Cable fault Detection Based on SVM and Wavelet Analysis

Liu Zhu\textsuperscript{1,a*}, Wanting Zheng\textsuperscript{1,b}

\textsuperscript{1} Electrical and control engineering, North China University of Technology, BeiJing, 100144, China
\textsuperscript{b} e-mail: zwt20001115@163.com,
\textsuperscript{a*}Corresponding author’s e-mail:18153010721@mail.ncut.edu.cn

Abstract. In reality, cable faults are highly concealed, so accurate detection and identification of cable faults plays an irreplaceable role in the stable operation of the power system. This article first uses SIMULINK in MATLAB to build a power system fault simulation model, through which the overcurrent under different faults is obtained, and then uses wavelet analysis to extract the characteristics of the overcurrent, and then uses four recognition algorithms to detect and identify different fault categories. By comparing the four recognition effects, it is found that SVM has the smallest error, fast calculation speed, and an accuracy of up to 0.997. Finally, a fault detection model based on SVM is determined.

1. Introduction
With the continuous progress of society and the rapid development of cities, the shortcomings of overhead transmission lines, such as covering a large area and being unsightly, are becoming more and more prominent. At the same time, the advantages of power cables such as safety and concealment are gradually being valued. In recent years, power cables have gradually replaced traditional overhead transmission lines and have become the development trend of urban distribution networks.

However, power cables have faults such as insulation aging, insulation dampness, and insulation damage during use. The fault points are highly concealed. If the location of the fault point cannot be found quickly and the fault is eliminated in time, the loss caused is immeasurable. Therefore, the accurate identification of cable fault categories is of great significance to the safe and stable operation of the power system.

In recent years, most scholars at home and abroad mainly use the time domain and frequency domain characteristics of voltage and current to detect them [1-6]. Literature [1] uses Fourier analysis of the voltage signal when the early fault occurs and judges the type of the fault according to the indicator of voltage distortion. Literature [2] uses the index of the maximum value of the modulus after wavelet transformation of the induced current to judge the fault. Literature [3] uses the morphological gradient wavelet method to decompose the phase current into approximate signal and detail signal. The detail signal is used to identify overcurrent disturbances, and the approximate signal is used for fault signals. Literature [4] uses AE to extract the features of the collected signals, and then builds a fault classifier based on the GRU neural network for fault detection and identification. Literature [5] uses multi-scale wavelet transform to extract features of fault current, and then recognize early faults through the gray correlation method. Literature [6] uses wavelet analysis to extract the characteristics of the overcurrent signal and then constructs a convolutional neural network to identify the fault.
This article first uses SIMULINK in MATLAB to build a power system fault simulation model and obtains the overcurrent under different faults through this model, and then uses wavelet analysis to extract the characteristics of the overcurrent. Finally, various recognition algorithms are used to detect and identify different fault categories.

2. Fault classification

Different fault phenomena will occur due to aging, dampness, and damage of cables during use. Generally, cable faults are divided into ground faults, open-circuit faults, and short-circuit faults. The specific classification is shown in the figure below.

![Fault classification diagram](image)

In this paper, Simulink in MATLAB is used to build the model of common cable faults in power system. The output voltage on the power side is 13.8kv and the frequency is 50Hz. The line is 20km and the simulation model is shown in the figure 2.

![Simulation model](image)

By changing the phase to phase and phase to phase ground fault circuit breaker, the fault current and fault voltage of four fault types, namely single-phase ground fault, two-phase short circuit fault, two-phase ground fault and three-phase short circuit fault, are obtained. The current and voltage waveform are shown in the figure below.
When the time is 0.02 seconds, the A phase to B phase short circuit fault occurs suddenly, and the fault is removed when the time is 0.1 seconds.

When the time is 0.1 second, the A-phase ground fault occurs suddenly, and the fault is removed when the time is 0.4 second.

3. Wavelet analysis

Wavelet transform is the detailed analysis of the signal through the expansion and translation operation in the time dimension and the frequency dimension [7], which can realize the feature extraction of the signal in any segment of time and frequency, and is especially suitable for the analysis of unstable signals.

The basic function is compared with the research signal after translation and expansion, and then the local characteristics of the research signal can be obtained. The continuous wavelet transform is defined as:

\[
WT_x(m,n) = \frac{1}{\sqrt{m}} \int_{-\infty}^{\infty} \psi(t) \frac{(t-n)}{m} dt = (\psi, \tilde{x}(t))
\]

Where \( \psi \) is the wavelet base, \( m \) is the scale parameter, \( n \) is the displacement parameter, and then \( M \) and \( N \) are discretized to obtain the discrete wavelet change of \( x(t) \):

\[
WT_x(j, c) = \int_{-\infty}^{\infty} \psi(t) \Phi (j, c, t) dt \quad j, c \in \mathbb{Z}
\]

Where \( m=m_x/n_x \), the following equation can be obtained by decomposing the signal step by step:

\[
x(t) = \sum_{j=1}^{k} d_j(t) + e_k(t)
\]
4. Detection method

4.1. SVM nonlinear multi-classification\[8-9]\n
The fault current identification in this paper is a nonlinear multi-value classification problem. The one-to-one classification method in SVM has a fast calculation speed and high accuracy of the results. Therefore, this paper adopts the one-to-one classification method.

In the case of N-classification, \(0.5N \times (N-1)\) training sets can be obtained by combining the sample data in pairs. SVM is used to learn between any two kinds of samples, and the following constraints are satisfied.

\[
\begin{align*}
\sigma^y_j \times \Psi(x_i) + c^y_i & \geq 1 - \xi_i \times f \times k_i \quad i, j = 1, 2, ..., l \\
\sigma^y_j \times \Psi(x_i) + c^y_i & \leq 1 + \xi_i \times f \times k_i
\end{align*}
\]

Then the minimum value of the objective function is obtained.

\[
ojb = \min\left[\frac{1}{2} \left\| \sigma^y \right\|^2 + N \sum_j \xi_j (\sigma^y_j)^2 \right]
\]

Then the class i and class j samples in the second classifier are voted. When all the training sets are classified, the category with the largest number of votes will be counted.

4.2. Other recognition algorithms

1) BP neural network

BP neural network\[11\] consists of two parts: forward propagation and backpropagation. When propagating forward, the signal passes through the input layer, hidden layer, and output layer in sequence. If the error is less than the set threshold, the algorithm is completed; otherwise, backpropagation is performed. In the backpropagation, the error is calculated by the backpropagation, and the weight and threshold of the corresponding layer are calculated through the gradient descent method, and the error is reduced at the same time, to obtain the classification label.

2) Random forest

The classifier is usually composed of two or more classification models\[12\]. First, a part of the sample is separated from the original training data, and then a part of the sample is randomly selected from the separated sample as a variable and the node of the classification tree is determined. According to the node, the difference of the decision tree model in the next step can be determined. In the first part of the sample, a classification model is established, the classification results can be obtained, and finally, the number of classifications is determined by voting.

3) Adaboost

This classifier is a strong classifier composed of weak classifiers\[13,14\]. Combining multiple weak meta-structures becomes a powerful classifier, which can effectively perform pattern recognition.

5. Case analysis

5.1. Sample Data

In this example, the power cable fault simulation model is used to simulate different types of cable faults, and the sample data of four fault currents are obtained, including single-phase short-circuit fault, two-phase short-circuit fault, two-phase grounding fault and three-phase short-circuit fault. There are 60 samples of each type, a total of 240 samples, and each sample has 2000 data points.
The wavelet analysis box in MATLAB is used to extract the characteristics of the fault current obtained by the simulation. The wavelet fundamental wave is selected as the demy filter, and the sample data is decomposed into three layers. The waveform after the decomposition of phase A current of AB short circuit fault is shown in the figure 5.

5.2. Detection and Identification
Firstly, the obtained feature data are standardized, then 50 data samples are taken as the training set, and the remaining 10 are taken as the test set. After comparing with the actual number, the accuracy of the recognition algorithm is obtained.

By comparing the four methods of Adaboost, Random Deep Forest, BP Neural Network, and SVM, their respective accuracy and running time are obtained. Each method is tested 5 times, and the effect is as follows.

| Accuracy | M1   | M2   | M3   | M4   |
|----------|------|------|------|------|
| 1        | 0.954| 0.946| 0.975| 0.625|
| 2        | 0.963| 0.895| 0.975| 0.75 |
| 3        | 0.970| 0.932| 0.95 | 0.55 |
| 4        | 0.965| 0.8908| 0.95 | 0.60 |
| 5        | 0.997| 0.896| 0.925| 0.725|

| time (s) | M1   | M2   | M3   | M4   |
|----------|------|------|------|------|
| 1        | 0.087| 0.262| 15.35| 22.528|
| 2        | 0.065| 0.2259| 15.26| 22.403|
| 3        | 0.0619| 0.2349| 15.37| 22.243|
| 4        | 0.0609| 0.232| 15.17| 22.286|
| 5        | 0.0709| 0.2339| 15.19| 22.244|
Among them, M1, M2, M3, and M4 represent SVM, BP, Adboost, and RF algorithms respectively. It can be seen from the table that compared to the other three methods, SVM has the highest accuracy and faster efficiency, so SVM is the best method for pattern recognition.

6. Conclusion
In this paper, a cable fault detection method based on SVM and wavelet analysis is proposed. The samples of different kinds of cable fault overcurrent are obtained by Simulink simulation in MATLAB, and then the features of all samples are extracted by wavelet transform. Finally, the proposed method is used to identify cable faults, and the accuracy is 99.7%. Compared with other methods, classification and recognition are more accurate and more efficient.

References
[1] ZHANG Wenhai, XIAO Xianyong, ZHOU Kai, et al. Multi-cycle incipient fault detection and location for medium voltage underground cable[J]. IEEE Transactions on Power Delivery, 2017, 32(3): 1450-1459.
[2] ZHANG Chao, KANG Xiaoning, MA Xiuda, et al. On-line incipient faults detection in underground cables based on single-end sheath currents[C] // IEEE PES-Pacific Power and Energy Engineering Conference, October 25-28, 2016, Xi’an, China: 795-799.
[3] HUANG Mingchang WU Qing hua, HAN Xing lei, et al. Detection of incipient faults in distribution cables based on mathematical morphology[C]/2019 Ninth International Conference on Power and Energy Systems (ICPES),2019. Perth, Australia: IEEE,2019:1-6.
[4] Xu Zihong, Ji TIANYAO, Deng Weimin, Xia Houkai Shun. Cable early fault detection and identification method based on AE and Gru neural network [J]. Guangdong electric power, 2020,33 (09): 27-34.
[5] Yao Haiyan, Zhang Jing, Liu Yi, Pan Shu, Xu Bei, Li tiyin, Zhou Niancheng. Cable early fault detection and identification method based on multi-scale wavelet criterion and time-frequency feature correlation [J]. Power system protection and control, 2015,43 (09): 115-123.
[6] Wang Ying, sun Jianfeng, Xiao Xianyong, LU Hong, Yang Xiaomei. Cable early fault classification and identification based on optimized convolutional neural network [J]. Protection and control of power system, 2020,48 (07): 10-18.
[7] Mallat S.A theory for multiresolution signal decomposition: the wavelet representation [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1989,11(7):674-693.
[8] Boser B, Guyon I, Vapnik V. A training algorithm for optimal margin classifiers, Fifth Annual Workshop on Computational Learning Theory. Pittsburgh: ACM Press, 1992
[9] Cortes C, Vapnik V. Support-vector networks. Machine Learning, 1995, 20: 273～297
[10] Scho ikopf B, Burges C, Vapnik V. Extracting support data for a given task. In: Fayyad U M, Uthurusamy R(eds.). Proc. of First Intl. Conf. on Knowledge Discovery& Data Mining, AAAI Press, 1995, 262～267
[11] Lakner P. Optimal trading strategy for an investor: the case of partial information[J]. Stochastic Processes and Their Applications, 1998, 76: 77-97.
[12] Breiman L. Random Forests[J]. Machine Learning, 2001,45(1).
[13] P. Viola, M. Jones. Rapid object detection using a boosted cas-cade of simple features. IEEE Conf. Computer Vision and Pattern Recognition, Kauai, Hawaii, USA, 2001
[14] Y. Freund, R. E. Schapire. Experiments with a new boosting algorithm. In: Proc. the 13th Conf. Machine Learning. San Fran-rewcisco: Morgan Kaufmann, 1996. 148～156