The Application of Compound Neural Network in Condition Estimate of Power Transformer

Zhang Weizheng[1], Wang Zhenggang[1], Fu Yingshuan[1], Liu Fazhan[1], Yang Lanjun[2], Li Yanming[2]
[1]Zhengzhou Power Supply Company
[2]Xi'an Jiao Tong University
xjtuzwz@163.com
China

Abstract: The analyses of light gases in transformer oil provide the basis of a diagnostic evaluation of equipment health. Two new normalized methods which named characteristic normalization and mix normalization are presented in this paper. The Fisher rule to evaluate the results of the two pretreatment methods is also introduced. The evaluation of the results indicates that both of the two data pretreatment methods can achieve the purpose of big difference in the value of mean between classes and small difference in dispersion of a class. The DGA data of the failure transformers are treated by different normalization methods as the training samples, and then the samples are trained in the compound neural networks which use the CP algorithm. The diagnosis results of the test samples indicate that the new methods may help to improve the precision of network diagnosis.

Key words: transformer; analysis of reliability data; CP compound neural networks; condition estimate

1 Introduction

Recent trends in the electrical power industry have resulted in more attention to diagnostic and prevention systems for the purpose of (a) minimizing maintenance and (b) prolonging the life of high cost equipment, In this environment, diagnostic systems for power transformers are very important. On the other hand, many vital transformer conditions are not directly measurable. In this case estimation methods provide the tool for monitoring vital conditions of transformer operation. State Estimation methods are valuable for monitoring complex systems with not directly measurable states. One case in point is the monitoring of power transformers. The analyses of light gases in transformer oil provide the basis of a diagnostic evaluation of equipment health. Based on the analysis results of dissolved gases' contents in transformer oil, the failure diagnosis problem can be solved effectively by using artificial neural network. In practical analysis of DGA, the sensitivity reflected by different typical gases has big difference. For example, the content of C_2H_2 in transformer oil is usually less, but it must be strictly monitored once it appears. Conversely, the content of H_2 is comparatively high, but its reflection to failure is less sensitive. Therefore, the potentially valuable information contained by some minor components will be neglected if the results of contents of DGA are directly as the input of neural network. It indicates that the differences between data must be decreased by using normalized treatment in order to reflect the failure type more accurately.

There are two universal neural network normalized methods: general normalization[1] and cumulative frequency normalization[2][3]. In the general normalization, the proportion of different gases in total gas concentrations is normalized. This method corresponds to the feature that gas content in oil reflects failure type and has real physical meaning. But the potentially valuable information contained by some minor components tends to be neglected as a result of gas content feature.

Cumulative frequency normalization strengthens the dispersion of samples which makes samples distribute in the interval [0, 1] and has strong mathematical meaning. However, this method only compares the same kind of gas longitudinally without considering gases' concentration in the same group. So the physical meaning is unclear. Consequently, finding a new normalized method which can not only strengthen the typical gas's characteristic of DGA but also differentiate samples in different conditions plays a key role in the improvement of accuracy of neural network diagnoses.

2 Characteristic Normalization And Mix

Normalization

Dealing with the advantages and disadvantages of general normalization, two new normalized methods which named characteristic normalization and mix normalization are presented in this paper. The process of characteristic normalization is:

The input data of such pattern $A_k = (H_2{}^k, C_2H_2{}^k, CH_4{}^k, C_2H_4{}^k, C_2H_6{}^k)$ can be normalized as follows:

$$\begin{cases} H_2{}^k = \dfrac{H_2^k}{H_2^k + C_2H_2^k + CH_4^k + C_2H_4^k + C_2H_6^k} \\ C_2H_2{}^k = \dfrac{C_2H_2^k}{C_2H_2^k + CH_4^k + C_2H_4^k + C_2H_6^k} \\ CH_4{}^k = \dfrac{CH_4^k}{C_2H_2^k + CH_4^k + C_2H_4^k + C_2H_6^k} \\ C_2H_4{}^k = \dfrac{C_2H_4^k}{C_2H_2^k + CH_4^k + C_2H_4^k + C_2H_6^k} \\ C_2H_6{}^k = \dfrac{C_2H_6^k}{C_2H_2^k + CH_4^k + C_2H_4^k + C_2H_6^k} \end{cases} \quad 1$$

Eq.(1) shows that after the treatment of characteristic normalization, H_2 reflects the content of H_2 in hydrogen hydrocarbon and various kinds of gases reflect their contents in hydrocarbon gas. Literature[4][5] indicates that the content of H_2 in hydrogen hydrocarbon and the contents of various kinds of gases in hydrocarbon gas are closely related with the failure type of power transformer. Therefore, the characteristic normalization has strong physical meaning. Meanwhile, because the content of H_2 in typical gas is likely to be comparatively high, Eq.(1) normalized hydrocarbon gas by using the ratio of hydrocarbon gas and total hydrocarbon gas instead of the ratio of hydrocarbon gas and hydrogen hydrocarbon. After such treatment, some valuable information won't be acquired because data of hydrocarbon gas is too tiny and be neglected. So the characteristic normalization also has strong mathematical meaning.

Because the cause of typical gas is complicated, it is difficult to well differentiate certain failure only by a single treatment of these three methods. So mix normalization is presented. The process of mix normalization is:

$$A_k = (H_2{}^k, C_2H_2{}^k, CH_4{}^k, C_2H_4{}^k, C_2H_6{}^k)$$

H_2 in input samples is probabilistically normalized; C_2H_2, C_2H_4, CH_4 and C_2H_6 are characteristically normalized as follows:

$$\begin{cases} C_2H_2{}^k = \dfrac{C_2H_2{}^k}{C_2H_2{}^k + CH_4{}^k + C_2H_4{}^k + C_2H_6{}^k} \\ CH_4{}^k = \dfrac{CH_4{}^k}{C_2H_2{}^k + CH_4{}^k + C_2H_4{}^k + C_2H_6{}^k} \\ C_2H_4{}^k = \dfrac{C_2H_4{}^k}{C_2H_2{}^k + CH_4{}^k + C_2H_4{}^k + C_2H_6{}^k} \\ C_2H_6{}^k = \dfrac{C_2H_6{}^k}{C_2H_2{}^k + CH_4{}^k + C_2H_4{}^k + C_2H_6{}^k} \end{cases} \quad 2$$

The definition of mix normalization shows that treating hydrocarbon gas by using the ratio of hydrocarbon gas and hydrogen hydrocarbon and H_2 by using probabilistic normalization can prevent the complexity generated by H_2. It is difficult to find the distinction between some failures only by characteristic normalization. Therefore, mix normalization enhances the mathematical meaning of the algorithm.

3 The CP Neural Networks

In view of the fact that different dissolved typical gases in oil react differently to various failure types, it is improper to identify all kinds of failures by uniform input vector for a single neural network model. To realize better classification, it's desirable to distinguish failures by using step-by-step division. Based on this consideration, the article refers to the gradual-judgment compound neural network pattern. The Fig.1 shows its structure. Because the problem that whether the nature of overheating failure refers to solid insulation certainty by analysis of gas in oil has not been satisfactorily solved[7], it is more suitable to take H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 as network model input.

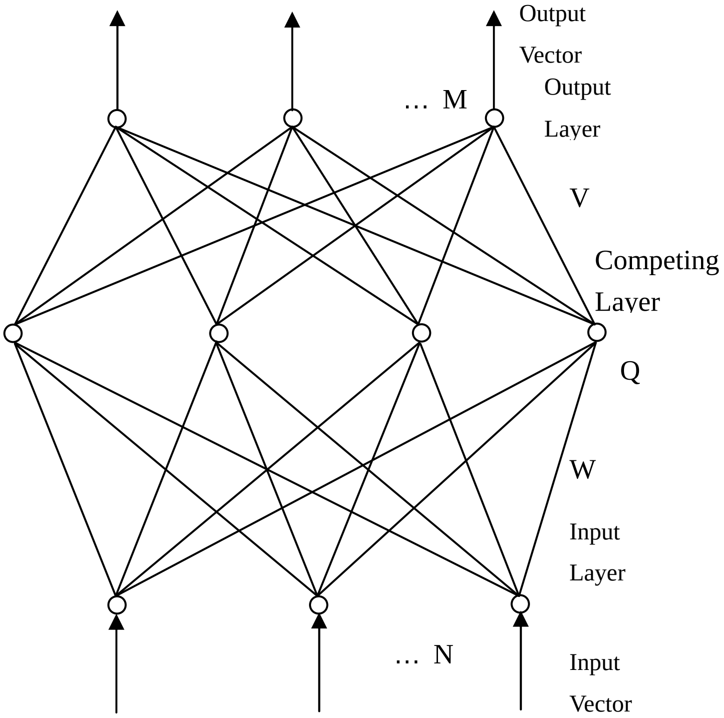

Fig.1 The structure figure for compound neural networks

To well reflect the physical premonition of failure point (e.g. overheating, discharge etc.), it is more proper to take high-temperature overheating, low-temperature overheating, arc discharge and spark discharge as network model output. The structure figure for compound neural networks model of power transformer is shown in Fig.2.

ANN1~ANN2-2: ANN1 fault diagnosis model. ANN2-1 is discharge fault diagnose model and ANN2-2 is thermal fault diagnose model. By studying their training samples set, these three fault diagnose models take CP network model[8] as essential classifier to obtain the description of concrete decision rule and select more directional input characteristic vectors for them regarding to different neural network modules.

RULE 0: According to the operation experience for many years, chromatographic observing value and factor of created gas specified in "Preventive Test Regulation for Electric Power Equipments"[9] is regarded as the limit to judge whether transformer is abnormal.

RULE1 and RULE2: Distinguishing discharge fault and thermal fault by the content of C_2H_2 in total hydrocarbon gas, this paper carries out a statistical analysis on 300 collected training samples and revealed the following findings: C_2H_2 accounts for 16.254% of total hydrocarbon gas in discharge faults and maximal 5.01% in thermal fault. Therefore, RULE 1 can be defined as whether C_2H_2's percentage of total hydrocarbon gas is greater than or equal to 16%; RULE2 can be defined as C_2H_2's percentage of total hydrocarbon gas is greater than or equal to 5%. The procedure of diagnose is as follows: For fault sample, the fault is considered a discharge fault if C_2H_2 accounts for more than 16% of total hydrocarbon gas, being diagnosed directly by discharge fault diagnose model ANN2-1; the fault is considered a thermal fault if C_2H_2 accounts for less than 5% of total hydrocarbon gas, being diagnosed by thermal fault diagnose model ANN2-2. And if the percentage of C_2H_2 falls between these two values, diagnose by thermal fault diagnose model ANN2-2 first. According to the result, diagnose whether it is arc discharge or spark discharge by discharge fault diagnose model ANN2-1 if the diagnose result is high-temperature thermal fault, if not, do not have to do this.

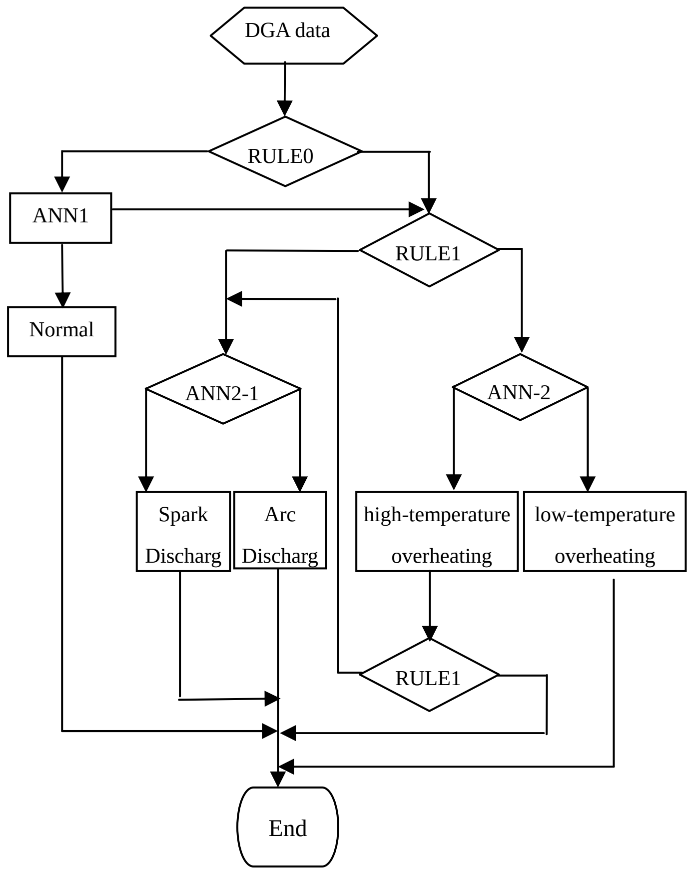

Fig.2 The structure figure for compound neural networks model of power transformer

4 Analysis Results

4.1 The Establishment of Data Samples Set

On the basis of statistic large fault examples, 420 times fault transformers which have comparatively clear factual conclusions are selected as fault samples set. The distribution of all the failure types in the set is shown in Tab.1.

Tab.1 The distribution of samples in networks

Item	ANN1	ANN2-1	ANN2-2
Trained samples	100	100	100
Evaluation samples	40	40	40

4.2 The Establishment of CPN's Parameters

Parameters of ANN1

This paper carries out a statistical analysis on 140 collected training samples for ANN-1 networks. There are two groups of samples, one has 70 normal samples and the other has 70 fault samples. 40 samples which selected from two groups separately are the evaluation samples; the other 100 samples are the training samples. In the ANN-1 network, the paper selects the training times form 3000 to 6000, and neural units from 10 to 20. The training results are Fig.2-fig.5.

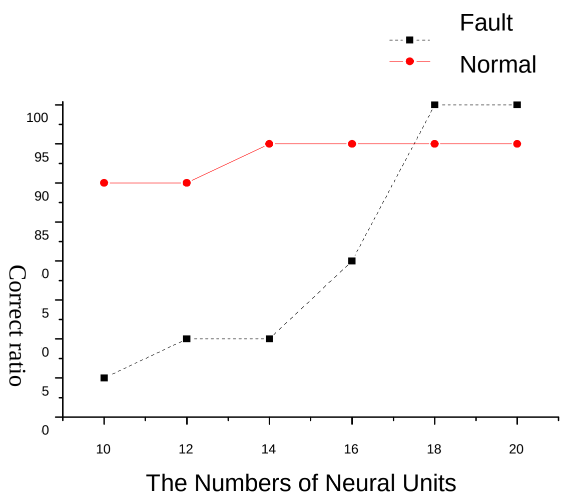

Fig.3 3000 Training times

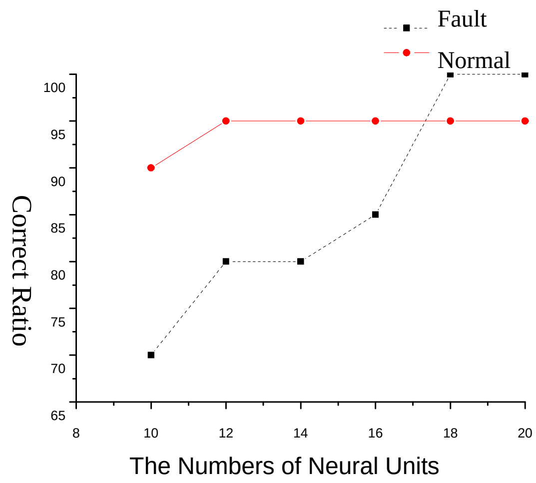

Fig.4 4000 Training times

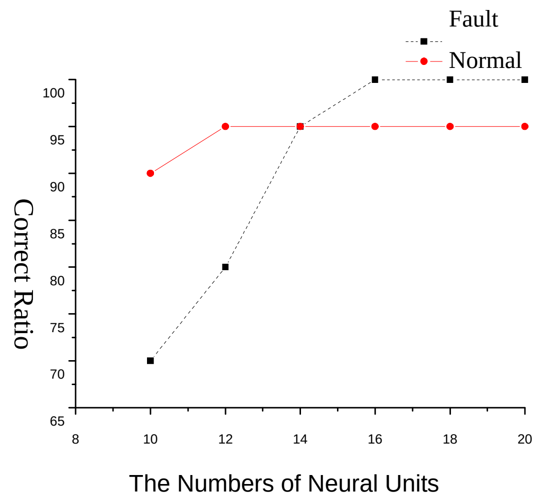

Fig.5 5000 Training times

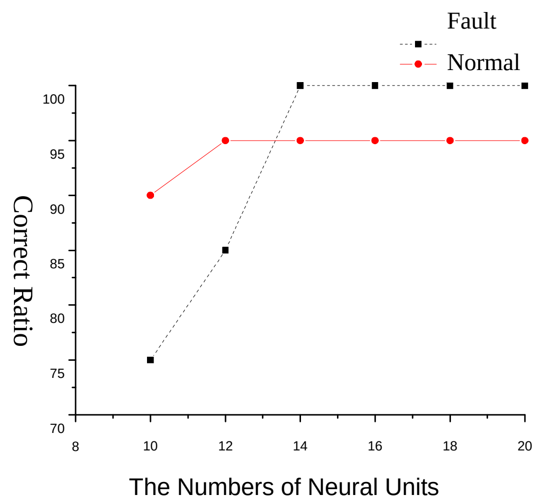

Fig.6 6000 Training times

From the figures, we can see that the training times and the numbers of neural units affect the judgments of fault conditions greatly. Integrated considering computing speed and the correct ratio of judgment, neural units which is 18 and training times which are 6000 are the optimum choices.

Parameters of ANN2-1

Just like ANN1, this paper also carries out a statistical analysis on 140 collected training samples for ANN2-1 networks. There are also two groups of samples, one has 70 normal samples and the other has 70 fault samples. 40 samples which selected from two groups separately are the evaluation samples; the other 100 samples are the training samples. In the ANN2-1 network, the paper selects the training times form 3000 to 7000, and neural units from 10 to 20. The training results are Fig.7-fig.11.

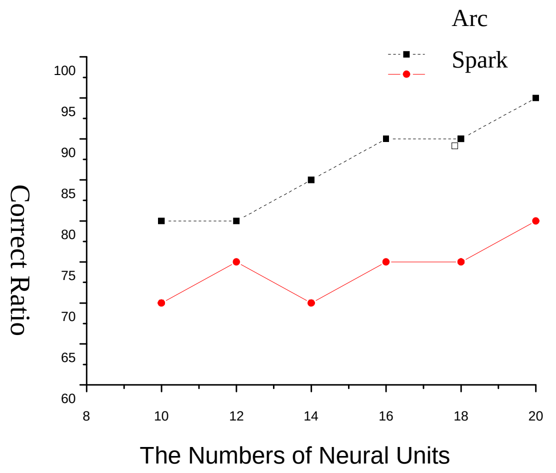

Fig.7 3000 Training times

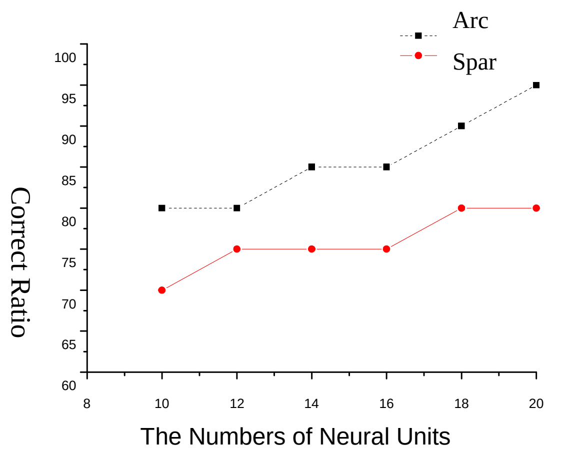

Fig.8 4000 Training times

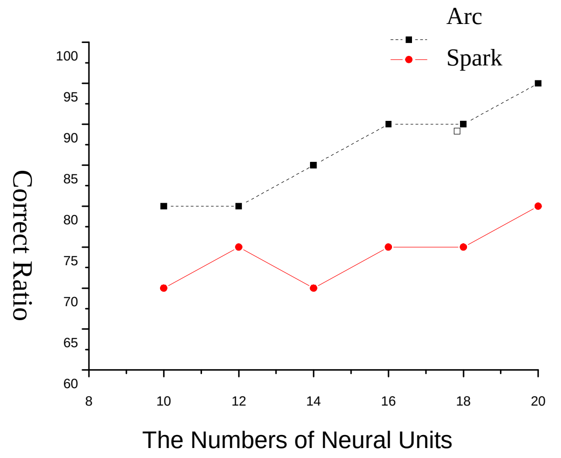

Fig.9 5000 Training times

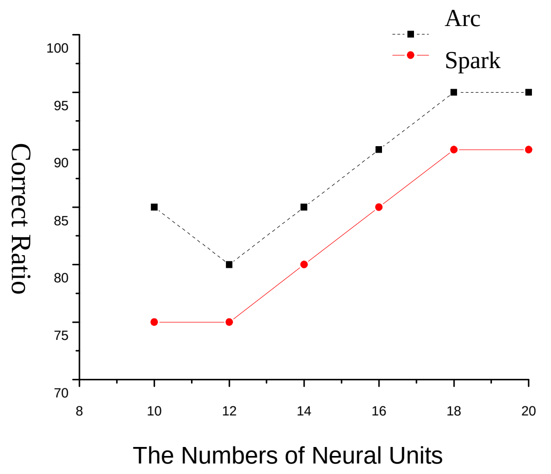

Fig.10 6000 Training times

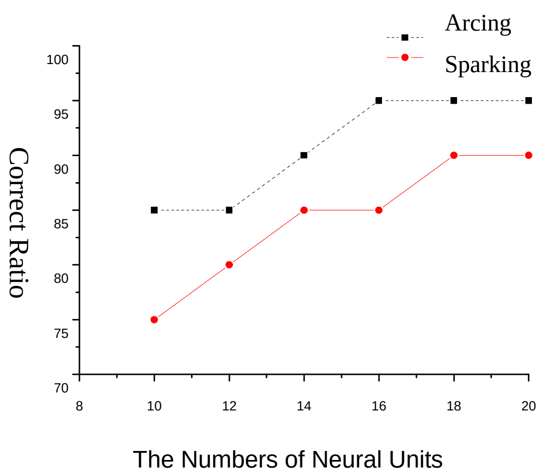

Fig.11 7000 Training times

From the figures, we can see that the correct ratio of judgments is increasing while the numbers of neural units increase and at the same time the training times are same. For the faults of sparking discharge and arcing discharge, when the correct ratio of judgments is 90% to 95%, the results can be accepted. Integrated considering computing speed and the correct ratio of judgment, neural units which is 19 and training times which are 6000 are the optimum choices.

Parameters of ANN2-2

Just like ANN1, this paper also carries out a statistical analysis on 140 collected training samples for ANN2-2 networks. There are also two groups of samples, one has 70 normal samples and the other has 70 fault samples. 40 samples which selected from two groups separately are the evaluation samples; the other 100 samples are the training samples. In the ANN2-2 network, the paper selects the training times form 3000 to 6000, and neural units from 10 to 20. The training results are Fig.12-fig.15.

Overheating faults mean the degradation of the transformer's insulation. There are two types of overheating faults. One is low temperature overheating faults; the other is high temperature overheating faults.

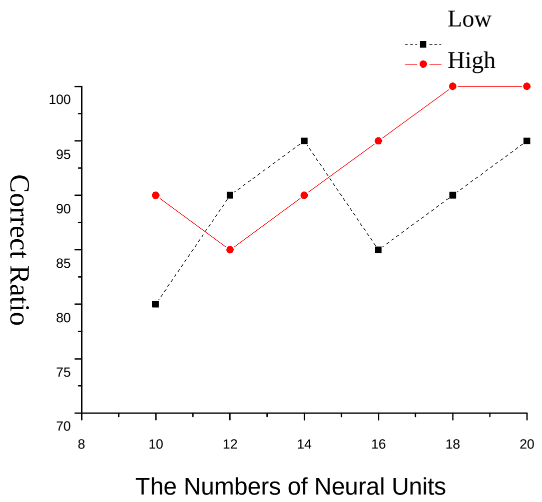

Fig.12 3000 Training times

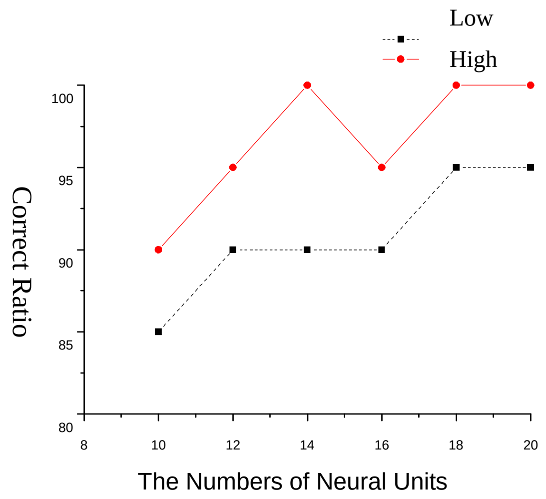

Fig.13 4000 Training times

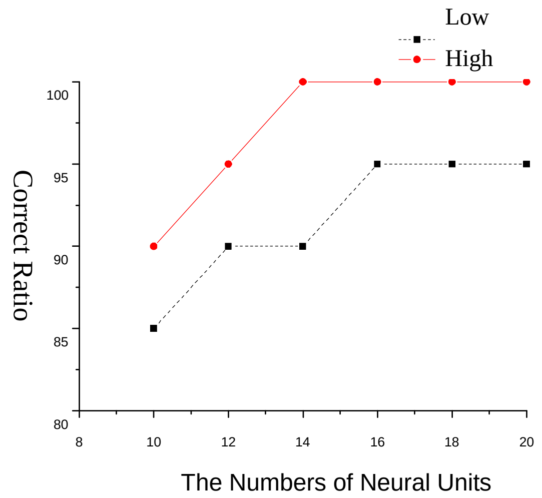

Fig.14 5000 Training times

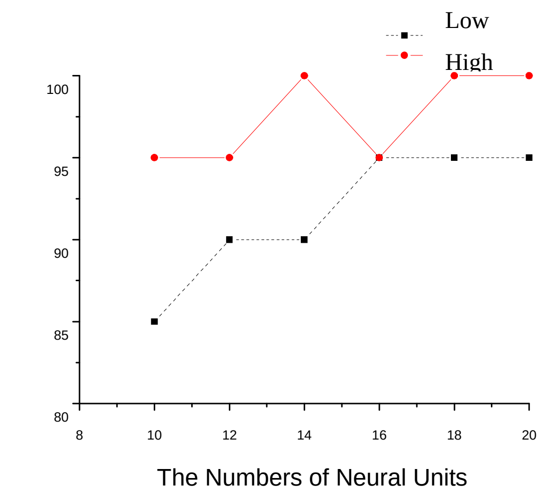

Fig.15 6000 Training times

From the figures, we can see that the training times and the numbers of neural units affect the judgments of fault conditions greatly. Integrated considering computing speed and the correct ratio of judgment, neural units which is 17 and training times which are 5000 are the optimum choices.

4.3 Different Normalization Treatment And Comparison For Samples

In order to test whether every kind of sample in one-dimension Y can distribute as much as possible after every kind of gas has been normalized, that is, the value of mean between two classes has big difference and the dispersion of a class has small difference, Fisher rule is introduced to evaluate the pretreatment results. The function of Fisher rule is as follows [10].

$$J_{F(W)} = \frac{(m_1 - m_2)^2}{S_1^{\ 2} + S_2^{\ 2}} \qquad 3$$

m_1,m_2 in Eq.(3) are mean of sample 1 and sample 2 and S_1,S_2 are RMS of sample 1 and sample 2. Pretreat training samples of each network by three different normalization methods and find out the Fisher rule values of their results. In order to state more clearly, general normalization, cumulative frequency normalization and characteristic normalization is referred to as Method-1, Method-2 and Method-3 respectively. For the results, see Tab.2-Tab.4.

Tab.2 The value of Fisher rule for training samples in ANN1

Normal-ization	H_2	C_2H_2	CH_4	C_2H_4	C_2H_6
Method-1	0.0010	0.0832	0.0140	0.0057	0.0002
Method-2	0.0002	0.0601	0.0034	0.0088	0.0003
Method-3	0.0010	0.1865	0.0657	0.0132	0.0020

Tab.3 The value of Fisher rule for training samples in ANN2-1

Normal-ization	H_2	C_2H_2	CH_4	C_2H_4	C_2H_6
Method-1	0.0011	0.0008	0.0003	0.0006	0.0035
Method-2	0.0721	0.0026	0.0126	0.0009	0.0007
Method-3	0.0011	0.0072	0.0205	0.0098	0.0043

Tab.4 The value of Fisher rule for training samples in ANN2-2

Normal-ization	H_2	C_2H_2	CH_4	C_2H_4	C_2H_6
Method-1	0.0664	0.0135	0.0038	0.0006	0.0035
Method-2	0.0170	0.0101	0.0005	0.0009	0.0143
Method-3	0.0664	0.0181	0.0711	0.0098	0.0042

Tab.2-Tab.4 indicates that training samples of each kind of gas which has been treated by characteristic normalization can achieve maximum value of Fisher rule for both model ANN1 and ANN2-2, and so does H_2 treated by cumulative frequency normalization and hydrocarbon gas treated by characteristic normalization for model ANN2-1. This illustrates that to treat training samples of model ANN1 and ANN2-2 by characteristic normalization and those of model ANN2-1 by mix normalization can not only strengthen the typical gas's characteristic of DGA but also differentiate samples in different conditions.

4.5 Comparative Analysis of Network Diagnose Results for Several Input Pattern

When the networks have the same perimeters, evaluate the results that treated by different normalization methods with CP compound neural network.

Tab.5 the results of diagnosis networks in the same parameters

Normal-ization	ANN1	ANN2-1	ANN2-2
Method-1	90	82.5	90
Method-2	75	70	77.5
Method-3	97.5	80	95
Method-4		92.5	

With no network diagnose of this method.
Method-4 is mix normalization.

Tab.5 indicates that when the networks have the same perimeters, the diagnosis rate of the network which has been treated by characteristic normalization is maximal for both model ANN1 and ANN2-2, and so does that treated by mix normalization for model ANN2-1. This shows that when the networks have the same perimeters, characteristic normalization can increase the diagnosis rate of model ANN1 and ANN2-2 and mix normalization can increase the diagnosis rate of model ANN2-1.

4.6 The Comparisons of CP Method IEC Three-ratio Codes and Improve Three-ratio Codes Methods

Some typical data are selecting to make comparisons between the cp method and IEC three-ratio and improve three-ratio methods. All the selected data are acquired from transformer in operation. Tan.6 shows the samples and Tab.7 shows the results.

Tab.6 the data of samples (μL/L)

samples	H_2	C_2H_2	CH_4	C_2H_4	C_2H_6
1	32.4	13.2	5.5	12.6	1.4
2	163	5.1	65.4	47	34
3	160	0	130	96	33
4	86	7.4	110	92	18
5	36	7.1	30	93	10
6	79.4	31.24	24.7	29.7	4.7
7	104.3	22.5	39	27	7.1
8	268	34	49	41	9.6
9	66	29.2	111	110	90.8
10	147.3	40.3	282.5	366.3	31.8
11	150	7.4	34	64	15
12	170	5.2	3.9	1.4	0.5
13	110.9	13.5	12.5	11.	7.5
14	12.88	12.39	4.2	2.6	1.5
15	130	4.7	39.4	32	0.6
16	167.12	3.1	84	34.6	13.6
17	72.3	0	35.7	10	9
18	153	0	59.4	32.4	9.8
19	15	0	38.2	4.9	41.6
20	175.7	8.5	253.9	119.2	84.1
21	198.4	2.1	30.4	25.2	13.1
22	50.4	2.3	25.3	77.1	11
23	86.7	1.7	148.6	82.2	141.1
24	97.4	0	103.5	237.1	46.2
25	250	3.2	100	100	330

Tab.7 the comparisons of cp method and IEC three-ratio and improve three-ratio methods.

S	IEC		Improve		CP	results
	code	type	code	type		
1	011		011		C	C
2	001	D	001	D	D	D
3	022	E	022	E	E	E
4	002		002	E	E	E
5	102	A	102	B	B	B
6	102	A	102	B	B	B
7	102	A	102	B	B	B
8	102	A	102	B	C	B
9	121		121	B	E	E
10	022	E	022	E	E	E
11	102	A	102	B	B	B
12	201		201	C	C	C
13	101		101	B	C	C
14	201		201	C	C	C
15	102	A	102	B	C	B
16	001	D	001	D	D	D
17	001	D	001	D	D	D
18	002		002	E	D	D
19	020	D	020	D	D	D
20	021	D	021	D	E	E
21	001	D	001	D	D	D
22	002		002	E	E	E
23	020	D	020	D	D	D
24	022	E	022	E	E	E

25	000	-	000		B	B

A: Discharge
B: Arc Discharge
C: Spark Discharge
D: Low Temperature Overheating
E: High Temperature Overheating
—: Normal

From tab.7, the correct judgment ratio of CP method is 88.0%. The CP method is better than IEC three-ratio method and also better than improve three-ratio method.

4.7 Example Calculation

Tab.8 The examples of DGA used to analysis

Sample	H_2	C_2H_2	CH_4	C_2H_4	C_2H_6
1	12.0	11.6	9.61	12.1	6.06
2	236	88	48	48	13.3
3	75	4.8	28.8	80.9	7

Sample 1 in Tab.8 is chromatographic data of a transformer's oil. After being treated by characteristic normalization, input data is [0.2337 0.2441 0.1539 0.3073 0.2946] and diagnose result of model ANN1 is [0.9067 0.0933]. According to maximum membership principle, this denotes malfunction. The percentage of C_2H_2 in total hydrocarbon gas is 29.46%, greater than 16% specified in RULE1, so it is discharge fault. After being treated by mix normalization, input data is [0.9900 0.2441 0.1539 0.3073 0.2946] and diagnose result of model ANN2-1 is [0.0827 0.9173]. According to maximum membership principle, this denotes spark discharge fault. Real test result is: there is a black spot which is about 250mm from equalizing ball measured approximately 7×7mm2 in porcelain surface under the sleeve of phase A. After careful inspection, find out that the black spot is an air hole with a carbon black inside. Hence, the fault is a spark discharge fault. CP network diagnose results correspond with the real test.

Sample 2 in Tab.8 is chromatographic data of a transformer's oil. According to the recommended concentration observing value of DGA in DLT/T722-2000, the transformer is faulty because the content of H2, C2H2 and total hydrocarbon gas has exceeded bid badly. The percentage of C2H2 in total hydrocarbon gas is 44.6%, greater than 16% specified in RULE1, so it is discharge fault. After being treated by mix normalization, input data is [0.98 0.2433 0.067 0.2433 0.4460] and diagnose result of model ANN2-1 is [0.8034 0.1966]. According to maximum membership principle, this denotes spark discharge fault. Real test result is: there are 8 breakdown spots in the first insulating carboard of the folding screen in the first layer from phase A's external and there is penetrating tree discharge in the second insulating carboard. Hence, the fault is an arc discharge fault. CP network diagnose results correspond with the real test.

Sample 3 in Tab.8 is chromatographic data of a transformer's oil. According to the recommended concentration observing value of DGA in DLT/T722-2000, the transformer is faulty because the content of total hydrocarbon gas has exceeded bid badly. The percentage of C2H2 in total hydrocarbon gas is 3.9%, less than 5% specified in RULE1, so it is thermal fault. After being treated by characteristic normalization, input data is [0.3817 0.2370 0.058 0.6658 0.040] and diagnose result of model ANN2-2 is [0.8094 0.1906]. According to maximum membership principle, this denotes high-temperature thermal fault. Real test result is: the tap switch moving contact of phase C and dial records differs 120°. The breakback contact is eroded seriously and moving contact is eroded slightly. Hence, the fault is a high-temperature thermal fault. CP network diagnose results correspond with the real test.

5 Conclusions

1 By comparative analysis, both the training samples' Fisher rule values under different normalizations and the results of different input patterns indicates that to treat training samples of discharge fault by mix normalization and training samples of thermal fault by characteristic normalization can well decrease the difference between samples and increase the diagnose rate.

2 With the characteristic of power transformer insulating fault, the network training efficiency and diagnose precision are both enhanced by judging different fault types with compound neural network. After the analysis of example calculation, to apply compound diagnose network constructed by artificial neural network model based on CP algorithm that taken as essential classifier to power transformer insulating fault diagnose that based on DGA can achieve better results. The test results of fault examples illustrate that

the algorithm is an accurate diagnose mode and has more precise diagnose result.

References

[1] Phase Technology Products Research and Development Center. Neural Network Theory and Implementation of MATLAB 7[M]. Beijing: Electric Industry Press, 2005.3.

[2] He Guofang, Xu Haibao. Reliability Data Collection and Analysis[M].Beijing: Defense Industry Press, 1995.

[3] Qian Zheng. Application Research of Artificial Intelligence Technology on Insulation Fault Diagnosis of Large-scale Electric Power Transformer [D]. Xi'an Jiaotong University Ph.D. Thesis, 2000.

[4] Yang Qiping, Xue Wude, Lan Zhida. Research on Diagnosing Techniques of Transformer Faults[J]. Transformer, 2002, 39 (10):38-42.

[5] Xu Wen, Wang Dazhong. Application of Artificial Neural Networks on Faults Diagnosis of Transformers. High Voltage Engineering.

[6] Yang Qiling. Analysis of Unusual Hydrogen Content in Transformer Oil[J]. Transformer, 2004,20(11):39(8).

[7] Ding Xiaoqun, Liu Y, Lin Zhongyun. Transformer Fault Diagnosis Using ANN[J]. Automation of Electric Power System, 1996, 20(2):32-35.

[8] Wang Zhe, Cai Weizheng, Chen Xueyun. Recognition of treeing discharges using CPN[J]. Journal of Harbin Institute of Technology, 2000, 32(5):5-8.

[9] Ministry of Power Industry of P.R China. Preventive Test Regulation for Electric Power Equipments DL/T 596 1996. Beijing: China Power Press, 1996.

[10] Bian Zhaoqi, Zhang Xuegong. Pattern Recognition[M]. Beijing: Tsinghua University Press, 1999.