Fault diagnosis of direct-drive wind turbine based on support vector machine

X L An, D X Jiang, S H Li and J Chen

State Key Laboratory of Control and Simulation of Power System and Generation Equipment, Department of Thermal Engineering, Tsinghua University, Beijing, 100084, China

E-mail: anxl614@163.com, jiangdx@tsinghua.edu.cn

Abstract. A fault diagnosis method of direct-drive wind turbine based on support vector machine (SVM) and feature selection is presented. The time-domain feature parameters of main shaft vibration signal in the horizontal and vertical directions are considered in the method. Firstly, in laboratory scale five experiments of direct-drive wind turbine with normal condition, wind wheel mass imbalance fault, wind wheel aerodynamic imbalance fault, yaw fault and blade airfoil change fault are carried out. The features of five experiments are analyzed. Secondly, the sensitive time-domain feature parameters in the horizontal and vertical directions of vibration signal in the five conditions are selected and used as feature samples. By training, the mapping relation between feature parameters and fault types are established in SVM model. Finally, the performance of the proposed method is verified through experimental data. The results show that the proposed method is effective in identifying the fault of wind turbine. It has good classification ability and robustness to diagnose the fault of direct-drive wind turbine.

1. Introduction

The wind energy is a clean, renewable new energy which can be large-scale developed and applied. Some faults may occur during wind turbine operation due to the more complex environment [1-5]. The failure number distribution for Swedish, Danish and German wind power plants is demonstrated in [1]. In order to ensure the security and stability of wind turbine in operation, it is necessary to install condition monitoring system (CMS) and fault detection system (FDS) in wind turbine [2]. Continuously monitor the condition of wind turbine systems is the most efficient way of reducing the costs of operational and maintenance costs [1].

Support vector machine (SVM) is a machine learning method, which is aimed at small sample, based on statistical learning theory and the principle of structural risk minimization. SVM has a simple mathematical form and a clear geometric interpretation. It has been successfully applied in pattern recognition, regression analysis, function approximation, signal processing and other fields [6-8]. In this paper, for a direct-drive wind turbine, five experiments with normal state, wind wheel mass imbalance fault, wind wheel aerodynamic imbalance fault, yaw fault and blade airfoil change fault are performed. The time-domain feature parameters are extracted from the horizontal vibration and the vertical vibration for the five experiments. Using an improved distance evaluation technique the sensitive features are selected from the original time-domain feature set. Finally, a based-SVM fault diagnosis model of direct-drive wind turbine is proposed.
2. Multi-class support vector machine algorithm

A standard support vector machine algorithm can transform a real problem to a convex quadratic programming problem. The main idea of using support vector machines to classify data is: firstly, through a nonlinear mapping, the data are mapped from the original space to a high dimensional feature space. Then, in high-dimensional feature space, the optimal separating hyperplane is sought to partition the points of training sample set. This makes the distance from the points of training sample set to the optimal hyperplane is far as possible, the specific algorithm can be seen in [6-8].

SVM is a binary classifier, however, wind turbine’s faults are more than two classes. Consequently, multiclass classifiers should be constructed in fault diagnosis. In this paper, the one-against-one method adopted to construct multiclass classifier [8]. The basic idea is: for the N-class classification problem, the N(N-1)/2 classifiers are constructed, each one from two classes can be separated by training. Given the data \( s = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\} \), where \( i = 1, \ldots, N \), and \( x_i \in R^m \) is the class, \( y_i \in \{+1, -1\} \) is the corresponding target. For the training samples from the \( k \)th and the \( l \)th classes, the binary classification can be transformed to multi-class classification [8]:

\[
\min \frac{1}{2} (w^{kl})^T (w^{kl}) + C \sum_{i=1}^{N} \xi_{kl}^i (w^{kl})^T \xi_{kl}^i \geq 0
\]  

If \( (w^{kl})^T \phi(x_i) + b^{kl} \geq 1 - \xi_{kl}^i \), then \( x_i \) belongs to \( k \)th class; if \( (w^{kl})^T \phi(x_i) + b^{kl} \leq -1 + \xi_{kl}^i \), then it belongs to \( l \)th class. The \( x_i \) is mapped to a higher-dimensional space by function \( \phi \) and \( C \). The \( C \) is penalty parameter. The N(N-1)/2 classifiers can be constructed to identify the test samples. The voting strategy is adopted to identify the most probable class. The detail explanation of this method can be seen in [8].

3. The fault features selection

Although the time domain features of vibration signal can identify the wind turbine faults to some extent, the features have different sensitivity for different objects and different faults. Therefore, the time domain features should be selected before they are input classifier. The selected features are closely related to the faults, and the classification performance can be improved. In this paper, an improved distance evaluation technique is used to select the sensitive time domain features. The detailed feature selection process is introduced in [9].

4. Direct-drive wind turbine fault diagnosis based on support vector machine

To verify the efficiency and reliability of the features selection and the SVM classifier, five experiments of direct-drive wind turbine are carried out. The five experiments including normal state (\( N \)), wind wheel mass imbalance fault (\( F_1 \)), wind wheel aerodynamic imbalance fault (\( F_2 \)), yaw fault (\( F_3 \)) and blade airfoil change fault (\( F_4 \)). The experimental setup of direct-drive wind turbine is displayed in figure 1. Based on these experiments, the fault diagnosis method of wind turbine is investigated. In experiments, the wind wheel mass imbalance fault is realized by affixing a light cardboard to a blade. The wind wheel aerodynamic imbalance fault is carried out by affixing a light cardboard to a blade. From blade tip to the root, 30% of the blade is covered to change the blade airfoil shape. The yaw fault is conducted by deviating the direction of the wind axle line to flow direction with 20°. The blade airfoil change fault is achieved through pasting light paper on three blades surface.

The normal state, four faults and their corresponding different characteristics of the direct-drive wind turbine are extracted as the samples of fault diagnosis. The time-domain feature parameters of vibration signal in horizontal and vertical direction are chosen to found the SVM fault diagnosis model. In this model, the peak-to-peak value (\( C_1 \)), peak value (\( C_2 \)), mean value (\( C_3 \)), average amplitude (\( C_4 \)), root mean square (\( C_5 \)), root mean square amplitude (\( C_6 \)), shape factor (\( C_7 \)), crest factor (\( C_8 \)), impulse factor (\( C_9 \)), clearance factor (\( C_{10} \)) and kurtosis (\( C_{11} \)) of main shaft vibration signal in the horizontal direction; the peak-to-peak value (\( C_{12} \)), peak value (\( C_{13} \)), mean value (\( C_{14} \)), average amplitude (\( C_{15} \)), root mean square (\( C_{16} \)), root mean square amplitude (\( C_{17} \)), shape factor
(C_{18}), crest factor (C_{19}), impulse factor (C_{20}), clearance factor (C_{21}) and kurtosis (C_{22}) in the vertical direction are preliminarily determined as the fault identification information. These feature parameters [10] are defined in table 1. For the $C_1$-$C_{22}$, the sensitive feature parameters are selected based on the improved distance evaluation technique [9]. The selected feature parameters are input the fault classification system to train and diagnose the faults. The rotational speed is an important parameter which has great influence on the vibration. In the same state, the vibration is different at different rotational speeds. Therefore, the vibration data of each state are analyzed and compared at the same speed. In this paper, the rotational speed is 270rpm.

(a) Experimental setup schematic diagram.

(b) Experimental setup picture.

Figure 1. Experimental setup of direct-drive wind turbine.
### Table 1. The time-domain feature parameters of wind turbine.

| Feature                   | Equation | Feature                   | Equation |
|---------------------------|----------|---------------------------|----------|
| Peak-to-peak value \((t_{p-p})\) | \(t_{p-p} = \max(t_i) - \min(t_i)\) | Shape factor \((S)\) | \(S = \frac{t_{rms}}{t}\) |
| Peak value \((t_p)\)       | \(t_p = \max(\left|t_i\right|)\) | Crest factor \((C)\) | \(C = \frac{t_p}{t_{rms}}\) |
| Mean value \((\bar{t})\)   | \(\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_i\) | Impulse factor \((I)\) | \(I = \frac{t_p}{t}\) |
| Average amplitude \((\bar{t'})\) | \(\bar{t'} = \frac{1}{N} \sum_{i=1}^{N} |t_i|\) | Clearance factor \((L)\) | \(L = \frac{t_p}{t_{rms}}\) |
| Root mean square \((t_{rms})\) | \(t_{rms} = \left(\frac{1}{N} \sum_{i=1}^{N} t_i^2\right)^{1/2}\) | Kurtosis \((K)\) | \(K = \frac{1}{N} \sum_{i=1}^{N} \frac{t_i^4}{t_{rms}^4}\) |
| Root mean square amplitude \((t_r)\) | \(t_r = \left(\frac{1}{N} \sum_{i=1}^{N} t_i^{1/2}\right)^2\) | Where the \(t_i\) is a signal series, for \(i=1, \ldots, N\), \(N\) is the number of data points. | |

### Table 2. The time-domain feature parameters of wind turbine main shaft.

(a)

| No. | States. | \(C_1\) | \(C_2\) | \(C_3\) | \(C_4\) | \(C_5\) | \(C_6\) | \(C_7\) | \(C_8\) | \(C_9\) | \(C_{10}\) | \(C_{11}\) |
|-----|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1   |         | 163    | 86     | 0.89   | 39     | 44     | 35     | 1.1    | 2      | 2.2    | 2.5    | 1.7    |
| 2   |         | 168    | 89     | 0.03   | 38     | 43     | 34     | 1.1    | 2.1    | 2.4    | 2.6    | 1.8    |
| 3   | N       | 163    | 86     | 0.53   | 37     | 43     | 34     | 1.1    | 2      | 2.3    | 2.6    | 1.7    |
| 4   |         | 171    | 88     | -0.3   | 38     | 43     | 34     | 1.1    | 2      | 2.3    | 2.6    | 1.8    |
| 5   |         | 162    | 83     | 0.62   | 38     | 43     | 35     | 1.1    | 1.9    | 2.2    | 2.4    | 1.7    |
| 6   |         | 161    | 90     | 0.72   | 36     | 41     | 32     | 1.1    | 2.2    | 2.5    | 2.8    | 1.8    |
| 7   |         | 171    | 94     | 0.88   | 37     | 42     | 33     | 1.1    | 2.2    | 2.5    | 2.8    | 1.7    |
| 8   |         | 174    | 91     | 0.32   | 37     | 42     | 33     | 1.1    | 2.2    | 2.5    | 2.7    | 1.8    |
| 9   |         | 164    | 87     | 0.31   | 36     | 41     | 33     | 1.1    | 2.1    | 2.4    | 2.7    | 1.8    |
| 10  |         | 160    | 86     | 0.18   | 37     | 42     | 34     | 1.1    | 2      | 2.3    | 2.5    | 1.7    |
| 11  |         | 153    | 79     | 0.41   | 34     | 38     | 30     | 1.1    | 2.1    | 2.4    | 2.6    | 1.7    |
| 12  |         | 152    | 79     | 0.46   | 33     | 38     | 30     | 1.1    | 2.1    | 2.4    | 2.6    | 1.8    |
| 13  |         | 149    | 78     | 0.43   | 33     | 37     | 30     | 1.1    | 2.1    | 2.4    | 2.6    | 1.7    |
| 14  |         | 151    | 78     | 0.62   | 33     | 38     | 30     | 1.1    | 2.1    | 2.4    | 2.6    | 1.8    |
| 15  |         | 150    | 79     | 0.09   | 33     | 38     | 30     | 1.1    | 2.1    | 2.4    | 2.6    | 1.8    |
| 16  |         | 164    | 84     | 0.69   | 39     | 44     | 35     | 1.1    | 1.9    | 2.2    | 2.4    | 1.7    |
| 17  |         | 164    | 83     | 1.02   | 39     | 44     | 36     | 1.1    | 1.9    | 2.1    | 2.3    | 1.7    |
| 18  |         | 173    | 88     | -0.25  | 38     | 44     | 34     | 1.1    | 2      | 2.3    | 2.6    | 1.8    |
| 19  |         | 166    | 88     | -0.67  | 39     | 44     | 35     | 1.1    | 2      | 2.3    | 2.5    | 1.7    |
| 20  |         | 166    | 86     | -0.57  | 38     | 43     | 34     | 1.1    | 2      | 2.3    | 2.5    | 1.7    |
| 21  |         | 133    | 70     | 0.53   | 25     | 29     | 23     | 1.2    | 2.4    | 2.8    | 3.1    | 2      |
| 22  |         | 130    | 69     | 0.61   | 25     | 29     | 22     | 1.2    | 2.3    | 2.7    | 3.1    | 2      |
| 23  |         | 130    | 72     | 0.33   | 25     | 29     | 22     | 1.2    | 2.4    | 2.9    | 3.2    | 2      |
| 24  |         | 128    | 67     | 0.08   | 25     | 29     | 22     | 1.2    | 2.3    | 2.7    | 3      | 2      |
| 25  |         | 130    | 66     | 0.11   | 24     | 28     | 22     | 1.2    | 2.3    | 2.7    | 3      | 2      |
The main shaft displacement signals of wind turbine in five conditions are collected. The 30 data sets are obtained for each condition; the sum of experimental data sets is $5 \times 30 = 150$. Firstly, the feature parameters, $C_1$-$C_{22}$, are all used (the sensitive feature parameters are not selected) to diagnose the faults. These parameters are shown in table 2. Owing to the limited paper space, only 25 data sets are given in this table. The 50 data sets (ten data sets are selected in each class) are used to training and the remaining 100 data sets are used to test the identification accuracy of the diagnoses model.
The results show that 23 data sets are erroneously diagnosed. The classification accuracy is 77%. The sensitive features should be selected for improving the classification performance.

The distance evaluation criteria of the 22 features ($C_1$–$C_{22}$) are calculated, the results are shown in figure 2. In this paper, the threshold is set as 0.5. It can be seen from figure 2 that the distance evaluation criteria of $C_4$, $C_5$, $C_6$, $C_{15}$, $C_{16}$, $C_{17}$ are greater than 0.5, so the $C_4$, $C_5$, $C_6$, $C_{15}$, $C_{16}$, $C_{17}$ are selected as sensitive features from the 22 features. The $C_4$, $C_5$, $C_6$, are average amplitude, root mean square, root mean square amplitude of main shaft horizontal displacement signal, respectively. The $C_{15}$, $C_{16}$, $C_{17}$, are average amplitude, root mean square, root mean square amplitude of main shaft vertical displacement signal, respectively. The six selected sensitive features ($C_4$, $C_5$, $C_6$, $C_{15}$, $C_{16}$, and $C_{17}$) of table 2 are input SVM to training and fault identification. The SVM is trained with 50 (10 data sets are selected in each class) sets and tested for 100 data sets. The results show that only 3 samples are misclassified, the classification accuracy rate is 97%. The results indicate that testing accuracy increases clearly when the sensitive features are selected.

5. Conclusions
In this paper, a fault diagnosis method of wind turbine based on features selection and SVM is presented. The sensitive time domain features of main shaft displacement signal are selected by an improved distance evaluation technique. The results show that using a sensitivity study to select the parameters used in the SVM improves the performance of wind turbine fault diagnoses. This paper provides a new idea to detect the fault type on the wind turbine in complex environment.

Acknowledgement
This work is supported by the National Basic Research Program (973 Program) (No.2007CB210304) and China Postdoctoral Science Foundation funded project (No.20090460273).

References
[1] Amirat Y, Benbouzid M, Al-Ahmar E, Bensaker B and Turri S 2009 A brief status on condition monitoring and fault diagnosis in wind energy conversion systems Renew. Sust. Energ. Rev. 13(9) 2629-36
[2] Hameed Z, Hong Y, Cho Y, Ahn S and Song C 2009 Condition monitoring and fault detection of wind turbines and related algorithms: A review Renew. Sust. Energ. Rev. 13(1) 1-39
[3] Barszcz T and Randall R 2009 Application of spectral kurtosis for detection of a tooth crack in the planetary gear of a wind turbine Mech. Syst. Signal Pr. 23(4) 1352-65
[4] An X, and Jiang D 2010 Chaotic characteristics identification and trend prediction of running state for wind turbine Electr. Power Autom. Equip. 30(3) 15-9
[5] Zaher A, McArthur S, Infield D and Patel Y 2009 Online wind turbine fault detection through automated SCADA data analysis Wind Energy 12 574-93
[6] Hou S and Li Y 2009 Short-term fault prediction based on support vector machines with parameter optimization by evolution strategy Expert Sys. Appl. 36 12383-91
[7] Mahadevan S and Shah S 2009 Fault detection and diagnosis in process data using one-class support vector machines J. Process Control 19 1627-39
[8] Widodo A and Yang B 2007 Support vector machine in machine condition monitoring and fault diagnosis Mech. Syst. Signal Pr. 21 2560-74
[9] Lei Y, He Z, Zi Y and Chen X 2008 New clustering algorithm-based fault diagnosis using compensation distance evaluation technique Mech. Syst. Signal Pr. 22 419-35
[10] Ding K, Li H and Zhu X 2005 Fault diagnosis practical technology for gear and gearbox (Beijing: China Machine Press)