Analysis of Transformer Operation State Based on Multi-dimensional Data Joint Analysis

Xu Zhendong¹, Lu Shuang¹, Xia Shijia¹ and Xu Fang¹

¹ Shanghai Heng Neng Tai Enterprise Management Co., Ltd., Shanghai 200437, China

E-mail: powerxzd@126.com

Abstract. It is of great significance to accurate grasp the operating status of the power transformer and to discover the latent faults existing in the transformer in a timely and effective manner. Aiming at the problems of fewer parameters and lower diagnosis accuracy in power transformer evaluation model, this paper proposes a transformer state prediction model based on multi-dimensional data joint analysis. Comprehensively considering the oil dissolved gas data, partial discharge data, infrared data and environmental information data, we first use the sample case database to train the wavelet neural network, and then use the deep long short-term memory network that considers the time series attention mechanism to predict the parameters. The results are fed into the classifier, and we can obtain the operation result of the transformer. Tested on the actual transformer data, the experimental results show that the method has high prediction accuracy, which provides a reliable basis for maintenance personnel to predict the operating state of the transformer, and has certain application value.

1. Introduction

As an important part of the entire power system, the power transformer is responsible for the transmission and distribution of power in the power grid, and determines the effective conversion of voltage, which plays a vital role in power systems. Maintaining the safe, stable and effective operation of power transformers is necessary to ensure the reliable and safe operation of the entire power systems [1]. If the transformer has an accident, it will bring huge consequences, such as a large-scale power outage, which will be accompanied by major economic losses, or it may cause a huge explosion, resulting in danger to the lives of surrounding people. Therefore, it is of great research significance to ensure the reliable and stable operation of power transformers, reduce the failures of power transformers, and reduce the probability of potential failures [2].

According to the report of the State Grid Corporation of China during the period 2000-2010, the power grid accidents caused by equipment failures accounted for about 31% to 61% of the total accidents of the year, basically ranked first in the proportion. Equipment failure is the most common cause of power grid accidents. Its failures and accidents will be affected by many factors, including external natural disasters, external force damage, defects in the manufacturing process and equipment aging [3]. In order to cope with this problem, it is necessary to establish a framework for predicting the operation status of transformers. Early warning and diagnosis of latent faults are required in order to improve the maintenance and repair capabilities of power transformers. The operating state of the equipment can be determined by analyzing the characteristic parameter information of the transformer. When different types of faults occur in the transformer, the faults appear in different parts and the degree of the fault is different, the characteristic parameters will also change accordingly. Therefore, we know that there is a
certain correlation between the changing trend of the characteristic parameters and the fault type of the transformer. By predicting and analyzing the changing trends and characteristics of these characteristic parameters corresponding to different types of faults, you can grasp the future operational changes of the transformer, and can obtain information about the operating state in advance [4-5].

The current research has carried out relevant state prediction work for dissolved gas in oil and transformer breakdown voltage. Current methods mainly include SVM [6], grey model (GM), neural network [7-8]. Liao used least squares SVM and particle swarm optimization (PSO) algorithm to forecast the dissolved gases in oil-filled transformers [6]. The radial basis function kernel was established to facilitate SVM regression model. Then the PSO working as a global optimizer was employed to optimize the hyper-parameters. Ghumem proposed a prediction model of furan content based on multiple neural networks. They took seven transformer oil parameters as inputs and the relationship between these parameters were also calculated. Furthermore, a stepwise regression was used to select the most important predictors [7]. Dai proposed a short-term dissolved gas forecasting model that can be found that the current parameter prediction method still has some shortcomings [8]. The neural network model depends on the selection of input features and is not easy to train when the training samples are too large. The support vector machine model is prone to misclassification, especially for data points located at the edge points. Although the gray correlation model solves the problem of insufficient information for a single input vector, it is necessary to use human experience to select a suitable threshold for the correlation during model training. In addition to the transformer oil chromatography data, more parameters need to be evaluated, such as partial discharge data and infrared data.

When performing fault diagnosis and analysis on power transformers, it is required that the constructed model can find the potential fault information existing in the transformer in time. Intelligent algorithms and machine learning models break through the traditional three-ratio method, threshold method and other evaluation and diagnosis methods, improving the accuracy and generalization ability of power transformer fault diagnosis [9-11]. Moreover, the multi-layer SVM classifier is trained by using ratios and graphical representation information to classify the faults [12]. In terms of operational status prediction, Abbasi built a model to detect the status of the transformer taking advantage of the high frequency response analysis [13], where clustering analysis and cross-correlation was also used.

At present, the evaluation of the operating state of the transformer is mainly based on the transformer state prediction and fault diagnosis, and the predicted state parameters are input into the evaluation and diagnosis model for state prediction. But relying solely on the transformer dissolved gas in oil data cannot completely estimate the operating state of the transformer. This paper considers the evaluation method of transformer operation status with multiple data dimensions. We take into account the changes in partial discharge, infrared detection data, and temperature and humidity changes. For the prediction of parameters, a long short-term memory network (LSTM) parameter prediction model based on the time series attention mechanism is proposed. This model can consider the influence of different moments on the prediction points and adaptively strengthen the influence weight of the data points with large correlation. At the same time, the historical fault case library is used to train the wavelet neural network (WNN) to classify the transformer fault. Finally, the double-layer combination network structure is used to evaluate the state of the transformer.

2. Long short-term memory network based on temporal attention mechanism
The network structure of LSTM network is shown in Figure 1.
Figure 1. The structure of LSTM network

Each value calculated by the LSTM network node has a weight value connected to itself at time t. The memory unit can retain the actual value of its own state and accumulated information from the outside, and this process can be controlled by the gate unit to control whether it is cleared or new memory is added during the learning process. When the far-end or near-end data flows in, it must first be analyzed by the gate unit to forget or retain the information, so that information that is farther away can also be retained. And during the training process of the network, the gate unit can also control the influence of the data flow on the weight of the node, thereby reducing the adverse effects caused by the gradient change. The main calculation formula (1-5) are as follows:

\[ f_t = \sigma(W_f[h_{t-1}; x_t] + b_f) \]  
\[ i_t = \sigma(W_i[h_{t-1}; x_t] + b_i) \]  
\[ o_t = \sigma(W_o[h_{t-1}; x_t] + b_o) \]  
\[ s_t = f_t e_s t_{t-1} + i_t e \tanh(W_s[h_{t-1}; x_t] + b_s) \]  
\[ h_t = o_t e \tanh(s_t) \]

where \( W_f, b_f \), \( W_i, b_i \), \( W_o, b_o \), \( W_s, b_s \) are the weighs and bias; \( \sigma \) and \( \tanh \) are activation function; \( e \) is multiplication of elements by matrix.

In order to find the degree of influence of the information of the relevant feature time series at each moment on the predicted state quantity at the current moment, this paper introduces a time series attention mechanism to autonomously extract critical moments. The attention weight of the hidden layer state at each moment of the encoder at the current LSTM iteration time \( t \) is calculated. The softmax function (6) is used to normalize the attention weight, and the importance of the hidden layer state of the encoder to the predicted gas at each time is recorded as \( \gamma_t \in [0,1] \) in (7).

\[ l_t = V_d^* \tanh(W_d[d_{t-1}; s_{t-1}] + b_d) + U_d h_t \]  
\[ \gamma_t = \frac{\exp(l_t)}{\sum_{t=1}^{T} \exp(l_t)} \]

where \( V_d \in R^{q \times l}, W_d \in R^{l \times 2q} \) and \( U_d \in R^{l \times p} \) is multi-layer perceptron weights and bias parameters for calculating attention weights.

3. Long short-term memory network based on temporal attention mechanism

WNN combines the characteristics of wavelet transform and neural network, and uses wavelet neurons to replace the neurons of traditional neural networks. Its network structure is shown in Figure 2, including input layer, hidden layer and output layer. First, the signal is decomposed by wavelet, and then the arbitrary function is approximated by neural network. Adding wavelet scale factor and
translation factor can make the series after wavelet transform more flexible, and have stronger function approximation ability and pattern recognition ability.

\[
\begin{align*}
\text{input layer} & \quad \text{hidden layer} & \quad \text{output layer} \\
X_1 & \quad \text{Wavelet function} & \quad Y_1 \\
X_2 & \quad \text{Wavelet function} & \quad Y_2 \\
\vdots & \quad \vdots & \quad \vdots \\
X_k & \quad \text{Wavelet function} & \quad Y_k
\end{align*}
\]

**Figure 2.** The network structure of wavelet neural network

Suppose the input vector is \( x = [x_1, x_2, \ldots, x_n]^T \), the output vector is \( y = [y_1, y_2, \ldots, y_m]^T \), the number of neurons in the hidden layer is \( N \), and the node weights between the two layers are \( a_j \) and \( b_j \). Let the translation factor and scale factor be \( a_j \) and \( b_j \), and the wavelet function expression is \( \Psi_{a,b}(t) \). Then at time \( t \), the output vector is calculated in (8):

\[
y(t) = \sum_{j=1}^{N} \omega_j \Psi_{a,b}(t) = \sum_{j=1}^{N} \omega_j x_j(t)
\]

(8)

Due to the structural characteristics of wavelet neural networks, different wavelet basis functions as hidden layer functions will have different effects on the network results. In this paper, Morlet wavelet function is selected as the basis function. The expression is indicated in (9).

\[
\Psi_{\text{Morlet}}(t) = \cos(1.75 \cdot t) \cdot e^{-t^2/2}
\]

(9)

4. Long short-term memory network based on temporal attention mechanism

The flow chart of transformer condition assessment considering multi-dimensional data joint analysis is shown in Figure 3.

**Figure 3.** Transformer status classification flow chart

The main parameters considered include oil chromatographic data, partial discharge signals, infrared parameters, and indoor temperature and humidity. The oil chromatography data includes \( H_2, CO, CO_2, C_2H_2, C_2H_4, C_2H_6 \) and \( CH_4 \). First, the sample case library is used to train the WNN network so that it can accurately distinguish the operating state of the transformer. Then train the LSTM network to predict the future change trend of the input parameters. Finally, the entire two networks input the predicted parameter results into the state classifier, and then estimate the future operating state of the transformer.
5. Case studies

5.1. Model evaluation index
For the evaluation of the accuracy of parameter prediction results, this paper mainly uses mean relative error and max relative error. The expressions (10-11) are as follows:

\[
\delta_{\text{mean}} = \frac{1}{N} \sum_{i=1}^{N} \frac{\hat{y}_i - y_i}{y_i} \times 100\% \tag{10}
\]

\[
\delta_{\text{max}} = \max \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{11}
\]

where \(N\) is the number of data in the test set, \(y_i\) is the true value, and \(\hat{y}_i\) is the predicted value.

For performance evaluation of transformer status classification, this paper mainly uses recall, accuracy and precision in (12-14).

\[
\text{recall} = \frac{TP}{(TP + FN)} \tag{12}
\]

\[
\text{accuracy} = \frac{TP + TN}{(TP + TN)} \tag{13}
\]

\[
\text{precision} = \frac{TP}{(TP + FP)} \tag{14}
\]

5.2. Parameter prediction performance evaluation
In order to verify the effect of the LSTM network considering the temporal attention mechanism on the prediction of transformer parameters, we take the methane data in the oil chromatography as an example. Figure 4 is the prediction result of methane gas. When the number of prediction points is small, the network has high prediction accuracy. As the number of predictions increases, the accuracy of predictions continues to decline.

![Figure 4. Prediction results of methane gas](image)

Further, this article selects BPNN and SVR for comparison. Among them, BPNN uses the network of three hidden layers, and the number of hidden elements is 100. The learning rate is set to 0.01, and the training period is 10000 times. SVR uses a radial basis kernel function with a penalty factor of 1000. Table 1 is the prediction result of different parameters. It can be seen from the table that the method proposed in this paper can make full use of the dependence of temporal information, and the mean relative error and the max relative error are significantly better than the other two network models.

| Parameters          | TA-LSTM | BPNN | SVR |
|---------------------|---------|------|-----|
|                     | \(\delta_{\text{mean}}\) | \(\delta_{\text{max}}\) | \(\delta_{\text{mean}}\) | \(\delta_{\text{max}}\) | \(\delta_{\text{mean}}\) | \(\delta_{\text{max}}\) |
| Dissolved Gas       | 1.31    | 4.21 | 3.47 | 8.03 | 6.78 | 10.08 |
| Partial discharge   | 4.58    | 8.66 | 8.41 | 13.11 | 12.52 | 15.67 |
| Surface temperature | 0.89    | 3.10 | 2.56 | 5.74 | 4.56 | 8.35 |
| Indoor temperature  | 1.56    | 4.01 | 3.98 | 6.51 | 5.31 | 9.92 |
| Indoor humidity     | 2.03    | 5.31 | 4.43 | 7.89 | 6.74 | 12.20 |

Table 1. Prediction error in percentage of different models on \(\text{CH}_4\) concentration (%)
5.3. Fault classification assessment
Multi-dimensional parameters are used as the input of the WNN network, and the 7 types of transformer status are used as the output. The classification results of DBN, SVM and BPNN during a certain test are shown in Table 2. WNN has a high accuracy in the classification of transformer faults, and it performs best in both the recall rate and accuracy. BPNN has the worst classification effect.

| States | WNN recall | WNN precision | BPNN recall | BPNN precision | SVR recall | SVR precision |
|--------|------------|---------------|-------------|---------------|------------|---------------|
| H      | 95.4       | 94.2          | 84.3        | 88.6          | 89.5       | 88.5          |
| PD     | 93.2       | 91.3          | 70.9        | 68.5          | 77.8       | 77.6          |
| LD     | 88.6       | 93.8          | 68.3        | 60.2          | 80.4       | 75.8          |
| HD     | 91.3       | 89.0          | 66.5        | 70.5          | 73.6       | 74.5          |
| LT     | 87.2       | 87.9          | 80.7        | 72.3          | 78.6       | 84.6          |
| MT     | 89.4       | 86.8          | 78.2        | 74.4          | 78.6       | 77.3          |
| HT     | 88.1       | 90.5          | 68.9        | 73.8          | 74.6       | 81.1          |

5.4. Operation state assessment performance
The current parameters and historical data are used as input to the combined network to predict the operating state of the transformer in a week. Among them, the state parameter prediction model uses LSTM network with time series attention mechanism, and the state classification uses WNN network. The prediction accuracy results on the training set and test set are shown in Figure 5. The network proposed in this paper has the highest accuracy. Among them, compared with BPNN and SVM on the training set, the accuracy is improved by 24.9% and 11.1%. In terms of test performance, it increased by 33.2% and 17.4%, respectively.

Figure 5. Transformer operation state prediction accuracy

We take a 110kV substation #2 main transformer as an example. The transformer was shipped in July, 2000 and officially put into operation in September 2001. Since the operation, the transformer has been in a relatively stable operating state, and the operating state is relatively good and stable. Use the data of transformers in July and August of 2007 to predict the operating status of the transformer in the next month, which is September 2007. The LSTM network is used to predict the state parameter data, and the calculation results are input to the WNN to classify the state prediction. The obtained operational state prediction results are [0, 0.031, 0.077, 0.101, 0.124, 0.613, 0.054]. The results indicate that high-temperature overheating faults will occur in September 2007.

The actual situation is: on September 30, 2013, the total hydrocarbon content of the main transformer was 170.24 μL/L, and the relative monthly gas production rate of total hydrocarbons from July 25 to
September 30 was 13.57%, exceeding the value required by the regulations. According to the preliminary analysis of the chromatographic test data, the main transformer has an internal overheating fault that does not involve solid insulation. After inspection, it was found that the two winding leads on the lower wiring board of the 35kV bushing of the main transformer were crimped under the same bolt, and the tightening bolts of the phase C wiring board had been loosened. The failure point was consistent with the internal overheating failure of the chromatographic data analysis. The transformer operating state prediction results based on machine learning are consistent with the actual transformer operating state.

6. Conclusion
During the operation of the transformer, it will be subjected to long-term effects such as external electric field and thermal force, which will gradually change from a normal state to a potential problem and then a failure. Once the transformer malfunctions, it will cause significant economic losses. Therefore, it is of the great significance to establish a predictive model for the operating state of the transformer, in order to timely detect potential problems of the transformer.

In recent years, with the improvement of computing ability and the optimization of algorithm network structure, more dimensional data can be obtained to evaluate the operating state of the transformer. This paper uses multi-dimensional data joint analysis to realize the prediction of transformer operation status. Comprehensively considering the gas concentration in the transformer oil, partial discharge data, infrared detection data and ambient temperature information, a transformer fault classifier is established using WNN, and then the characteristic parameters are predicted using LSTM considering the time attention mechanism, and the obtained results are input to the trained classifier. Finally, the operating state of the transformer can be obtained.

Acknowledgments
This work was supported by the project “Research and application of state integrated perception and diagnosis for substation equipment” (No. P19-085).

References
[1] Pramanik S, Duvvury V S B C, Sahoo S. Tank Current Measurement of Three-Phase Transformer: Its Resonance Behavior and Sensitivity to Detect Mechanical Faults[J]. IEEE Transactions on Power Delivery, 2019, 34(6): 2211-2218.
[2] Gerdroodbari Y Z, Davarpanah M, Farhangi S. Remanent Flux Negative Effects on Transformer Diagnostic Test Results and a Novel Approach for Its Elimination[J]. IEEE Transactions on Power Delivery, 2018, 33(6): 2938-2945.
[3] Kazerooni M, Overbye T J. Transformer protection in large-scale power systems during geomagnetic disturbances using line switching[J]. IEEE Transactions on Power Systems, 2018, 33(6): 5990-5999.
[4] Jiao Z, Li Z. Novel magnetization hysteresis-based power-transformer protection algorithm[J]. IEEE Transactions on Power Delivery, 2018, 33(5): 2562-2570.
[5] Kari T, Gao W, Zhao D, et al. Hybrid feature selection approach for power transformer fault diagnosis based on support vector machine and genetic algorithm[J]. IET Generation, Transmission & Distribution, 2018, 12(21): 5672-5680.
[6] Liao R, Zheng H, Grzybowski S, et al. c[J]. Electric Power Systems Research, 2011, 81(12): 2074-2080.
[7] Ghunem R A, Assaleh K, El-Hag A H. Artificial neural networks with stepwise regression for predicting transformer oil furan content[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2012, 19(2).
[8] Dai J, Song H, Sheng G, et al. LSTM networks for the trend prediction of gases dissolved in power transformer insulation oil[C]/2018 12th International Conference on the Properties and Applications of Dielectric Materials (ICPADM). IEEE, 2018: 666-669.
[9] Meng K, Dong Z Y, Wang D H, et al. A self-adaptive RBF neural network classifier for transformer fault analysis[J]. IEEE Transactions on Power Systems, 2010, 25(3): 1350-1360.

[10] Sun Y, Zhang S, Miao C, et al. Improved BP neural network for transformer fault diagnosis[J]. Journal of China University of Mining and Technology, 2007, 17(1): 138-142.

[11] Duval M, Lamarre L. The duval pentagon-a new complementary tool for the interpretation of dissolved gas analysis in transformers[J]. IEEE Electrical Insulation Magazine, 2014, 30(6):9-12.

[12] Bacha K, Souahlia S, Gossa M. Power transformer fault diagnosis based on dissolved gas analysis by support vector machine[J]. Electric power systems research, 2012, 83(1): 73-79.9

[13] Abbasi A R, Mahmoudi M R, Avazzadeh Z. Diagnosis and clustering of power transformer winding fault types by cross-correlation and clustering analysis of FRA results[J]. IET Generation, transmission & Distribution, 2018, 12(19): 4301-4309.