Bayesian Network and Association Rules-based Transformer Oil Temperature Prediction

Wei Rao1,2*, Lipeng Zhu1,2, Sen Pan1,2, Pei Yang1,2 and Junfeng Qiao1,2

1Global Energy Interconnection Research Institute Co. Ltd, Beijing, 102209, China
2Artificial Intelligence on Electric Power System State Grid Corporation Joint Laboratory, Beijing 102209, China
3Corresponding author’s e-mail: 310714175@qq.com

Abstract. The oil temperature prediction of transformer is very important for the operation stability and life evaluation of transformer. As the oil temperature prediction of transformer is still short of a comprehensive and efficient method with combining various information of transformer such as operation data and meteorological data, this paper proposed a Bayesian network and association rules-based transformer oil temperature prediction method, which can improve the prediction accuracy of RBF-NN for transformer oil temperature prediction. The proposed method first mines all association rules among transformer state data and transformer operation data and environmental meteorological information by combining the Bayesian network and the Apriori algorithm and then uses the association rules to improve the prediction accuracy of RBF-NN based on only transformer state data. A case study with a 500kV transformer is conducted to test the effectiveness of the proposed method, and the result shows that the proposed method can improve the prediction accuracy of RBF-NN about 10%.

1. Introduction
Transformer is one of the most important equipment in power transmission network. In order to ensure the safe and stable operation of transformer, it is very necessary to predict the oil temperature of transformer. The existing transformer oil temperature prediction methods are mainly based on a single or few state parameters. Most of them are limited to threshold diagnosis and the inherent relationship between transformer state variables and transformer operation data and environmental meteorological information to conduct a comprehensive prediction, which lead to the analysis and prediction results one-sided and inaccurate. Therefore, how to make full use of the inherent law between various state data of transformers to increase the prediction accuracy of transformer oil temperature becomes a key problem. Obviously, association rule mining among transformer state variables with combining transformer operation data and environmental meteorological information will help to improve the prediction accuracy of transformer oil temperature.

In this paper, the Bayesian network is introduced to mine the association rules between transformer oil temperature and its related state variables. An association rule mining method is proposed by combining the Bayesian network and the Apriori algorithm, and the results of association rule mining are used to improve the prediction accuracy of transformer oil temperature. First, the proposed method uses the Apriori algorithm to find out all frequent 3-itemsets. Second, the Bayesian network is introduced to mine all association rules and calculates the support and confidence of each association rule. Finally, the association rules mined by the proposed method is used to improve the prediction
accuracy of transformer oil temperature obtained by a radial basis function neural network (RBF-NN). A case study on a 500kV transformers is conducted to verify the effectiveness of association rule mining with combining the Bayesian network, and the results show that the proposed method can efficiently mine all association rules among transformer state variable and improve the prediction accuracy of transformer oil temperature with RBF-NN.

2. Related works
At present, the calculation methods of transformer oil temperature mainly include hot spot temperature calculation method, heat circuit model calculation method, numerical calculation method and intelligent prediction method. Among them, the empirical model for calculating hot spot temperature of transformer windings recommended by GB/T 15164-1994 and IEEE Std C57.91-1995 was first used. In view of the fact that the first derivative model in two guidelines does not take environmental temperature into account, many researchers improved the first derivative model and proposed a top oil temperature model including environmental temperature and a heat circuit model for solving hot spot temperature [1-3]. Numerical calculation method is widely applied in transformer internal temperature calculation with its continuous development, for example finite volume method is often used to solve the flow and heat transfer problem [4]. With the rapid development of artificial intelligence, intelligent algorithms for predicting transformer oil temperature based on historical monitoring data are emerging [5,6]. Li constructed a hot spot temperature and top oil temperature prediction model of transformer winding by using an artificial neural network, and obtained a lot of good prediction results [6]. In addition, support vector machine and Kalman filter are also applied to oil temperature prediction [7,8].

In summary, the direct measurement method is simple to operate and the measurement results are the most accurate, but in practical application, the high cost of measurement equipment and the difficulty of operation and maintenance increase the cost of measurement. The hot spot temperature calculation method often produces non-negligible errors when transformer load changes. The thermal circuit model calculation method is simple, but the calculation results need to be optimized. The numerical calculation method involves many parameters. The process of oil temperature prediction using intelligent algorithm is simple and the prediction accuracy is high. It will certainly develop into an important tool to solve the prediction problem of transformer oil temperature in the future.

3. Bayesian network and association rules-based transformer oil temperature prediction
In the prediction of transformer oil temperature based on Bayesian network and association rules, all association rules among various states of transformer are firstly found by using the Apriori algorithm and the Bayesian network, and then the model parameters of RBF-NN are optimized according to the association rules. Finally, the accurate prediction of transformer oil temperature is achieved by RBF-NN with the optimized parameters.

3.1 The construction of the Bayesian network
A Bayesian network is a directed acyclic graph that can describe the relationship of state variables and achieve the probabilistic reasoning. For a set of variables $X = \{X_1, X_2, ..., X_n\}$, the Bayesian network represents the joint probabilities based on conditional probability chains as shown in equation (1).

$$P(X_1, X_2, \cdots, X_n) = \prod_{i=1}^{n} P(X_i | X_1, X_2, \cdots, X_{i-1})$$

A Bayesian network consist of two parts: qualitative description and quantitative description.
A directed acyclic graph consists of several nodes and directed edges. Nodes represent random variables such as phenomena, states or attributes in a target problem. Directed edges represent the relation of dependence or causality between nodes and the arrow of a directed edge represents the direction of dependence or causality, and undirected edges between nodes represent the conditional independence of the corresponding variables.

A conditional probability table represents the correlation degree between a child node and its parent node, and the probability of a node without the parent node is its prior probability. Assuming that \( G \) denotes a directed acyclic graph, \( V \) is a set of \( n \) random variables, \( L \) denotes a set of edges, \( P \) denotes a set of conditional probability distributions, the Bayesian networks can be described as follows:

\[
BN = (G, P) = (V, L, P)
\]  

Herein, \( G, V \) and \( P \) can be described with the following equations.

\[
G = (V, L)
\]  

\[
V = \{V_1, V_2, \ldots, V_n\}
\]  

\[
P = \{P(V_i | V_1, V_2, \ldots, V_{i-1}, V_i \in V)\}
\]

According to the chain rule (1) and conditional independence assumption, if \( P_{a_i} \) is the parent node set of a variable \( V_i \), then the joint probability distribution can be described as follows:

\[
P(V) = P(V_1, V_2, \ldots, V_n) = \prod_{i=1}^{n} P(V_i | P_{a_i})
\]  

3.2 The key steps of association rule mining

The purpose of association rule mining is to find out the relationship between different items in a dataset. Association rule mining has the following important concepts:

1. The dataset of association rule mining is denoted as \( D = \{t_1, t_2, \ldots, t_m\} \), where \( t_k = \{i_1, i_2, \ldots, i_n\} \) is called an item. Each transaction has a unique identifier \( T_{id} \).

2. Let \( I = \{i_1, i_2, \ldots, i_n\} \) is a set of all items in \( D \). Any subset \( X \) of \( I \) is called an itemset of \( D \). If \( |X| = k \), \( X \) is a \( k \)-itemset. If \( X \in t_k \), the transaction \( t_k \) contains the itemset \( X \).

3. The number of the transactions containing the itemset \( X \) in \( D \) is called the support number of \( X \), which is denoted as \( \sigma_X \). The support of \( X \) is denoted as \( \text{support}(X) \):

\[
\text{support}(X) = \frac{\sigma_X}{|D|} \times 100\%
\]  

Where \( |D| \) is the element number of \( D \). If \( \text{support}(X) \) is not less than the minimum support threshold \( \text{minSup} \), then \( X \) is called a frequent itemset. Otherwise, \( X \) is called a non-frequent itemset.

4. If \( X \) and \( Y \) are two itemset and \( X \cap Y = \phi \), \( X \Rightarrow Y \) is called an association rule. The support of \( X \cup Y \) is called the support of \( X \Rightarrow Y \), which is denoted as \( \text{support}(X \Rightarrow Y) \).

\[
\text{support}(X \Rightarrow Y) = \text{support}(X \cup Y)
\]  

The confidence of \( X \Rightarrow Y \) is denoted as \( \text{confidence}(X \Rightarrow Y) \):

\[
\text{confidence}(X \Rightarrow Y) = \frac{\text{support}(X \cup Y)}{\text{support}(X)} \times 100\%
\]

The minimum confidence threshold should be specified in association rule mining and denoted as \( \text{minConf} \). The support is used to measure the statistical importance of an association rule in \( D \), and the confidence is used to measure the statistical credibility of an association rule in \( D \).

5. If \( X \Rightarrow Y \) satisfies \( \text{support}(X \Rightarrow Y) \geq \text{minSup} \) and \( \text{confidence}(X \Rightarrow Y) \geq \text{minConf} \), \( X \Rightarrow Y \) is called a strong association rule. If not, \( X \Rightarrow Y \) is called a weak association rule.

The purpose of association rule mining is to mine all strong association rules in \( D \). Association rule mining can be divided into two steps:

Step 1: Find out all frequent itemsets in \( D \) according to the minimum \( \text{support} \).

Step 2: Mine all association rules according to the frequent itemsets and the minimum \( \text{confidence} \).
3.3 Mining association rules with Bayesian network

The procedure of mining all association rules between state variables based on the Bayesian network is as follows:

Stage 1: Use the Apriori algorithm to find out all frequent 3-itemsets. The Apriori algorithm uses the iterative method of layer-by-layer search to generate the frequent itemsets from the candidate itemsets, and finally obtain all association rules from the frequent itemsets. The specific steps of finding out all frequent 3-itemsets are listed as follows:

Step 1. Scanning all transactions and calculating the occurrences number of each item to generate the frequent 1-itemsets \( L_1 \);

Step 2. Generating the frequent 3-itemsets \( L_3 \) from \( L_1 \) with the following method:

- Joining two patterns that have one common item to get \( C'_2 \) and pruning \( C'_2 \) according to the anti-monotonicity of frequent itemset to get candidate 3-itemsets \( C_2 \).
- Scanning the dataset \( D \) and calculating the \( \text{support}(C_2,t) \) of the candidate 2-itemsets contained in each transaction \( t \) and storing it in a hash table.
- Deleting the itemsets that the \( \text{support} \) is lower than \( \text{minSup} \) and getting the frequent 3-itemsets \( L_3 \).

Stage 2: Selecting the frequent 3-itemsets according to the concept of interest proposed by Piatetsky-Shapiro. When \( \text{support}(X \cup Y) \approx \text{support}(X) \times \text{support}(Y) \), we consider that the 2-itemset \( (X,Y) \) are independent and the association rule \( X \Rightarrow Y \) is not interesting. The interest of the 2-itemset \( (X,Y) \) is defined as follows:

\[
\text{interest}(X,Y) = \frac{\text{support}(X \cup Y)}{\text{support}(X) \times \text{support}(Y)} - 1 = \frac{P(X|Y) - P(X)}{P(X)}
\]

- If \( \text{interest}(X,Y) > 0 \), \( X \) and \( Y \) are positively correlated.
- If \( \text{interest}(X,Y) = 0 \), \( X \) and \( Y \) are independent and the 2-itemset \( (X,Y) \) will be deleted from the frequent 2-itemsets.
- If \( \text{interest}(X,Y) < 0 \), \( X \) and \( Y \) are negatively correlated.

Stage 3: According to the 2-itemsets \( L_2 \) in Stage 2, the Bayesian network can be constructed by taking each state as a node and taking \( L_2 \) and conditional probability distribution as the edge.

Stage 4: According to the Bayesian network, the association rule set \( R \) is generated.

Stage 5: Listing the probability table of all association rules and optimizing the parameters of RBF-NN with the association rules.

4. The case study

4.1 The associated rule mining of transformer stage variables based on Bayesian network

To test the effectiveness of the proposed method, we take a 500kV transformer in a substation as an example to mine all association rules of transformer state variables. The daily average value probability density of transformer operation data such as \( \text{operation current} \), \( \text{active power} \) and \( \text{reactive power} \). The daily average value probability density of transformer operation data such as \( \text{operation current} \), \( \text{active power} \) and \( \text{reactive power} \).
power are shown in figure 2. The probability density of environmental meteorological data such as air temperature, ground temperature, humidity and average wind velocity are shown in figure 3.

Figure 3. The probability density of environmental meteorological data.

The probability density of transformer on-line monitoring states such as various dissolved gas content and oil temperature are shown in figure 4.

Figure 4. The probability density of transformer on-line monitoring state data.
The association rules in all frequent itemset that the item number is more than 3 and their maximum support and maximum confidence are shown in table 1.

Table 1. The association rule mining result of transformer state data.

| Rules | LHS                                    | RHS                     | Support | Confidence |
|-------|----------------------------------------|-------------------------|---------|------------|
| R1    | Average Temperature, Average Operation Current and Maximum Operation Current | C₂H₃Content            | 0.55    | 0.72       |
| R2    | C₂H₃Content, CH₄ Content, C₂H₄ Content | Total Hydrocarbon Content | 0.63    | 0.81       |
| R3    | Average Temperature, Average Operation Current, C₂H₃Content | Oil Temperature         | 0.51    | 0.74       |

4.2 Using the result of association rule mining to improve the prediction accuracy of RBF-NN

RBF-NN is a forward three-layer neural network. The first layer is the input layer, the second layer is the hidden layer and the third layer is the output layer. As a neuron transformation function in the hidden layer, RBF is a radially symmetric and attenuated non-linear function for the central point, which changes a low-dimensional space input vector into a high-dimensional space input vector and makes a linear inseparable problem in a low-dimensional space become linear separable in a high-dimensional space. RBF-NN not only has the advantages of global approximation, good generalization, small computation and fast learning velocity, but also has no local minimum problem. Therefore, RBF is widely used in many fields such as time sequence analysis and pattern recognition. The 1200 days’ data of oil chromatogram, oil temperature and air temperature about the 500kV transformer from March 21, 2010 to June 28, 2013 are shown in figure 5, and the load current data are shown in figure 6.

![Figure 5. The 1200 days' state data of the 500kV transformer.](image-url)
Figure 6. The 1200 days' operation current data of the 500kV transformer.

The data in figure 5 and figure 6 are used as the input of RBF-NN to predict the oil temperature change of the transformer in the next 60 days. The output layer weights of RBF-NN are manually tuned by using the association rules shown in table 1 as prior knowledge, and the predicted results of transformer oil temperature are shown in figure 7.

Figure 7. The prediction result comparison of transformer oil temperature.

From figure 7, it can be observed that the average prediction error of RBF-NN with association rules is reduced about 10%, which proves that the proposed method can optimize the parameters of RBF-NN and improve the prediction accuracy of RBF-NN effectively.

5. Conclusion and future work
The changing trend of transformer oil temperature is affected by many factors such as load conditions and meteorological conditions. Using a large number of transformer historical data to mine the data law and predict the changing trend of transformer oil temperature is a key content of electric power big data analysis. In this paper, the correlation relationship among numerous state variables of transformer oil temperature is studied, and Bayesian network and association rules are introduced to improve the efficiency of association rule mining. It is applied to the prediction of transformer oil temperature, and improves the prediction accuracy of transformer oil temperature. In the future, we will predict the oil temperature of transformer with more machine learning algorithms.

Acknowledgments
This work was funded by the 2017 Science and Technology Project of SGCC (GYB17201700204): "Research of Fault Diagnosis and Maintenance Assistive for Transmission and Distribution Equipment Based on Big Data Technology." All the authors of this paper are appreciated for that.
References

[1] Swift G., Molinski T.S., Lehn W. (2001) A fundamental approach to transformer thermal modelling-part I: Theory and equivalent circuit. IEEE Transactions on Power Delivery, 16(2): 171-180.

[2] Susa D., Lehtonen M. (2006) Dynamic thermal modelling of power transformers: Further Development-Part II. IEEE Transactions on Power Delivery, 21(4): 1961-1969.

[3] Chen W.G., Li M.L., Sun C.X., et al. (2010) Analogue calculation model simulation of hot spot temperature of transformer winding. Journal of Chongqing University, 33(12):8-13.

[4] Chen W.G., Su X.P., Sun C.X., et al. (2011) Winding temperature distribution calculation of oil-immersed transformer based on finite volume method. Power automation equipment, 31(6): 23-27.

[5] Du W.X., Lu F., Ju X.Y. (2007) Fault Diagnosis of Power Transformer Based on BP Neural Network. Transformer, 3(44):46-47.

[6] Li M.L. (2010) Research on prediction method of transformer winding hot spot temperature based on BP neural network. Chongqing University, Chongqing.

[7] Li T. (2012) Research on the prediction of heat circuit calculation and hot spot temperature prediction of oil-immersed transformer by genetic support vector machine. Chongqing University, Chongqing.

[8] Wang Y.Q., Yue G.L., He J., et al. (2014) Research on top oil temperature prediction of power transformer based on Kalman filter algorithm. High Voltage Electrical Appliances, 50(8): 74-79.