Power Transformer Fault Diagnosis Based on MPSO-SVM

Zhiqiang Yang
College of information engineering, Nanchang University, Nanchang 330031, Jiangxi, China

Abstract — This paper builds the power transformer diagnosis model based on the improved particle swarm optimization—support vector machine (MPSO-SVM). Speed update and particle self-adaptation self-variation are introduced to optimize the standard particle swarm algorithm, thus overcoming the defect of the standard particle swarm optimization algorithm and increasing the power transformer fault diagnosis accuracy rate of SVM. Through the analysis of the relationship between the transformer fault and the dissolved gas, the volume content of the dissolved gas of the transformer is adopted as the fault feature index. Through the experiment numerical analysis, results suggest that: the test sample recognition accuracy of the model parameters acquired by MPSO-SVM is higher than that acquired by the standard PSO by 17.86%.

Keywords - improved PSO, SVM (support vector machine), power transformer, fault diagnosis

I. INTRODUCTION

Power transformer is a critical part to the power system [1, 2]. A correct diagnosis of power transformer faults, especially latent faults of large-scale oil immiscible power transformer, is of vital significance to improve operation safety and reliability of the power system. With the rapid development of China’s national power industry, the power system has been widely applied to various fields. However to ensure safe operation of the power system has become an important measure to safety people’s life safety and national property safety. As a core part of the power system, the transformer can cause great inconvenience to people’s life or even threaten humans’ life and property safety once going wrong. At the same time, the regular check of transformers can result in the waste of human and material resources, and even introduce new safety dangers [3-4]. Besides, there are many causes for transformer faults [5-6], such as circuital factors, magnetic circuit factors, physical factors, chemical factors or even artificial factors. All these factors might not only bring transformer faults, but also cause interactive influence and increase the complexity of fault causes.

Currently, characteristic gas method, four-ratio method and three-ratio method formed based on dissolved gas analysis (DGA) are typical methods to diagnose power transformer faults. However, there are defects including erroneous judgment, wrong judgment, low diagnosis accuracy and so on. Consequently, the traditional diagnosis methods cannot meet the diagnosis requirements.

Concerning the above problems, this paper focuses on studying the intelligent fault diagnosis system based on the support vector machine (SVM) model. The improved particle swarm optimization (MPSO) is introduced to conduct parameter optimization the SVM model so as to overcome the defect of its quick convergence and easy trapping into the local extreme, increase the model’s fault diagnosis accuracy and provide theoretical guidance for the state evaluation and maintenance strategies of transformers.

II. PARAMETER OPTIMIZATION OF THE SVM MODEL BASED ON MPSO

A. SVM model

The SVM model is built on the VC dimension theory and the SRM principle. It attempts at acquiring the best generalization ability by seeking the optimal compromise according to the complexity of the limited sample information in terms of learning ability and model. Its basic idea can be interpreted with the optimal classification shown in Fig. 1. The solid points and the empty points in Fig. 1 stand for different data samples. H stands for the optimal classification hyperplane; H1 and H2 stand for the data sample lines in parallel to the classification line and nearest to the classification line, respectively. The distance between H1 and H2 is called the class interval.

Figure 1. Optimal classification plane schematic diagram

\[
\begin{align*}
\omega^T x_i + b & \geq 1, \quad y_i = +1 \\
\omega^T x_i + b & \leq -1, \quad y_i = -1
\end{align*}
\]

In other words,

\[
y_i (\omega^T x_i + b) \geq 1, i = 1,2,\ldots,N
\]

The corresponding decision-making function is shown in Eq. 3 below:

\[
y(x) = \text{sgn}[\omega^T x + b]
\]
f(x) = sign(α^T x + b)  \quad (3)

The optimal classification plane is in fact the superplane meeting the minimization of Eq. 4 below.

\[ \Phi(\alpha) = \| \alpha \|^2 \quad (4) \]

When data are linearly indivisible, the above models can compromise the fewest wrongly-divided samples and the maximum classification interval. Therefore, the model introduces \( \xi \) and the model is transformed into the quadratic programming problem with the constraint condition. (See Eq. 5 below)

\[
\min \Phi(\alpha, \xi) = \frac{1}{2} \| \alpha \|^2 + C \sum_{i=1}^{N} \xi_i \\
\text{subject to} \quad \begin{cases} 
    y_i (\alpha^T x_i + b) \geq 1 - \xi_i \\
    \xi_i \geq 0, i = 1, 2, \ldots, N
\end{cases} \quad (5)
\]

Where, \( \xi_i \) stands for the slack variable, which can control the wrong division rate to some extent; \( C \) for the model’s punitive factor, which can control the punishment degree on wrong samples.

B. Standard PSO and MPSO

When SVM is used for fault diagnosis, the SVM kernel function type should be clarified first, and parameters and punitive parameters of the kernel function should be chosen. If the proper parameters are chosen, the diagnosis effect of the SVM model can greatly improve. Therefore, this paper conducts parameter optimization of the SVM model through the PSO algorithm [11] so as to obtain the optimal fault diagnosis accuracy rate. However, the PSO algorithm is a heuristic optimization calculation model. It assumes that there is a particle swarm which can contain \( M \) particles; the number of dimensions of the searching space for the particle swarm is \( D \); and the status attribute value of Particle \( i \) at the \( t \) moment is calculated in Eq. 6 to Eq. 8, respectively:

1: Position status:
\[
X'_i = (X'_{i1}, X'_{i2}, X'_{i3}, \ldots, X'_{in})' \quad (6)
\]

Where, \( X'_{\text{min}} \) stands for the lower limit of the coordinate position; \( X'_{\text{max}} \) for the upper limit of the coordinate position.

2: Velocity status:
\[
V'_i = (V'_{i1}, V'_{i2}, V'_{i3}, \ldots, V'_{in})' \\
V'_{id} \in (V'_{\text{min}}, V'_{\text{max}}) \quad (7)
\]

Where, \( V'_{\text{min}} \) stands for the lower limit of the velocity; \( V'_{\text{max}} \) for the upper limit of the velocity.

3: Individual optimal position:
\[
P'_i = (P'_{i1}, P'_{i2}, P'_{i3}, \ldots, P'_{in})' \quad (8)
\]

4: Global optimum position:
\[
P'_g = (P'_{g1}, P'_{g2}, P'_{g3}, \ldots, P'_{gn})' \quad (9)
\]

All the above is the status attribute value of the particle at the \( t \) moment; while the status property of the particle at the “\( t+1 \)” moment can be updated and iterated through Eq. 10.

\[
V'_{id} = wV'_{id} + c_1r_1(P'_i - X'_{id}) + c_2r_2(P'_g - X'_{id}) \\
X'_{id} = X'_{id} + V'_{id} \quad (10)
\]

Where, \( w \) stands for the inertia weight value of the PSO algorithm; \( c_1, c_2 \) for the acceleration constants of the PSO algorithm; \( r_1, r_2 \) for the random variables and obey the even distribution within the region \((0, 1)\). When the status attribute value of the particle at the \( t \) moment is iterated into that at the \( “t+1” \) moment, there contain three core parts. First, the velocity of the particle of the former moment, meaning that the velocity of the particle in the latter moment keeps the velocity in the former moment; second, self-cognition, meaning that the particle will take its flying experiences into consideration and adjust its flying state according to the swarm’s optimal position; third, the social factor, meaning that the particle will also refer to the flying information of the whole population apart from referring to its own experiences, and adjust its flying state according to the swarm’s optimal position.

2.1 Linear declination of the iteration factor based on the particle velocity

Based on the above standard PSO, it can be seen that, during the iteration process, \( w \) stands for the inertial weight value of the PSO algorithm. The higher the weight value is, the stronger the global searching capability is. When the weight value is relatively small, it has a relatively strong local searching capability. Therefore, in the standard PSO algorithm, \( w \) is set at a fixed value, and its algorithm convergence is relatively poor, which might fail to get the optimal solution. Therefore, in order to improve the convergence of the standard PSO algorithm and endow the algorithm with a good global searching capability in the early operation period and a good local searching capability in the latter operation period, \( w \) gradually changes with the model’s iteration number. The iteration equation of \( w \) is shown in Eq. 11 below:

\[
\]
w(t) = w_2 + (w_1 - w_2) \frac{T_t - t}{T_t} \tag{11}

Where, \( w_1 \) stands for the initial inertial weight value; \( w_2 \) for the inertial weight value at the beginning and the end; \( T_t \) for the iteration number; \( T \) for the maximum iteration number.

Besides, from the standard PSO, it can be seen that, when \( c_1, c_2 \) are relatively small, the particle can partially adjusted within the optimal target region; when \( c_1, c_2 \) are relatively large, particles far away can move quickly to the target region. Therefore, proper adjustment of the value of \( c_1, c_2 \), can help the model stay close to the optimal value more easily. The iteration equation for the acceleration constants, \( c_1, c_2 \), is shown in Eq. 12:

\[
c_1(t) = c_{1i} + (c_{1f} - c_{1i}) \frac{t}{T_t}
\]

\[
c_2(t) = c_{2i} + (c_{2f} - c_{2i}) \frac{t}{T_t}
\]

Where, \( c_{1i} \) and \( c_{2i} \) are initial acceleration constants; \( c_{1f} \) and \( c_{2f} \) are final acceleration constants; \( t \) is the iteration number; and \( T_t \) is the maximum iteration number.

To sum up, the PSO velocity and position update equation based on the linear declination of the velocity iteration is shown below: (See Eq. 13)

\[
V_{id}^{t+1} = w(t)V_{id}^t + c_1(t)\eta_1(P_{id}^t - X_{id}^t) + c_2(t)\eta_2(P_{gd}^t - X_{id}^t)X_{id}^{t+1} = X_{id}^{t+1} + V_{id}^{t+1}
\]

\[
\eta_1, \eta_2 \sim \text{gauss}(0,1).
\]

2.2.2 Variation based on the particle self-adaptation

From the standard PSO, it can be seen that all particles stay close to the optimal position, \( P_g \). When the optimal position becomes a local optimal point, the standard PSO cannot re-search in the solution space and becomes trapped in the local optimal solution. Therefore, it is necessary to conduct variation operation of the global optimal position, \( P_g' \), change the forward direction of parameters during the iteration process, and enter the other regions to keep on searching for the global optimal solution. Certain dimension is randomly chosen for the variation operation. The variation probability, \( P \), is shown in Eq. 14 below:

\[
p = \begin{cases} k \sigma^2 < \sigma_0^2, f(P_g') > f_g \\ 0 \end{cases}
\]

Where, \( k \) is a random value within the section of \([0, 1, 0.3]\); \( \sigma^2 \) for the group fitness variance; \( f_g \) for the optimal target solution. The particle’s position variation is shown in Fig. 15 below, where \( \eta_p \) is a random variable with the section of gauss(0,1).

\[
P_g' = (P_g^1, P_g^2, P_g^3, ..., P_g^d)^T
\]

\[
P_g' = P_g^d (1 + 0.5\eta_p), i = 1, 2, ..., d
\]

III. MODEL CONSTRUCTION AND RESULT ANALYSIS

A. Relationship between the transform internal fault and the dissolved gas

The internal structure of the oil immersible power transformer is complex, and its fault occurrence rate is higher than that of other power devices. However, lots of research results suggest that DGA can find the latent faults and their development degree of power transformers at an earlier date. According to relevant literatures and judgment standards, the fault-related gases of the internal insulated materials of transformers form bubbles, which are dissolved into the oil through convection, propagation and dissolution. The composition and content of these fault-related gases are closely related to the type and severity degree of faults. It is one of the most efficient and widely-accepted method for the fault diagnosis of transformers [12].

The content of the major characteristic gases and the secondary characteristic gases is different, which is shown in Table 1 below respectively:

<table>
<thead>
<tr>
<th>Fault types</th>
<th>Major gases</th>
<th>Secondary gases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oil overheating</td>
<td>( \text{H}_2, \text{C}_2\text{H}_4 )</td>
<td>( \text{H}_2, \text{C}_2\text{H}_6 )</td>
</tr>
<tr>
<td>Oil and paper overheating</td>
<td>( \text{CH}_4, \text{C}_2\text{H}_6, \text{CO}, \text{CO}_2 )</td>
<td>( \text{H}_2, \text{C}_2\text{H}_4 )</td>
</tr>
<tr>
<td>Local discharge in the paper oil insulation</td>
<td>( \text{H}_2, \text{C}_2\text{H}_4, \text{CO} )</td>
<td>( \text{C}_2\text{H}_6, \text{CO}_2 )</td>
</tr>
<tr>
<td>Spark discharge in dissolved oil</td>
<td>( \text{H}_2, \text{C}_2\text{H}_6 )</td>
<td>( \text{H}_2, \text{C}_2\text{H}_4, \text{CO}_2 )</td>
</tr>
<tr>
<td>Arc in dissolved oil</td>
<td>( \text{H}_2, \text{C}_2\text{H}_4, \text{CO} )</td>
<td>( \text{CH}_4, \text{C}_2\text{H}_6, \text{C}_2\text{H}_6 )</td>
</tr>
<tr>
<td>Oil and paper arc</td>
<td>( \text{H}_2, \text{C}_2\text{H}_6 )</td>
<td>( \text{C}_2\text{H}_6, \text{C}_2\text{H}_6 )</td>
</tr>
</tbody>
</table>

B. Collection of sample index data

Through the analysis of the relationship between the transformer faults and the dissolved gas relationship, it can be seen that the gas volume content is usually adopted as the fault diagnosis index during the transformer fault diagnosis.

Dissolved oil data is of vital importance to the performance of the power transformer. Under general conditions, the samples and the sample indexes used for fault diagnosis and analysis should meet the following three principles:

DOI 10.5013/IJSSST.a.16.02.06 6.3 ISSN: 1473-804x online, 1473-8031 print
1) Choose the eigenvectors related to the fault characteristics and avoid the disturbance of a large number of redundant components;
2) Reflect the fault type as much as possible and ensure the comprehensiveness of the fault diagnosis;
3) Take the influence of operation environment and surrounding factors into consideration, collect the multi-group samples under different operation environments and ensure the diversity of samples.

Based on relevant research results, this paper adopts the volume content of five gases, namely $H_2$, $CH_4$, $C_2H_6$, $C_2H_4$, and $C_2H_2$, as the input samples. There are four fault types among the samples, namely high-energy discharge, low-energy discharge, high-temperature over-heating and medium-and low-temperature heating. This paper collects 80 groups of sample data. The sample data are divided into the training ones and the test ones before the fault diagnosis. The sample type distribution and the partial sample data of the training and test samples are shown in Table 2 and Table 3, respectively:

### Table 2. Sample type distribution

<table>
<thead>
<tr>
<th></th>
<th>Training data</th>
<th>Test data</th>
<th>Total data</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-energy discharge</td>
<td>16</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>Low-energy discharge</td>
<td>12</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>High-temperature overheating</td>
<td>14</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>Medium-and low-temperature overheating</td>
<td>10</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>28</td>
<td>80</td>
</tr>
</tbody>
</table>

### Table 3. Index data of partial samples (volume content)

<table>
<thead>
<tr>
<th>Fault type</th>
<th>$H_2$</th>
<th>$CH_4$</th>
<th>$C_2H_6$</th>
<th>$C_2H_4$</th>
<th>$C_2H_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-energy discharge</td>
<td>117</td>
<td>5.8</td>
<td>15.9</td>
<td>7</td>
<td>12.8</td>
</tr>
<tr>
<td>Low-energy discharge</td>
<td>730</td>
<td>750</td>
<td>190</td>
<td>1300</td>
<td>7</td>
</tr>
<tr>
<td>High-energy discharge</td>
<td>421</td>
<td>135</td>
<td>27.7</td>
<td>351</td>
<td>374</td>
</tr>
<tr>
<td>Low-energy discharge</td>
<td>51.1</td>
<td>7.1</td>
<td>1.7</td>
<td>5.9</td>
<td>9.2</td>
</tr>
<tr>
<td>Low-energy discharge</td>
<td>3.3</td>
<td>1.1</td>
<td>0.9</td>
<td>3.1</td>
<td>3.3</td>
</tr>
<tr>
<td>High-temperature heating</td>
<td>63</td>
<td>0.29</td>
<td>0.59</td>
<td>3.64</td>
<td>16.8</td>
</tr>
<tr>
<td>Low-energy discharge</td>
<td>220</td>
<td>502</td>
<td>178</td>
<td>1017</td>
<td>7.8</td>
</tr>
<tr>
<td>High-temperature heating</td>
<td>7.8</td>
<td>111</td>
<td>129</td>
<td>701</td>
<td>2.8</td>
</tr>
<tr>
<td>Medium-and low-temperature heating</td>
<td>58</td>
<td>103</td>
<td>51</td>
<td>251</td>
<td>6</td>
</tr>
<tr>
<td>Medium-and low-temperature heating</td>
<td>4.9</td>
<td>21</td>
<td>15</td>
<td>69.3</td>
<td>3.8</td>
</tr>
<tr>
<td>Medium-and low-temperature heating</td>
<td>140</td>
<td>173</td>
<td>45</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Medium-and low-temperature heating</td>
<td>72</td>
<td>442</td>
<td>221</td>
<td>461</td>
<td>0.7</td>
</tr>
</tbody>
</table>

C. Fault diagnosis simulation within the transformer

After normalization of the transformer’s sample index, this paper divides models into different kinds for simulated calculation during the simulation process so as to reflect the difference between the optimized PSO and the ordinary PSO. The parameters of various models are shown in Table 4 below:

### Table 4. Setting results of various model simulated parameters

<table>
<thead>
<tr>
<th>Standard PSO algorithm</th>
<th>Velocity update and optimization PSO algorithm</th>
<th>Self-adaptation variation optimization PSO algorithm</th>
<th>Speed update+self-adaptation optimization PSO algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialize the inertial weight value, $w_1$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Terminate the inertial weight value, $w_2$</td>
<td>1</td>
<td>0.1</td>
<td>1</td>
</tr>
<tr>
<td>Initialize the acceleration constants, $c_{11}$ and $c_{21}$</td>
<td>1, 1</td>
<td>1, 1</td>
<td>1, 1</td>
</tr>
<tr>
<td>Terminate the acceleration constants $c_{1T}$ and $c_{2T}$</td>
<td>1, 1</td>
<td>2.5, 2.5</td>
<td>1, 1</td>
</tr>
<tr>
<td>Maximum iteration number</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>Variation probability, $k$</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td>Particle scale</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

Conduct simulated calculation of model parameters in Table 4. The optimal fitness and average fitness curve iteration process of various models are shown in Fig. 2 below. The model’s convergence parameters and model precision are shown in Table 5 below:
From the optimal fitness and average fitness iteration curve of various PSO models in Fig. 1, it can be seen that:

1) The optimal fitness curve of the PSO algorithm converges quickly from the initial 10 to 20 iteration periods. The average fitness curve is at a fluctuating status and the fluctuation is within the range of 40%~50%, and shows weak descending trend;

2) In terms of the iterative process of the velocity update PSO algorithm model and the self-adaptation variation PSO algorithm model, their optimal fitness curve converge quickly and the fitness curve quickly converges and stays at a steady value, 82.69%, and the overall average fitness curve first increases and then stabilizes at 70%~75%.

3) In terms of the iterative process of the velocity update and self-adaptation PSO algorithm model, the optimal fitness curve quickly stabilizes at 82.69%. The optimal fitness curve increases to 83.21% near the 70th iteration. When the iteration number reaches 130, the optimal fitness curve increases to 84.61% once again. The optimal fitness curve’s stable value is higher than that of the former three models.

From Table 5, through the test of training and test samples, the recognition accuracy of the standard PSO algorithm is relatively low, while the recognition accuracy of training samples by the velocity update PSO algorithm and the self-adaptation variation PSO model can reach as high as 94.23%, but their recognition accuracy of the test samples is just 82.14%. The recognition accuracy of training and test samples by the velocity update and self-adaptation PSO algorithm is 96.16% and 89.29%, respectively. It is higher than that by the standard PSO algorithm by 13.47% and 17.86%, respectively.
IV. CONCLUSIONS

This paper first introduces the relationship between power transformer faults and dissolved gas and clarifies the power transformer fault indexes at an attempt to efficiently recognize the power transformer fault types. First, the basic theory of the SVM model is analyzed, and is optimized through the standard PSO. In order to overcome the defect that the standard PSO might be easily trapped into the local extreme, this paper introduces the dynamic inertial weight iteration factor and the acceleration coefficient iteration factor to optimize the particle velocity update process and the particle generation self-adaption variation, optimize the standard PSO model and control the exploration capability of the standard PSO algorithm. Results suggest that: MPSO put forward in this paper can well solve the problem that the standard PSO might be easily trapped into the local extreme, and boasts a favorable stability. The test sample recognition accuracy of the model parameters acquired by MPSO-SVM is higher than that acquired by the standard PSO by 17.86%.

REFERENCES