

# Application of Similarity Technology in Transformers State Early Warning

Feng Li & Hongbin Wang

*Electric Power Research Institute of Guangdong Power Grid Corporation, Guangzhou, Guangdong, China*

**ABSTRACT:** This paper presents the application of similarity mining technology in transformers state early warning. Depending on the complexity of power transformer characteristics, similarity mining technology provides a transformer fault diagnosis model based on mass data. The analysis of historical data based on a large number of operating states is the foundation of the transformer normal state model which is derived by similarity mining technology. This paper describes the modeling process and the application of early warning in detail. The model also can be improved in diagnostic effect by rich training samples. The example also demonstrates the effectiveness of this method.

**Keywords:** similarity technology; status early warning; transformer

## 1 INTRODUCTION

Transformer is one of the most important equipment in power system. Though equipped with lightning arrester, multiple protections, but because the internal structure is complex, the electric field and thermal field is not equal in many factors, and accident rate is still high.

In the long term of safety monitoring of transformer, many scholars propose effective fault diagnosis method. With further increasing security requirements for transformers, people not only want to provide fault detection and isolation in the failure, but also can be able to know before the device fails. It is called the failure warning. This will avoid unplanned power outage losses. Fault maintenance changes from a reactive to a proactive prevention. Until advanced forecasting and planning management is done, to a certain extent, the development direction of the fault diagnosis methods is changed.

The basic idea of modeling transformer equipment based on similarity technology in this paper is: use a large amount of historical data to establish the normal state model of transformer, and the assessment of the transformer is done with the real time data. The transformer equipment status is judged by comparing the real-time status of transformer and historical status. The professionals can take appropriate measures when the system sends early warning message.

## 2 SIMILARITY TECHNOLOGY

Similarity calculation is used to measure the degree of similarity between objects. It is a basic calculation in data mining and natural language processing. The key technology is divided into two parts, the repre-

sentation of the object's feature, and the similar relationship between the feature sets. In information retrieval, pages sentence, recommendation systems, and so on, the calculation of similarity between the objects is important. And for different application scenarios, subject to limitations of data size, time and space overhead and the similarity calculation method will be different.

There are two vectors  $X = \{x_1, x_2, \dots, x_n\}$ ,  $Y = \{y_1, y_2, \dots, y_n\}$ . When its similarity function  $\text{sim}(X, Y) < \varepsilon$ ,  $\varepsilon$  is a given similarity threshold, the vectors bounded by  $\varepsilon$  are similar. Similarity function  $\text{sim}(X, Y)$  satisfy the positive definiteness, symmetry and triangle inequality.

The positive definiteness,  $\text{sim}(x, y) \geq 0$ , if and only if  $X = Y$ ,  $\text{sim}(x, y) = 0$ ;

Symmetry,  $\text{sim}(x, y) = \text{sim}(y, x)$ ,  
Satisfies the triangle inequality  $\text{sim}(a, c) \leq \text{sim}(a, b) + \text{sim}(b, c)$ .

When the similarity function  $\text{sim}(X, Y)$  is equal to the Euclidean distance  $d(X, Y)$ , and  $d(X, Y) < \varepsilon$ , the vector are similar under the boundary of  $\varepsilon$ .

Euclidean distance between two vectors reflects the degree of difference between them, the greater the value, the bigger the difference degree. But the Euclidean distance of vector is affected by the original sequence specific data dimensionless unit, whose value will be more than one. In many cases of the research, it is difficult to accurately compare similarities and closeness between vectors. Thus, this study needs to use similarity to measure the similarity of the two objects. The similarity varies between zero and one. The method of the study is to find the nearest distance between current state and model state to complete the equipment measuring point state estimation.

For any two vectors  $X = \{x_1, x_2, \dots, x_n\}$  and  $Y = \{y_1, y_2, \dots, y_n\}$

$\{y_1, y_2, \dots, y_n\}$ ,  $\text{sim}(X, Y)$  is the similarity function between  $X$  and  $Y$ . Given threshold  $\epsilon > 0$ , if  $X \approx Y$ , then we define  $s(X, Y) = 1 - \text{sim}(X, Y) / \epsilon$ . It is similarity of sequence  $X$  and  $Y$ .

For any set of vectors  $X$ , and vector group  $Y$  which is composed of a cluster of vector  $Y_n$ , it can be defined that the max similarity between  $X$  and  $Y_n$  is the similarity between  $X$  and  $Y$ .  $\text{sim}(X, Y) = \text{Max}(\text{sim}(X, Y_n))$

### 3 TRANSFORMER DATA MODELING

Similarity theory is based on a large amount of historical data, and the modeling sample data determines the accuracy of the model of equipment. Therefore the historical data which is used in modeling has become extremely important. The historical data used to generate the equipment state model should meet the following requirements:

Data should be long enough to cover a period of running time

Each data have expressed a normal operating status of the device;

Each set of data values must be sampled in the same time;

For example, there are 12 measuring point variables. The equipment is sampled once every minute. 10080 samples constitute the historical data for device modeling.

Every running state of the equipment is represented by a group of sampling value. A collection of historical data is the actual operation of the device. The function model generation program is to extract the most representative state points of the process objects. After the extraction, the model is constructed by these state points.

From the perspective of understanding the space, model generation process is to use historical data from the device state points for super ball model construction, as shown in Figure 1.

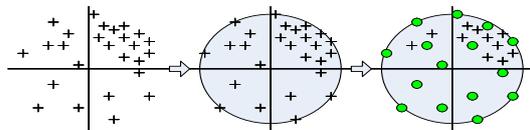


Figure 1. Generation process of dynamic object model

First, the normal work of the equipment history data is mapped into a state space. The circumscribed super ball is constructed by boundary points within the space. This super ball is a primary device object model. In order to improve the accuracy of the dynamic model, it is necessary to compare the state point inside the super ball. According to the requirements of the model, the state points are screened from all the points whose difference is bigger than a certain similarity index. These state points are associated reference

points which are inside the super ball. The super ball of dynamic model is divided detailed enough.

Through the calculation of the relevance similarity between the state points, the boundary and inside reference points together constitute the dynamic state of an accurate device model. It is the super ball dynamic model.

When the model receives real-time data, it will judge the location of the current data measurement points in the dynamic model. The distance between the state and the current state of the model is got to determine the current similarity. Using the similarity and model internal reference point, it can generate a prediction of the operation state. Predictive value depends not only on the measured operating rules long-term point of history, simultaneously also depends on the measurement of correlation similarity between points and other measuring points. The effect of interfering signals on predictive value is shielded. It greatly enhances the accuracy and reliability of the predicted values.

### 4 TRANSFORMER CONDITION MONITORING AND EARLY WARNING

The fault early warning system which is based on similarity modeling fault provides users with devices similarity graph. The failure warning system automatically recognizes the nuances of the current operating status of the device between the state and its history. The difference is displayed by a similarity line which varies between 0~100%. When the current operation state of the system equipment changes and deviates from the historical operation rule, the fault alarm system recognizes this kind of change and continuous monitoring. When the early warning monitoring line equipment state continued deterioration and crossed the "similarity curve", the system automatically releases early warning of potential failures. Also, the data of the failure early warning system come from the massive real-time data. It can predict the early failure and offer a real-time prediction of measuring points. There is a broken line in the similarity graph. It is the similarity line of equipment monitoring. Monitoring line is not manually set. However, the similarity line is derived by calculation of history data. Similarity curve is characterized by equipment operating conditions. When the similarity line is above the monitoring line, it suggests that the equipment is fine. But if the similarity line is jumping over the monitoring line, it suggests that the equipment deviates from the normal state. So the early warning comes. The users can read the related analysis, as shown in Figure 2.

### 5 FIELD APPLICATIONS

The following is a transformer state warning model established by similarity mining in a substation. The

application and effects will verify the method suggested in this paper.

5.1 Parameters selection of transformer modeling

Reference [5] points out that the parameter is selected to reflect the state of the transformer. The parameter is broadly divided into four categories:

- 1) On-line monitoring data
- 2) Measuring the amount of scheduled data
- 3) Test Data
- 4) The relevant qualitative indicators

Based on the similarity index, we initially selected the first three parameters for modeling. The detail

parameters are shown in Table 1.

5.2 Transformer modeling

After data selection is completed, the analysis and modeling is done by computer. The related data is stored in database. The parameters are shown in Table 2.

5.3 Display of transformer state assessment

Through the calculation of real-time transformer's



Figure 2. The similarity line of equipment

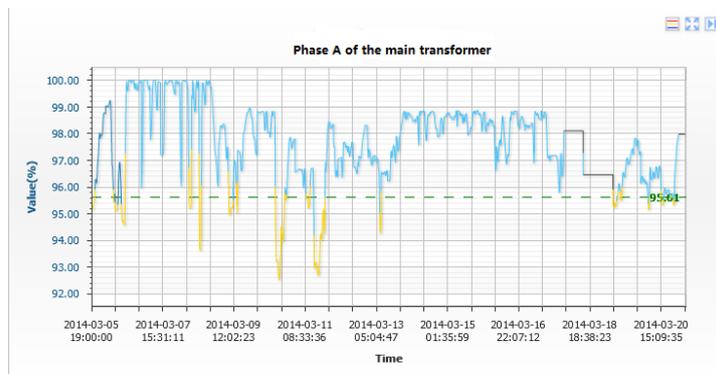


Figure 3. Similarity curve of main transformer

佛山顺德/#1主变A相 详细页面								
time	similarity value	associated measuring points	the number of abnormal points					
2014/3/8 14:38:45	0.93723	3	0					
. associated measuring points								
1	H2_A	#1	u/l	4.37	3-5	3.61-4.4	on line	2014-03-08 14:38:45
2	CO2_A	#1	u/l	729	652-749	666-735	on line	2014-03-08 14:38:45
3	ToHydC_A	#1	u/l	3.54	1-4	2.27-3.69	on line	2014-03-08 14:38:45

Figure 4. Detail interface of the abnormal state of main transformer

Table 1. Transformer modeling parameters

Number	Description	Unit
1	Main transformer current	A
2	Main transformer current of high voltage side	A
3	Main transformer current of low voltage side	A
4	Main transformer current of middle voltage side	A
5	Main transformer active power	KW
6	Main transformer active power of high voltage side	KW
7	Main transformer active power of low voltage side	KW
8	Main transformer active power of middle voltage side	KW
9	Main transformer reactive power	kVar
10	Main transformer reactive power of high voltage side	kVar
11	Main transformer reactive power of low voltage side	kVar
12	Main transformer reactive power of middle voltage side	kVar
13	Main transformer A, B, C phase's Total hydrocarbons	ul/l
14	Main transformer A, B, C phase's total composite gas	ul/l
15	Main transformer A, B, C phase's total amount of flammable gas	ul/l
16	Main transformer A, B, C phase's oxygen content	ul/l
17	Main transformer A, B, C phase's nitrogen content	ul/l
18	Main transformer A, B, C phase's hydrogen	ul/l
19	Main transformer A, B, C phase's containing gas concentration	ul/l
20	Main transformer A, B, C phase's carbon dioxide	ul/l
21	Main transformer A, B, C phase's carbon monoxide	ul/l
22	Main transformer A, B, C phase's methane	ul/l
23	Main transformer A, B, C phase's ethane	ul/l
24	Main transformer A, B, C phase's ethylene	ul/l
25	Main transformer A, B, C phase's acetylene	ul/l
26	Main transformer's neutral current	mA

Table 2. The main parameters

Sampling time	Total hydrocarbon	Hydrogen	Carbon monoxide	Carbon dioxide	Methane	Ethylene	In the side of the current	High side current
2014/03/06 18:30	3.53	3.79	146	700	2.57	0.47	959.79	419.54
2014/03/06 22:20	3.53	3.79	146	700	2.57	0.47	785.76	339.85
2014/03/07 02:28	3.29	3.86	141	699	2.26	0.53	479.89	207.42
2014/03/07 07:30	3.29	3.86	141	699	2.26	0.53	683.81	298.84
2014/03/07 10:30	3.29	3.86	141	699	2.26	0.53	1004.62	439.46
2014/03/07 13:30	3.29	3.86	141	699	2.26	0.53	818.28	355.08

Table 3. The abnormality of transformer parameters

Sampling time	Total hydrocarbon	Hydrogen	Carbon monoxide	Carbon dioxide	Methane	Ethylene	In the side of the current	High side current
2014/03/08 14:27	3.54	4.37	143	729	2.57	0.47	1510.88	655.09
2014/03/08 14:38	3.54	4.37	143	729	2.57	0.47	1520.55	659.78
2014/03/08 15:00	3.54	4.37	143	729	2.57	0.47	1531.98	666.81

operating data, it realizes real-time monitoring and assesses the state of transformer. In this study, we use the B/S mode, and let the user understand the running status of transformer through the browser. When the user opens a browser and logs in system, it is first to see the current state of the device. The user can click

on a device, into the similarity curve of the interface of the device. So the running state of equipment in a certain period of time can be seen in Figure 3.

When the similarity value is lower than the monitoring line, the user can enter by clicking on the similarity curve to the detail screen. It is easy to know the

main associated measurement information and the abnormal state, as shown in Figure 4.

In Figure 4, the user can know the specific parameters which cause equipment malfunction, and the actual operation of these parameters. The analysis of equipment fault is convenient. It can be seen from Figure 4, the abnormal problem is mainly composed of A phase's Hydrogen, A phase's hydrogen dioxide and total hydrocarbon of A phase, as shown in Table 3.

It can be seen from Table 2 and Table 3, when the transformer warns, the main parameters of the state are on the rise, and especially the acetylene has been slow climbed from 0 to 1.5. It needs more attention.

## 6 CONCLUSION

With the development of artificial intelligence technology, the fault diagnosis technology of transformer is gradually mature. The traditional "Application guide" type diagnostic mode gradually transition to the application of artificial intelligence technology comprehensive diagnosis method. At present, the scientific research personnel and operation personnel do a lot of research and practical work of intelligent fault diagnosis, and have made considerable achievements.

This paper adopts iEM "super ball" algorithm and the similarity of technology. A large amount of data of transformer is analyzed. The state is assessed by real time data, and the result is satisfactory. The algorithm has strong robustness, good convergence performance. The results of its diagnostic can guide the practical work.

## REFERENCES

- [1] Bridgman M S. 2002. Relating failure prognostics to system benefits. *Aerospace Conference Proceedings*, (7): 3521-3526.
- [2] Chang Shuping, Wu Ruitao. 2011. The application of failure prognostic system in state monitoring of power plant generation equipments. *GEESD, 2011 International Conference*, Jilin: [s.n.].
- [3] Chang Shuping, Guo Jianglong, Lv Yukun. 2011. The application of nonlinear state estimation method in fault warning system. *Software*, 32(7): 57-60.
- [4] Mo Juan, Wang Xue, Doong Ming, Yan Zhang. 2004. Power transformer fault diagnosis based on rough set theory, *Chinese Society for Electrical Engineering*, 24(7): 162-167.
- [5] Xiong Hao, Sun Caixing, Du Peng. 2006. Integrated condition assessment of power transformer based on matter element theory. *Journal of Chongqing University (Natural Science Edition)*, 29(10): 24-28.
- [6] Jianyuan, Ji Yancliao. 2003. Application of fuzzy Petri knowledge representation in electric power transformer fault diagnosis. *Proceedings of the CSEE 2003*, 23(1): 121-125.
- [7] Sun Hui, Li Weidong, Sun Qizhong. Electric power transformer fault diagnosis using decision tree. *Proceedings of the CSEE 2001*, 21(2): 50-55
- [8] Qian Zheng, Yan Zhang, Luo Chengmu. 2001. Fault diagnosis method of power transformer by integrating case-based reasoning with fuzzy theory and neural network. *High Voltage Engineering*, 27(6): 1-5.
- [9] Shu Hongchun, Sun Yiangtei, Si Dajun. ARS approach to founding and maintaining ES knowledge base for fault diagnosis of power transformer. *Proceedings of the CSEE 2002*, 22(2): 31-35.
- [10] Yang Li, Qian Zheng, Zhou Yueteng, et al. 1999. Parsimonious covering theory applied for identifying power transformer malfunction. *Journal of Yi'an Jiaotong University*, 33(4): 13-16.