{+1,-1\}$, $i = 1,2,\ldots,k$ can be separated by an optimal hyperplane with smallest generalization error and the distance between the closest data point and the hyperplane is maximal (here y has two values that stand for two classes). The following optimization problem is used to find the hyperplane:

$$\begin{aligned} &\min \quad \frac{1}{2}(\boldsymbol{w}\bullet\boldsymbol{w}) + C\left(\sum_{i=1}^{k}\xi_i\right) \\ &s.t. \quad \begin{cases} y_i\left[\left(\boldsymbol{w}\bullet\boldsymbol{x}_i\right)+b\right] \geq 1-\xi_i \\ \xi_i \geq 0, \quad i=1,\cdots,k. \end{cases} \end{aligned} \tag{6}$$

where the ξ_i are slack variables and the variable $C > 0$ gives the trade-off between maximal distance and the number of errors.

The Lagrange multiplier method can transform (6) into a simpler dual problem:

$$\begin{aligned} &\max \quad \sum_{i=1}^{k}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{k}\alpha_i\alpha_j y_i y_j (\boldsymbol{x}_i \bullet \boldsymbol{x}_j) \\ &s.t. \quad \begin{cases} \sum_{i=1}^{k} y_i\alpha_i = 0 \\ 0 \leq \alpha_i \leq C \\ i = 1,\cdots,k. \end{cases} \end{aligned} \tag{7}$$

where α_i are Lagrange multipliers. The optimal hyperplane can be obtained by solving this quadratic programming problem. The support vectors of the hyperplane are the data points whose α_i are non-zero.

The points in the original input space are mapped into a feature space by a kernel function when this kernel function $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ satisfies Mercer's theorem. In the feature space, the quadratic programming problem (7) for hyperplane can be expressed as

$$\begin{aligned} &\max \quad \sum_{i=1}^{k}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{k}\alpha_i\alpha_j y_i y_j K(\boldsymbol{x}_i, \boldsymbol{x}_j) \\ &s.t. \quad \begin{cases} \sum_{i=1}^{k} y_i\alpha_i = 0 \\ 0 \leq \alpha_i \leq C \\ i = 1,\cdots k. \end{cases} \end{aligned} \tag{8}$$

The decision function for two-class classification problem is described as

$$f(\boldsymbol{x}) = \operatorname{sgn}\Big(\sum_{i \in sv} y_i\alpha_i K(\boldsymbol{x}_i, \boldsymbol{x}) + b\Big). \tag{9}$$

where sv is the set of support vectors.

Now, most algorithms of the multi-class classification SVMs are based on the two-class classification SVMs above. One of them is the hierarchy structure SVMs.

This algorithm decomposes the multi-class classification problem into a series of binary classification sub-problems. One of the hierarchy schemes is the binary tree with a single root node at the top of the graph and with terminal nodes (leaves) at the bottom. Individual classes are represented in the leaves, and the other nodes within the graph are sub-classifiers performing a binary decision task (see Figure. 2). This scheme is applied to SVMs in this paper. Hierarchy structure SVMs are fast and the number of binary sub-classifiers constructed is small.

The fault diagnosis of rotating machine is a multi-fault problem and should be solved by the multi-class classification algorithm. We apply hierarchy structure SVMs for the fault diagnosis in this paper.

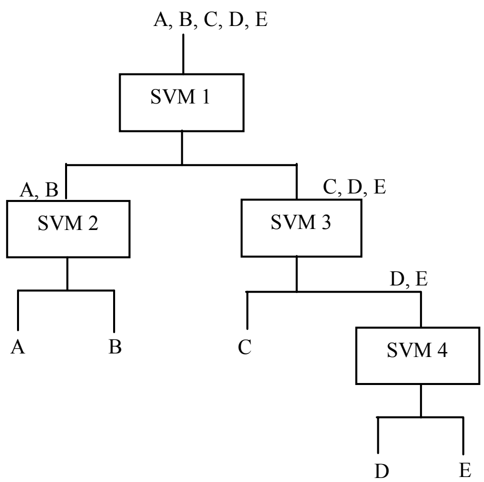

Figure 2. E Hierarchy structure SVMs

V. THE NEW METHOD OF FAULT DIAGNOSIS AND NUMERICAL EXPERIMENTS

A. A new method of Fault Diagnosis

In this section, a new method of fault diagnosis for rotating machinery based on ISVDD and hierarchy structure SVMs is firstly given. The procedure is described as follows:

Step 1 Acquire known fault data for rotating machinery as training sample data. Usually, the fault sample data is obtained from mechanical vibration signals.

Step 2 Train the recognizer used to pick out the new fault data based on ISVDD.

Step 3 Train the classifier used to classify known fault data based on hierarchy structure SVMs.

Step 4 Sift fault data by the trained recognizer. If the data are known fault data, they can enter the next step.

Step 5 Classify the known fault data sifted by the Step 4.

Step 6 Finish fault diagnosis for rotating machinery.

B. Numerical Experiments

Numerical experiments have been carried out to verify the performance of the proposed algorithm and the method of fault diagnosis. There are two experiments.

The fault dataset of the numerical experiments are from [11]. There are five fault types which are rotor misaligned fault, rotor imbalance, surge, oil film whirl, oil film whip. The number of their attributes is 22. These attributes include 15 frequency features as well as speed, oil temperature, the volume of flow, pressure, load, axis trace, the direction of axis precession. There are 55 fault data and each fault type consists of 11 data.

For each experiment, the data are ordered at random first and then 3-cross validation is used.

1) Experiment 1

The purpose of this experiment is to verify the effectiveness of the recognizer based on ISVDD for fault diagnosis of rotating machinery.

The fault data are divided into five experimental groups. The data of one fault type are taken as the new fault data and the data of other four types as known fault data in each group. The new fault in group 1 is rotor misaligned fault. The new fault in group 2 is rotor imbalance. The new fault in group 3 is surge. The new fault in group 4 is oil film whirl. The new fault in group 5 is oil film whip. The recognizer is trained by the known fault data in each group. The data of total five fault types are used as test data to check out the performance of the recognizer.

The contrast experiments are carried out between recognizers based of SVDD and ISVDD in this part. The final result is the average value of three experiments in Table I.

TABLE I. CONTRAST RESULTS OF CORRECT RECOGNITION RATE

	Recognizer based on SVDD	***Recognizer based on ISVDD***
Group 1	92%	100%
Group 2	92%	100%
Group 3	87%	100%
Group 4	92%	100%
Group 5	87%	100%
Average	90%	100%

The experimental results in Table I show that the correct recognition rate of the new fault data based on ISVDD is higher than that based on SVDD and the recognizer based on ISVDD is able to pick out the new fault data and prepare the eligible data for next step classification.

2) Experiment 2

The purpose of this experiment is to verify the method of fault diagnosis proposed in this part. The experimental data in this part have been pretreated by the recognizers based on SVDD and ISVDD in experiment 1. The classifier based hierarchy structure SVMs classifies the pretreated data.

The experimental results in Table II show that the correct classification rate of the method based on ISVDD and SVMs is higher than that based on SVDD and SVMs. The final classification rate relies on the performance of the recognizer greatly. In practice, the performance of the recognizer affects the result of fault diagnosis obviously.

TABLE II. CORRECT CLASSIFICATION RATE

	The method based on SVDD and SVMs	*The method based on ISVDD and SVMs*
Group 1	92%	100%
Group 2	92%	100%
Group 3	87%	100%
Group 4	92%	100%
Group 5	87%	100%
Average	90%	100%

VI. CONCLUSION

The essence of the fault diagnosis is pattern recognition. When SVMs are applied in fault diagnosis, error decisions would often be made if new unknown fault types appear in field test. In this paper, an improved algorithm of support vector domain description (ISVDD) is proposed to pretreat test data, thus enhancing fault diagnosis precision of SVMs. The recognizer based on ISVDD can effectively avoid classification errors by preventing some new fault data entering the classifier. A new method of fault diagnosis based on ISVDD and hierarchy structure SVMs is given in section V. The experiments are performed on a real fault dataset. The results verify that the performance of ISVDD is better and the correct classification rate of the new method of fault diagnosis based on ISVDD and hierarchy structure SVMs is higher.

REFERENCES

[1] B.C. Chun, X.H. Wu, Q. Qian and Q.K. Han, Research and Experiments on Nonlinear Dynamics of Rotating Machinery with Faults. Beijing, China: Science Press, 2004.

[2] Z.Y. Yang and W. Zheng, State Monitoring and Fault Diagnosis of Machines. Beijing, China: China Planning Press, 2006.

[3] V.N.Vapnik, “An overview of statistical learning theory,” IEEE Transactions on Neural Networks, vol.10, pp. 988–999. September 1999.

[4] S. Theodoridis, K. Koutroumbas, Pattern Recognition, 3rd ed.. Beijing, China: China Machine Press, 2006.

[5] B. Samanta , K.R. Al-Balushi, S.A. Al-Araimi, “Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection,” Engineering Applications of Artificial Intelligence,vol. 16, 657-665, 2003.

[6] S.B. Zheng, H.J. Tang, Z. Z. Han, et al., “Sensor Fault Diagnosis Method in ESP System with Support Vector Machines,” Journal of System Simulation, ,vol. 17, 682-717, 2005.

[7] S. F. Yuan, F. L. Chu, “Support vector machines-based fault diagnosis for turbo-pump rotor,” Mechanical Systems and Signal Processing, , vol. 20, 939-952, 2006.

[8] D.M.J. Tax, R.P.W. Duin, “Support vector domain description.” Pattern Recognition Letters, vol. 20: 1191-1199,1999.

[9] S. Friedhelm, P. Günther, “Tree-structured support vector machines for multi-class pattern recognition,” Lecture Notes in Computer Sciences, vol. 2096: 409-417,2001.

[10] Z.Q. Bian, X.G. Zhang, Pattern Recognition. Beijing, China: Tsinghua University Press, 2000.

[11] Z.G. Fei, The Application of Bayes Networks in Fault Diagnosis. Zhengzhou, China: Zhengzhou University, 2004.