Paper ID APEN-MIT-2022_1182
Applied Energy Symposium: MIT A+B
July 5-8, 2022 • Cambridge, USA

Prediction Method of Dissolved Gas in Transformer Oil Based on Firefly Algorithm - Random Forest

Xiu Zhou
Power Science Research Institute of State Grid Ningxia Power Co.
Yinchuan, China
915406702@qq.com

Yunlong Ma
Power Science Research Institute of State Grid Ningxia Power Co.
Yinchuan, China
myl8858@qq.com

Yan Luo
Power Science Research Institute of State Grid Ningxia Power Co.
Yinchuan, China
444701910@qq.com

Tian Tian
Power Science Research Institute of State Grid Ningxia Power Co.
Yinchuan, China
titan9012@163.com

Weifeng Liu
Power Science Research Institute of StateGrid Ningxia Power Co.
Yinchuan, China
1106307501@qq.com

Xiuguang Li
Power Science Research Institute of StateGrid Ningxia Power Co.
Yinchuan, China
13895677195@126.com

Ninghui He
Power Science Research Institute of StateGrid Ningxia Power Co.
Yinchuan, China
178858519@qq.com

Zhenghua Yan
Power Science Research Institute of StateGrid Ningxia Power Co.
Yinchuan, China
154460982@qq.com

Hui Ni
Power Science Research Institute of StateGrid Ningxia Power Co.
Yinchuan, China
nihui1989@foxmail.com

Abstract—When a transformer fault occurs, the transformer oil will decompose and produce a large amount of dissolved gas in the oil, based on the dissolved gas in the oil to diagnose whether there is a fault in the transformer, known as dissolved gas analysis (DGA), in order to effectively predict whether a transformer fault will occur in the future, so as to prevent the development of the fault in time at the early stage of the fault, proposed A model for predicting the dissolved gas concentration in transformer oil based on the firefly algorithm (FA) optimized random forest (RF), which uses the random forest as the prediction model and adjusts the parameters in the RF by means of the firefly algorithm. The experimental results show that the FA algorithm can effectively optimize the parameters in the RF and improve the prediction accuracy of the model, overcoming the shortcomings of the traditional RF algorithm which uses random parameters with low accuracy, and the model can predict the dissolved gas concentration in oil more accurately than the existing methods.

Keywords—Transformer, Prediction of dissolved gas in oil, Firefly algorithm, Random forest

I. INTRODUCTION

Power transformers are the pivotal equipment for the conversion of electrical energy in the power system. Once a transformer has failed, it will cause irreversible damage to the power system, so it is particularly important to find out whether there is a fault in the transformer in time and to arrange a reasonable maintenance plan. The insulation system of power transformers mainly adopts an oil-paper insulation structure, in which, the insulating oil in the transformer mainly undertakes the role of insulation and cooling, in the good working condition of the transformer, its insulating oil in the gas content is less, but with the power transformer aging and insulation failure, the insulating oil in the arc, overheating and other effects will be decomposed into a large number of decomposition into a large number of low molecular hydrocarbons dissolved in the oil, these characteristics The gases mainly include hydrogen (H_2), methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4), acetylene (C_2H_2), carbon monoxide (CO), carbon dioxide (CO_2), etc. The analysis of transformers for faults based on the dissolved gases in the oil is called dissolved gasses analysis (DGA). Based on the dissolved gas content in the oil, it is possible to assess whether the transformer is operating under adverse operating conditions and whether there are latent faults inside the transformer, and by predicting the dissolved gas content in the transformer oil, it is possible to provide technical support for fault prediction and condition maintenance of transformers, which has certain engineering application value.

The common models for predicting dissolved gas in transformer oil are time series models, regression analysis, and grey prediction models, but such statistical analysis models are often influenced by the distribution pattern of the data itself and are often difficult to achieve high accuracy. With the development of artificial intelligence technology, machine learning algorithms are widely used in the prediction of dissolved gas in oil due to their powerful non-linear fitting ability, of which the most representative ones are Artificial neural networks, and support vector machines [8-13]. There are also integrated algorithmic prediction models, such as the random forest model represented by the bagging set, which has high accuracy in predicting dissolved gases in oil [14-16]. With the hot development of deep learning, some scholars have tried to use long and short-term memory networks and deep confidence networks to predict the content of dissolved gas in oil, although the prediction

accuracy can be improved, its network training time is long, which is difficult to meet the actual engineering needs

In view of this, this paper proposes a combined prediction model based on the firefly algorithm and RF, in which the parameters of RF are optimized using the firefly optimization algorithm on the basis of RF as a prediction model for dissolved gas in oil, and the optimal parameters of RF are obtained through a continuous iterative search of the firefly population, overcoming the shortcomings of the traditional RF model using empirical parameters with low accuracy. In order to verify the superiority of the correlation method proposed in this paper, oil chromatography data of a transformer is used to verify the performance of the model, and the dissolved gas content at time t is used as the input of the model, and the target gas content at time t+1 is used as the output, and the results show that the correlation method proposed in this paper has higher prediction accuracy compared with the existing methods, and is a more suitable The results show that the proposed method is a more suitable model for the prediction of dissolved gas in transformer oil.

I. RANDOM FOREST

The random forest (RF) is an intelligent algorithm that has been used in various fields in recent years. From a practical point of view, RF has high accuracy in predicting almost any category of data. The main idea is to use the decision tree algorithm as the basic structural unit, the Bagging algorithm as the model framework, and to introduce randomness in the composition of the "forest", that is, the random forest algorithm.

A. Decision Tree

The decision tree is the basic building block of the RF model and is the basis of the RF algorithm. A decision tree classification is a method of inferring classification rules in the form of a decision tree representation from a disordered set of training samples, which contains a number of split attributes (features) and target attributes (labels), with only three main components: the root node, non-leaf nodes, and leaf nodes. The trunk specifies the result of the classification by the current attribute; the leaf node represents the marker after the classification has been completed. It uses a top-down principle to select split attributes at each node starting at the root and branches down the training sample set according to all possible values of that attribute until each subset of data has a unique corresponding target. When a test sample is input, the decision tree determines a unique path from the root node to a leaf node, and the class of the leaf node of that path is the class to which the sample to be classified belongs, whose classification model is shown in Figure 1.

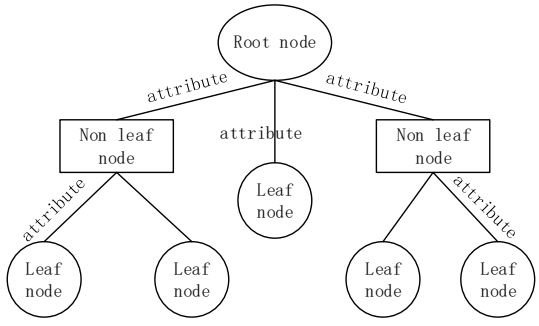

Fig. 1. Decision tree model

The key to constructing a decision tree is the choice of the optimal partitioning attribute. The choice of the partitioning attribute for a decision tree in a random forest is based on the "Gini index":

$$Gini(D) = 1\text{-}\sum_{i=1}^{n} p_i^{\,2} \tag{1}$$

Where P_i and P_j are the proportion of class *i* or *j* samples in the sample set in the total sample.

From equation (1), the Gini index captures the likelihood that any two members of the sample set will be taken into different categories, so that the higher the purity of the sample set, the smaller the Gini value will be instead. Denote the Gini index for attribute *a* as :

$$\text{G}ini\_index(X,a) = \sum_{m=1}^{M} \frac{|Xm|}{|X|} Gini(Xm) \tag{2}$$

B. Bagging Algorithm

Bagging (bootstrap aggregating) is an integrated learning framework that can improve the accuracy of the algorithm, the main idea of which is: in the sample set *D*={*X1, X2.....Xm*}, *m* samples (bootstrap sampling) are extracted in a randomized and put-back manner to form a new set of samples, and this step is repeated *n* times to obtain n weak learners.

Since Bagging takes a random sample with replacement, the probability that for a given sample in the original sample set, this sample will not be taken once in *m* samples is:

$$P = (1-\frac{1}{m})^m \tag{3}$$

When the original samples are infinite and the number of acquisitions is also infinite, there is $P \to \frac{1}{e} \approx 0.368$, which results in approximately one-third of the samples not being acquired in each acquisition from the original sample set, which becomes the out of bag (OOB) data, and these out of bag data can be used as the validation set to test the generalization ability of the learner.

C. The Principle of Random Forest

This combination of decision trees and Bagging is the RF algorithm, which is essentially a variation of the Bagging framework using decision trees as the base learner, with randomization added to the construction of the decision trees to achieve better generalization. The algorithmic process is rough as follows:

Step 1: For the sample set *D*={*X1, X2.....Xm*}, use bootstrap sampling to draw *n* samples to form a new set D1 and use *D1* as the training data for the decision tree. This is called the number of trees in the random forest *Ntree*.

Step 2: If each sample has *M* features (attributes), then take a constant $m \leqslant M$. Each time the decision tree splits and grows, *m* features are randomly selected as candidates for the division. This evaluation uses the Gini index as a principle to select the best segmentation feature among the candidate

features for the current sample. The value of *m* remains constant throughout the growth of the "forest", and the optimal attribute is divided by the value of these *m* sub-attributes, called *mtry*.

Step 3: After repeating step 1 and step 2 to obtain *N* decision trees, for the classification model, the final RF model is combined using the minority rule of thumb voting method, while for the regression model, the predicted values of the output of the *N* decision trees are counted using the mean, for this paper, the prediction of the dissolved gas content in the oil is a regression problem of the random forest algorithm.

In general, the more trees there are in a forest, the greater the diversity. In the RF algorithm, as the number of decision trees increases, the higher the generalization performance of RF will also be achieved. On the other hand, in random forests, since each split of the decision tree at this point does not take into account all sample attributes, but rather examines a random set of attributes, each decision tree is no longer deterministic, which also increases the generalization ability of the RF model. In general, the tree *Ntree* of a decision tree ranges between the interval [1,500] and, in this paper, the parameter *mtry* ranges between the interval [1,8] because there are eight characteristic gases in the sample set of dissolved gases in the oil.

II. FIREFLY ALGORITHM

The firefly algorithm (FA)[22] is a heuristic algorithm, and the main idea of this algorithm is to simulate the flickering behavior of fireflies in nature. The fireflies attract other fireflies by fluorescence, and the fireflies with low fluorescence will move towards the fireflies with high fluorescence. In this algorithm, the position coordinates of the fireflies represent the solution of the objective function, and the fireflies with bright fluorescence represent the more optimal solution of the objective function. In this case, the expression for the light intensity of the fireflies is shown in equation (4):

$$\mathrm{I}(\mathrm{r}) = I_0 e^{-\gamma r^2} \tag{4}$$

Where I_0 is the initial light intensity, r is the spatial distance between the fireflies, and γ is the light absorption coefficient.

The expression for the attractiveness of the firefly versus the light intensity is shown in equation (5)：

$$\beta(\mathrm{r}) = \beta_0 e^{-\gamma r^2} \tag{5}$$

where β_0 is the absorbance at r=0.

$$\mathrm{r}_{ij} = \| X_i - X_j \| = \sqrt{\sum_{k=1}^{d} (x_{i,k} - x_{j,k})^2} \tag{6}$$

where $x_{i,k}$ are the coordinate values of the *i*th firefly.

Firefly *i* will move in the direction of firefly *j*, which is brighter than it：

$$\mathrm{x}_i = x_i + \beta_0 e^{-\gamma r_{ij}^2}(x_j - x_i) + \alpha(rand - \frac{1}{2}) \tag{6}$$

where α is the step size factor.

A. *The Flow of Optimizing Random Forest with Firefly Algorithm*

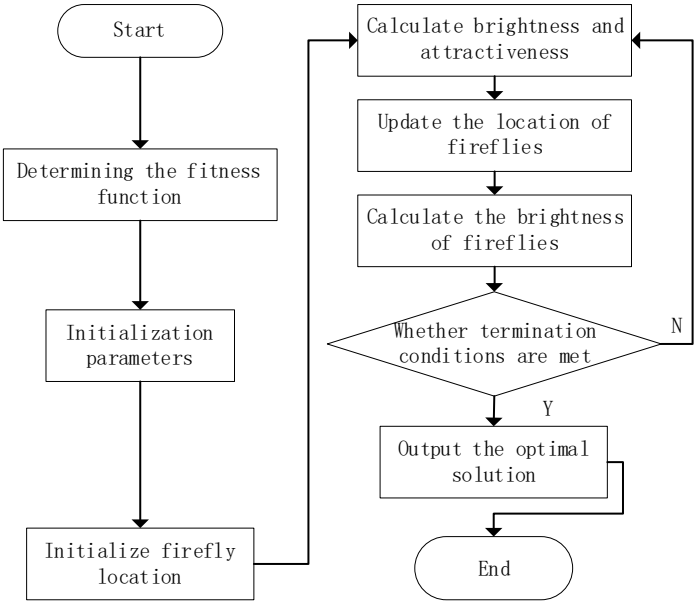

Fig. 2. Flow chart of FA-RF based model for predicting dissolved gas in transformer oil

According to the above principle, the specific flow chart of the FA-RF model for predicting dissolved gas in transformer oil is shown in Figure 2, as follows:

(1) Collect transformer oil chromatography data as the training data of the RF model to predict the content of CH_4, for example, the content of gas at time *t* as the input of the RF model, and the content of CH_4 at time *t+1* is the output of RF. Randomly divide the training set and test set, and take the minimum value of mean square error as the objective function of FA optimization, to be optimized as the number of decision trees *ntree* and decision tree splitting attribute *mtry* in the RF model;

(2) Initializing the parameters of the FA algorithm, the initial parameters of FA in this paper are the number of populations n=10, the maximum number of iterations T=500, the step factor α=0.25, the degree of attraction β=1, and the light absorption coefficient γ=1;

(3) Calculation of firefly brightness and attractiveness;

(4) Updating the location of fireflies;

(5) judge whether the termination condition is satisfied, if not, go back to step (4), if satisfied, the optimal parameters are output and passed to the RF model so that the FA-RF transformer oil dissolved gas prediction model is established.

III. EXPERIMENTAL RESULTS AND ANALYSIS

In order to verify that the model proposed in this paper has higher prediction accuracy, the actual data of transformer oil chromatography in a certain area were used for analysis.

The oil chromatography online monitoring device of this transformer has a collection period of 4 hours and an accuracy of 0.01 μL/L. The total characteristic gases collected include methane, ethylene, acetylene, hydrogen, ethane, total hydrocarbons, carbon dioxide, and carbon monoxide. In a certain period, the transformer from the oil chromatography online monitoring device uploaded to the back-end part of the data as shown in Table 1, all the gas content in uL / L.

Table 1 Sample example

Collection time	CH_4	C_2H_4	C_2H_2	H_2	C_2H_6	Total hydrocarbons	CO	CO_2
5/1 2:12	16.9	26.5	1.2	7.3	4.8	49.4	518.4	266
5/1 6:12	17.1	26.5	1.1	6.4	5	49.7	520	266.5
5/1 10:12	17.2	26.8	1.2	6.6	5.2	50.4	523.6	266.5
5/1 14:12	16.5	26.5	1.3	7	4.8	49.1	521.7	265.4
5/1 18:12	16.5	26.3	1.2	6.8	5	49	522.7	265.4

In the following section, taking the prediction of H_2 content as an example, the historical data of this transformer is first collected, and the gas content at moment *t* is taken as the input of RF, which is denoted as *X*, and the content of CH_4 at moment *t+1* is taken as the output of RF, which is denoted as *Y*. The data span from May 1, 2019, to August 17, 2019, with a total of 688 sets of data. These 688 sets of data were randomly divided into training and testing sets in the ratio of 9:1, and the minimum value of the mean square error was used as the fitness function of FA, then the fitness curve of the FA-optimized RF model is shown in Figure 3.

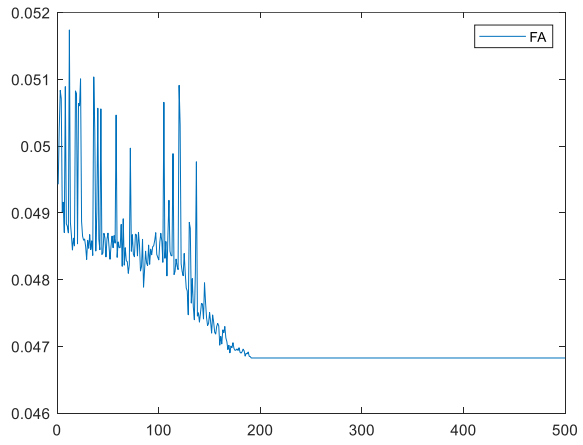

Fig. 3. Adaptation curve of FA

In figure 3, the horizontal coordinate is the number of iterations of the FA algorithm and the vertical coordinate is the average fitness error of the firefly algorithm. From figure 4, it can be seen that as the number of FA iterations increases, the prediction error gradually decreases and the firefly population converges and finally reaches a high accuracy. To prove the effectiveness of the related method proposed in this paper, the FA-RF model in this paper is compared with other models proposed in the literature, such as random forest (RF) and support vector machine (SVM), and the *RMSE* and *R2* (coefficient of determination) is calculated on the test set of the group, respectively, then the *RMSE* and *R2* of each model are shown in Table 2.

Table 2 Performance of each model

Model	RMSE	R2
FA-RF	0.02370	0.987422
RF	0.03615	0.956537
SVM	0.03912	0.95373

As can be seen from Table 2, compared with the unoptimized RF prediction model, the RF optimized by FA has higher prediction accuracy, which proves that the FA algorithm can effectively optimize the parameters in the RF and improve the performance of the model, and the related method proposed in this paper has higher accuracy compared with the existing models such as RF, SVM, etc.

IV. CONCLUSION

Dissolved gas analysis in oil is one of the effective means to detect whether a transformer is malfunctioning. Analyzing the dissolved gas concentration in oil and predicting the development trend, can provide an important basis for transformer fault diagnosis and condition assessment. In this paper, in order to improve the accuracy of dissolved gas content prediction in transformer oil, an FA-optimized RF-based dissolved gas prediction model in oil is proposed, analyzed with actual transformer cases, and compared with RF, SVM, and other methods, and the conclusions are drawn are specifically as follows:

FA can effectively optimize the parameters in RF, so as to improve the accuracy of RF, compared with RF, SVM , and other models, FA-RF has higher prediction accuracy, the model is a more suitable model for the prediction of dissolved gas in transformer oil, which can provide technical support for subsequent fault diagnosis and condition assessment of transformers.

REFERENCES

[1] Dai JJ, Song F, Yang Y, et al. A deep belief network-based method for predicting dissolved gas concentration in transformer oil[J]. Power Grid Technology, 2017,41(08):2737-2742.

[2] Wang K, Li JZ, Zhang SHQ, et al. New characteristic parameters of dissolved gas in oil for transformer fault diagnosis[J]. Chinese Journal of Electrical Engineering, 2016,36:6570-6578.

[3] Mirowski P, Lecun Y. Statistical Machine Learning and Dissolved Gas Analysis: A Review[J]. IEEE Transactions on Power Delivery, 2012,27(4):1791-1799.

[4] Rong Zhihai, Qi Bo, Li Chengrong, et al. A combined DBN diagnosis method for dissolved gas analysis in transformer oil[J]. Power Grid Technology, 2019,43(10):3800-3808.

[5] Wei Yunbing, Wang Donghui, Shao Qiuhua, et al. A new method for calibration of dissolved gases in power transformer oil[J]. Electrical Measurement and Instrumentation, 2016,53(01):122-128.

[6] Duval M. A review of faults detectable by gas-in-oil analysis in transformers[J]. IEEE Electrical Insulation Magazine, 2002,18:8-17.

[7] Liu KZ, Gou JZ, Luo ZHAO, et al. A particle swarm optimization-long and short-term memory network model for predicting dissolved gas concentration in transformer oil[J]. Power Grid Technology, 2020,44:2778-2784.

[8] Lin Yi, Yan Hongsen, Zhou Bo. Nonlinear time series prediction method based on multidimensional Taylor network and its application[J]. Control and Decision Making, 2014,29(05):795-801.

[9] Deng Yurong, Zhang Wei, Liang Junbin, et al. Trend measurement of gas production data in oil based on time series analysis[J]. Smart Grid, 2016,4:744-748.

[10] Zhao WQ, Zhu YL, Zhang SQ. A model for predicting dissolved gas in transformer oil based on improved gray theory[J]. Power Automation Equipment, 2008,28(9):23-26.

[11] Lin Xiangning, Huang Jing, Xiong Weihong, et al. Interval prediction of dissolved gas concentration in transformer oil[J]. Power Automation Equipment, 2016,36(04):73-77.

[12] Xiao Huaishuo, Li Qingquan, Shi Yalin, et al. Application of grey theory-variational modal decomposition and NSGA-II optimization of support vector machine for gas prediction in transformer oil[J]. Chinese Journal of Electrical Engineering, 2017,37(12):3643-3653.

[13] Zheng H, Zhang Y, Liu J, et al. A novel model based on wavelet LS-SVM integrated improved PSO algorithm for forecasting of dissolved gas contents in power transformers[J]. Electric Power Systems Research, 2018,155:196-205.

[14] Bian Jianpeng, Liao Ruijin, Yang Lijun. Prediction of dissolved gas concentration in transformer oil by applying weak buffer operator and least squares support vector machine[J]. Power Grid Technology, 2012,36(02):195-199.

[15] Lin J, Sheng G, Yan Y, et al. Prediction of Dissolved Gas Concentrations in Transformer Oil Based on the KPCA-FFOA-GRNN Model[J]. Energies, 2018,11(1):225.

[16] P. H, M. K, M. H, et al. Joint Optimization of Masks and Deep Recurrent Neural Networks for Monaural Source Separation[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2015,23(12):2136-2147.

[17] Lian W-L, Geng B, Zhou Zhou, Si G-Q. Prediction of dissolved gas concentration in transformer oil based on adaptive particle swarm optimization LSSVM[J]. New Technology of Electrical Power,2021,40(05):42-49

[18] Huang X-B, Jiang W-T, Zhu Y-Can, Tian Y. Transformer fault prediction based on time series and support vector machine[J]. High Voltage Technology,2020,46(07):2530-2538.

[19] Wang X, Rong HN, Wang J, Zhang GX. Prediction of dissolved gas concentration in transformer oil based on wavelet decomposition and long and short-term memory networks[J]. Electrical Engineering Technology,2020(09):24-29+33.

[20] Lu Peiwen, Huang Yongjing, Zhang Heng, Dong Fengzhu. Research on gas fault diagnosis in transformer oil based on Bagging improvement algorithm [J]. Electrical and mechanical engineering technology,2020,49(04):13-15+111.

[21] Wang Ke, Gou Jiazhu, Peng Jing, Liu Kezhen, Tian Xiaohang, Sheng Ge Hao. Prediction of dissolved gas concentration in transformer oil based on LSTM networks[J]. Electronic Measurement Technology,2020,43(04):81-87.

[22] Song Ani,Bao Xianzhe,Quan Yi. Chaotic adaptive firefly algorithm-based allocation strategy for UAVs[J]. Computer Applications and Software,2022,39(02):300-306.