Prediction of dissolved gas in power transformer oil based on LSTM-GA

Xin Zhang¹, Shengyuan Wang², Yijun Jiang¹,³, Feifei Wu¹ and Chenming Sun¹

¹ Hangzhou Power Equipment Manufacturing Co., Ltd., Hangzhou 310018, China; ² Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Ministry of Education(Northeast Electric Power University), Jilin 132012, China
³ Corresponding author’s e-mail: 1084610669@qq.com

Abstract. Because the change process of the content of dissolved gas in transformer oil is fluctuating and is affected by many factors such as oil temperature and external environment, the gas content does not increase exponentially and is nonlinear and non-stationary. This feature determines the current prediction technology cannot accurately predict the concentration of dissolved gas in transformer oil. To improve the prediction accuracy of the dissolved gas content in the transformer oil, more accurately evaluate the transformer status and customize the entire preventive plan under the alarming development trend of the transformer, to minimize the prediction error, this paper proposes a transformer combining genetic algorithm and long short-term memory (LSTM) neural network Prediction model of dissolved gas content in oil. The genetic algorithm (GA) is used to optimize look back (lb), LSTM nets (ls), epochs (ep), and the dropout (dp), and then the genetic algorithm is combined with the long short-term memory neural network. The gas content is predicted. This model overcomes the problem of low prediction accuracy caused by selecting parameters based on experience. The analysis result of the calculation example shows that compared with the traditional prediction algorithm, the proposed method can better track the change law of the dissolved gas concentration in the oil, improve the prediction accuracy, and provide a strong guarantee for the safe and stable operation of power transformers.

1. Introduction

The power transformer is one of the most critical equipment in the power system. Its stable, safe, and trouble-free operation is the vital factor in ensuring the reliable power supply of the entire power system [1]. During the regular process of the transformer, due to the appearance of internal insulation aging and other phenomena, a small amount of gas is dissolved in the insulating oil inside the transformer, mainly including Hydrogen(H₂), Methane(CH₄), Ethane(C₂H₆), Acetylene(C₂H₂), Ethylene(C₂H₄), Carbon monoxide(CO), carbon dioxide(CO₂) seven gases. In response to this phenomenon, the internationally recognized DGA method can be used to know the current fault status of the transformer or some latent faults and their degree of development in the early stage [2]. The online oil chromatographic monitoring device can form a historical gas monitoring sequence, and then predict related gases, which can be used as an essential basis for evaluating the status of a transformer, and has important reference significance for customizing exclusive preventive programs under the adverse development trend of transformers. There are several mainstream methods for predicting the gas content of transformers:
The statistical forecasting method mainly conducts time series analysis, such as using gray model prediction [3-5] and differential autoregressive moving average method prediction [6]. However, the gray model has advantages in dealing with index and linear data series. But its accuracy mainly depends on the distribution characteristics of the experimental data set. The difference autoregressive moving average method in the forecasting aspect is too simple to fit the function. It requires empirical support, which leads to the low complexity of the established model, and the forecast result is difficult to reflect the actual situation.

Artificial intelligence prediction uses sensors and other devices for data collection, computer technology for data processing and analysis, and building models for prediction. Standard synthetic intelligence methods are random forest [7-9], support vector machine [10-13], and recurrent neural network [14] and other forms. Due to the shortcomings of traditional artificial intelligence algorithms in dealing with time series, the predicted value and the actual value have a large error. Recurrent neural network (RNN) overcomes this problem, but it has short-term memory problems and cannot handle long sequences, and the training time is too long. LSTM is a select recurrent neural network model that adds a memory module to RNN and is widely used in time series forecasting problems. The related parameters of \( lb, ls, ep, \) and \( dp \) are set only by manual experience, which is uncertain, which reduces the prediction effect of the model.

Because the content of dissolved gas in transformer oil has fluctuating characteristics and is affected by factors such as oil temperature, partial oil pressure, fault properties and its development speed, the sequence of dissolved gas content in transformer oil will also change. Affected by load and operating time, the gas content data does not alter strictly according to the exponential growth law. Therefore, a certain degree of non-linear and non-stationary characteristics of the gas content sequence in the transformer oil will have a more significant impact Forecast effect. For faulty and faulty transformer oil, the dissolved gas concentration data also has the characteristics of small gas data samples, and less information. This feature determines that the current prediction technology cannot accurately predict the dissolved gas concentration in transformer oil. Therefore, it is necessary to use appropriate methods to improve the prediction accuracy. This paper proposes a long-term and short-term neural network combined with a genetic algorithm to predict dissolved gas in the oil. The genetic algorithm is used to optimize the \( lb, ls, ep, \) and \( dp \) parameters, which overcomes the low accuracy of the prediction model caused by the traditional LSTM using experience to determine the optimal parameters. The calculation example shows that the LSTM-GA combined prediction model proposed in this paper has higher accuracy than the traditional prediction model and can better track the changing trend of dissolved gas in the oil.

2. LSTM-GA algorithm principle

2.1. LSTM

LSTM was first proposed by Hochreiter and Schmidhuber based on RNN. RNN is superior to other neural networks in processing time series. LSTM solves RNN based on RNN and only has a short-term memory, and RNN training time exists.

Compared with RNN, LSTM adds a memory unit to judge whether the information is useful and has a complex dynamic structure. Each memory unit includes input gate \( i_t \), output gate \( o_t \), and forget gate \( f_t \). The gating unit controls the degree of influence of the current time information on the previous statement. The model has a long-term memory function and can solve the problem of long-term sequence prediction. Besides, LSTM can also solve gradient disappearance and gradient explosion during long sequence training [15-16]. Put merely, LSTM can perform better in longer sequences than ordinary RNN. The LSTM process is shown in Figure 1. At time \( t \), the cell state calculation process is as follows:
\begin{align*}
    i_t &= \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \\
    f_t &= \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \\
    o_t &= \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \\
    c'_t &= \tanh(W_c h_{t-1} + W_r x_t + b_c) \\
    c_t &= f_t \cdot c_{t-1} + i_t \cdot c'_t \\
    h_t &= o_t \cdot \tanh(c_t)
\end{align*}

In the above formula, \(i_t, o_t, f_t\) represents the state calculation results of the input gate, output gate, and forget gate. \(W_{ix}, W_{fx}, W_{ox}\) represent the weight matrix, \(b_i, b_f, b_o\) define the bias term and \(\sigma\) represents the sigmoid activation function. \(c'_t\) represents the unit state-input at time \(t\). \(\tanh\) is the hyperbolic tangent activation function; \(W_c\) and \(b_c\) respectively represent the state weight matrix and bias term of the input layer.

During LSTM training, the output value of the LSTM unit is calculated according to the forward propagation, the error value of the LSTM unit is calculated and backpropagated; the weight gradient is calculated according to the error value; the gradient descent is performed using the optimization algorithm, and the weight is recursively updated in real-time.

2.2. GA

A genetic algorithm is a computational model that simulates the biological evolution process of natural selection and the genetic mechanism of Darwin's biological evolution theory. It is a method of searching for the optimal solution by simulating the natural evolution process. The realization process of the genetic algorithm is roughly divided into four steps: coding, initializing population, calculating fitness value, and genetic operation. The genetic algorithm has been developed rapidly because it can be combined with other algorithms to optimize and improve different algorithms, and is widely used in various optimization problems.

The advantages of the genetic algorithm include the following three aspects: First, it has an excellent global search ability, through crossover, mutation, and other means to prevent falling into the optimal local solution, to obtain the optimal global solution; second, it has good compatibility, It can be combined with other algorithms for optimization; third, compared with general optimization
problems, it has lower requirements on mathematics, and does not need to establish conditions such as objective functions, and only needs to be solved according to the problem. Based on the above three advantages, this paper selects a genetic algorithm to optimize the neural network, find the optimal global solution, and perform modeling.

3. The realization process of gas concentration prediction based on LSTM-GA

3.1. DGA data collection and processing
The data source of this article is the 81 groups of five gas contents monitored by the No. 2 main transformer of a substation (produced by Chongqing ABB Transformer Co., Ltd.) from June 1, 2018, to August 20, 2018, which are used as data for the prediction model. Part of the data is shown in Table 1, and the unit of gas solubility is μL/L.

| ID | Date      | Methane | Ethane | Ethylene | Acetylene | Hydrogen |
|----|-----------|---------|--------|----------|-----------|----------|
| 1  | 2018/6/1  | 1.7     | 0.13   | 0.34     | 0.22      | 2.1      |
| 2  | 2018/6/2  | 1.67    | 0.12   | 0.35     | 0.21      | 2        |
| 3  | 2018/6/3  | 1.7     | 0.1    | 0.3      | 0.2       | 2        |

In order to make the optimization process of the optimal solution smoother and more comfortable to correctly converge to the optimal solution, the sample data of each characteristic parameter is normalized by the deviation standardization formula min-max and mapped to [0,1] Between, the conversion function is:

\[ X^* = \frac{X - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \]  

Among them: \( x_{\text{min}} \) is the minimum value of the sample data; \( x_{\text{max}} \) is the maximum value of the sample data; \( X \) is the sample data before conversion; \( X^* \) is the sample data after conversion. Part of the data after using min-max for normalization is shown in Table 2.

| ID | Date      | Methane | Ethane | Ethylene | Acetylene | Hydrogen |
|----|-----------|---------|--------|----------|-----------|----------|
| 1  | 2018/6/1  | 0.3     | 0.13   | 0.34     | 0.22      | 0.19     |
| 2  | 2018/6/2  | 0.2     | 0.12   | 0.35     | 0.21      | 0        |
| 3  | 2018/6/3  | 0.3     | 0.1    | 0.3      | 0.2       | 0        |

3.2. Model evaluation index
The evaluation index selected in this paper is the root mean square error (RMSE), root mean square error, which is the square root of the ratio of the square of the deviation between the observation and the actual value and the number of words n. In actual measurement, the number of comments is n It is always limited, and the real value can only be replaced by the best deal. The root square error is susceptible to the very large or tiny mistake in a set of measurements, so the root mean square error can well reflect the precision of the measure. The heart represents square error demonstrates the degree of deviation of the measurement data from the actual value. The smaller the root mean square error, the higher the measurement accuracy. Therefore, RMSE(Root Mean Square Error)can be used as the standard for evaluating the accuracy of this measurement process. The calculation formula is as follows:

\[ e_{\text{RMSE}} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \]  

(8)
In the formula: $y_i$, $\hat{y}_i$ respectively, represent the actual value and the predicted value of the dissolved gas concentration in the oil; $n$ represents the number of prediction verification data; $i$ represents the sequence number of the prediction point.

3.3. LSTM-GA network construction process

The LSTM-GA transformer oil dissolved gas prediction model was established. Including data processing module, parameter optimization module, network training module, network prediction module, using GA to data time window step $l_b$, network hidden layer $l_s$ in the LSTM model, training times $e_p$, $d_p$, four parameters are optimized in the search space and determined Optimal parameter combination, the dissolved gas content in transformer oil is used as input data. The predicted value of dissolved gas content in transformer oil in the next stage is used as an output matrix, and the model weight is adaptively adjusted through a long, and short-term memory loop neural network to form a fit. The LSTM-GA model inputs test data for prediction, compares the prediction results with actual test data for error, and outputs.

4. Example analysis

The computer configuration and software environment used in this experiment are as follows: the processor is Intel i5-8400, the memory is 16GB; the computer system is Windows 10 (64-bit); the programming language version is Python 3.6.2; the development environment is the integrated development environment Anaconda; Keras has the advantages of modularization and supports the free combination of training model layers. Keras is used in the experiment to predict dissolved gas in transformer oil. Set the individuals in the population to 50, the number of iterations to 200, the mutation probability to 0.1, and the crossover probability to 0.6. Finally, the optimal time window step size is set to 13, the number of hidden layers of the network is 100, the epochs is 200, the dropout is 0.37, the random seed is selected 7, and the prediction results of the four data samples are drawn on the two prediction models.

As shown in Table 3, it is obvious that the prediction effect of the model optimized by the genetic algorithm is significantly better than that of the unoptimized model, and the predicted values of the five gases after optimization are better than those before optimization. The root mean square error has been significantly reduced, and the prediction accuracy of methane, ethane, acetylene, ethylene, and hydrogen content has increased by 2.4%, 4.7%, 4.5%, 4.1%, and 5%, respectively. LSTM-GA performs better in predicting gas in transformer oil, with rapid trend change, strong model adaptability, higher prediction accuracy, and smaller errors. Compared with literature [17], this method has better prediction effect. The improvement of gas prediction accuracy has better state assessment capabilities for transformers and makes a significant contribution to the safe and stable operation of the entire power grid.

| GAS | CH₄ | C₂H₆ | C₂H₂ | C₂H₄ | H₂ |
|-----|-----|------|------|------|----|
| LSMT-GA | 0.265 | 0.254 | 0.248 | 0.321 | 0.335 |
| LSTM | 0.289 | 0.301 | 0.293 | 0.362 | 0.385 |

5. Conclusions

Based on the current rapid development of artificial intelligence technology, this paper proposes a network model based on LSTM-GA to track the change law of dissolved gas concentration in the oil for the requirement of increasing prediction accuracy of dissolved gas concentration in transformer oil. Get the following conclusions:

(1) The GA optimization algorithm is used to optimize the parameters in the LSTM model, which solves the problem of insufficient model fitting ability and low prediction accuracy caused by selecting parameters based on experience.
(2) Utilizing the characteristics of LSTM network suitable for time series, compared with other models, the LSTM-GA combined prediction model constructed in this paper has a more significant improvement in prediction accuracy.

(3) The progress of the prediction accuracy of dissolved gas in transformer oil and the customized complete preventive plan under the alarming development trend of transformers has essential reference significance, and at the same time has essential value for the safe and stable operation of the power system.

(4) The rapid development of computer technology combined with the comprehensive application of the big data platform, the application of the model to other prediction fields, may dig out more practical information, and then improve the prediction accuracy, which can provide theories for subsequent fault diagnosis and state evaluation of transformers guide.

Acknowledgment
This work was supported by Hangzhou Power Equipment Manufacturing Co., Ltd. The project number is YF200501.

References
[1] Jiang Xiuchen, Sheng Gehao 2018 Research and Application of Big Data Analysis of Power Equipment State[J] High Voltage Technology 44(04) 1041-1050
[2] Li Changhai 2019 Research on fault diagnosis method of oil-immersed power transformer based on support vector machine[J] Electrical Application 38(05) b67-72
[3] Sun Jianmei, Qian Xiuting, Wang Yongqing 2019 Medium and long term power load forecasting based on improved grey model [J] Electrical technology (19) 28-31
[4] Xu Ming, Fang Yangyang, Yang Peng 2018 Prediction of power transformer oil temperature based on grey model algorithm [J] Journal of electric power 33(05) 359-364
[5] Zhang Qingping, Yan Zhenhua, Zhou Xiu, Li xiuguang, Niu Bo 2019 Reliability analysis of power transformer based on grey model [J] Ningxia electric power (05) 62-65
[6] He Xingping, Geng Yuan, Guo Zhiwei, Wang Ting, Duan Shousheng 2016 Temperature prediction of electrical equipment based on differential autoregressive moving average model [J] Automation and instrumentation (12) 96-98
[7] He Jianzhang, Wang Haibo, Ji Zhixiang, Meng Xiangjun, Zhang Tao 2017 Prediction of heavy overload of distribution transformer based on random forest theory[J] Power System Technology 41(08) 2593-2597
[8] Luo Yan, Xiao Fusheng, Wang tinggang, Zhou Zhihai 2020 Risk assessment method of power grid real-time operation based on random forest [J] Information technology 44(04) 23-26
[9] Qiao Liwei, Wang Jingyi, Guo Wei, Li Guowen, Han Junjie 2020 Medium and short term power consumption prediction based on Stochastic Forest algorithm [J] Journal of electric power science and technology 35(02) 150-156
[10] Gu Hongxia, Huang Zhihua 2018 Identification of winding deformation type of power transformer based on SVM [J] Internal combustion engine and accessories (13) 131-134
[11] Sun Zhipeng, Cui Qing, Zhang Zhilei, Wang Tao, Zhang Tianwei 2019 Application of multi classification support vector machine in power transformer fault diagnosis [J] Electrical technology 20(10) 25-28
[12] Liu Chen Fei, Cui Hao, Yang, Li Xin, Shu Jiang, Li Ya 2019 Transformer fault diagnosis based on support vector machine under asymmetric samples [J] High voltage apparatus 55(07) 216-220
[13] Zhou Yanzhen, Wu Junyong, Ji Luyu, Yu Zhihong, Hao Liangliang 2018 Power system transient stability prediction and preventive control based on a two-stage support vector machine[J] Proceedings of the Chinese Society of Electrical Engineering 38(01) 137-147
[14] Yuan Jiabo, Xu Pengcheng, Li Lei, Liu Yanwen, Wang Xin, Zheng Yihui 2020 Prediction method of transformer oil-paper insulation aging based on chicken flock optimization BP
neural network[J] Journal of Electric Power Science and Technology 35(04) 33-41

[15] Tang Chengshun, Sun Dan, Tang Wei, Cao Shiyu, Li Jianzhao, Jing Jianping 2020 Surface stress prediction model of steam turbine rotor based on LSTM recurrent neural network[J/OL] Proceedings of the Chinese Society of Electrical Engineering:1-15[2020-08-29]

[16] Gaiping Sun, Chuanwen Jiang, Xu Wang, Xiu Yang 2020 Short - term building load forecast based on data - mining feature selection and LSTM - RNN method[J] IEEJ Transactions on Electrical and Electronic Engineering 15(7)

[17] Wang Qingjun, Bai Yang, Liang Hao, Yan Da 2015 Gas prediction in transformer oil based on grey theory[J] Agricultural Science and Technology and Equipment (07) 47-48