Research on Abnormal Signal Identification Algorithm of Distribution Network Operation Based on Big Data Mining

Yang Li1*, Sheng Liu2, Ziyi Zhao3, Lin Yan4, Zhaohua Hu5

1828937@qq.com1*, 309503096@qq.com2, 850583975@qq.com3, yanlinll@163.com4, 412301613@qq.com5

Shenzhen Power Supply Co., Ltd. Shenzhen, Guangdong, 518000, China

Abstract. Conventional algorithm for identifying abnormal signals of distribution network mainly uses ESD (extreme studied deviate test) to obtain signal time series, which is easily influenced by proximity clustering relationship, resulting in poor identification performance. Therefore, it is necessary to design a new algorithm for identifying abnormal signals of distribution network based on big data mining. That is, the abnormal signal of distribution network operation is collected, and the characteristics of abnormal signal of distribution network operation are extracted by big data mining, and the optimization algorithm of abnormal signal identification of distribution network operation is generated, thus realizing the identification of abnormal signal of distribution network operation. The experimental results show that the designed algorithm for identifying abnormal signals of distribution network operation has good recognition performance, high accuracy, recall and F1 score, and it is reliable and has certain application value, which has made certain contributions to reducing the operation risk of distribution network.

Keywords: Big data mining; Power distribution; Power grid; Operation; Abnormal signal; Identification; algorithm.

1 Introduction

With the acceleration of urbanization and the development of economy and society, the scale and complexity of distribution network are increasing, and its development and importance are increasingly prominent. With the growth of power demand and the adjustment of energy structure, the load structure and distribution of distribution network are also changing, so it is necessary to upgrade and transform the distribution network to meet the new demand [1-3]. In addition, how to improve the intelligence and automation level of distribution network, reduce loss and improve efficiency is also a problem that needs attention in the development of distribution network. Under the background of rapid energy structure transformation, the distribution network structure in China is becoming more and more complex [4-6], which is often affected by abnormal signals, resulting in serious operation accidents. Therefore, effective identification of abnormal signals in distribution network operation is needed.

Identification of abnormal operation signals of distribution network is one of the important tasks in power system. Due to the influence of various factors, such as load fluctuation,
equipment failure, external interference, etc., the power grid operation may be abnormal, such as voltage fluctuation, current imbalance, frequency deviation, etc. [7-10]. If these abnormal signals are not found and handled in time, it may have a serious impact on the stable operation of power grid and the safety of power equipment. Therefore, it is necessary to design an effective algorithm for identifying abnormal signals in distribution network operation to support the identification of abnormal signals in distribution network. The identification of abnormal signals in distribution network operation mainly includes the following steps. First, it is necessary to collect real-time data of distribution network, including parameters such as current, voltage and power factor [11-13]. These data can be obtained through on-site monitoring equipment or remote monitoring system of power system. Secondly, the collected data need to be preprocessed such as denoising and filtering to remove interference and outliers and ensure the accuracy and reliability of the data. Extract features from the preprocessed data, which can be time domain features, frequency domain features, waveform features, etc. These characteristics can reflect the operation state and abnormal situation of distribution network. Finally, anomaly classification and recognition [14-17] is carried out, and the extracted features are input into classifiers, such as support vector machine, neural network and decision tree, for classification and recognition. By comparing the training samples with the test samples, the accuracy and recall rate of the algorithm are evaluated. According to the classification results and the characteristics of fault types, fault diagnosis and location are carried out. In order to solve the current identification problem, this paper designs a new algorithm for identifying abnormal signals in distribution network operation based on big data mining.

2 Design of abnormal signal identification algorithm for distribution network operation based on big data mining

2.1 Collect abnormal signals of distribution network operation

In order to improve the accuracy of abnormal signal identification, it is necessary to ensure the acquisition quality of the original abnormal signal of distribution network. Therefore, this paper first collects the abnormal signal of distribution network operation [18-20]. Firstly, an abnormal signal set \( T \) can be generated according to the coordinate relationship of abnormal signals, as shown in the following (1).

\[
T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}
\]  

In the formula (1), \((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\) are the abnormal signal representative of the distribution network operation collected for the first time, respectively. According to the above-mentioned abnormal signal sets, the fault signals can be linearly converted, and the distribution network abnormal signal collection kernel function \( X_n \) can be generated by using the relationship between the horizontal and vertical coordinate sets, as shown in the following (2).
In the formula (2), $w$ represents an angular variable between elements, $\theta$ represents the relaxation factor, $\pi$ represents the transformation coefficient of kernel function, $D$ represents the penalty parameter, according to the above-mentioned kernel function transformation coefficient [21-22], an effective abnormal signal recognition filter factor $s(k)$ can be obtained, as shown in the following (3).

$$s(k) = k y_{\text{max}}$$  \hspace{1cm} (3)

In the formula (3), $k$ represents the number of abnormal signals, $y_{\text{max}}$ represents the maximum abnormal signal value. Based on this, the generated abnormal signal acquisition flow chart is shown in Figure 1 below.

![Flow chart of abnormal signal acquisition in distribution network.](image)

**Fig. 1.** Flow chart of abnormal signal acquisition in distribution network.

As can be seen from Figure 1, according to the above-mentioned abnormal signal collection flow of distribution network, abnormal signal collection nodes can be effectively arranged, and the quality of collected abnormal signals can be improved to the greatest extent, which can be used as a reference for subsequent abnormal signal feature extraction.

### 2.2 Based on big data mining, extract abnormal signal characteristics of power grid operation

Big data mining technology is a high-performance data analysis and processing technology, which can obtain the required abnormal features from massive data, and combine the correlation between features for comprehensive mining analysis. Therefore, this paper extracts...
the abnormal signal features of power grid operation based on big data mining technology. Firstly, we need to denoise the signal and calculate the local mean value of the signal, as shown in the following (4).

\[ m_i = \frac{n_i + n_{i+1}}{2} \]  

(4)

In the formula (4), \( n_i \) and \( n_{i+1} \) respectively represent the time-frequency distribution weights of the signals before and after the estimation, and according to the above-mentioned local mean value of the signals, abnormal signals can be separated, and the original signals are subdivided. At this time, the separated abnormal signals \( h(t) \) is shown in the following (5).

\[ h(t) = x(t) - m(u) \]  

(5)

In the formula (5), \( x(t) \) represents a status signal, \( m(u) \) represents the locally separated signal, after the above steps are completed, the feature extraction feature of abnormal signal in distribution network can be obtained, that is, the collected high-quality signal is decomposed by LMD method, the components containing abnormal features are selected, and then the energy values of different components \( E \) are calculated, as shown in the following (6).

\[ E = \sum_{p=1}^{n} E_p(t) \]  

(6)

In the formula (6), \( E_p(t) \) represents the energy component representing the main fault information can be used to process the calculated energy values of different signal components. At this time, the abnormal signal characteristics identification processing formula \( f(x) \) based on hidden layer neural network can be used, as shown in the following (7).

\[ f(x) = \sum B_i G(X) \]  

(7)

In the formula (7), \( B_i \) represents the hidden-layer output matrix, \( G(X) \) represents the sample output matrix. According to the above processing procedure, the representative sample output matrix can be used to mine big data, output the extreme learning function, and describe it in a matrix way. By combining the output connection relationship between the hidden layer and the sample layer, the abnormal signal feature recognition result can be obtained, which improves the comprehensive performance of the abnormal signal recognition algorithm.

2.3 Generation of distribution network operation abnormal signal identification optimization algorithm

The identification of abnormal signals in actual distribution network operation can be roughly divided into several steps, that is, generating a similarity matrix according to the sample set, extracting the characteristic vectors of abnormal signals according to the similarity matrix, and standardizing them to obtain the final identification results of abnormal signals. However, the
conventional abnormal signal identification algorithm ignores the multidimensional relationship of feature vectors, which is prone to identification distortion and has poor overall performance. Therefore, based on big data mining and the principle of adaptive spectral clustering, this paper uses Euclidean distance to determine the identification relationship of abnormal signals. Based on this, the generated distribution network operation abnormal signal identification optimization algorithm is shown in Figure 2 below.

![Optimization algorithm of abnormal signal identification in distribution network operation.](image)

As can be seen from Figure 2, using the above-mentioned optimization algorithm for identifying abnormal signals in distribution network operation can quickly obtain feature vectors, classify them according to the permutation matrix, accurately identify the threshold of abnormal signals, and reduce the possible interference in identifying abnormal signals.

3 Experiment

In order to verify the recognition effect of the designed abnormal signal recognition algorithm for distribution network operation based on big data mining, this paper configures a basic experimental platform, and compares it with the conventional abnormal signal recognition algorithm for distribution network operation based on random matrix and the abnormal signal recognition algorithm for distribution network operation considering strategy fusion, and carries out experiments as follows.

3.1 Experimental preparation

According to the performance verification and identification requirements of the abnormal signal identification algorithm for distribution network operation, this paper selects Lab Windows/CVI as the experimental platform to carry out the performance verification experiment. First, it is necessary to collect the abnormal data of distribution network operation with a data acquisition card, which is relatively large and has many types. Therefore, this paper selects SK2010 data acquisition card to perform the collection task and carry out effective interface matching. At this time, an effective experimental data collection process can be generated, as shown in Figure 3 below.
As can be seen from Figure 3, this experiment mainly uses multi-channel signal acquisition method to collect experimental data signals. In order to meet the signal measurement and receiving requirements of the experimental platform, this paper uses PXI software to preprocess the signals, converts the basic receiving frequency of the signals, and installs brand-new Active experimental components. At this time, the experimental self-checking database can be generated. The attributes and data type parameters of different types of abnormal signals are shown in Table 1 below.

**Table 1. Database Parameters of Experimental Self-inspection.**

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Data type</th>
<th>Data examples</th>
<th>Source</th>
<th>Structured data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial Number</td>
<td>INT</td>
<td>5</td>
<td>automatic</td>
<td>Y</td>
</tr>
<tr>
<td>area</td>
<td>VARCHAR</td>
<td>XX market</td>
<td>OMS</td>
<td>Y</td>
</tr>
<tr>
<td>Factory Station</td>
<td>VARCHAR</td>
<td>XX stand</td>
<td>D5000</td>
<td>Y</td>
</tr>
<tr>
<td>protection equipment</td>
<td>VARCHAR</td>
<td>busbar protection</td>
<td>D5000</td>
<td>Y</td>
</tr>
<tr>
<td>Alarm element</td>
<td>VARCHAR</td>
<td>intelligent terminal</td>
<td>D5000</td>
<td>Y</td>
</tr>
<tr>
<td>Alarm content</td>
<td>VARCHAR</td>
<td>Device abnormality</td>
<td>D5000</td>
<td>Y</td>
</tr>
<tr>
<td>Alarm time</td>
<td>TIME</td>
<td>2020.1.6/22:33</td>
<td>D5000</td>
<td>Y</td>
</tr>
<tr>
<td>Faulty component</td>
<td>VARCHAR</td>
<td>terminal</td>
<td>OMS</td>
<td>N</td>
</tr>
<tr>
<td>Action situation</td>
<td>VARCHAR</td>
<td>maloperation</td>
<td>OMS</td>
<td>Y</td>
</tr>
</tbody>
</table>

According to the data parameters in Table 1, the abnormal signals are stored, and the structured data attributes extracted from the experimental database are Region, Factory and Alarm Time. At this time, the recognition accuracy, recall rate can be selected, and F values are used as experimental preparations, and the calculation formulas are shown in the following (8)–(10).
\[
P = \frac{\sum |R(u) \cap T(u)|}{\sum |R(u)|}
\]
(8)

\[
R = \frac{\sum |R(u) \cap T(u)|}{\sum |T(u)|}
\]
(9)

\[
F = \frac{2PR}{P + R}
\]
(10)

In formulae (8) to (10), \(R(u)\) represents the actual accurate identification signal, \(T(u)\) represents a predicted abnormality recognition signal. The higher the recognition accuracy, recall and F value represents the predicted abnormal signal recognition, it proves that the recognition performance of the abnormal signal recognition algorithm is better, on the contrary, it proves that the recognition performance of the abnormal signal recognition algorithm is relatively poor.

3.2 Experimental results and discussion

According to the above experimental preparation, the performance verification of distribution network abnormal signal identification algorithm can be carried out, that is, the distribution network abnormal signal identification algorithm based on big data mining, the conventional distribution network abnormal signal identification algorithm based on random matrix and the distribution network abnormal signal identification algorithm considering strategy fusion are used to identify different operation abnormal signals respectively, and the identification performance of the three algorithms is calculated by formulas (8)–(10). The experimental results are shown in the following Table 2.

<table>
<thead>
<tr>
<th>Identification signal category number</th>
<th>The recognition accuracy of the distribution network abnormal signal recognition algorithm based on big data mining (%)</th>
<th>Identification accuracy of distribution network abnormal signal recognition algorithm based on random matrix (%)</th>
<th>Recognition accuracy of abnormal signal recognition algorithm for distribution network operation considering strategy fusion (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FA01, FA02, FA03, FA04, FA05, FA06, FA07, FA08, FA09, FA10</td>
<td>98.747, 94.841, 95.285, 95.475, 96.369, 96.173, 99.245, 95.854, 98.256, 97.584</td>
<td>65.157, 61.251, 64.368, 76.545, 68.269, 75.875, 76.478, 67.352, 65.145, 67.284</td>
<td>68.185, 64.148, 65.355, 62.276, 51.985, 66.175, 79.124, 67.284, 65.269, 52.278</td>
</tr>
</tbody>
</table>
As can be seen from Table 2, the recognition accuracy, recognition recall and recognition F value of the distribution network operation abnormal signal recognition algorithm designed in this paper are high under different types of abnormal signals, while the recognition accuracy, recognition recall and recognition F value of the conventional distribution network operation abnormal signal recognition algorithm based on random matrix and the distribution network operation abnormal signal recognition algorithm considering strategy fusion are relatively low under different types of abnormal signals. The above experimental results prove that the abnormal signal identification algorithm of distribution network based on big data mining designed in this paper has good performance, reliability and certain application value.

4 Conclusion

Distribution network is an important part of power system, which undertakes the important task of transmitting power from power generation side to user side. The identification of abnormal signals in distribution network operation is of great significance in ensuring the stable operation of power system, improving the operation efficiency of power equipment, optimizing the planning and construction of power system, ensuring social economy and people's life, preventing and reducing accidents and improving the service level of power supply. This paper introduces the methods and steps of identifying abnormal signals in distribution network operation, including data acquisition, data preprocessing, feature extraction, abnormal classification identification and fault diagnosis. At the same time, this paper also introduces the importance and significance of abnormal signal identification of
distribution network operation, as well as the existing problems and future development trends, and designs a brand-new algorithm for abnormal signal identification of distribution network operation based on big data mining. The experimental results show that the designed algorithm based on big data mining has good performance, reliability and certain application value, and has made certain contributions to improving the operation reliability of distribution network.

References


