Fault diagnosis of transformer based on fuzzy clustering and the optimized wavelet neural network

Wenhui Tenga, Shuxian Fanb, Zheng Gongb, Wen Jianguob and Maofa Gonga

aCollege of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao, People’s Republic of China; bState Grid Shandong Weifang Electric Power Company, Weifang, People’s Republic of China

ABSTRACT

In order to solve the disadvantages of the traditional wavelet neural network (WNN) algorithm applied in transformer fault diagnosis, such as uneven sample distribution of training samples and high diagnostic error rate and long training time, an improved fault diagnosis method is proposed based on fuzzy clustering and the flower pollination algorithm. Firstly, fuzzy clustering is applied to deal with transformer fault sample data so as to remove the bad data; secondly, the flower pollination algorithm is applied to obtain the optimal parameters of the WNN. The example analysis results show that WNN based on the flower pollination algorithm (FPA-WNN) has better convergence, lower diagnosis error rate and shorter training time compared with WNN based on the particle swarm algorithm (PWA-WNN) and it is more suitable for transformer fault diagnosis.

ARTICLE HISTORY

Received 29 June 2018
Accepted 29 December 2018

KEYWORDS

Transformers; fault diagnosis; wavelet neural network; fuzzy clustering; flower pollination algorithm

1. Introduction

The power transformer is the core equipment in power transmission and distribution, and the safe and reliable operation of power transformer has an important impact on the power grid and national economy. It is significant to the stable operation of the power system if transformer faults or latent faults can be diagnosed and predicted accurately and quickly.

Dissolved gas analysis (DGA) technology is one of the most convenient and effective methods for fault diagnosis of oil-immersed transformers. It can diagnose the latent fault which may cause serious damage accurately and reliably. In recent years, various criteria for transformer fault diagnosis have been proposed at home and abroad using the DGA gas ratio as a characteristic parameter, such as the IEC ratio (Yang, Liu, Li, & Hu, 2007) and the improved Rogers ratio (Rogers, 1978). However, most of the fault diagnosis criteria are based on field experience, and there is misdiagnosis occurring in practical applications.

Artificial intelligence (AI) technology has been widely used because the complex nonlinear relationships between DGA gas content and transformer faults have been established based on these methods and it has the advantages of continuous learning and updating compared with conventional DGA ‘hard criteria’. Among them, the support vector machine and artificial neural network are widely used in transformer fault diagnosis and have made good effects.

Xue, Zhang, Li, and Peng (2015) proposed a method which optimized the parameters of the support vector machines using the cuckoo algorithm to obtain the best diagnosis model. However, the method would lead to a local optimum. Ma (2008) proposed a method based on the genetic algorithm and WNN, but the method had the disadvantages of a relatively complicated network structure, long training time and low accuracy. Song and Wang (2015) proposed a method which optimized the weights and scaling factors of WNN using the PSO algorithm, and the method also had the disadvantages of high diagnostic error rate and long training time. Moreover, there were fault diagnosis methods based on grey theory (Li, Sun, Chen, Zhou, & Du, 2003), Bayesian classifiers (Wang, Zhang, Jin, & Guo, 2018) and the expert system (Shi, Shi, Mu, Li, & Liu, 2014), but most of them had the disadvantages of over fitting, long training time and low diagnostic accuracy.

In this paper, an improved fault diagnosis method is proposed based on fuzzy clustering and the flower pollination algorithm. Firstly, the fuzzy c-means clustering algorithm (FCM) is used to deal with the collected fault sample data in order to filter out the isolated sample and avoid uneven distribution of the sample, and it can influence the correct rate of transformer fault diagnosis.
the number of input layer nodes is \( n \); the number of hidden layers is \( m \); the number of output layer nodes is \( N \); and the kth input sample of the input layer is \( x_k \); the actual output value of the \( i \)th node of the output layer is \( y_i \); the expected output value of the \( i \)th node of the output layer is \( \hat{y}_i \); the connection weight between the input layer node \( k \) and the hidden layer node \( j \) is \( w_{kj} \) and the connection weight between the output layer node \( i \) and the hidden layer node \( j \) is \( w_{ji} \); the telescopic translation coefficients of the \( j \)th hidden layer nodes are \( a_j \) and \( b_j \). The hidden layer wavelet neurons use the Mexican Hat wavelet function, and the output layer nodes use the Sigmoid function. The input of the \( j \)th wavelet element of the hidden layer is

\[
 s_j = \sum_{k=1}^{m} w_{kj} x_k, \quad j = 1, 2, \ldots, n. \tag{2}
\]

The output of the \( j \)th wavelet element of the hidden layer is

\[
 h_j = \frac{1}{\sqrt{a_j}} \left( 1 - \frac{(s_j - b_j)^2}{a_j^2} \right) e^{-(s_j - b_j)^2/2a_j^2}. \tag{3}
\]

Then the output of the \( i \)th node of the network output layer is

\[
 y_i = \frac{1}{1 + \exp \left( - \sum_{j=1}^{n} w_{ji} h_j \right)}, \quad i = 1, 2, \ldots, N. \tag{4}
\]

The network output error function is

\[
 \text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2. \tag{5}
\]

Therefore, when the input node, hidden layer node and output nodes are determined, the key to construct a suitable WNN is to determine the parameters in the network. The selection of these parameters depends on the optimization of the WNN training algorithm and thus it is especially important to find a suitable WNN training algorithm.

### 3.2. FPA optimize the parameters of WNN

The British scholar X.S. Yang proposed a flower pollination algorithm (Wang, 2016; Wang et al., 2018) in 2012. Wang...
et al., (2018) described four ideal conditions of flower pollination.

There are two key steps in this algorithm, such as self-pollination and cross-pollination.

The self-pollination (local pollination) is shown in the following equation:

\[ x_{t}^{i+1} = x_{t}^{i} + L(x_{t}^{i} - g^{*}), \]  

(6)

where \( x_{t}^{i} \) is the pollen \( i \) or solution vector \( x_{t} \) at iteration \( t \) and \( g^{*} \) is the current best. The parameter \( L \) is a step size which obeys Levy distribution.

The cross-pollination (global pollination) is shown in the following equation:

\[ x_{t}^{i+1} = x_{t}^{i} + \varepsilon(x_{t}^{j} - x_{k}^{j}), \]  

(7)

where \( x_{t}^{i} \) and \( x_{t}^{j} \) are pollens from the different flowers of the same plant species and \( \varepsilon \) is a proportional coefficient between 0 and 1.

For selected transformer fault data, the process of a wavelet neural network fault diagnosis model based on the flower pollination algorithm are as follows:

Step 1: Process transformer fault data gathered from related references.

Step 2: Initialize the basic parameters of the FPA algorithm.

Step 3: Make each pollen position correspond to a set of parameters \( \{w_{ji}, a_{i}, b_{j}, w_{ji}\} \) and calculate the best solution \( g^{*} \) with the initial pollen and assign \( F_{\text{min}} \) to the fitness at \( g^{*} \).

Step 4: For each pollen, if \( \text{rand} < p \), // A switch variable \( p \in [0, 1] \), the global agent is updated via Equation (6), else the local agent is updated via Equation (7).

Step 5: Evaluate a new solution.

\[ f(x_{t}^{j+1}) < f(x_{t}^{j}), x_{t}^{j+1} = x_{t}^{j}. \]

Compare with the best pollen, if \( f(x_{\text{new}}) < f(g^{*}) \)

\[ g^{*} = x_{\text{new}}. \]

End if

\[ t = t + 1; \]

Step 6: Check the termination condition, and if it is not safe, please go back to Step 3.

4. Transformer fault diagnosis process

4.1. Transformer insulation fault gas treatment using the FCM method

In this paper, 241 sets of transformer fault data from related references (Li, 2014; Sun, Li, & Sun, 2001; Xiong et al., 2007; Yin, 2007; Yin, 2013; Zhou, 2010; Zeng et al., 2011) are collected and each set of transformer fault data has a clear fault conclusion. Due to the transformer insulation gases are highly dispersive, the data quality may not be ideal. Therefore, the accuracy of the algorithm trained by these samples is poor, and the diagnostic accuracy of the algorithm is not ideal. To solve this problem, the FCM method is used to deal with transformer fault gas before the algorithm is trained.

The specific process is as follows:

Step 1: Select the appropriate number of FCM clusters according to the fault data.

Step 2: Cluster fault data.

Step 3: Delete data in categories that have less than 4% of the selected failure data.

Taking the low-temperature (LT) superheat data as an example, the results after using the FCM method are shown in Table 1. The fault data which have less than 4% in categories of the selected example should be removed. As seen from Table 1, the percentage Cluster 3 in categories has less than 4% of the selected data, so the gas of this category should be removed.

### 4.2. Determination of WNN topology and failure sample data

When the transformer fails, the transformer insulation oil will decompose and produce different types of insulting gas. Different faults will produce different types and volumes of insulting gas. Therefore, depending on the type and volume fraction of the insulting gas, the type of failure can be determined by a certain method. Insulating gases such as H₂, CH₄, C₂H₆, C₂H₄, and C₂H₂ are mainly caused by abnormal faults in the transformer. However, the volume fraction of the fault gas may vary greatly. Large fluctuations in the fault data will have a large effect on programme convergence. Therefore, the fault data which should be normalized is shown in the following formula:

\[ y_{ix} = 1/(1 + \exp(-x_{ix}/c_{i})), \]  

(8)

### Table 1. FCM clustering results of LT samples.

| Number of clusters | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 | Cluster 5 |
|--------------------|-----------|-----------|-----------|-----------|-----------|
| Number of samples  | 15        | 10        | 1         | 5         | 9         |
| Percentage/%       | 37.5      | 25        | 2.5       | 12.5      | 22.5      |

### Table 2. The attention concentration of insulating gas.

| X₁ (H₂) | X₂ (CH₄) | X₃ (C₂H₆) | X₄ (C₂H₄) | X₅ (C₂H₂) | X₆ (all) |
|---------|----------|------------|------------|------------|----------|
| 100 µL/L  | 50 µL/L  | 100 µL/L  | 3 µL/L  | 150 µL/L |


where $c_i$ is the attention concentration of insulating gas and $x_{ix}$ is the actual value of the insulating gas. The attention concentration of insulating gas is shown in Table 2.

The topology of the WNN of this paper is shown in Figure 2.

The input value $X_1, X_2, \ldots, X_6$ represents the volume fractions of the normalized H$_2$, CH$_4$, C$_2$H$_6$, C$_2$H$_4$, C$_2$H$_2$, and total hydrocarbons (sum of the total insulation gas volume normalized by the five gases) respectively. The output value $Y_1, Y_2, \ldots, Y_6$ represents the failure of partial discharge, low energy discharge, high energy discharge, low-temperature overheat, medium temperature overheat and high temperature overheat. The training speed and accuracy of the algorithm can be affected by the number of hidden layer nodes. According to relevant references, the topological structure of WNN has 6 outputs and 6 inputs. The best number of hidden layer nodes in WNN is about 12 (Fang, Peng, Li, & Shu, 2011).

In this paper, the mean square error of the algorithm is compared by changing the number of hidden layer nodes in WNN to determine the best hidden layer node number. When the input and output data of the training set are the same, the mean square error of different hidden layer nodes is shown in Table 3. As seen from Table 3, the number of hidden layer nodes is 12 and the mean square error is the minimum. So the hidden layer number of WNN is 12.

### 4.3. Case analysis

Transformer fault data that were collected from related references are randomly divided into two parts: 180 sets of transformer fault data were selected as training data and the remaining 61 sets of transformer fault data were used as test data. The structure of WNN of 6-12-6 based on PSO and FPA is trained and the best training result is shown in Figure 3.

It can be seen from Figure 3 that FPA-WNN has a faster convergence rate than PSO-WNN. FPA-WNN reaches the target mean square error value in about 200 iterations while PSO-WNN reaches in about 270 iterations.

In order to verify the superiority of the proposed method, 61 sets of test samples were brought into the trained PSO-WNN and FPA-WNN. The accuracy of the algorithm and the mean square error are shown in Table 4. It can be seen from Table 4 that the diagnostic

| Network structure | Training set MSE |
|-------------------|------------------|
| 6-10-6            | 0.047            |
| 6-11-6            | 0.035            |
| 6-12-6            | 0.01             |
| 6-13-6            | 0.068            |
| 6-14-6            | 0.073            |

In this paper, the mean square error of the algorithm is compared by changing the number of hidden layer nodes in WNN to determine the best hidden layer node number. When the input and output data of the training set are the same, the mean square error of different hidden layer nodes is shown in Table 3. As seen from Table 3, the number of hidden layer nodes is 12 and the mean square error is the minimum. So the hidden layer number of WNN is 12.

### 4.3. Case analysis

Transformer fault data that were collected from related references are randomly divided into two parts: 180 sets of transformer fault data were selected as training data and the remaining 61 sets of transformer fault data were used as test data. The structure of WNN of 6-12-6 based on PSO and FPA is trained and the best training result is shown in Figure 3.

It can be seen from Figure 3 that FPA-WNN has a faster convergence rate than PSO-WNN. FPA-WNN reaches the target mean square error value in about 200 iterations while PSO-WNN reaches in about 270 iterations.

In order to verify the superiority of the proposed method, 61 sets of test samples were brought into the trained PSO-WNN and FPA-WNN. The accuracy of the algorithm and the mean square error are shown in Table 4. It can be seen from Table 4 that the diagnostic

| Network structure | Training set MSE |
|-------------------|------------------|
| 6-10-6            | 0.047            |
| 6-11-6            | 0.035            |
| 6-12-6            | 0.01             |
| 6-13-6            | 0.068            |
| 6-14-6            | 0.073            |

In this paper, the mean square error of the algorithm is compared by changing the number of hidden layer nodes in WNN to determine the best hidden layer node number. When the input and output data of the training set are the same, the mean square error of different hidden layer nodes is shown in Table 3. As seen from Table 3, the number of hidden layer nodes is 12 and the mean square error is the minimum. So the hidden layer number of WNN is 12.

### 4.3. Case analysis

Transformer fault data that were collected from related references are randomly divided into two parts: 180 sets of transformer fault data were selected as training data and the remaining 61 sets of transformer fault data were used as test data. The structure of WNN of 6-12-6 based on PSO and FPA is trained and the best training result is shown in Figure 3.

It can be seen from Figure 3 that FPA-WNN has a faster convergence rate than PSO-WNN. FPA-WNN reaches the target mean square error value in about 200 iterations while PSO-WNN reaches in about 270 iterations.

In order to verify the superiority of the proposed method, 61 sets of test samples were brought into the trained PSO-WNN and FPA-WNN. The accuracy of the algorithm and the mean square error are shown in Table 4. It can be seen from Table 4 that the diagnostic

| Network structure | Training set MSE |
|-------------------|------------------|
| 6-10-6            | 0.047            |
| 6-11-6            | 0.035            |
| 6-12-6            | 0.01             |
| 6-13-6            | 0.068            |
| 6-14-6            | 0.073            |

In this paper, the mean square error of the algorithm is compared by changing the number of hidden layer nodes in WNN to determine the best hidden layer node number. When the input and output data of the training set are the same, the mean square error of different hidden layer nodes is shown in Table 3. As seen from Table 3, the number of hidden layer nodes is 12 and the mean square error is the minimum. So the hidden layer number of WNN is 12.

### 4.3. Case analysis

Transformer fault data that were collected from related references are randomly divided into two parts: 180 sets of transformer fault data were selected as training data and the remaining 61 sets of transformer fault data were used as test data. The structure of WNN of 6-12-6 based on PSO and FPA is trained and the best training result is shown in Figure 3.

It can be seen from Figure 3 that FPA-WNN has a faster convergence rate than PSO-WNN. FPA-WNN reaches the target mean square error value in about 200 iterations while PSO-WNN reaches in about 270 iterations.

In order to verify the superiority of the proposed method, 61 sets of test samples were brought into the trained PSO-WNN and FPA-WNN. The accuracy of the algorithm and the mean square error are shown in Table 4. It can be seen from Table 4 that the diagnostic

| Network structure | Training set MSE |
|-------------------|------------------|
| 6-10-6            | 0.047            |
| 6-11-6            | 0.035            |
| 6-12-6            | 0.01             |
| 6-13-6            | 0.068            |
| 6-14-6            | 0.073            |

In this paper, the mean square error of the algorithm is compared by changing the number of hidden layer nodes in WNN to determine the best hidden layer node number. When the input and output data of the training set are the same, the mean square error of different hidden layer nodes is shown in Table 3. As seen from Table 3, the number of hidden layer nodes is 12 and the mean square error is the minimum. So the hidden layer number of WNN is 12.

### 4.3. Case analysis

Transformer fault data that were collected from related references are randomly divided into two parts: 180 sets of transformer fault data were selected as training data and the remaining 61 sets of transformer fault data were used as test data. The structure of WNN of 6-12-6 based on PSO and FPA is trained and the best training result is shown in Figure 3.

It can be seen from Figure 3 that FPA-WNN has a faster convergence rate than PSO-WNN. FPA-WNN reaches the target mean square error value in about 200 iterations while PSO-WNN reaches in about 270 iterations.

In order to verify the superiority of the proposed method, 61 sets of test samples were brought into the trained PSO-WNN and FPA-WNN. The accuracy of the algorithm and the mean square error are shown in Table 4. It can be seen from Table 4 that the diagnostic
Table 4. Comparison of the diagnostic accuracy of the two algorithms.

| Algorithm          | PSO-WNN algorithm | FPA-WNN algorithm |
|--------------------|--------------------|--------------------|
| Diagnostic accuracy| 49/61 = 80.3%      | 54/61 = 88.5%      |
| Final mean square error | 0.02586        | 0.00824           |

Accuracy rates of FPA-WNN and PSO-WNN, respectively, are 80.3% and 88.5%. The accuracy of FPA-WNN diagnosis is higher than that of PSO-WNN. The final mean square errors of the two algorithms are 0.02586 and 0.00824, and the final mean square error of PSO-WNN is 3.13 times that of FPA-WNN. The results indicate that FPA-WNN has better convergence accuracy. Therefore, FPA-WNN can obtain better diagnostic results and reduce training errors.

5. Conclusion

A WNN fault diagnosis model based on fuzzy clustering and the flower pollination algorithm can avoid the problems existed in the other fault diagnosis methods, such as long training time and high diagnostic error rate. Simulation results show that the FPA-WNN algorithm has a lower fault diagnosis error rate and shorter training time, which provides a better fault diagnosis method for transformer fault diagnosis compared with the PSO-WNN algorithm.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Chen, W. G., Pan, C., Wang, Y. Y., & Yun, Y. X. (2007). Fault diagnostic method using wavelet neural network for power transformers. High Voltage Engineering, 33(8), 52–55.
Fang, J., Peng, H., Li, Z. P., & Shu, N.-Q. (2011). An improved wavelet network with application to the fault diagnosis of oil-immersed transformer. High Voltage Apparatus, 47(11), 115–120.
Li, Y. H. (2014). Transformer fault diagnosis based on support vector machine optimized by simulated annealing. Lanzhou: Lanzhou Jiaotong University.
Li, J., Sun, C. X., Chen, W. G., Zhou, Q., & Du, L. (2003). A method of synthesis based on the grey cluster and fuzzy cluster about internal fault diagnosis of transformers. Proceedings of CSEE, 23(2), 112–115.
Ma, W. J. (2008). Power transformers fault diagnosis based on the wavelet neural network. Techniques of Automation and Applications, 27(12), 17–19.
Rogers, R. R. (1978). IEEE and IEC codes to interpret incipient faults in transformers using gas in oil analysis. IEEE Transactions on Power Apparatus and Systems, 135(5), 348–354.
Shi, R. F., Shi, Y. F., Mu, J., Li, H., & Liu, X. H. (2014). Power transformer fault diagnosis expert system with dissolved gas analysis in oil. Proceedings of the CSU-EPSA, 26(12), 49–54.
Song, Z. J., & Wang, J. (2015). Transformer fault diagnosis based on BP neural network optimized by fuzzy clustering and LM algorithm. High Voltage Apparatus, 2013(5), 54–59.
Sun, H., Li, W. D., & Sun, Q. Z. (2001). Electric power transformer fault diagnosis using decision tree. Proceedings of the CSEE, 21(2), 50–55.
Wang, R. (2016). Application research of flower pollination algorithm. Nanning: Guangxi University for Nationalities.
Wang, Z. Q., Zhang, M. X., Jin-Long, Z., & Guo-Hai, Z. (2018). Optimization of Bayesian classifier based on flower pollination algorithm. Microelectronics & Computer, 35(3), 47–51.
Xiong, H., Zhang, X. X., Liao, R. J., et al. (2007). Fault diagnosis of power transformer using dynamic clustering algorithm. Chinese Journal of Scientific Instrument, 28(3), 456–459.
Xue, H. R., Zhang, K. H., Li, B., & Peng, C. (2015). Fault diagnosis of transformer based on the cuckoo search and support vector machine. Power System Protection and Control, 43(8), 8–13.
Yang, T. F., Liu, P., Li, J. L., & Hu, Y. (2007). New fault diagnosis method of power transformer by combination of FCM and IEC three-ratio method. High Voltage Engineering, 33(8), 66–70.
Yin, G. L. (2013). Study on oil-immersed power transformer fault diagnosis based on relevance vector machine. Beijing: Electric Power University.
Yin, Z. (2007). Fault diagnosis system of gas chromatograph in transformer oil based on artificial neural network. Tianjin: Tianjin University.
Zeng, L. P., Wang, H. B., & He, Y. H. (2011). Simulated annealing – wavelet network to power transformer fault diagnosis. Computer Simulation, 28(7), 316–321.
Zhou, Y. Y. (2010). Power transformer fault diagnosis based on dissolved gas analysis. Changsha: Changsha University of Science and Technology.