Corrosion fault identification model of substation grounding grid based on PSO-LS-SVM

Jun Qin1, Jin Miao1, Xi Wu1, Zibo Wang1 and Shuai Shao2,3,4,5

1 State Grid Wuxi Power Supply Company, Wuxi, Jiangsu, China
2 Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang, China
3 Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China
4 Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang, China
5 E-mail: shaoshuai@sia.cn

Abstract. As one of the important equipment to ensure the normal operation of substation, the performance of grounding grid has been highly concerned. In recent years, researchers propose that the theory of electromagnetic induction can be used to diagnose and identify the corrosion fault of substation grounding grid. In this paper, a fault classification model based on LS-SVM optimized by PSO is proposed to identify the corrosion fault of grounding grid. Firstly, the wavelet packet transform principle is used to filter the original data, and the wavelet packet energy is used to construct the fault eigenvalue as the input of the fault classification model. Based on LS-SVM, a corrosion fault classification model is constructed, and particle swarm optimization method is used to optimize the parameters of the model, which solves the problem of traditional SVM parameter optimization. Through the practical application in substation, the model proposed in this paper can identify the corrosion of grounding grid conductor without excavation and power failure, which provides an effective scheme for engineering application.

1. Introduction

As one of the important equipment to ensure the normal operation of substation, the performance of grounding grid has been highly concerned. Because the grounding grid is deeply buried underground, it is highly susceptible to the influence of chemical, electrochemical and microbiological effects, resulting in corrosion and thinning. When it is serious, it even breaks, threatenning the normal operation of the substation and the safety of personnel. At present, in engineering, it generally relies on experience to find the corrosion section of grounding grid through large-area excavation. This kind of method has the advantages of heavy workload, long detection time, high cost and lack of pertinence [1]. Therefore, it is urgent to find a method to identify the corrosion fault of grounding grid without power cut and blind excavation.

In recent years, relevant experts have proposed that the electromagnetic induction theory can be used to diagnose and identify the corrosion fault of substation grounding grid [2-4]. Researchers can judge the corrosion of grounding grid by analyzing the excitation magnetic induction signal of grounding grid surface. However, due to the complex electromagnetic environment and other factors in the substation site, it is difficult to extract and identify the effective signal in this method, which leads to the lack of this method in practical engineering. In this paper, wavelet packet analysis method
is used to solve this problem. Through multi-scale approximation and refinement of the original signal, fault eigenvalue extraction is realized.

For fault signal analysis and recognition, common methods include expert system, decision tree, artificial neural network, deep learning and support vector machine, etc. Among them, the artificial neural network and deep learning scheme need a large number of samples for training, and the learning speed is slow. The portability of expert system is usually poor. The decision tree model is suitable for the processing of discrete features, but it is prone to over fitting, which affects the analysis effect. Support vector machines (SVM) is a classification model scheme based on statistical learning, using supervised machine learning algorithm, it can effectively solve the problems of small sample, nonlinear and high dimension. SVM is widely used in various fields, including human image recognition, text classification, biological information recognition and handwritten character recognition. The application effect of standard SVM model depends on the selection of kernel function and parameters. As the most commonly used parameter optimization method, cross test has the problems of slow speed and poor effect.

In view of the above problems, combined with the nonlinear characteristics of grounding grid corrosion data and the limited number of samples, this paper proposes a grounding grid corrosion fault identification method based on PSO-LS-SVM classification model. In the model, wavelet packet transform is used to filter the original electromagnetic induction data. Wavelet packet transform is used to extract the eigenvalue of weak signal to solve the problem of strong interference in substation. And the wavelet packet energy is used to construct the fault eigenvalue as the input of the fault classification model. A corrosion fault classification model is constructed based on LS-SVM to solve the problem of non-linearity and small sample of grounding grid detection signal. However, the traditional model still has the disadvantage of relying on experience to select kernel function parameters. Therefore, the particle swarm optimization method is used to filter the parameters of the LS-SVM classification model, which solves the problem of traditional SVM parameter optimization. Finally, the model of this paper is verified in a substation. The verification results show that this method can effectively identify the corrosion fault of grounding grid.

2. Feature extraction of the signal

Since the birth of wavelet transform, it has been widely used because of its outstanding performance in signal processing [5]. Wavelet transform is very important for signal processing. It uses time domain and frequency domain localization to realize multi-resolution analysis [6, 7]. However, due to the limitation of the algorithm, the wavelet transform only decomposes the low frequency part of the signal, but it cannot process and approximate the high frequency part of the signal more precisely. The wavelet packet transform is an improvement of the standard wavelet transform. It not only decomposes low-frequency signals, but also decomposes high-frequency signals more finely. By decomposing and reconstructing the signal, the wavelet packet transform maps the signals with different characteristics to different frequency bands, and realizes the fine analysis of the signal.

Taking three-layer wavelet packet transform as an example, the basic process of wavelet packet transform can be shown in Figure 1. S represents the original signal, H represents the low-frequency part of the signal, G represents the high-frequency part of the signal, and the number represents the number of layers of signal decomposition.

![Figure 1. The three-layer tree structure of wavelet packet transform.](image-url)
Then the decomposition of wavelet packet transform can be expressed by the Formula (1).

\[
\begin{align*}
    d_{j,2n}^{i} &= \sum_{k} h_{k,2^{-j}n} d_{k}^{i+1,n} \\
    d_{j,2n+1}^{i} &= \sum_{k} g_{k,2^{-j}n} d_{k}^{i+1,n}
\end{align*}
\]

Among it, \(d_{j,2n}^{i}\), \(d_{j,2n+1}^{i}\) and \(d_{j,2n+1}^{i+1,n}\) are wavelet packet decomposition coefficients. \(h_{k,2^{-j}}\) and \(g_{k,2^{-j}}\) are the low-pass filter and the high-pass filter coefficients of wavelet packet decomposition. The essence of wavelet packet transform decomposition is to use conjugate quadrature filter to decompose the signal into different frequency bands, and use different frequency bands to characterize the different characteristics of the signal.

The reconstruction process of wavelet packet transform is the process of using \(d_{j,2n}^{i+1,n}\) and \(d_{j,2n+1}^{i+1,n}\) to solve \(d_{j,2n}^{i+1,n}\). The recursive formula is as Formula (2).

\[
    d_{j,2n}^{i+1,n} = \sum_{k} \left[ h_{k,2^{-j}} d_{k,2n}^{i} + g_{k,2^{-j}} d_{k,2n+1}^{i+1,n} \right]
\]

\(h_{k,2^{-j}}\) and \(g_{k,2^{-j}}\) are the low-pass filter and high-pass filter coefficients of wavelet packet reconstruction.

Through the reconstructing coefficients of the wavelet packet decomposition, the wavelet coefficients of corresponding frequency bands of each layer can be obtained, and the energy of each frequency band signal can be calculated. In this paper, we use the band energy to represent the fault characteristics. So we select the band which is closely related to the fault in the analysis and calculate its energy value. The frequency band energy calculation formula is as follows.

\[
    E_{m,j} = \int |H_{j}(\omega)|^{2} dt = \sum_{n=0}^{N} |d_{j,m,n}^{j}|^{2}
\]

\(E_{m,j}\) is the energy corresponding to the band signal \(j\) of layer \(m\). \(d_{j,m,n}^{j}\) is the wavelet packet coefficient of the band signal \(j\) of layer \(m\). \(N\) is the number of wavelet packet coefficients contained in the current band.

The frequency band energy value is used to construct fault feature vector.

\[
    E = \left[ E_{m,0}, E_{m,2}, \cdots, E_{m,2^{j-1}} \right]
\]

We use wavelet packet transform to decompose and reconstruct the original signal, extract the energy signals in different frequency bands to construct the fault feature vector, and use it as the input signal of fault identification.

3. Corrosion fault identification of LSSVM based on PSO optimization

After extracting the characteristic data of corrosion fault, this paper constructs an improved PSO-LS-SVM classification diagnosis model to classify the characteristic data and realize the recognition of grounding grid corrosion fault [8].

3.1. Fault diagnosis model based on LS-SVM

Support vector machine is one of the common methods in fault classification, which has good robustness and generalization ability [9, 10]. The essence of support vector machine is to find an optimal hyperplane, which divides the input data into two classes, and keep the maximum distance between classes. The principle of SVM is shown in the Figure 2.

Least squares support vector machines (LS-SVM) is an improvement of SVM [11-13]. It uses the equality constraints to replace the inequality constraints in the standard SVM algorithm. It transforms the quadratic programming problem into a system of linear equations and solves it, thus reducing the difficulty of solving the problem. Compared with the standard SVM algorithm, LS-SVM has better advantages in solving small sample and nonlinear problems.
Suppose the sample signal to be processed is \((x_i, y_i)\), where \(x_i\) is the input and \(y_i\) is the output of \(x_i\). Then the following boundary equation can be constructed.

\[
y_i = \omega^T \varphi(x_i) + b
\]  (5)

Starting from the loss function, LS-SVM optimizes the Function (5) with the help of two norm. At the same time, some outliers may appear in the region between \(H_1\) and \(H_2\) in Figure 2, so the relaxation variable \(\xi_k\) is introduced, and \(\xi_k > 0\). The following objective functions and constraints can be derived.

\[
\begin{align*}
\min J(\omega, \xi) &= \frac{1}{2} \|\omega\|^2 + \gamma \sum_{k=1}^{N} \xi_k^2 \\
\text{s.t. } y_k \left[\omega^T \varphi(x_k) + b\right] &= 1 - \xi_k, \quad k = 1, \ldots, N
\end{align*}
\]  (6)

Among it, \(\gamma\) is the penalty parameter, which is a constant greater than zero. It is used to limit the complexity and generalization of the model.

The LS-SVM equation can be simplified by using Lagrange function.

\[
L(\omega, b, \xi, \alpha) = J(\omega, \xi) - \sum_{k=1}^{N} \alpha_k \left[ y_k \left[\omega^T \varphi(x_k) + b\right] - 1 + \xi_k \right]
\]  (7)

\[
= \frac{1}{2} \|\omega\|^2 + \gamma \sum_{k=1}^{N} \xi_k^2 - \sum_{k=1}^{N} \alpha_k \left[ y_k \left[\omega^T \varphi(x_k) + b\right] - 1 + \xi_k \right]
\]

\(\alpha_k\) is a Lagrange multiplier, then the original optimization problem is transformed into solving the single parameter \(\alpha_k\). The partial derivatives of \(\omega\), \(b\), \(\xi_k\) and \(\alpha_k\) are 0.

\[
\begin{align*}
\frac{\partial L}{\partial \omega} &= 0, \quad \frac{\partial L}{\partial b} = 0, \quad \frac{\partial L}{\partial \xi_k} = 0, \quad \frac{\partial L}{\partial \alpha_k} = 0
\end{align*}
\]  (8)

Finally, the LS-SVM classification decision equation is obtained by simplification.

\[
y(x) = \text{sign} \left[ \sum_{k=1}^{N} \alpha_k y_k K(x, x_k) + b \right]
\]  (9)

\(K(x, x_k)\) is the kernel function, which can map the low dimensional space input to the high dimensional space. It could find the optimal classification hyperplane in the high dimensional space. The commonly used kernel functions include linear kernel function, polynomial kernel function, radial basis function and sigmoid kernel function, etc. Considering the computational complexity and nonlinear approximation, radial basis kernel function is selected.
\[ K(x, x') = \exp\left(-C|x - x'|^2\right) \]

\( C \) is the parameter of the kernel function.

### 3.2. Parameter optimization based on PSO

Through the above derivation, we know that the actual classification effect of LS-SVM classification decision equation is mainly determined by the penalty parameter \( \gamma \) and kernel function parameter \( C \).

Traditional parameter selection mostly uses cross validation method, which has strong experience and general effect [14]. In order to improve the accuracy of fault diagnosis, the SVM model based on artificial bee colony (ABC) algorithm optimization appears, but the algorithm is easy to fall into the local optimal solution when optimizing parameters [15]. Some scholars put forward LS-SVM model based on genetic algorithm (GA) optimization to predict the corrosion rate of grounding grid, but it has strong randomness and blindness, and the efficiency is not good [16].

In this paper, particle swarm optimization algorithm is used to optimize the parameters of LS-SVM classification equation. Particle swarm optimization algorithm has the advantages of fast search speed, simple in structure and easy to implement in engineering [17, 18]. PSO has the characteristics of robustness and parallelism, and has strong global search ability. And this search ability does not depend on specific solution model. Therefore, PSO algorithm is used to optimize parameters automatically.

The parameter optimization process of LS-SVM based on PSO is shown in Figure 3.

![Figure 3. The process of LS-SVM parameter optimization based on PSO.](image)

1) The particles in the particle swarm represent the parameter combination \((\gamma, C)\) of LS-SVM, and initialize the parameter optimization interval;
2) The initial particle swarm size is \(N\);
3) Randomly initialize the speed and position of each particle;
4) Calculate the fitness function of the particles;
5) According to the fitness function, the optimal position of individual particle and the global optimal position are calculated and updated;

6) If the optimal solution is obtained within the number of iterations, the LS-SVM parameter optimization result will be output. If the optimal solution is not obtained, the speed and position of some particles are initialized with probability P. At the same time, it performs a linear search in the negative direction of the negative gradient with probability (1-P) according to the gradient information, and determines the particle moving step. Then it performs step 4).

4. The test experiment in substation

In order to verify the actual effect of this method, we choose a 330kV high voltage substation for field verification, which has been running for more than 15 years. Based on the condition of the substation, we choose a group of underground leads located on the surface diagonally as the current injection and extraction locations, and use a multi-strand copper core cable with a diameter of $\Phi=16\text{mm}^2$ to inject the different frequency excitation current into the tested grounding grid. The connection method of the experimental device is shown in Figure 4.

![Figure 4. Schematic diagram of experimental device connection.](image)

The transformation experiment is shown in Figure 5. During the experiment, a mobile electromagnetic induction signal acquisition device is used to collect the excitation electromagnetic induction intensity signal of the grounding grid. In the process of signal acquisition, by observing the numerical change of the electromagnetic induction intensity, the topological distribution of the grounding grid conductor can be judged. Adjust the moving direction of the acquisition device, we could keep the acquisition device moving along the direction of strong signal distribution, that is, we always keep it moving directly above the grounding grid conductor. After the collection is completed, the electromagnetic induction data is processed by wavelet packet transform to form the eigenvalue vector of corrosion fault. The eigenvalues are input into the fault recognition model, the eigenvalues are analyzed and processed, and finally the analysis results are obtained.

![Figure 5. Corrosion detection of grounding grid in a 330kV power station.](image)

In the data analysis results of this test, it is found that there is a corrosion alarm in the tested area. During the power outage of substation, the manual excavation detection was carried out at this
location. It was found that the surface of grounding grid conductor at this location was seriously rusted, and the phenomenon of thinning and thinning could be observed at the local location of the conductor. It is proved that this method can identify the corrosion fault of grounding grid.

5. Conclusions
Based on the principle of electromagnetic induction, this paper designs a model of substation grounding grid corrosion fault identification based on PSO-LS-SVM. In this model, the wavelet packet transform theory is used to extract the feature vector of corrosion fault, and the LS-SVM is used to create the corrosion diagnosis model. In order to improve the accuracy and generalization of the model, we use PSO method to optimize the model parameters. Through the field test of the grounding grid in a 330KV substation, it is proved that the theory proposed in this paper can refine and approximate the high frequency part of the original signal, which is beneficial to the extraction of fault features. The LS-SVM fault diagnosis model optimized by PSO not only balances the efficiency and accuracy, but also solves the problem of small samples of grounding grid corrosion. The model has good recognition function for early corrosion. In the future, we will continue to study the equipment miniaturization scheme to solve the problems encountered by workers on site.

Acknowledgements
The authors gratefully acknowledge the anonymous reviewers for their constructive comments. This work is supported by State Grid Corporation Science and Technology Project (No. SGJSWX00KJJS2002066).

References
[1] Zhang S L, Liu D L, Xu H F, Song W L and Wang L 2020 Detection and evaluation method of breakpoint and corrosion magnetic field of substation grounding grid Insulators and Surge Arresters 05 190-193
[2] Liu Y, Cui X, Zhao Z B, Qi L and Chen J J 2009 Method of corrosion diagnosis of substations grounding grids based on electromagnetic induction theory Proceedings of the Chinese Society for Electrical Engineering 29(4) 97-103
[3] Liu Y, Cui X, Zhao and Z B 2010 Method of structure estimation and fault diagnosis of substations grounding grids Proceedings of the CSEE 30(24) 113-118
[4] Wang F H, Wang S J, Liu Y D, Huang R H and Zhang H Z 2016 Comparison of substation grounding grid fault diagnosis results using surface potential and magnetic induction intensity High Voltage Engineering 42(07) 2281-2289
[5] Wu F, Yang Z and Liu Y 2017 Current similarity based open-circuit fault diagnosis for induction motor drives with discrete wavelet transform Microelectronics Reliability 75 309-316
[6] Awal A, Sheikh S M, Mohiuddin A and Mohd A R 2014 An adaptive level dependent wavelet thresholding for ecg denoising Biocybernetics and Biomedical Engineering 34(4) 238-249
[7] Liu Z H, Miao X G and Tang B Y 2019 Improvement of speckle noise suppression algorithm for sar image based on wavelet transform Robot Technique and Application 04 43-45
[8] Wang H Y, Li J H and Yang F L 2014 Overview of support vector machine analysis and algorithm Application Research of Computers 31(05) 1281-1286
[9] Zhang X G 2000 Introduction to statistical learning theory and support vector machines Acta Automatica Sinica 01 36-46
[10] Vladimir C and Yunqian M 2004 Practical selection of svm parameters and noise estimation for svm regression Neural Networks 17(1) 113-126
[11] Lawi A and Aziz F 2018 Classification of credit card default clients using ls-svm ensemble 2018 Third International Conference on Informatics and Computing (ICIC) (Palembang, Indonesia) 1-4
[12] Jafar N, Thamrin S A and Lawi A 2016 Multiclass classification using least squares support vector machine 2016 International Conference on Computational Intelligence and
[13] Lawi A and Adhitya Y 2018 Classifying physical morphology of cocoa beans digital images using multiclass ensemble least-squares support vector machine *Journal of Physics: Conference Series* **979**

[14] Kecman V 2015 Iterative k data algorithm for solving both the least squares svm and the system of linear equations *SoutheastCon 2015* (Fort Lauderdale FL USA) 1-6

[15] Liu Y G and Chen C 2019 Corrosion rate prediction model of grounding grid based on support vector machine optimized by artificial bee colony algorithm *Electric Power Automation Equipment* **39**(5) 182-186

[16] Huang H, Liu Y C and Gao X 2019 Corrosion rate prediction of grounding grid based on combined model *Computing Technology and Automation* **38**(3) 79-83

[17] Kennedy J and Eberhart R 1995 Particle swarm optimization *IEEE International Conference on Neural Networks* **4** 1942-1948

[18] Wang D M, Lu C H, Jiang W W, Xiao M X and Li B R 2015 Study on pso-based decision-tree svm multiclass classification method *Journal of Electronic Measurement and Instrumentation* **29**(04) 611-615