Power quality disturbance detection method based on optimized kernel extreme learning machine

Lin Xu¹, Chun Zhao², Lisha Guo², Jiayu Xiong¹, Chang Liu¹, Zhao Wang², Zhen Wei², Bo Liu²

¹Electric Power Research Institute State Grid Sichuan Electric Power Company, Chengdu, China
²Wuhan NARI Limited Liability Company State Grid Electric Power Research Institute Liability Corporation, Wuhan, Nanjing, China

Abstract. In order to improve the accuracy of rapid detection of power quality, a power quality disturbance (PQD) classification method based on kernel-based extreme learning machine (KELM) is proposed, and chaos optimization is used to improve the global optimization performance of the particle swarm algorithm. This method first uses KELM to establish a classification model, and then uses an improved chaotic particle swarm optimization (CPSO) to optimize the parameters of KELM. Comparative analysis of example simulation results shows that the algorithm has higher classification accuracy and improves the reliability of power quality disturbance detection.

1 Introduction

With the large-scale penetration of non-traditional main units such as new energy sources and power electronic equipment into the power grid, power quality issues have become increasingly prominent[1]. Power quality disturbances usually manifest as sudden changes in the amplitude, phase, and frequency of voltage or current, including sags, harmonics, oscillations, and other forms of disturbances[2]. Correct identification of power quality disturbances is the basis for power quality analysis, such as power quality disturbance (PQD) location, disturbance data correlation analysis, dynamic load modeling, and comprehensive assessment of voltage sags.

With the proposal and development of theories such as neural networks, genetic algorithms, and fuzzy theory, artificial intelligence has been applied to research on PQD classification[3]. Artificial neural networks (ANNs), support vector machines (SVMs), etc., are commonly used methods for PQD detection and analysis[4]. They have achieved certain results in practical applications, but they also have shortcomings. For example, the ANN method requires a large number of parameters for transformer fault diagnosis training, and its training speed is slow. If the learning time is too long, it is easy to fall into a local optimum, hindering the desired learning outcome. The output result of the SVM-based classification method is a hard segmentation boundary, and selecting the appropriate kernel function can be challenging[5]. Complex disturbance identification of classified power quality is prone to misjudgment. The single hidden layer feedforward neural network algorithm learned through gradient descent can easily converge to local extreme values and requires a large number of training samples to achieve satisfactory classification results[6]. Additionally, training neural networks may lead to overfitting, resulting in unsatisfactory test results.

To address the aforementioned problems, this paper proposes a rapid detection method of power quality using chaotic particle swarm optimization (CPSO) with kernel-based extreme learning machine (KELM). Comparative analysis of example simulation results demonstrates that the algorithm achieves higher classification accuracy.

2 Methodology

2.1 Kernel extreme learning machine

ELM is a single hidden layer feedforward neural network. It improves the training speed of the algorithm by replacing the iterative process during network training with randomly generated weights and thresholds. This approach reduces the need for parameter settings. Additionally, it avoids the issue commonly encountered with gradient descent-based training, where the algorithm can easily converge to local extreme values. The structural model of ELM is illustrated in Figure 1.
where $W$ is the connection weight, $B$ is the threshold, $G$ is the nonlinear activation function, $\beta$ is the weight of the output link, $X$ is the input matrix, $H$ is the hidden layer, and $T$ is the output expectation value. The objective of ELM is to select appropriate parameters in order to minimize the difference between the actual output and the expected value. The model can be expressed as:

$$f(X) = h(X)\beta = H\beta$$

where $f(X)$ is the category vector output result, $h(X)$ and $H$ are the matrices of hidden layer feature mapping.

The calculation is as shown in equation (2):

$$\beta = H^T (HH^T + I/\lambda)^{-1}T$$

where $I$ is the unit diagonal matrix, $\lambda$ is the regularization coefficient.

To avoid the random generation of weights and thresholds in the ELM algorithm, the kernel function is introduced to replace the original hidden layer output matrix, resulting in the establishment of KELM. By utilizing the advantages of the kernel function, the algorithm ensures efficient computation speed. The calculation process is as follows:

$$K(\mu, \nu) = \exp\left(-\frac{||\mu - \nu||^2}{2s^2}\right)$$

where $s$ is the width of RBF, $\nu$ is the center of RBF.

When $\nu$ is set to 0.5 and $s$ takes on values of 0.1, 0.3, 0.5, and 0.7 respectively, the characteristic curve of the radial basis kernel function shown in Figure 2 is obtained. As observed from Figure 2, sample data that are in proximity to the test point are more strongly influenced by the radial basis kernel function. This influence increases as the radius of the kernel function decreases, indicating its effectiveness in extracting local features of the samples.

The analysis reveals that the classification accuracy of KELM is influenced by the regularization coefficient $\lambda$ and the setting of the kernel function parameter $s$, which can lead to the algorithm easily converging to a local minimum. The incorporation of the kernel function renders the KELM algorithm highly sensitive to the parameter setting $(\lambda, s)$. Furthermore, the range of $(\lambda, s)$ that enables the KELM algorithm to achieve optimal generalization capabilities is very narrow. Chaos possesses characteristics such as randomness, ergodicity, and sensitivity to initial conditions, which can help prevent convergence to local extrema. This paper addresses the parameter selection issue in optimizing KELM by combining chaos theory and the particle swarm algorithm.

### 2.2 Chaotic particle swarm optimization

PSO is an evolutionary computing algorithm based on population search. The particles maintain their positions while flying within the search range and aim to find the best value among all positions in order to achieve the search for the optimal solution in multi-dimensional space. If a particle swarm consists of $M$ random particles with a population size of $M$ and a dimension of $D$, let $x_i = (x_{i1}, x_{i2}, \ldots , x_{iD})$ represent the position of the $i$th particle, which is a random solution to the optimization problem. Through iteration, the algorithm aims to find the optimal solution. The particle updates its position by tracking two extreme values: the local optimal position $pbest$ of the particle and the optimal position $gbest$ of all particles, as shown in equations (4) and (5):

$$V_{i,d}^{k+1} = V_{i,d}^k + c_1 \times r_1 \times (pbest_{i,d} - x_{i,d}^k) + c_2 \times r_2 \times (gbest - x_{i,d}^k)$$

$$x_{i,d}^{k+1} = x_{i,d}^k + V_{i,d}^{k+1}$$

where $V_{i,d}$ is the speed of the particle, $pbest_{i,d}$ is the local optimal position of the particle swarm, $G$ is the global optimal position of the particle swarm, $\mu$ is the current algebra, $c_1$ and $c_2$ are acceleration constants, $x_{i,d}^k$ is the current position of the particle. $i=1,2,\ldots , M, d=1,2,\ldots , D, r_1$ and $r_2$ are random numbers in $(0,1)$.

When PSO is applied to discretized optimization problems, it is prone to getting stuck in local extrema, leading to decreased convergence speed and accuracy in the later stages of evolution. To overcome this limitation, chaos search is introduced. Chaos search traverses all states within a given range by following its own rules, ensuring that it avoids getting trapped in local optima.

To enhance convergence, the particle swarm velocity update formula after introducing the inertia weight factor $\omega$ is:

$$V_{i,d}^{k+1} = \omega \times V_{i,d}^k + c_1 \times r_1 \times (pbest_{i,d} - x_{i,d}^k) + c_2 \times r_2 \times (gbest - x_{i,d}^k)$$

### 3 Power quality disturbance identification

This paper refers to the IEEE Std and utilizes the multi-classification capability of the KELM algorithm to address the multi-classification problem of PQD. Table

---

![Fig. 2. Radial basis kernel function characteristic curve diagram](image-url)
1 presents the corresponding state codes for the seven disturbance signals, namely swell, sag, interrupt, flicker, harmonics, pulse transient, and oscillating transient.

<table>
<thead>
<tr>
<th>Status category</th>
<th>Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swell</td>
<td>(0,0,0,0,0,1)</td>
</tr>
<tr>
<td>Sag</td>
<td>(0,0,0,0,1,0)</td>
</tr>
<tr>
<td>Interrupt</td>
<td>(0,0,0,0,0,0)</td>
</tr>
<tr>
<td>Flicker</td>
<td>(0,0,1,0,0,0)</td>
</tr>
<tr>
<td>Harmonics</td>
<td>(0,0,0,0,0,0)</td>
</tr>
<tr>
<td>Pulse transient</td>
<td>(0,0,0,0,0,0)</td>
</tr>
<tr>
<td>Oscillating transient</td>
<td>(0,0,0,0,0,1)</td>
</tr>
</tbody>
</table>

A total of 350 sets of power quality disturbance sample data are utilized in this article, with 50 sets dedicated to each type of disturbance. Among the 350 sample data, 280 sample sets are used for training cross-validation, while the remaining 70 sets are used for validation. To evaluate the effectiveness of the algorithms proposed in this chapter, namely the KELM algorithm, PSO-KELM algorithm, and CPSO-KELM algorithm, simulation verification tests are conducted on the 70 sets of sample data. The results are presented in Figures 3-5.

In Figures 3-5, the blue "o" indicates the actual classification output result, while the red "*" indicates the corresponding predicted output result of each algorithm. The analysis reveals that among the 70 sets of test sample data, the KELM algorithm has a total of 13 errors, the PSO-KELM algorithm has 8 errors, and the CPSO-KELM algorithm proposed in this article exhibits the lowest error rate with only 6 errors. Furthermore, the CPSO-KELM algorithm achieves the highest accuracy rate among the three algorithms.

To verify the classification performance of the CPSO-optimized KELM fault diagnosis method, the same fault sample data is compared using three different methods: CPSO-KELM, ELM, and SVM. For ELM, the default Sigmoid function is utilized, and the number of hidden layer nodes is set to 100. Each result represents the average of 10 experiments. The perturbation classification accuracy (Acc) comparison curve is depicted in Figure 6.

As can be seen from Fig. 6, the overall performance of CPSO-KELM is more stable compared with ELM and SVM.

### 4 Conclusions

This paper combines chaos theory, PSO, and nuclear extreme learning machine to optimize KELM using CPSO and applies it to PQD classification. In this method, chaos theory is initially integrated with PSO to enhance the local search capabilities of PSO. The resulting mixed CPSO is then employed in the training of the neural network to optimize the parameters of
KELM. This optimization aims to improve the network structure of KELM and enhance the reliability of classification. Additionally, the classification results of SVM, ELM, and the proposed CPSO-optimized KELM method are compared and analyzed using actual data. The experimental results demonstrate that the PQD classification method based on CPSO-optimized KELM proposed in this paper exhibits better learning ability and higher accuracy.

This work was supported by the Science and Technology Project of State Grid Corporation of China (Research and application of fusion perception, intelligent diagnosis, and service enhancement technologies at the end of rural power network) under grant No. 52199922000M.

References