Research on Fault Diagnosis Model of Rotating Machinery Vibration Based on Information Entropy and Improved SVM

Hankun Bing\(^1\)\(^*\), Yuzhu Zhao\(^2\), Le Pang\(^3\), Minmin Zhao\(^4\)

\(^1\)ENGINEER, Huadian Electric Power Research Institute, NO. 2Xiyuan 9 Road, Hangzhou Zhejiang Province, 310030, China
\(^2\)SENIOR ENGINEER, Huadian Electric Power Research Institute, NO. 2 Xiyuan 9 Road, Hangzhou Zhejiang Province, 310030, China
\(^3\)SENIOR ENGINEER, Huadian Electric Power Research Institute, NO. 2 Xiyuan 9 Road, Hangzhou Zhejiang Province, 310030, China
\(^4\)ENGINEER, Huadian Electric Power Research Institute, NO. 2 Xiyuan 9 Road, Hangzhou Zhejiang Province, 310030, China

Abstract. Based on the concept of information entropy, this paper analyzes typical nonlinear vibration fault signals of steam turbine based on spectrum, wavelet and HHT theory methods, and extracts wavelet energy spectrum entropy, IMF energy spectrum entropy, time domain singular value entropy and frequency domain power spectrum entropy as faults. The feature is supported by a support vector machine (SVM) as a learning platform. The research results show that the fusion information entropy describes the vibration fault more comprehensively, and the support vector machine fault diagnosis model can achieve higher diagnostic accuracy.

1 Introduction

In the research of turbine vibration fault diagnosis, the extraction of fault features directly affects the accuracy of fault diagnosis. In the previous feature extraction, vibration signal analysis is usually performed by a single signal analysis method, and the extracted fault features exhibit multi-dimensional state, which is targeted in fault diagnosis. Different fault diagnosis accuracy rates are divided.

Information fusion first appeared in the early 1970s, when the United States first used a variety of sensors to collect battlefield information in the military C3I system, and automatically analyzed the information they obtained using computers. The more precise definition of information fusion can be summarized as: Using computer technology to automatically analyze and comprehensively judge the multi-source observation information obtained by time series to complete the corresponding decision-making and estimation tasks\(^[1]\).

Information entropy describes the overall characteristics of the source in an average sense, and it characterizes the average uncertainty of the source. It can also be regarded as the uncertainty measurement of the amount of information in the system. A single information entropy usually cannot meet the characteristics of multiple faults in steam turbine vibration. Based on information fusion, the information entropy analysis of time domain, frequency domain and energy spectrum is carried out for different fault vibration signals, and information entropy is integrated. As a fault feature, vibration fault diagnosis is performed to form a fault diagnosis method based on fusion information entropy.

2 Information entropy feature definition of vibration signal

2.1 Wavelet energy spectrum entropy\(^[2]\)

Based on wavelet transform theory, the definition of wavelet energy spectrum entropy is proposed. The function \(f(t)\) with finite energy is conserved before and after wavelet transform:

\[
\int_{-\infty}^{\infty} |f(t)|^2 dt = \frac{1}{C_\psi} \int_{0}^{\infty} a^{-2} E(a) da
\]

among them: \(C_\psi = \int_{-\infty}^{\infty} |\psi(\omega)|^2 d\omega\) is the permissible conditions of the wavelet function, \(E(a) = \int |W(a,b)|^2 db\) is the energy value (also called wavelet energy spectrum) of the function \(f(t)\) when the scale is \(a\), \(\psi(\omega)\) is the Fourier transform of the wavelet basis function, \(W(a,b)\) is the amplitude of the wavelet transform.

Assume \(E = \{E_1, E_2, \ldots, E_n\}\) is the wavelet energy spectrum of the signal \(f(t)\) on \(n\) scales, \(E\) can form a division of the signal energy in the scale domain, and define the wavelet energy spectrum entropy as

\[
H_N = -\sum_{i=1}^{n} S_i \log_2 S_i \]

among them \(S_i = E_i / \sum_{i=1}^{n} E_i\) is the proportion of the i-th spectral value in the entire wavelet energy spectrum.

\(^*\) Corresponding author: Hankun Bing: bing_hankun@163.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
2.2 IMF energy spectrum entropy

American Chinese NE. Huang proposed HHT theory in 1996. The Empirical Mode Decomposition (EMD) algorithm is an adaptive time-frequency analysis tool that can adaptively decompose complex nonlinear signals to obtain the sum of the Intrinsic Mode Function (IMF) components[3]. The spectral components contained in each IMF component can be derived from the marginal spectrum by performing a Hilbert transform on each IMF component, so that each IMF component can represent the spectral distribution of the fault signal.

Define the IMF energy spectrum entropy as:

$$H_{IMF} = \sum_{i=1}^{n} S_i \log S_i$$

where $S_i$ is the proportion of the $i$-th spectral value in the entire IMF energy spectrum.

2.3 Time domain singular value entropy

For each discrete vibration signal sequence $X = \{x_1, x_2, \ldots, x_N\}$, the delay embedding technique is used to map the signal to the embedding space. If the length of the space is $M$, then a trajectory matrix $A$ of $N$ rows and $M$ columns is obtained, and the singular value decomposition of the matrix $A$ is performed. The singular value $\delta_i$ obtained by the design calculation $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_M$, then $\delta_i$ constitutes the singular value spectrum of the vibration signal[4]. Defining the singular spectral entropy of the signal in the time domain as:

$$H_S = -\sum_{i=1}^{n} S_i \log S_i$$

among them $S_i = n_i \sum_{j=1}^{p} p_j$ is the proportion of the $i$-th singular value in the entire singular value spectrum.

2.4 Frequency domain power spectrum entropy

$X(\omega)$ is discrete Fourier transform of signal $x(t)$, its power spectrum is $|X(\omega)|^2 / 2\pi N$. The energy is conserved[3] when the signal is transformed from the time domain to the frequency domain, ie:

$$\sum_{t} x^2(t) \Delta t = \sum_{\omega} |X(\omega)|^2 \Delta \omega$$

Therefore, the power spectrum can be regarded as a division of the vibration signal in the frequency domain. Define the signal power spectral entropy as:

$$H_L = -\sum_{i=1}^{n} S_i \log S_i$$

among them $S_i = p_i \sum_{j=1}^{p} p_j$ is the proportion of the $i$-th spectral value in the overall power spectrum.

Information entropy describes the spectral structure of vibration signal. The more uniform the distribution of vibration signal in time domain, frequency domain and time-frequency domain, the more complex the signal is, the greater the degree of uncertainty, and the higher the entropy value. On the contrary, the entropy goes down.

3 Support vector machine theory foundation[5]

SVM is developed from the optimal classification in the case of linear separability. The basic idea can be expressed in Figure 1 below:

![An optimal superplanar schematic](image)

In $n$-dimensional space, let the linear separable sample set be $(x_i, y_i), i = 1, 2, \ldots, n$. $x \in \mathbb{R}^d, y \in \{+1, -1\}$. For category labels, the linear discriminant function is generally in the form $w^T x + b = 0$. If all samples are separated by the equation, the following conditions must be met:

$$y_i[(w^T x_i) + b - 1] \geq 0, i = 1, 2, \ldots, n$$

At this point the classification interval is equal to $2 / \|w\|$. Maximizing the interval is equivalent to minimizing $\|w\|$. It satisfies the conditional expression and makes the minimum classification surface is the optimal plane, and the sample points on $H1$ and $H2$ are called support vectors.

The optimal classification surface has the best generalization performance, so that the problem of finding the optimal classification surface is transformed into an optimization problem. It is usually discussed in three cases: linearly separable, linearly inseparable, and non-linear. The first two cases are based on a linear function and are transformed into a typical quadratic programming problem:

$$Q(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

The constraints are: $\sum_{i=1}^{n} y_i \alpha_i = 0, \alpha_i \geq 0$. $\alpha_i$ is the Lagrange multiplier corresponding to the sample.

For the nonlinear problem, the kernel function is used to transform the nonlinear transform into a linear problem in a high-dimensional space, and the optimal classification surface is found in the transformed space. In other words, the input sample $x$ is projected into the high-dimensional feature space $L$, and then the generalized optimal classification surface is calculated in the feature space $L$.

Therefore, an appropriate inner product function $K(x, x')$ can be used in the optimal classification surface to realize the linear classification after a nonlinear
transformation. At this time, the optimization function becomes:

\[ L(w, b, a) = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i,j=1}^{n} y_i y_j a_i a_j K(x_i, x_j) \]  (4)

among them, \( K(x, x) \) is the satisfy the kernel function of the Mercer condition, there are currently three kinds of kernel functions applied:

(1) A polynomial kernel function of order q, namely:

\[ K(x, x) = (x \cdot x + 1)^q \]  (5)

(2) Radial basis function kernel function, namely:

\[ K(x, x) = \exp\left(-\frac{|x - x|^2}{\sigma^2}\right) \]  (6)

(3) Neural network kernel function, namely:

\[ K(x, x) = \tanh[c(x \cdot x) + c_2] \]  (7)

4 Application of information entropy feature extraction in signals

The vibration signals in this paper are generated from Bently simulated rotor test bench (RK4), and three vibration faults such as imbalance, collision and oil film oscillation are performed on the rotor test bench. The sampling points are 2560 and the sampling frequency is 1280Hz. 10 groups of samples are selected for each of the three faults.

For each fault signal sample obtained from the experiment, the above four kinds of information entropy calculation are carried out, and each calculated information entropy is regarded as one dimension. Then each final fault sample can be represented by a four-dimensional information entropy. In this paper, fault diagnosis will be conducted based on the characteristics of this four-dimensional information entropy as the fault sample.

In the process of wavelet energy spectrum entropy extraction, db6 was selected as the wavelet basis for wavelet analysis through optimization selection, and the vibration signal was decomposed into six layers of wavelet, and then the coefficients of each node in the layer were obtained, and then calculated according to the calculation method in 1.1. In the process of IMF energy spectrum entropy extraction, Hilbert transformation was performed on each IMF component to determine the spectral distribution of each IMF component and extract the energy of each component for entropy calculation based on the definition of 1.2. In the singular value entropy extraction, the embedded number of vibration signal is selected as 100 dimension. Then the singular decomposition of the time domain signal is carried out, and the singular value entropy is calculated according to the calculation method defined in 1.3. In the power spectrum entropy extraction, the indirect power spectrum estimation analysis is carried out on the signal, and the solution is carried out according to the calculation method defined in 1.4. The final calculation results are shown in Table 1:

From Table 1, we can see the basic law of the existence of information entropy. Firstly, for the unbalanced fault, in the unbalanced fault, the doubling frequency generated by the unbalanced quantity is the main component in the signal, and the vibration mode is single and the energy is concentrated. Therefore, the uncertainty of the energy distribution is low, so it is the smallest among the four-dimensional entropy values of the other two faults. In the IMF energy entropy, the Hilbert transform of each IMF component can be seen from the marginal spectrum that the low-frequency components are concentrated in one. In the frequency band, the IMF energy distribution is relatively concentrated compared to the rubbing, so the IMF energy entropy is smaller than the rubbing. Since the oil film oscillation signal mainly contains the low frequency region of (0.4-0.9) frequency doubling, the rubbing fault mainly includes the power frequency and the high frequency doubling component. Therefore, the power spectrum entropy can be seen that the power of the oil film oscillation is compared with the dynamic static rubbing. The spectral entropy is small. From the wavelet energy spectrum entropy, the oil film oscillation entropy is stronger than the rubbing instability. The reason for the analysis is that when using binary wavelet analysis, the wavelet decomposition has strong sensitivity to low frequency components.

| Sample | Information entropy | Fault type          |
|--------|---------------------|---------------------|
|        | Imf energy entropy  | Wavelet energy entropy | Singular value entropy | Power spectrum entropy |
| 1      | 0.9158              | 1.8469              | 2.0050              | 8.5318               | Oil film oscillation  |
| 2      | 1.0017              | 1.9003              | 1.9417              | 8.3122               | Oil film oscillation  |
| 3      | 1.1562              | 2.6540              | 1.9414              | 8.6223               | Oil film oscillation  |
|        |                     |                     |                    |                     |                     |
| 10     | 1.1524              | 2.4849              | 1.9521              | 8.0560               | Oil film oscillation  |
| 1      | 1.3633              | 1.8753              | 1.4424              | 9.5500               | Static and static rubbing |
| 2      | 1.3332              | 1.8007              | 1.4248              | 9.5814               | Static and static rubbing |
| 3      | 1.0594              | 1.8758              | 1.4374              | 9.5777               | Static and static rubbing |
|        |                     |                     |                    |                     |                     |
| 10     | 1.3548              | 1.8642              | 1.4294              | 9.6044               | Static and static rubbing |
| 1      | 0.1807              | 0.7043              | 1.0342              | 7.9856               | unbalanced           |

Table 1 Information entropy characteristics of fault samples
5. Support vector machine fault diagnosis

It can be seen from the above analysis that the four kinds of information entropy values of each fault can basically reflect the complexity degree of the fault signal. In fact, regardless of the time domain, frequency domain or time-frequency domain features, it only reflects one aspect of the signal characteristics. If the four kinds of information entropy of the fault entropy point are comprehensively analyzed to analyze the vibration signal, the distinction of each fault is more obvious.

The 30 sets of samples, the first 6 samples of each fault were used as training samples, and the remaining 4 samples were used as test samples for fault diagnosis tests.

5.1 Parameter optimization selection method

In the training process, due to different penalty coefficients and the choice of parameters of the kernel function, the accuracy of the training samples is determined, which affects the accuracy of the model prediction. Therefore, two methods are used to optimize the parameters when samples was trained. A cross-validation method is used to select the optimal training parameters. The cross grid is selected as 10 by 10 and the grid step is 0.5. The second is to use particle swarm optimization to select parameters, the evolutionary algebra is set to 100 generations, and the population size is chosen to be 20.

The kernel function is selected as the radial basis kernel function. After cross-validation, the parameter optimization selection results are shown in figure 2. The penalty factor C is 2.8284, the gama is 1, and the test sample accuracy is 100%. The classification result is shown in figure 3. After particle swarm optimization parameter selection, the parameter optimization selection results are shown in figure 4. The penalty factor C is 8.2327, the gama is 0.7666, and the diagnostic accuracy is 100%. The classification result is shown in Fig. 5.

6. Conclusion

In this paper, the vibration fault samples simulated by the rotor test bench are taken as the research object. Based on the wavelet and HHT signal analysis basic theory, the time domain, frequency domain and time-frequency information entropy extracted from the vibration signal can be used to describe the faults in different aspects of the signal. The experimental results show that based on the cross-validation and particle swarm optimization parameter selection methods, the fault feature samples combined with a single information entropy have high
diagnostic accuracy on the support vector machine fault diagnosis platform. This method can be used as steam turbine vibration. A new method of fault diagnosis, the follow-up work will be based on the actual vibration signal of the site for fault diagnosis and verification.

References

1. Yuan Xiaohong, Qu Liangsheng. Research on information fusion and utilization in mechanical fault diagnosis[J]. Vibration, Testing and Diagnosis, 19,187-192(1999)

2. Hou Jinghong, Huang Shuhong, Shen Wei et al. Quantitative characteristics of vibration signals of rotating machinery based on wavelet analysis[J]. Journal of Mechanical Engineering, 2004, 40, 131-135 (2004)

3. WU Z, HUANG N E. Ensemble empirical mode decomposition: a noise assisted data analysis method[R]. Calverton: Center for Ocean-Land—Atmosphere Studies, (2005)

4. Zhang Yanping, Huang Shuhong, etc. Turbine Vibration Fault Diagnosis Based on Information Entropy-Gray Correlation Method[C]. Proceedings of Steam Turbine Professional Symposium of Electrical Engineering Society of Hubei, Jiangsu, Jiangsu and Anhui Provinces, (2007)

5. Wei Yufan. Research on Support Vector Machine in Intelligent Fault Diagnosis[D]. North China Electric Power University. (2007)

6. Chen Fei, Huang Shuhong et al. Vibration fault diagnosis method for rotating machinery based on information entropy distance[J]. Vibration, Testing and Diagnosis, 28, 9-13(2008)