Research of Fault Diagnosis of Rolling Bearing Based on Wavelet Packet and FSVM

Shan YuXiang¹
¹Information Center of China Tobacco Zhejiang Industrial Co., Ltd, No.77, Zhongshan South Road, Hangzhou 310008, China

Abstract: Aiming at reducing the classification accuracy caused by unbalanced sample data in bearing fault diagnosis, we present a bearing fault diagnosis method based on wavelet packet and fuzzy support vector machine (FSVM). In the actual sampling environment, the bearing fault sample is difficult to obtain, the data between the normal sample and the fault sample are not balanced, and the bearing is easily caused by the difference of the factors such as the personnel operation, the environment temperature, and the inherent noise of the machine. Therefore, the wavelet packet decomposition technique is used to extract the energy of the bearing vibration signal, and adaptive FSVM is used as a fault diagnosis algorithm to solve the unbalancing problem of training samples. The simulation results show that compared with the standard SVM and FSVM, the method can improve the classification accuracy of bearing fault diagnosis.

1. Introduction
Rolling bearing is one of the important components in mechanical equipment, especially rotating machinery. Whether it can work normally is related to the normal operation of mechanical equipment and even industrial production lines. Therefore, it is necessary to study the rolling bearing fault diagnosis methods. According to the different fault diagnosis technology mechanism, the common methods of rolling bearing fault are vibration analysis, noise analysis, temperature analysis, oil film resistance, acoustic emission diagnosis and so on [1]. The vibration analysis method uses the characteristics of the mechanical vibration signal, combined with the time frequency analysis and the machine learning algorithm to judge whether the machine has fault. As the vibration signal can characterize the abnormality and fault information of the equipment, it is the research hotspot of the bearing fault diagnosis technology at present [2] [3].

SVM [4] is a kind of machine learning algorithm developed on the basis of statistical theory. Compared with traditional machine learning algorithm, it has higher learning and generalization ability under the condition of limited sample. It is suitable to be used as a fault diagnosis classification algorithm for bearing. Considering the actual sampling conditions, the fault samples are difficult to obtain, and the number of normal samples and fault samples is unbalance. It is easy to cause SVM training less learning and less classification accuracy for fewer samples. In addition, due to the interference of personnel operation, ambient temperature and inherent noise of the machine, the sampling signal is noisy, and the observability and classification effect of the signal become worse. Therefore, the fuzzy technology is introduced on the basis of SVM to design a fuzzy factor based on the hyperplane distance for the sample.

By automatically adjusting the fuzzy factor size, the influence of noise and sample unbalance on SVM classification is reduced, and the accuracy of fault diagnosis algorithm is improved.
2. The Process of fault diagnosis

The rolling bearing fault diagnosis process is shown in figure 1, which can be divided into three stages: the feature extraction stage of vibration signal based on wavelet packet decomposition, FSVM training stage and FSVM recognition stage. Wavelet packet decomposition and fuzzy factor calculation is the key to fault diagnosis. The wavelet packet decomposition can transform the time domain signal of bearing vibration into the energy characteristics of different frequency domain. It can effectively extract the fault characteristics of the bearing and facilitate the learning and training of the FSVM. The FSVM training process is trained to get the classifier. Compared to the standard SVM, the appropriate weight is calculated for each sample according to the sample's contribution to the classifier in training, so as to improve the anti-interference of the classifier to the unbalanced sample and noise. Finally, using the classifier trained by FSVM, the test samples can be identified to judge whether the rolling bearing is faulty or not.

![Figure 1. Rolling bearing fault diagnosis flow chart](image)

3. Wavelet packet decomposition principle

The traditional method of vibration signal analysis is generally adopted by Fu Liye analysis. It is a fixed and invariant analysis method of window function, which cannot reflect the non-stationary, time-domain and frequency domain localization characteristics of the signal. Wavelet packet decomposition can provide finer decomposition for multiple frequency bands, especially high frequency signals, which can extract the details of the non-stationary signal of mechanical vibration signals. Therefore, three-layer wavelet packet decomposition is used to extract bearing fault signal characteristics. Using the gradually subdivided sampling steps for different frequency components in the signal, any local details of the signal can be observed. This function is very suitable for the analysis of vibration signals.

The sampling signal is decomposed by \( j \)-layer wavelet packet, and the wavelet packet coefficients of \( j \)-layer are extracted respectively from low frequency to high frequency. Finally, the energy of each band is obtained according to the wavelet packet coefficients. The following three-layer of wavelet packet decomposition will be described.

\[
E_{3,j} = \int |S_{3,j}(t)| dt = \sum_{k=1}^{n} |d_{j,k}|^2
\]

In the formula, \( d_{j,k} (j = 0, 1, \ldots, 7, k = 1, 2, \ldots, n) \) represents \( S_{3,j} \) wavelet packet coefficients.

4. Design of FSVM

4.1. Algorithm Ideas

The algorithm of FSVM[5] is to use fuzzy membership function to fuzzily deal with training samples, and assign different membership values to different samples, also called fuzzy factor [8]. For noise or outlier samples, the membership degree is usually given a small degree. Then, the fuzzy membership function is used to construct the quadratic optimization problem, and obtain the optimal classification hyperplane of the SVM.

In the FSVM, all the training samples need to be fuzzified, that is, each sample \((x_i, y_i)\) is assigned a fuzzy factor \( s_i \), and sets \( S = \{(x_1, y_1, s_1), (x_2, y_2, s_2), \ldots, (x_i, y_i, s_i)\}, 0 \leq s_i \leq 1 \) for a fuzzy sample.
Under non-linear conditions, the nonlinear mapping is introduced: $\Phi: \mathbb{R}^n \rightarrow H$. The two categories of classification problems after fuzzing can be solved by the following second optimization problem:

$$\begin{align*}
\min_{w, b, \xi} & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{l} s_i \xi_i \\
y_i(w \cdot \Phi(x_i) + b) & \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, \ldots, l
\end{align*}$$

(2)

In the formula, the coefficient $C$ is constant, $\xi_i$ is the relaxation variable, $w$ is the normal vector of the hyperplane, and $b$ is the displacement.

After calculation, the general form of the classification decision function established in the high-dimensional feature space is

$$f(x) = Sgn\left(\sum_{i=1}^{l} a_i^* y_i K(x_i, x) + b^*\right)$$

(3)

In the formula, $Sgn(\cdot)$ is symbolic function, $K(x_i, x_j)$ is kernel function, $x$ is support vector obtained by FSVM training. According to the positive and negative types of classification decision functions, the categories of samples $x_i$ can be identified.

### 4.2. The Design of Membership Function

The key of FSVM classification accuracy is the design of membership function. At present, most of the common membership functions are based on geometric function, sample center distance and sample compactness\[6-8\]. The method based on geometric function does not take into account the actual distribution of samples, which obviously has some limitations. The method of sample center distance is that the distance from the center of the class is larger, and the smaller the sample membership value is, the smaller the contribution to the classifier is. As shown in figure 2, $H$ is a classified hyperplane of two classes of samples, $x_1$ and $x_2$ is support vector. According to the center distance method, the sample $x_0$ membership value will be greater than the membership value of the support vector sample $x_1$. But according to the formula (3), it can be seen that the classifier is mainly constructed by the combination of support vectors, so the support vector in the training sample has the greatest contribution to the classifier, and the membership value of the corresponding sample should be given a larger value, and the membership value of the sample $x_1$ will be greater than the membership value of the support vector sample $x_0$. Therefore, the method of membership function determination based on the center distance cannot reflect the importance of samples correctly. Similarly, the method based on compactness is also true.

![Figure 2. Classified hyperplane sample distribution diagram](image)

For this reason, the distance between samples and classified hyperplanes is used as a measure of membership. First, the training samples are trained by SVM to get the hyperplane $f_H(x)$ of classification. Then, the sample $x_i$ to the hyperplane $d_i$ is calculated, where the maximum values of the positive and negative samples from the hyperplane are $d_{max}^+$ and $d_{max}^-$ respectively. The membership function of the sample is defined as:

$$s^\pm(x_i) = 1 - \frac{x_i}{d_{max}^\pm + \varepsilon}$$

(4)

In the formula, $\varepsilon$ is Minimal positive number.

In addition, in order to adjust the imbalance of the number of two types of samples, the membership
function of formula (4) needs to be reformed. When the number of samples is not balanced, fewer samples are less likely to learn, leading to lower classification accuracy. On the classified hyperplane, the hyperplane will be shifted to a few classes. In order to eliminate the classification surface offset, the penalty factor \( C \) is usually given to the positive and negative samples. The number of two samples is \( l^+ \) and \( l^- \) respectively, and the penalty factor can be set to \( C^+ = \delta \frac{l^-}{l^+}, C^- = \delta \frac{l^+}{l^-} \). Combined with formula (2), the membership function can be designed as:

\[
s^\pm(x_i) = C^\pm(1 - \frac{x_i}{d_{\max}^\pm + \epsilon})
\] (5)

5. Simulation

The simulation data are derived from the data of the rolling bearing fault simulation test-bed of Case Western Reserve University electrical engineering laboratory [9]. The four kinds of signal samples of inner ring, outer ring, ball fault and normal sample were selected as training test samples, and the energy features were obtained by wavelet packet decomposition. The DB4 wavelet in the Daubechies wavelet system is chosen as the wavelet basis function. According to the formula (1), the 8 dimensional energy feature vector \( E = \{E_1, E_2, \ldots, E_8\} \) can be obtained. In order to eliminate the dimension, the energy feature elements need to be normalized to the (0, 1) interval.

In the MATLAB simulation environment, the Libsvm-weights-3.17[10] toolbox, developed by Dr. Lin Zhiren from National Taiwan University, was used to write training and testing classifier algorithms. The toolbox supports standard multi-classification FSVM. Kernel function uses radial basis Gaussian kernel function \( \exp(-g \|u-v\|^2) \), selecting the kernel parameter \( g = 0.6 \) and the error penalty factor \( C = 8 \) by using the grid search method and the cross validation method [11]. The experimental results are shown in table 1.

| Table 1. Comparison of fault diagnosis accuracy of rolling bearing | SVM | center distance based on class FSVM |
|---------------------------------------------------------------|-----|-----------------------------------|
| algorithm in the paper                                       | inner ring fault | 82.4% | 83.3% |
|                                                              | outer ring fault | 80.5% | 82.7% | 84.1% |
|                                                              | ball fault        | 79.2% | 79.8% | 82.6% |
|                                                              | normal sample     | 86.1% | 87.5% | 84.2% |

From the diagnosis results of three fault samples in Table 1, it can be seen that compared with the standard SVM and center distance based on class FSVM, the algorithm in the paper improves the diagnostic accuracy of the fault samples, and shows that the algorithm is feasible. For normal samples, compared with the other two algorithms, the recognition accuracy of normal samples is reduced. It shows that by introducing different penalty factors for the unbalanced sample class, the precision of the sample can be improved effectively, and the recognition accuracy of the fault samples and the normal samples tends to balance.

6. Conclusion

A method of combining wavelet packet decomposition and FSVM is designed to diagnose the fault of rolling bearings. The method of determining the membership function of FSVM is based on the distance from the hyperplane. This method can effectively give different subordinate degrees to the samples with different degree of contribution to the hyperplane, and introduce the penalty factors in the membership function. The penalty factor improves the interference immunity of the classifier to noise and unbalanced samples. The experimental results show that, although the recognition accuracy of normal samples is reduced, this method improves the diagnostic accuracy of three typical faults of rolling bearings, and has certain practical value.
References
[1] Ten LiLi, Tang Tao Wang MingFeng 2011 Overview and development trend of rolling bearing fault diagnosis technology based on Vibration Analysis. Science & Technology Information, Vol.23, pp 123-126
[2] Yang LiuSong 2013 Research on Rolling Bearing Fault Diagnosis based on Wavelet Analysis and Neutral Networks(ShenYang:Northeast Forestry University)
[3] Chen Chao, Shen Fei 2017 Yan RuQiang. Enhanced leastsquares support Vector machine-based transfer learning strategy for bearing fault diagnosis. Chinese Journal of Scientific Instrument, Vol.38, No.1, pp 33-39
[4] Vapnik V 2000 The Nature of Statistical Learning Theory (New York: Springer-Verlag)
[5] C.F.Lin, S.D.Wang 2002 Fuzzy support vector machines. IEEE Transactions On Neural Network, 13(3), pp 466-471
[6] Yang ZhenZhen 2011 Research on FSVM theory and algorithm based on information geometry(Nanjing: Nanjing University Of Posts And Telecommunications)
[7] Xu CuiYun, Ye Ning 2014 A novel fuzzy support vector machine based on the class centripetal degree. Computer Engineering & Science, Vol.36, No.8, pp 1623-28
[8] Zhang Xiang, Zi Xiao Lin,Xu GuangGu 2006 Fuzzy Support Vector Machine Based on Affinity Among Samples. Journal of Software , 17(5), pp 951-958
[9] Case Western Reserve University 2018 Bearing Data, http://www.eec$_.cwru.edu/laboratory/bearing
[10] Chih-Chung Chang, Chih-Jen Lin 2011 LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2(3), pp 1-27.
[11] Shi Feng, Wang Xiaochun, Yu Lei, etc 2010 Analysis of 30 cases of MATLAB neural network(Beijing: Beihang University Press)