Case-based reasoning adaptive optimization algorithm for power transformer fault diagnosis

Wei Zhang, Qiu-li Wu, Yu-rong Deng, Ze-cheng Lv*

Guangxi Electric Power Research Institute, 530023 Nanning Guangxi Zhuang Autonomous Region, China

Received 1 March 2014, www.cmnt.lv

Abstract

The adaptive learning rate for the introduction of case-based reasoning transformer fault type identification. The adaptive learning rate theory, through improved data normalization, typicality and best filtering diversity to extract the original example and optimal neural network. In the sample processing and analysis process to be solved according to the type of fault feature automatically adjusts the data processing methods, processes, boundary conditions and constraints to adapt statistical distribution, the probability characteristics. Examples show that this method can overcome the DGA data ambiguity and dispersion problems in the recognition accuracy and convergence speed advantage.

Keywords: Adaptive, Case-Based reasoning, neural network, fault type, normalization, data filtering

1 Introduction

Dissolved gas analysis in oil technology has proven to be simply and effectively on transformer fault diagnosis. In recent years, scholars have put forward comprehensive diagnostic methods, which is mainly of dissolved gas analysis, combined with other electrical test results, such as pteri [1-2], information fusion network [3-4], decision tree [5-6], Expert System [7], etc. At present transformer faults diagnosis is by building more between fault symptoms and the mechanism of the mathematical model for fault type recognition, but because of the complex relationship between fault symptoms and mechanism, it is difficult to extract the input feature effectively according to the fault type [9]. Making fault recognition through the deterministic model is difficult.

In recent years, based on the DGA, some more efficient ways of dealing with the uncertain fault rough set theory and Bayesian network. But the methods need lots of statistical information and prior knowledge, which make calculating and training complex; By the way, the complexity of the actual transformer faults and the running environment that cause failure to grasp whether the model in the learning process fully or excessive[9], when type key information incomplete, fault fuzzy, rough knowledge intensive rule base too large, which affect the diagnosis of applicability.

In view of the need to overcome ambiguity problem in transformer fault type identification, and accurately reveal the key fault information; In this article, through improved typical and sample screening classification method and the network training process, set up Case-Based reasoning model based on vector optimization of adaptive adjustment. Case analysis indicates that the reasoning model can reduce redundant diagnostic information, avoid the tedious calculating deduction, overcome data discreteness and fuzziness of gases dissolved in transformer oil, which make fault recognition neural network model more practical.

2 Case-Based matching method of fault type

2.1 CASE-BASED REASONING PROCESS

Case-Based reasoning make the problem to be solved as the goal, and the existing problems of samples is called the source example. Case-Based reasoning, in a nutshell, is by the target sample tips to known source paradigm, and directed by source sample solving target Case [10], the process is shown in figure 1.

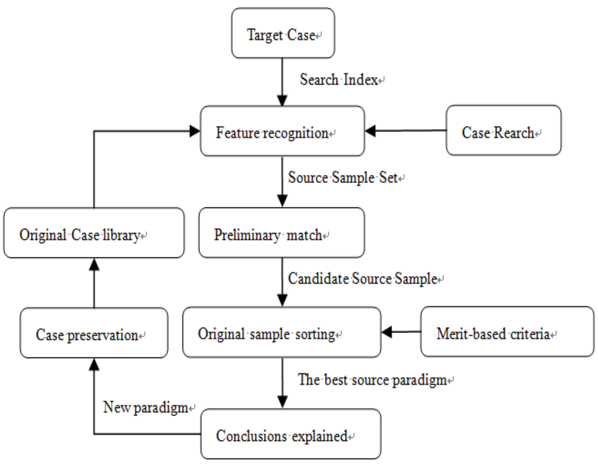

FIGURE 1 the basic principle of case-based reasoning

* ***Corresponding author*** e-mail: wu_ql.sy@gx.csg.cn

Process, the sample retrieval decided to source sample quality of solving the key problem. Sample retrieval contains characteristics identification, best source example, preliminary matches the selected three sub-processes:

(1) feature recognition is a new problem in the process of feature extraction.

(2) Initial matching is based on the new problems from the source of case base retrieval related candidate source example.

(3) Selected from the best refers to the selected candidate for example a and failure type of the current source example of the best match.

As a result of the example reasoning knowledge basis is sample, and sample acquisition relatively easy; At the same time, will be the best source sample results and answer for new solution of the fault type, improve the efficiency of solving the new problems. It is hard to according to the procedures for identification of fault types, using example reasoning may have better effect.

2.2 HIERARCHICAL DIAGNOSTIC CLASSIFICATION AND CASE BASE

Through hierarchical diagnosis for fault feature data necessary to simplify and reasonable assumptions, can make the fault diagnosis process linearization, homogenization. The hierarchical diagnosis can step by step a thorough and detailed strategies, achieve a detailed and accurate fault type recognition [11]. It step by step by step a course assigned to the segment to 11 class using the neural network algorithm to calculate the possible fault types and similar cases. In the hierarchical diagnosis combined with the prior data of the sample reasoning, can improve diagnosis effectiveness; its premise is to must classify source case base.

2.3 SAMPLE CLASSIFICATION RETRIEVAL

Model retrieval based on hierarchical diagnostic process, coarse assigned to the segment classification retrieval method. The method has the characteristics of heuristic algorithm, can quickly find the optimal solution; And avoid the classification carefully caused by different fault types of interference.

Based on heuristic rules to retrieve, not all conditions are necessary, may be only part of the sample to infer the conclusion. Retrieval process is the process of identification of target sample, in the first level recognition classification using three ratio method to coarse points, the second and third level by using BP neural network (BPNN) segmentation step by step. The basic process is shown in figure 2.

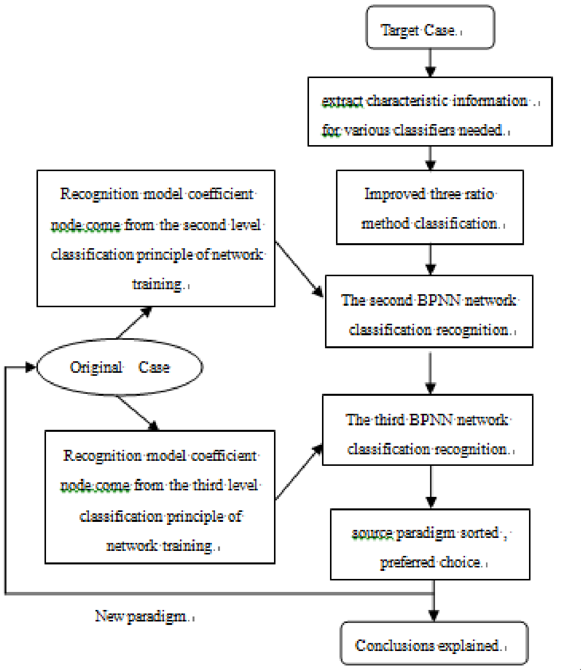

FIGURE 2 case retrieval process

2.4 IDENTIFICATION MODEL OF NETWORK STRUCTURE

BPNN is according to the error backward propagation algorithm training of the multilayer feedforward network, is currently one of the most widely used neural network model [12]. BPNN can learn and store a lot of input - output model mapping, without prior mathematical equation describing the mapping relation, for example reasoning. According to example reasoning of causality, structure to solve the fault type and source of case base relationships between each vector equation model, can map the complex nonlinear relationship, make it simple and applicable. The network structure as shown in figure 3.

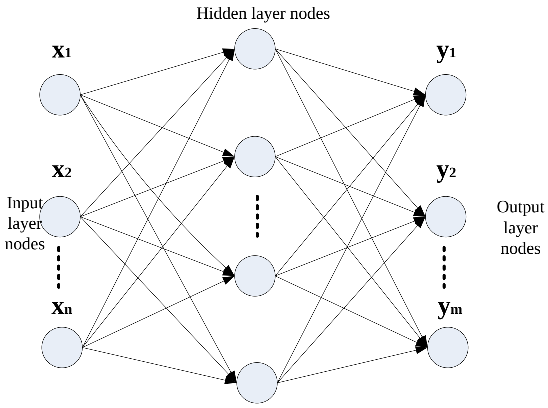

FIGURE 3 Schematic diagram of BPNN

3 Adaptive adjustment of vector model building

BPNN model was constructed based on sample classification retrieval algorithm, and the coefficient of

BPNN trained by analytic network nodes in each layer, to determine the network identification model. BPNN network training contains positive transmission output and back propagation adjustment of two parts, in the process of its positive transmission output, the input value by the input layer through the hidden layer nodes and send to the output layer; If the output values reach the desired effect, the output value of error back propagation along the original connection path to the input layer, automatic correction the connection weights between nerve cells, reduce the error gradually narrowed.

3.1 ADAPTIVE ADJUSTMENT VECTOR ALGORITHM

Network identification model first need to training the model, the training of the BPNN model steps as follows:

(1) the beginning of the weight coefficient of each layer of Wij, I said layer index, j said node number index.

(2) From the standard input value and initialized weights coefficient, network positive propagation path is used to calculate the output.

(3) According to the difference of the output and the expected output as a feedback factor adjustment coefficient of node weights Wij.

(4) Repeat steps 2 and 3 until the calculation error of the output and the expected value to achieve the ideal range.

(5) Set the input vector X = (x1, x2… xn), toward the output of Y = (y1, y2…ym) and expected output for T = (t1, t2…tm), node weights coefficient is W = (wi1, wi2…wis). The first k times back propagation error is: $E(k) = T - Y(k) = T - XW(k)$.

In order to improve the convergence of the network training algorithm, using the adaptive vector adjustment of network learning algorithm, its weight adjustment process can be represented as:

$$\begin{cases} w(k+1) = w(k) + \Delta w(k+1) \\ \Delta w(k+1) = mc \times \Delta w(k) + (1-mc) \times \alpha(k) \times D(k) \\ D(k) = -\dfrac{\partial E}{\partial W} \end{cases}. \quad (1)$$

Type, was the weight variation; The MC for momentum factor MC (0;1); (k) is a vector of k time; D(k) for the gradient values of k time.

Vector among them, the adjustment has the adaptive ability, when training network output error is greater than the last time in the process of the output of the error, namely E E (k) (k - 1), the vector is automatically reduced, namely(k + 1) = (k) by dm (constant) of dm for less than 1. < E and E (k) (k - 1), the vector will increase, namely (k + 1) = (k) by im (im constant of greater than 1).

Network training and network to identify the difference between network training input layer and output layer parameters as known variables, and the coefficient of network nodes as unknown variables, should be calculated through the network training; Network to identify the parameters of input layer and the network node coefficient is known variables, output layer output parameters as unknown variables, to identify calculated through the network. Network training, first of all, to do the normalized processing data, ensure the different characteristics of the measuring data consistency; And then to the training sample selection, remove the borderline sample; Finally, the training sample data after processing.

2.2 DATA NORMALIZATION PROCESS

Due to the different characteristics of the components of the content of dissolved gas in transformer oil and its sensitive reflect the fault degree vary; So directly to the characteristics of the component gases neural network data input, will lost some small volume contains the key information [13, 14]. At the same time, if the source of BPNN model concentrated too discrete sample data will lead to the neural network convergence difficulties, so the concept of cumulative frequency of the data of gas dissolved in transformer oil are normalized processing, the process is:

(1) collecting and confirmed by the verification, and the conclusion is clear transformer accident before chromatography detection data of the training sample set.

(2) In view of the seven characteristics of gas (H2, CH4, C2H6, C2H4, C2H2, CO, CO2) content size sorting, respectively, and size are grouped according to its content.

(3) Statistics of seven characteristics of gas in the frequency of each packet rj and frequency wj = rj/n, where n is number of observations for gas.

(4) Using the concept of cumulative frequency calculation after each group of data, such as in the first set of data the cumulative frequency I to $F_i = \sum_{j=1}^{i} w_j = \sum_{j=1}^{i} \frac{\Delta r_j}{n} = \frac{r_i}{n}$.

Of them by the end of the first group I to the ri cumulative frequency. The following will have to calculate the cumulative frequency value of Fi instead of the group characteristics of the gas content as the input of the neural network.

(5) By the training sample set of the seven characteristics of the gas content is replaced by their corresponding cumulative frequency, of the original sample set.

Using this method in 699 cases of chromatography test sample data normalization processing, H2 and cumulative frequency change relations as shown in figure 4.

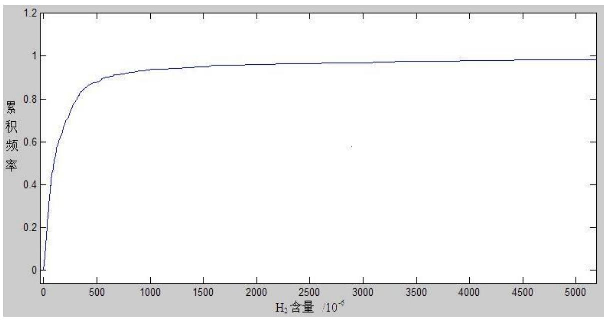

FIGURE 4 the relationship between H2 content and cumulative frequency

2.3 DATA SCREENING CLASSIFICATION

Because it is difficult to describe professional disciplines in a precise numerical definition and complex expert experience a lot of words in [15]; Examples so as to ensure that the reasoning can effectively identify the fault type, a source of case base training sample should satisfy certain boundary conditions or constraints, namely has typicality and differences. Identification model of typicality to guarantee training effectively identify the unknown sample data; Difference is to reduce the redundancy of the training database.

Define a generic function is the key to the typical samples selected, N training samples respectively under K fault type of solution, the evaluation of the individual sample generic function can be represented as:

$$F_{class} = \max(M_1/l, M_2/l, \cdots M_i/l \cdots M_k/l)\,. \tag{2}$$

Type of l for the nearest distance is evaluating individuals training sample, the amount can be set according to the size of the existing data artificially, I said the failure mode of the subscript, Mi is l study samples belong to the first class I fault sample.

Visible, according to the evaluated samples in specific degree of fault type of the relative merits of the calculation results, can be assessed individual sample belongs to the barrier type of possibility. Therefore, a generic function values can be as the measure is to assess individuals sample representativeness index; In all of the source case base sample after a generic function value calculation, preferable one of the biggest generic function values as a representative of the combination of the training sample data sample.

On the basis of the typical filter, to eliminate differences in borderline example of small samples. The measure of diversity index, through correlation analysis to solve the phase relationship between numerical implementation. Numerical calculation of phase relationship can be represented as:

$$r = \sum_{i=1}^{7}(x_i - \bar{x})(y_i - \bar{y}) / [(\sum_{i=1}^{7}(x_i - \bar{x})^2) \times (\sum_{i=1}^{7}(y_i - \bar{y})^2)]^{0.5}$$

Type in the x, y, respectively in the oil dissolved gas analysis results of two different training sample; Whereas, respectively mean {xi} and {yi}.

After the normalization processing of the original sample set after screening, available capacity is not big but have typical source example of training sample set.

2.4 NETWORK TRAINING PROCESS

According to the classification of fault type recognition process (figure 2), in the second classification in overheating fault and discharge two types, and each fault type contains two branch fault types. Therefore, the second neural network training, the need to build corresponding overheating fault, discharge two branch network model. At the same time, by back propagation to constantly adjust the network weights and threshold, reduce the error sum of squares of the network. Network model training and the corresponding fault type source case base.

Network training input layer data for gas concentration data after normalization processing. If conditions permit, the transformer can be electrical test data as input parameters, also facilitate neural network identification. The network training process as shown in figure 5.

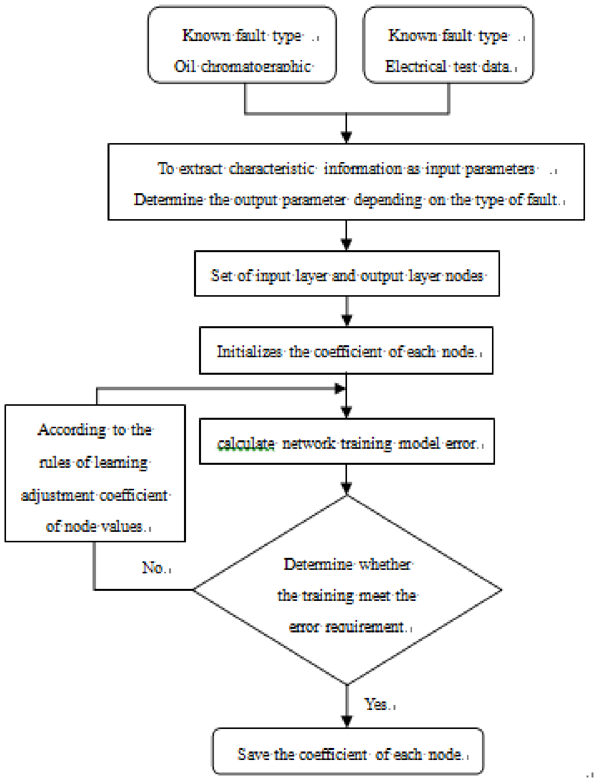

FIGURE 5 the training process of neural network

Will get the network node coefficient were input to the training BPNN (figure 4) nodes in each layer, can establish the corresponding fault type recognition model. Its output is type matching similarity, paradigms and specific fault types of target is obtained matching similarity, the higher the similarity is to determine the accuracy of the higher. If sample retrieval using a variety of retrieval way, all kinds of retrieval methods can get a classification as a result, the final target is obtained by using the weighted average of the sample and the failure types of comprehensive matching similarity, ensure that the minimum error sum of squares of the network. The comprehensive evaluation index as:

$$syn = \sum Rig_{ij} \times Sim_{ik} / 2 . \quad (3)$$

Type, Rigij for a large number of failure data (DGA) calculation for each sample retrieval algorithm for each source sample correct classification rate index of the search space, it is using the known failure data to judge each classification retrieval algorithm is sentenced to rate, Simik is the i retrieval algorithm retrievals

4 The example analysis

A 110 kv good feng varying oil chromatographic data as the target of the 2012-11-23 07:48 time paradigm, monitoring data as shown in table 1,

TABLE 1 Oil chromatographic monitoring data of liangfeng station

Monitoring time	H2	CH4	C2H2	C2H4	C2H6	CO
2012.11.23 07:48:00	19.87	13.75	1.23	45.67	4.11	3016.26

Step 1: sample retrieval

Known from the analysis of three ratio encoding C2H2 / C2H4, CH4 / H2, C2H4, C2H6 corresponding code of 0, 0, 2. Therefore, through the example to retrieve the first level classification target sample can be classified as overheating fault.

Sample retrieval in the second classification and the third level is used to identify the BPNN network in the classification, so need to source and examples of typical network training samples, to obtain network node coefficient. Source at the same time, the case base from 699 cases after selecting the type of known fault cases of original samples, so the second layer neural network training samples for 383 cases of thermal fault samples. Among them, the thermal circuit fault samples of 171 cases of magnetic circuit thermal failure of 212 cases.

The source case base in the magnetic circuit fault and heat circuit fault data samples respectively do normalization processing and typical sample data filtering, with typical samples as training BPNN network learning samples. Figure 7 for the gas concentration and the mapping relationship between cumulative frequency, by the mapping relationship between the sample data normalization.

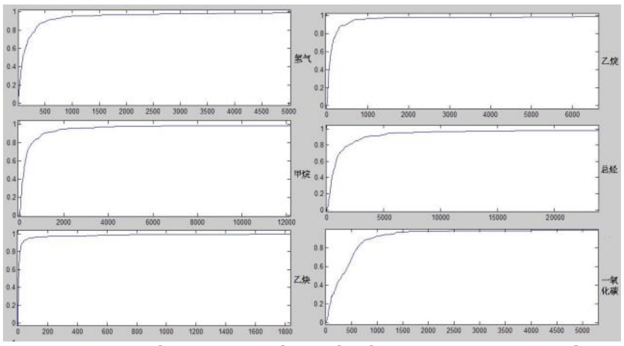

FIGURE 6 The mapping relationship between gas content and cumulative frequency

Data normalization typicality after screening, screening after 50 cases received circuit fault and 50 cases of magnetic circuit hot fault. The second layer network training input data, a total of 100 groups, namely the $X_i=\{x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6}, x_{i7}\}$, $Y_i=\{y_{i1}, y_{i2}\}$, $i\in[1,100]$.

Type in the Xi for the normalized oil chromatographic data, Yi for fault classification results (i.e. $y_{i1}=1$, $y_{i2}=0$ means heat circuit fault; $y_{i1}=0$, $y_{i2}=1$said magnetic circuit thermal failure), the subscript i said the first group of data, I said the size of the y value and the fault type of compatibility. Network model are obtained by training, identification of target model, the calculation results as follows: $y_{11}=0.999999439228727$, $y_{12}=3.30988595178941*10^{-7}$.

Similarity matching with heat circuit fault close to 100%, thus can be considered as heat circuit malfunction.

Carried out in accordance with the same method in the third layer neural network recognition, the result is: $y_{11}=0.99999980962573$, $y_{12}=4.57039595983834*10^{-5}$, $y_{13}=1.92874984796392*10^{-22}$.

Similarity matching with tap-changer fault also close to 100%, thus can be considered a fault tap-changer.

Step 2: source sample ordering, the preferred choice of fault type

According to the sample, matching similarity retrieval failure is calculated, sample selection and source library failure types match the highest failure type as the target sample identification results, namely tap-changer of failure.

Step 3: keep sample.

Examples given recognition as a result, the target as a source of known fault types to keep new paradigm case base, namely source similar fault samples increase 1 case in case base.

5 Conclusions

Case reasoning is a kind of artificial intelligence methods based on prior knowledge reasoning, it according to the problem requirements or features, matching sample source, and get the optimal original sample under its guidance. By comparing the fault identification probability, select from course to fine hierarchical structure, so establish the diagnosis mode must give full consideration to the edge the effect on the stability of the sample of the model.

In this paper, the adaptive adjustment algorithm have great inhibition effect to solve unstructured edge data influence. At the same time to the associated classification pattern extraction and classification of input vector, the adaptive adjustment of vector optimization is also has strong flexibility, makes the establishment of the diagnosis model is more accurate. From the hierarchical diagnosis, the case of adaptive adjustment vector reasoning algorithm achieved ideal result.

References

[1] Wang Jianyuan, Ji YanChao 2003 Fuzzy Petri network knowledge representation method and application in transformer fault diagnosis *Proceedings of the CSEE* **23**(1) 121-5

[2] Wang Nan, L Fangcheng, Liu Yunpeng, etc. 2003 Based on rough set theory and fuzzy Petri network comprehensive oil-immersed power transformer fault diagnosis *Proceedings of the CSEE* **23**(12) 128-32

[3] Peng Jjian, Luo An, Zhou Ke, etc. 2007 The application of information fusion technology in the transformer fault diagnosis *High voltage technology* (3) 144-7

[4] Shang Yong, Rui Chunjiang, Yan Zhang, etc. 2002 Based on information fusion of large oil-immersed power transformer fault diagnosis *Proceedings of the CSEE* **22**(7) 115-8

[5] Sun Hui, Li Weidong, Sun Qizhong 2001 Decision tree method is used for transformer fault diagnosis research *Proceedings of the CSEE* **2**(2) 50-5

[6] Wang Yingying, Luo Yi, Tu GuangYu 2008 Distribution network fault diagnosis based on rough set and decision tree method *High voltage technology* (4) 794-8

[7] Mo Juan, Zhang Yan, Li Hua, etc. 2005 Comprehensive transformer fault diagnosis model based on multiple intelligence method *Automation of electric power systems* **29**(18) 85-9

[8] Dong Xiucheng, Tao JiaGui, Wang Haibin, etc. 2010 The adaptive fuzzy support vector machine increment algorithm in the application of transformer fault diagnosis *Electric power automation equipment* **30**(11) 48-52

[9] Li Hua, Mo Juan, Zhang Rui, etc. 2005 The transformer insulation fault diagnosis system based on Web mode *High voltage technology* **31**(6) 12-4

[10] Shi Haojie, Xing QingHua, Liu Fuxian, etc. 2009 Based on the introduction of probability sample matching new method *Computer engineering and application* (10) 62-4

[11] Chen Jiangbo, Wen XiShan, Lan Lei, etc. 2007 Based on a new radial basis function networks *Journal of transformer fault diagnosis of high voltage technology* (3) 140-3

[12] Zhang Quanming, Liu Huijin 2008 Least squares support vector machine in the application of power quality disturbance classification *Proceedings of the CSEE* **28**(1) 106-10

[13] Zhang Weizheng, Wang Xiaoming, Wu Xiaohui, etc. 2007 Combination of neural network in the transformer fault diagnosis application *High voltage electrical appliances. Lancet* (5) 364-7

[14] Lan Hua, Zheng Guangrui, Liao Zhimin *Based on the fuzzy TOPSIS and the BP network* 15:26 2014/4/13

Authors

Wei Zhang

Current position, grades: Senior Engineer
University studies: High voltage technology
Scientific interests: Online monitoring
Publications: none
Experience: Be engaged in On-line monitoring of electric equipment of radio frequency identification technology, Excellence at warning centre maintenance management, condition evaluation, reliability analysis.

Qiu-li Wu

Current position, grades: Engineer
University studies: High voltage technology
Scientific interests: Online monitoring
Publications: none
Experience: Be engaged in On-line monitoring of electric equipment of radio frequency identification technology, Excellence at warning centre maintenance management, condition evaluation, reliability analysis.

Yu-rong Deng

Current position, grades: Professor level senior engineer
University studies: High voltage technology
Scientific interests: Online monitoring
Publications: none
Experience: Be engaged in On-line monitoring of electric equipment of radio frequency identification technology, Excellence at warning centre maintenance management, condition evaluation, reliability analysis.

Ze-cheng Lv

Current position, grades: Senior Engineer
University studies: High voltage technology
Scientific interests: Online monitoring
Publications: none
Experience: Be engaged in On-line monitoring of electric equipment of radio frequency identification technology, Excellence at warning centre maintenance management, condition evaluation, reliability analysis.