



An Algorithm of Intelligent Classification For Rotating Mechanical Failure Based on Optimized Support Vector Machine

Yun-sheng Chen^(✉)

Guangzhou Huali Science and Technology Vocational College,
Guangzhou 511325, China
pofjha@sina.com

Abstract. The classification algorithm of rotating machinery fault cannot effectively recognize the false components and true components in fault signal of rotating machinery. Therefore, an intelligent classification algorithm of rotating machinery fault based on optimized support vector machine was put forward. The K-L divergence was used to measure the nonlinear and symmetry of probability distribution of two processes in rotating machinery, and the error of information in the process of rotating machinery was measured to eliminate the false component of fault signal of rotating machinery. Meanwhile, the multi-value classification support vector machine algorithm based on decision directed acyclic graph was used to process the signal that only had a true component. Moreover, the value of each node in support vector machine decision function was calculated. Finally, based on calculation results, the fault categories were excluded. Thus, the intelligent classification of rotating machinery fault was completed. According to experimental results, the proposed algorithm can accurately eliminate false components in the rotating machinery fault signal. Meanwhile, the classification result is accurate.

Keywords: Support vector machine · Rotating machinery · Intelligent classification · Fault signal

1 Introduction

With the continuous growth of importance on innovation technology and the rapid development of modern science and technology, the scale of modern production is expanding increasingly and the degree of automation of industrial equipment is increasing. The structure of industrial equipment is more and more complex in the meantime [1]. Based on this background, the basis of modern industrial production is the safe operation of equipment. In the operation, a minor issue will cause casualties, economic losses, and baneful influence to the society and other consequences [2]. The improvement of the degree of automation and the integration level of equipment in modern industry makes the safe operation become the main effort way. With the large scale of modern industrial system and the complexity of industrial equipment, the problem of safety, reliability, maintainability and usability of industrial equipment and machinery has gradually emerged, which promotes the research on diagnosis technology

of rotating machinery fault [3]. To classify the fault of rotating machinery is an important step in the intelligent diagnosis technology of rotating machinery fault. Therefore, it is necessary to analyze and study the fault classification algorithm of rotating machinery in depth [4].

Yang et al. have proposed a classification algorithm of rotating machinery fault based on FSVM. When classifying rotating machinery fault feature, this algorithm extracts characteristic index of rotating machinery fault signal and solves the fuzzy membership degree of sample. The classification of rotating machinery fault is completed through fuzzy support vector machine model. This algorithm cannot remove false component in fault signal of rotating machinery [5]. Wang et al. propose a kind of classification algorithm of rotating machinery fault based on KSLPP and RWKNN. This algorithm extracts information in rotating machinery fault feature set through KSSLPP method and uses category information to maximize the degree of separation between fault feature categories in the process of dimension reduction. Then, the feature set is identified by RWKNN to complete the classification of rotating machinery fault. The classification results obtained by the proposed algorithm has error [6]. Chen et al. propose a fault classification algorithm based on point-to-point principal component analysis. This algorithm uses amplitude-frequency characteristic technology to extract feature of rotating machinery fault. Then, the algorithm uses the point-to-point principal component analysis to construct three-layer neural network and classify the fault of rotating machinery. Thus, this algorithm cannot accurately classify the category of rotating mechanical fault [7]. Wang et al. propose a classification algorithm of rotating mechanical fault based on RSCNMF. This algorithm improves the decomposition algorithm of coefficient constraint non-negative matrix and classifies the rotating machinery fault through the fault model of robustness sparsity constraint. The proposed algorithm cannot accurately identify the real and false components existing in fault signal of rotating machinery and cannot accurately eliminate false components [8].

In conclusion, an algorithm of intelligent classification of rotating mechanical fault based on optimized support vector machine is proposed, which solves the problem existing in current algorithm. The specific steps are as follows:

- (1) K-L divergence is used to eliminate the false components in the fault signal of rotating machinery and improve the classification accuracy.
- (2) The multi-value classification support vector machine algorithm based on decision directed acyclic graph is used to complete the intelligent classification of rotating mechanical fault.
- (3) Experimental results and analysis: the effectiveness of proposed algorithm is verified through the elimination of false components and the accuracy of classification results.
- (4) Conclusion: summarize the full text and put forward the future direction of development.

2 Elimination of False Components

To eliminate the false components in the fault signal of rotating machinery can improve the accuracy of classification of rotating machinery fault [9]. The K-L divergence is used to measure the nonlinear and symmetry of probability distribution of two processes in rotating machinery. That is to say, similarity and information of two process are measured. Meanwhile, the threshold value is set to judge false components existing in the mechanical fault information, so as to eliminate the false ingredient in fault signal of rotating machinery.

Supposing that the probability density functions of the process $P = (p_1, p_2, \dots, p_n)$ and the process $Q = (q_1, q_2, \dots, q_n)$ are represented by $p(x)$ and $q(x)$, $\delta(P, Q)$ denotes the K-L distance from process P to process Q . The formula of $\delta(P, Q)$ is as follows:

$$\delta(P, Q) = \sum_{x \in n} p(x) \log \frac{p(x)}{q(x)} \quad (1)$$

According to the formula (1), the K-L distance $\delta(Q, P)$ from process Q to process P is obtained, and the formula of $\delta(Q, P)$ is as follows:

$$\delta(Q, P) = \sum_{x \in n} q(x) \log \frac{q(x)}{p(x)} \quad (2)$$

The K-L divergence between P and Q through formula (1) and formula (2) is obtained.

$$D = (P, Q) = \delta(P, Q) + \delta(Q, P) \quad (3)$$

The nonparametric estimation is used to calculate probability density, and the formula is as follows:

$$p(x) = \frac{1}{nh} \sum_{i=1}^n k \left[\frac{x_i - x}{h} \right] \quad (4)$$

The formula (4) is the density estimation of density function $p(x)$. x_i and x are the parameters in the density function $p(x)$. In the formula, $k(\cdot)$ denotes the kernel function, and h is known as the smooth parameter or the window width. Usually, the Gauss function is used to express the kernel function $k(\cdot)$. The formula is:

$$k(u) = \frac{1}{\sqrt{2\pi}} e^{-2} \quad (5)$$

In formula (5), $k(u)$ is Gauss function. the information difference in two processes of rotating machinery can be measured by K-L divergence. That is to say, the similarity between two processes is measured [10]. The bigger the K-L divergence between two processes, the greater the difference between two processes, and the smaller the similarity. The smaller the K-L divergence between two processes, the smaller the

difference between two processes, and the greater the similarity. For the fault signal $x(t)$ of rotating machinery, the specific algorithm is:

The signal $x(t)$ is decomposed to get n components $c_i(t)$ of the signal $x(t)$. Where, $i = 1, 2, \dots, n$. The probability density function $p_x(t)$ and $p_i(t)$ are calculated by the method of nonparametric estimation [11]. Through formula (1), formula (2) and formula (3), the K-L divergence value $D_i(x, c)$ between $x(t)$ and $c_i(t)$ is calculate. Then, $D_i(x, c)$ is normalized to get K-L divergence value λ_i . The formula of λ_i is:

$$\lambda_i = \sqrt{\frac{D_i^2}{\sum_{i=1}^n D_i^2}} \tag{6}$$

The value of λ_i is $0 < \lambda_i < 1$. The smaller the value of λ_i is, the stronger the correlation between the original signal and component is. The bigger the value of λ_i is, the weaker the correlation between the original signal and component is, which may be a false component [12]. When judging false components, we set threshold value λ , and compare λ with λ_i . When $\lambda_i > \lambda$, it represents that component is the false components in the rotating machinery fault signal, and we need to remove it. When $\lambda_i < \lambda$, this shows that component is the true constituent of signal [13].

3 Support Vector Machine Classification Algorithm Based on Decision Directed Acyclic Graph

The rotating machinery fault after eliminating false components was classified. Then a k class of classification problem of rotating machinery fault was set. The training samples are $(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots, (x_n, y_n)$. In training samples, $i = 1, 2, \dots, n$, $x_i \in R^n, y_i \in [1, 2, \dots, k]$. The support vector machine classifier based on decision directed acyclic graph is equivalent to a triple $\langle F, SVM, ST \rangle$. Where, F is the leaf node of directed acyclic graph, and it is a set including k fault types of normal state. SVM denotes is the set of internal nodes of directed acyclic graph algorithm composed of $k(k - 1)/2$ support vector machines. The formula of SVM is:

$$SVM = (SVM_{12}, SVM_{13}, \dots, SVM_{pq}, \dots) \tag{7}$$

In the formula, SVM_{pq} denotes a sub-classifier, which is used to differentiate the p -th category and q -th category, $p, q \in [1, k], p < q$. ST denotes the training set, which includes k categories of rotating machinery failure mode. The formula of ST is:

$$ST = (ST_1, ST_2, \dots, ST_i, \dots, ST_k) \tag{8}$$

In the formula, ST_i denotes is the i -th type of composition of sample. All the training samples are consist of $x_i, y_i, \sum n_i = n$.

The classifier SVM_{pq} is used to classify training sample S_{pq} by formula (9):

$$\begin{cases} S_{1k} = \text{All the training samples} \\ S_{pq} = \{ST_p, ST_{p+1}, \dots, ST_q\}, p, q \in [1, k], p < q \end{cases} \quad (9)$$

According to formula (9), the number of training samples gradually decreases with the continuous increase of the number of layers during training.

The support vector machine classifier of directed acyclic graph corresponds to a table operation. All fault categories of rotating machinery are included in the table. After operating a node, we delete a class of rotating mechanical faults in the table. All of possible rotating machinery categories are arranged in a certain order under the initial state [14]. The first category and last category in the table are compared when they are classified. Then, the most impossible category is deleted to reduce the number of categories in the table. That is to say, for the inputting sample of rotating machinery fault category, we start from the root node, then calculate the value of each node by support vector machine decision function based on decision directed acyclic graph [15, 16]. When the value is one more than the original value, we enter the next node from the left. When the value is one less than the original value, we enter the next node from the right. And so on, after the $k - 1$ -th calculation, the only category remaining in the table is the category of rotating machinery fault. Thus, the intelligent classification of rotating machinery fault is completed.

4 Experiment and Analysis

In order to prove the overall effectiveness of intelligent classification algorithm of rotating machinery fault based on optimized support vector machine, the intelligent classification algorithm of rotating machinery fault based on optimized support vector machine is tested. The operation system in this test is Windows 7.0. The instantaneous frequency of Chirp signal in different time is different. It is a typical non-stationary signal. The Chirp signal is decomposed into 6 components. As shown in Fig. 1, only the component 2 is the true component.

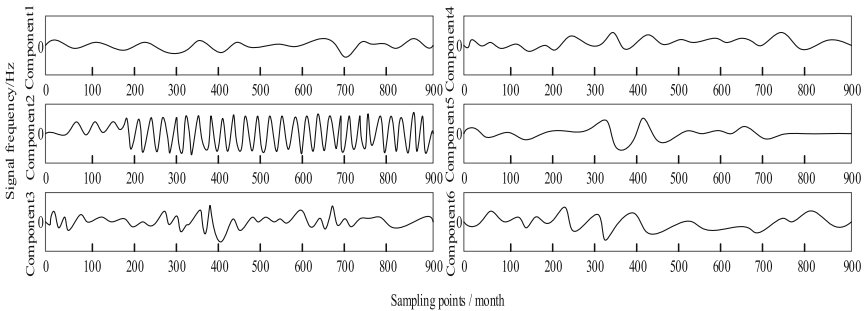


Fig. 1. Components of chirp signal

Many features can be extracted from vibration signals, such as various dimensionless indexes in time domain and amplitude spectrum, power spectrum and Cepstrum in frequency domain. The total amount of different frequency peaks in several frequency bands of vibration signal spectrum is the most commonly used characteristic quantity in rotating machinery fault diagnosis, which is widely used in rotating machinery system fault diagnosis. Extract the total amount of spectrum peaks of different frequencies in 9 frequency bands of vibration signal spectrum, including $0.01-0.39f$, $0.4-0.49f$, $0.5f$, $0.51-0.99f$, $1f$, $2f$, $3-5f$, odd multiple F , high-power F , a group of 9 features, forming a multi-dimensional feature vector. The specific division of frequency band is shown in Table 1, where f represents the working frequency of rotor.

Table 1. Frequency band division

Frequency band	1	2	3	4	5
Frequency range	$0.01-0.39 f$	$0.4-.049 f$	$0.5 f$	$0.51-0.99 f$	$1 f$
Frequency band	6	7	8	9	
Frequency range	$2 f$	$3-5 f$	Odd multiple f	High fold f	

In the experiment, 80 groups of vibration signals of normal, rotor unbalance, rotor misalignment, rotor bending, oil whirl and oil film oscillation were collected respectively, and the data were processed and normalized as training data set.

The intelligent classification algorithm of rotating machinery fault based on optimized support vector machine (algorithm 1) and the classification algorithm of rotating machinery fault based on FSVM (algorithm 2) are used to test the false components in Chirp signal through K-L divergence value. The test results are shown in Table 2.

Table 2. Test results of two methods

K-L divergence value and threshold λ	Component1	Component2	Component3	Component4	Component5	Component6
Algorithm 1	$>\lambda$	$<\lambda$	$>\lambda$	$>\lambda$	$>\lambda$	$>\lambda$
Algorithm 2	$<\lambda$	$<\lambda$	$<\lambda$	$>\lambda$	$>\lambda$	$>\lambda$

From the analysis of the data in Table 2, we can see that using the intelligent classification algorithm of rotating machinery fault based on optimized support vector machine to test, only the K-L divergence of component 2 is less than the given threshold λ . However, by using the classification algorithm of rotating machinery fault based on FSVM, the K-L divergence values of component 1, component 2 and component 3 are less than the given threshold λ . It is known that when K-L divergence value is greater than the threshold value λ , the component is a false component, and when K-L divergence value is less than the threshold value λ , the component is a true component. The false components obtained by the intelligent classification algorithm of rotating machinery fault based on optimized support vector machine are consistent with

the fact. To verify the intelligent classification algorithm of rotating machinery fault based on optimized support vector machine can remove the false components in fault signal of rotating machinery.

The intelligent classification algorithm of rotating machinery fault based on optimized support vector machine (algorithm 1) and the classification algorithm of rotating machinery fault based on FSVM (algorithm 2) are used to test. Then, classification results of two methods are compared. The test result is shown in Fig. 2. In the diagram, different shapes represent different types of rotating machinery faults.

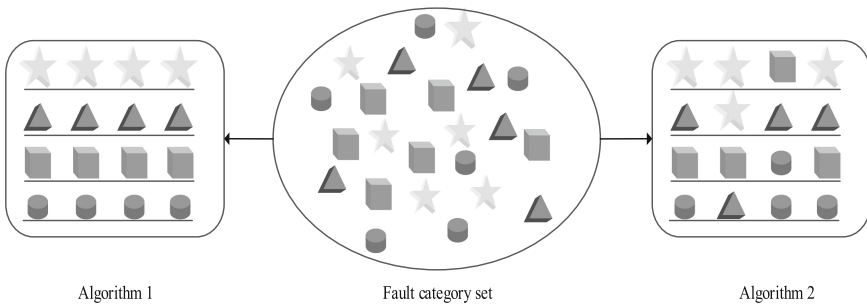


Fig. 2. Classification results of two algorithms

It can be seen from the analysis in Fig. 2 that different fault types can be distinguished when the rotating machinery fault intelligent classification algorithm based on optimal support vector machine is used for centralized fault classification of rotating machinery fault. So that the classification results are more accurate. The intelligent fault classification algorithm of rotating machinery based on FSVM can not accurately classify the fault types when classifying the concentrated fault types of rotating machinery. Therefore, the classification results are not accurate.

5 Conclusions

To accurately classify rotating machinery fault is the basis of the safe operation of rotating machinery. Using current algorithm to classify the fault categories of rotating machinery has a problem about inaccurate identification of false components and inaccuracy of classification results. An intelligent classification algorithm for rotating machinery fault based on optimal support vector machine is proposed. The first step is to remove false components in rotating machinery fault signal according to K-L divergence. The second step is to use support vector machine classification algorithm based on decision directed acyclic graph for the intelligent classification of rotating machinery fault.

To classify the faults of rotating machinery based on the support vector machine is a new technology, which has attracted the extensive attention of experts and scholars.

The support vector machine has many advantages and some disadvantages. Therefore, it is necessary to research on the following two points in future work.

- (1) The selection problem of the kernel function parameters and the kernel function.
- (2) Research on the learning method of support vector machine.

References

1. Shom, P.D., Sudarsan, P.: A novel hybrid model using teaching–learning-based optimization and a support vector machine for commodity futures index forecasting. *Int. J. Mach. Learn. Cybernet.* **9**(1), 97–111 (2018)
2. Zhou, Y.H., Zhang, H.L., Li, F.F., et al.: Local focus support vector machine algorithm. *J. Comput. Appl.* **38**(4), 945–948 (2018)
3. Björn, J.P., van der Ster Frank, C.B., Tammo, D., et al.: Support vector machine based monitoring of cardio-cerebrovascular reserve during simulated hemorrhage]. *Front. Physiol.* **8**(8), 1057–106 (2018)
4. Lyndia, C.W., Calvin, K., Jesus, L., et al.: Detection of american football head impacts using biomechanical features and support vector machine classification. *Sci. Reports* **8**(1), 15–21 (2018)
5. Liu, S., Li, Z., Zhang, Y., et al.: Introduction of key problems in long-distance learning and training. *Mobile Networks Appl.* **24**(1), 1–4 (2019)
6. Wang, X.D., Zhao, R.Z., Deng, L.F.: Rotating machinery fault diagnosis based on KSLPP and RWKNN. *J. Vibration Shock* **35**(8), 219–223 (2016)
7. Shuai, L., Weiling, B., Nianyin, Z., et al.: A fast fractal based compression for MRI images. *IEEE Access* **7**, 62412–62420 (2019)
8. Wang, S.L., Niu, Y.G., Han, P., et al.: RSCNMF algorithm and its application in fault detection of industrial process. *Comput. Simul.* **34**(8), 386–390 (2017)
9. Zhao, M., Lin, J.: Health assessment of rotating machinery using a rotary encoder. *IEEE Trans. Ind. Electron.* **12**(99), 1–7 (2017)
10. Sasaki, H., Igarashi, H.: Topology optimization using basis functions for improvement of rotating machine performances. *IEEE Trans. Magn.* **14**(99), 1–4 (2017)
11. Patel, D., Chothani, N.G., Mistry, K.D., et al.: Design and development of fault classification algorithm based on relevance vector machine for power transformer. *IET Electric Power Appl.* **12**(4), 557–565 (2018)
12. Fu, W., Liu, S., Srivastava, G.: Optimization of big data scheduling in social networks. *Entropy* **21**(9), 902 (2019)
13. Han, T., Liu, C., Yang, W., et al.: A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults. **165**(1), 474–487 (2019)
14. Ertuncay, D., Costa, G.: An alternative pulse classification algorithm based on multiple wavelet analysis. *J. Seismol.* **23**(4), 929–942 (2019)
15. Brankovic, A., Falsone, A., Prandini, M., et al.: A feature selection and classification algorithm based on randomized extraction of model populations. *IEEE Trans. Cybernet.* **48**(4), 1151–1162 (2018)
16. Zheng, P., Shuai, L., Arun, S., Khan, M.: Visual attention feature (VAF): A novel strategy for visual tracking based on cloud platform in intelligent surveillance systems. *J. Parallel Distr. Comput.* **120**, 182–194 (2018)