A Diagnosis Method of Vibration Fault of Steam Turbine Based on Information Entropy and Grey Correlation Analysis

Fei Chen¹, ²*, Hao Cao¹, Bolin Zhang¹, ², Jilong Wei¹, ²

¹Hunan Province Key Laboratory of High Efficient & Clean Thermal Power Technology (State Grid Hunan Electric Power Corporation Limited Research Institute), Changsha, 410007, China
²Hunan Xiangdian Test & Research Institute Co., Ltd., Changsha, 410004, China

*Corresponding author: cfhust123@sgcc.com.cn

Abstract. Based on the faulty signals collected from the rotor test platform, the grey correlation analysis was studied for vibration fault diagnosis of steam turbine shafting. The reference faulty matrix and the calculation model of grey correlation degree were established based on three kinds of information entropy. The analysis shows that grey correlation analysis is a useful method for fault diagnosis of shafting and can be used as a quantitative index for fault diagnosis.

Keywords: Steam turbine, vibration fault diagnosis, information entropy, grey correlation analysis.

1. Introduction
The vibration fault, one of the common faults in the steam turbine generator unit, brings great damage to the production and the running process. The information entropy, which describes the degree of the uncertainty in the system, can be used to measure the vibration condition of the steam turbine. The document analyzes the fault identification capability of three different information entropies, which are respectively the singular spectrum entropy, power spectrum entropy and wavelet space feature spectrum entropy. Despite their efficiency, one kind of information entropy is just enabled to identify certain part of the faults. In order to make up for this limitation, this thesis proposes the grey correlation degree as the identification feature and establishes a calculation model of grey correlation degree based on the information entropies. Through the analysis of the fault test signals, it is proved that the grey correlation analysis based on the three information entropies is an effective way to identify the faults of the steam turbine shafting.

2. The rotor faults simulation experiment
The experimental rotor is shown schematically in Fig.1. Five group faulty signals were collected during the process of rising speed of the rotor described above through eddy current probe. The five faults are rotor imbalance (RI), shafting misalignment (SM), oil whipping (OW), pedestal looseness (PL) and radial friction (RF). This platform mounts four eddy current probes close to the shaft and an accelerometer on the bearing housing. The signals from the sensors are fed into a signal pre-processing device, and then sampled by synchronous sampling. The sampling frequency is 32 times the basic
frequency of rotor, the sampled signals, each containing 512 data points, are carefully stored in computer for further analysis. In the signal pre-processing device, the signals are amplified, filtered, and the direct current component of signals is removed too.

Figure 1. Rotor test platform

3. Information entropy features of vibration signal

3.1. Basic concept of information entropy

According to Shannon's entropy concept, the system information entropy can be defined as follows [1].

Supposing $K$ be a Lebesque space with a $\sigma$ algebra created by a measurable set, which has $\mu$ measure with the property $\mu(K)=1$. The space $K$ can be denoted $A$ composed of some finite incompatible partitions $A_i$, depicted as (1).

$$K = \bigcup_{i=1}^{n} A_i \quad \text{And} \quad A_i \cap A_j = \emptyset, \forall i \neq j$$

Then, the information entropy of $A$ is given by (2).

$$S(A) = -\sum_{i=1}^{n} \mu(A_i) \log \mu(A_i)$$

Where $\mu(A_i)$ is defined as the measure of the partition $A_i$ ($i = 1, 2, \ldots, n$).

3.2. Singular spectrum entropy

The data from the vibration sensor is a discrete time sequence. Singular-spectrum analysis is a toolkit for short, noisy chaotic signals.

For each time sequence $X = [x_1, x_2, \ldots, x_N]$, the signal is mapped to the built-in space by the time-lapse embedding technique. Supposing the length of the space is $M$, the track matrix $A$ with $N-M$ rows and $M$ columns is defined as (3).

$$A = \begin{bmatrix}
  x_1 & x_2 & \cdots & x_M \\
  x_2 & x_3 & \cdots & x_{M+1} \\
  \vdots & \vdots & \ddots & \vdots \\
  x_{N-M} & x_{N-M+1} & \cdots & x_N
\end{bmatrix}$$
The singularity decomposition is made to the matrix $A$. Let the value of the singularity be $\delta_1 \geq \delta_2 \geq \ldots \geq \delta_n$, then the singular spectrum of vibration is composed of the $\{\delta_i\}$. Let $r$ denote the number of the nonzero singularity, the $r$ represents the number of different mode included in the matrix $A$. We think that the singular spectrum $\{\delta_i\}$ be one partition to the vibration in time field. So, the singular spectrum entropy (SSE) in time field is defined as (4).

$$S_1 = -\sum_{i=1}^{M} p_i \log p_i$$

Where $p_i = \frac{\delta_i}{\sum_{i=1}^{M} \delta_i}$ denotes the ratio between the $i$th singularity and the whole singular spectrum.

3.3. Power spectrum entropy
Let $X(\omega)$ be the Fourier transformation of the signal $\{x_i\}$. The power spectrum is denoted by (5).

$$L(\omega) = \frac{1}{2\pi N} |X(\omega)|^2$$

The energy keeps conservation when the signal changes from the time field to the frequency field, which can be expressed as follows (6).

$$\sum x^2(t) \Delta t = \sum \left| X(\omega) \right|^2 \Delta \omega$$

Thus, $L = \{L_1, L_2, \ldots, L_N\}$ is considered as one partition to the original signal where $L_i$ devotes some power spectrum of the signal. The power spectrum entropy (PSE) is defined as (7).

$$S_2 = -\sum_{i=1}^{N} q_i \log q_i$$

Where $q_i = \frac{L_i}{\sum_{i=1}^{N} L_i}$ denotes the ratio of the $i$th spectrum in all the spectrums.

3.4. Wavelet space feature spectrum entropy
The function $f(t)$ with finite powers keeps conservation during the wavelet transform, which verifies the following property:

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

Where $C_\psi = \int_{-\infty}^{+\infty} \left| \hat{\psi}(\omega) \right|^2 d\omega$ and $E(a) = \int_{-\infty}^{+\infty} |W_f(a, b)|^2 db$. 

3
\( C_\psi \) is a finite value and called the allowed condition of wavelet function, \( E(a) \) is called wavelet energy spectrum denotes the power of the function \( f(t) \) in a scale, \( W_j(a,b) \) is called the amplitude of the wavelet transform.

From (8), wavelet transform is used for mapping the signal from one dimension to two dimensions, where \( W = \left[ \frac{W_j(a,b)^2}{C_\psi a^2} \right] \) is called the power distribution matrix in two dimensions space. Similar to SSE in time field, by the singularity decomposition of \( W \), wavelet space feature spectrum entropy (WSFSE) is defined as (9).

\[
S_3 = -\sum_{i=1}^{n} p_i \log p_i
\]  

Where \( p_i = \frac{\delta_i}{\sum_{i=1}^{n} \delta_i} \) denotes the ratio of the \( i \)th spectrum to the whole spectrums.

4. The grey correlation analysis
Grey correlation analysis (which is also called correlation degree analysis) describes and compares the changes in the system in a quantitative way. On the mathematical basis of the space theory and consistent with the principles of normalization, even symmetry, integrity and proximity, establish the correlation coefficient and correlation degree between the reference data sequence and several comparison data sequence. By finding main relationship between all kinds of elements in the system and identifying the important element that affect the target value, the grey correlation analysis aims to grasp the major features of the system and further to rapidly improve the efficiency the system [2].

The general steps of grey correlation analysis are as follows:
1) Confirm the reference data sequence: \( X_0 = \{X_0(k) | k = 1, 2, \ldots, m\} \)
2) Confirm the comparison data sequence: \( X = \{X(k) | k = 1, 2, \ldots, m\} (i = 1, 2, \ldots, n) \)
3) Average the above data sequences:
   The reference data sequence: \( Y_0 = \{X_0(k)/X_0 | k = 1, 2, \ldots, m\} \)
   The comparison data sequence: \( Y_i = \{X_i(k)/X_i | k = 1, 2, \ldots, m\} (i = 1, 2, \ldots, n) \)
4) Draw the correlation coefficient:
   \[
   \xi_i = \frac{\min_{i \in k} \Delta_i(k) + \rho \max_{i \in k} \Delta_i(k)}{\Delta_i(k) + \rho \max_{i \in k} \Delta_i(k)}
   \]  

In (10), \( \Delta_i(k) = | X_0(k) - X_i(k) | ; \rho \) is the resolution ratio, generally, \( \rho \in (0, 1) \), and the value is set as 0.5; \( \min_{i \in k} \Delta_i(k) \) is the two level minimal value; \( \max_{i \in k} \Delta_i(k) \) is the two level maximum value.

4) Draw the correlation degree
   \[
   r_i = 1/n \sum \xi_i(k)
   \]
5) Arrange the correlation degree by its value:

The higher the correlation degree between \( X_i \) and \( X_0 \) is, the more similar their development trend is, and \( X_i \) will exert more influence on \( X_0 \).

5. Case calculation

Three spectrum entropies, which are the singular spectrum entropy in time domain, the power spectrum entropy in frequency domain, and the wavelet space feature spectrum entropy in time-frequency domain, reflect the degree of the complexity of the vibration fault signal of the rotor from different angles respectively. By combining these three information entropy features, this thesis forms a reference data sequence, namely the reference fault model, as well as a comparison data sequence, namely the model to be diagnosed, and establishes a grey correlation calculation model.

According to the analysis of the five fault test signals\([3, 4]\), choose the signal record which shows the most distinct fault symptom as the reference fault signal, and correspondently choose its information entropy feature as the reference fault models. The reference fault model matrix is as follows in Table 1.

Calculate the information entropy values of 15 groups of the fault test signals respectively. The results constitute the comparison data sequence, namely model to be diagnosed. Then calculate their grey correlation degree respectively, and see the result in Table 2 in which \( r_{01}, r_{02}, \ldots, r_{05} \) represent the correlation degree between the comparison models and 5 reference models in Table 1.

| \( X_{01} \) | SSE | PSE | WSFSE | Fault type |
|-------|-----|-----|-------|------------|
| \( X_{02} \) | 39.67 | 45.96 | 15.36 | RI |
| \( X_{03} \) | 66.38 | 72.53 | 25.56 | SM |
| \( X_{04} \) | 72.64 | 74.89 | 32.36 | OW |
| \( X_{05} \) | 40.35 | 45.83 | 12.68 | PL |

Table 1. Reference data sequence of information entropy

As shown in Table 2, in the diagnosis results of 15 comparison models with the usage of grey correlation analysis, 9 models to be diagnosed with the fault of SM, OW and PL have all been rightly identified, while among the 6 models to be diagnosed with the fault of RI and RF, only two of them have been misidentified. Accordingly, a conclusion can be drawn that the grey correlation analysis is an effective way to identify the vibration fault of steam turbine shafting with comparatively high accuracy.

| Number | \( r_{01} \) | \( r_{02} \) | \( r_{03} \) | \( r_{04} \) | \( r_{05} \) | Result of grey correlation analysis | Actual fault type |
|--------|-----|-----|-----|-----|-----|----------------|----------------|
| 1 | 0.995 | 0.536 | 0.589 | 0.943 | 0.879 | RI | RI |
| 2 | 0.975 | 0.568 | 0.538 | 0.928 | 0.872 | RI | RI |
| 3 | 0.785 | 0.720 | 0.745 | 0.725 | 0.801 | RF | RI |
| 4 | 0.478 | 0.963 | 0.845 | 0.468 | 0.586 | SM | SM |
| 5 | 0.448 | 0.962 | 0.789 | 0.485 | 0.534 | SM | SM |
| 6 | 0.454 | 0.968 | 0.668 | 0.435 | 0.564 | SM | SM |
| 7 | 0.512 | 0.834 | 0.965 | 0.456 | 0.564 | OW | OW |
| 8 | 0.431 | 0.816 | 0.902 | 0.415 | 0.465 | OW | OW |
| 9 | 0.456 | 0.786 | 0.863 | 0.488 | 0.442 | OW | OW |
| 10 | 0.926 | 0.546 | 0.567 | 0.981 | 0.836 | PL | PL |
| 11 | 0.929 | 0.558 | 0.549 | 0.983 | 0.826 | PL | PL |
| 12 | 0.932 | 0.562 | 0.534 | 0.963 | 0.803 | PL | PL |
| 13 | 0.911 | 0.568 | 0.572 | 0.891 | 0.938 | RF | RF |
| 14 | 0.933 | 0.579 | 0.581 | 0.906 | 0.891 | RI | RF |
| 15 | 0.915 | 0.561 | 0.567 | 0.881 | 0.935 | RF | RF |

Table 2. The grey correlation degree of all fault test signals
6. Conclusions
Based on the simulation experiment of typical faults in the steam turbine shafting, the theory of grey correlation analysis is introduced to identify the vibration faults in the steam turbine shafting. Through analyzing the experiment data, the grey correlation reference fault matrix composed by three information entropies is confirmed and the correspondent calculation model of the grey correlation degree is established. The grey correlation degree calculations of different groups of fault signal prove that the correlation degree, as a quantitative feature reflecting the similarity between the signal to be diagnosed and the reference fault matrix, can realize a good identification to parts of the vibration faults in the shafting. In the future analysis, the actual fault signal of the steam turbine should be used to establish the standard fault model for the specific steam turbine shafting, and gradually apply the grey correlation analysis into the actual production.

References
[1] X. H. Yuan and L. S. Qu, Utility of Information Fusion in Mechanical Fault Diagnosis, Journal of Vibration, Measurement & Diagnosis, 19 (1999) 187 - 192.
[2] X. G. Cai and P. Ma, Study on Simultaneous Fault Diagnosis Based Information Fusion Technique, Proceedings of the CSEE, 23 (2003) 112 - 115.
[3] T. Shen, S. H. Huang, S. M. Han, et al., Extracting Information Entropy Features for Rotating Machinery Vibration Signals, Chinese Journal of Mechanical Engineering, 37 (2011) 94 - 98.
[4] F. Chen, S. H. Huang, Y. P. Zhang, et al., Methodology and development of fault diagnosis based on information fusion in power generation unit, Journal of Vibration, Measurement & Diagnosis, 25 (2005) 17 - 22.