Fuzzy Neural Classifier for Fault Diagnosis of Transformer Based on Rough Sets Theory

Hongsheng Su, Qunzhan Li
Institute of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China

Abstract—**Due to enduring more disturbance such as environment varieties and surveying interference and information transmission mistakes as well as arisen error while processing data in surveying and monitoring state information of transformer, thus uncertain and incomplete information and ill data may be produced. So the study how to apply these data to achieve the approving effect is a very significant job for fault diagnosis of transformer. Moreover, real time is another important characteristic so as to meet high-speed diagnosis requirements. Based on points, a fuzzy neural classifier is proposed based on rough sets theory in this paper, the method firstly considers all sorts of gas capacities in transformer oil to form Rogers ratio diagnosis table, then rough sets is applied to implement attributes reduction and a simplified decision table is got, fuzzy algorithm with Gauss subjection function makes attribute values fuzzy, afterwards, fuzzy attributes are connected to input neurons of neural classifier to make patterns classified, finally, a fuzzy neural classifier is formed for fault diagnosis for transformer. The practical results show the approach can effectively minimize the problem-solving scale and improve real time properties, and owns high anti-inference capabilities, and is an effective method for fault diagnosis of transformer.**

Index terms—**Fuzzy neural classifier, rough sets, fault diagnosis, transformer.**

I. INTRODUCTION

In electrical power systems, transformer is one of the most important equipments, whose normal operation is a basic guarantee for reliable power supply. Presently, to study on the data of DGA is a widely applied method for transformer fault diagnosis [1],[2],[3]. References [1], [2] aim at finding out the relationships between fault sources and fault symptoms through samples based on fuzzy methods, which overcomes experts knowledge acquisition puzzle to some extents, but fuzzy subjection degree function and the related power weights are artificially given out, the methods therefore own evident subjectivity. Reference [3] applies rough sets methods to implement knowledge reduction, diagnosis rule sets with diverse belief levels are acquired for incomplete information and knowledge library is also constantly updated, but the method can't get a sound interpretation from information theory angle. Texts [4],[5] adopt neural networks methods to execute malfunctioning diagnosis, the aim is to resolve self-learning as well as bottle neck puzzle in knowledge acquisition, but while samples space distribution is complex, convergence property of neural networks is difficult to be ensured. texts [6],[7] diagnose faults through expert systems, effectively simulating experts reasoning decision process, but when experts knowledge base is not incomplete or immature, reasoning error can't be avoided. A remote state monitoring and fault diagnosis method based on multi-agent is also studied in [8], [9], but in the distributed heterogeneous environment, where the new and the old system coexist, how to coordinate their job is still a difficult thing. Based on points, the paper proposes a fuzzy neural classifier for transformer malfunctioning diagnosis based on rough sets theory, which can effectively overcome the above defects and deficiencies in transformer fault diagnosis, automatically adapt itself to variable environment, and exhibit excellent anti-inference characteristics, practical results show that it is an effective method for fault diagnosis of transformer.

II. ROUGH SETS THEORY

Rough sets theory was proposed by Pawlak [10] in 1982, as a new mathematics tool, it is widely applied to dispose incomplete and uncertain information, whose main aim is that under the precondition without any change in keep classification capabilities, the classification rules of the concept can be acquired through rough sets reduction. Rough sets therefore have been in an increasingly diverse range of applications including artificial intelligence (AI), patter recognition (PR) and decision support systems (DSS) and many other related fields.

A. Decision Table and Reduction

In rough sets theory, knowledge denotation system may be described by

$$S=<U,A,V,F>. \tag{1}$$

where U is universe and expresses a set with finite objects, A is attribute set composed of condition attribute C and decision attribute D, $A= C\cup D$, $C\cap D=\phi$, $a\in A, V=V_a, V_a$ is range of a, $f:U\times A\rightarrow V$ is a information function, it specifies attribute values of every object in U.

Information systems based on rough sets definition can be denoted by the use of table format, where columns express attributes and rows represent objects, and every row describes information of an object. The table therefore is called decision table, which can generalize the relationships among data and educe the classification rules of the concepts. In rough sets, binary indivisible relationship $ind(R)$ determined by $R\subseteq A$ can be expressed by

$$ind(R)=\{(x,y)\in U\times U|\ \forall a\in A, f(x,a)=f(y,a)\}. \tag{2}$$

It is very clearly that if $(x,y)\in ind(R)$, then x and y can not be differentiated according to existing information, they are an equivalent relation in U.

Set S=<U, $C\cup D$>, if $C1\in C$, $C1\neq\phi$, and the following two conditions hold.

A1): $ind_{C1}(D)=ind_C(D)$.

A2): $\forall C2\subseteq C1$, $ind_{C2}(D)\neq ind_{C1}(D)$.

Based on **A1)-A2)** we can say $C1$ is a reduction of C with regard to D, the intersection of all these reductions is called core, defined as $\mathrm{core}_D(C)=\cap\mathrm{red}_D(C)$.

B. Rough Subjection Degree Function

Another denotation means of rough sets is rough subjection degree function $u_x^R(x)$, and is expressed by

$$u_X^R(x)=\frac{card(X\cap[x]_R)}{card([x]_R)}. \quad (3)$$

Where $u_x^R(x)$ expresses degree that the element x belongs to X based on indivisible relationship R. Obviously, those elements of same equivalent class in U own equal function value of rough subjection degree, rough subjection degree function $u_x^R(x)$ satisfies $0\leq u_x^R(x)\leq 1$.

III. Fuzzy neyral classifier

In transformer malfunction diagnosis, incomplete and imprecise information as well as ill data are often met due to environment alteration, measuring interference and other impacts. To resolve the problem, according to fuzzy sets theory, exact input information can be mapped as fuzzy variable information through fuzzy subjection degree function, whose primary aim is to resolve ill data problems in process of classification. Gauss subjection degree function is often selected as fuzzy membership function, and is defined as

$$u_A(x)=\exp(-\frac{1}{2}(\frac{x-c_i}{\sigma_i})^2). \quad (4)$$

In the above equation (4), Gauss subjection degree function is determined by $\{C_i,\sigma_i\}$, parameter C_i establishes the centre of function and σ_i determines width of function.

Let one characteristic variable of input vector is mapped as C fuzzy variables, if precise value of one characteristic variable is x,range is from $x_{\min}$ to $x_{\max}$, then the centre c_i of each subjection functions is determined by

$$c_i=\frac{x_{\max}-x_{\min}}{C-1}(i-1)+x_{\min}, \quad i=1,2,\ldots,C. \quad (5)$$

Selection of σ_i should fully consider adequate overlapping between each subjection function.

Set objects set of classification is composed of N input vectors, and every input vector consists of n-dimensional characteristic variables, s- dimensional characteristic variables are remained after rough sets reduction, after fuzzification, $C\times s$-dimension input vector is gained. For convenience analysis, here let C=3, we have

$$\boldsymbol{u}(X_j)=[u_H(x_{1j}),u_M(x_{1j}),u_L(x_{1j}),\ldots,$$
$$u_H(x_{sj}),u_M(x_{sj}),u_L(x_{sj})], \quad j=1,2,\ldots,N. \quad (6)$$

In the above equation (6) X_j expresses input vector, u_H, u_M and u_L represent Gauss subjection degree function defined by (4) respectively.

As to transformer fault diagnosis, the main aim is to fuzz condition attribute values of decision table, fuzzification attribute values are connected to input neurons of multi-layer sensor for pattern classification. Due to fast exact traits the classifier is very suitable to on-line fault diagnosis, and is an ideal pattern classifier.

Assume that after rough sets reduction, dimension of condition attribute drops to s from initial n, s-dimension vector $X_j=(x_{1j},x_{2j},\ldots,x_{sj})\in \boldsymbol{R}^s$ is input of fuzzy neural networks, after input exactness values are fuzzed, fuzzy input information is yielded by

$$\boldsymbol{F}_j=[\boldsymbol{u}_H(x_{ij}),\boldsymbol{u}_M(x_{ij}),\boldsymbol{u}_L(x_{ij})], \quad i=1,2,\ldots,s. \quad (7)$$

Fuzzification input vector is shown in (6), accordingly, $3s$ neurons in all are required in input layer of networks. The following introduced is batch learning method of fuzzy neural classifier with tutor.

For drill information set S=<U,A>, set N objects are divided into Q classes, according to every output classification, set n_k input objects are mapped as output class l, and meet

$$\sum_{l=1}^{Q} n_{lk}=N. \quad (8)$$

Thus input objects and output l (l=1,2,…,Q) establish a mapping relationships. While characteristic attribute i of n_k input objects is mapped to l, we define centre of l with regard to i by

$$u_{il}=\frac{\sum_{j=1}^{n_k} f_l(X_j)x_{ij}}{\sum_{j=1}^{n_k} f_l(X_j)}, \quad i=1,2,\ldots s;\ l=1,2,\ldots,Q. \quad (9)$$

Similarly, based on fuzzy subjection function of input data, fuzzy variable vector of U_l may be expressed by

$$\boldsymbol{u}(U_l)=[\boldsymbol{u}_H(u_{1l}),\boldsymbol{u}_M(u_{1l}),\boldsymbol{u}_L(u_{1l}),\ldots,$$
$$\boldsymbol{u}_H(x_{sl}),\boldsymbol{u}_M(x_{sl}),\boldsymbol{u}_L(x_{sl})], \quad l=1,2,\ldots,Q. \quad (10)$$

Distance between X_j and l is defined by

$$D_{jl}=\frac{\|\boldsymbol{u}(X_j)-\boldsymbol{u}(U_l)\|}{\sum_{i=1}^{s}\alpha_{x_i}(l)x_{ij}}=\frac{\sum_{i=1}^{s}\sum_{p=1}^{3}[\boldsymbol{u_p}(x_{ij})-\boldsymbol{u_p}(u_{il})]^2}{\sum_{i=1}^{s}\sum_{p=1}^{3}\alpha_{x_i}(l)\boldsymbol{u_p}(x_{ij})},$$
$$l=1,2,\ldots,Q; \quad j=1,2,\ldots,n_k. \quad (11)$$

where subscript P=1,2, 3 respectively expresses H,M,L, $\alpha_{xi}(l)$ is importance factor of x_i related to l, and noting that defuzzification of fuzzy variables is also done, thus, output results can be exactly expressed.

Function $f_l(X_j)$ is used to express a level that input vector X_j belons to l ,and defined by

$$f_l(X_j)=(1+D_{lj})^{-1}=\frac{\sum_{i=1}^{s}\alpha_{x_i}(l)\boldsymbol{x}_{ij}}{\sum_{i=1}^{s}\alpha_{x_i}(l)x_{ij}+\|\boldsymbol{u}(X_j)-\boldsymbol{u}(X_l)\|},$$

$$l=1,2,\ldots,Q;\quad j=1,2,\ldots,n_k. \tag{12}$$

Equation (12) indicates that if distance D_{lj} equals zero, then $f_l(X_j)$ equals one, otherwise if D_{lj} equals infinite, then $f_l(X_j)\to 0$.

In batch learning algorithm of networks with tutor, the aim function of classifier is defined by

$$\boldsymbol{E}=\frac{1}{2}\sum_{l=1}^{Q}\sum_{j=1}^{n_k}f_l(\boldsymbol{X}_j)\|\boldsymbol{X}_j-\boldsymbol{U}_l\|. \tag{13}$$

Where

$$\|\boldsymbol{X}_j-\boldsymbol{U}_l\|=\sum_{i=1}^{s}(x_{ij}-u_{il})^2=\sum_{i=1}^{s}\sum_{p=1}^{3}[\boldsymbol{u}_p(x_{ij})-\boldsymbol{u}_p(x_{il})]^2. \tag{14}$$

The centre of network output classification is acquired by minimizing aim function (13). Weight of each classification centre is attained through the following equation.

$$\frac{\partial \boldsymbol{E}}{\partial u_{il}}=\frac{1}{2}f_l(\boldsymbol{X}_j)(\|\boldsymbol{X}_j-\boldsymbol{U}_l\|)+\frac{1}{2}\|\boldsymbol{X}_j-\boldsymbol{U}_l\|f_l'(\boldsymbol{X}_j). \tag{15}$$

According to (11)(12)(13)(14) (15), we have

$$\frac{\partial \boldsymbol{E}}{\partial u_{il}}=-(x_{ij}-u_{il})f_l^2(X_j). \tag{16}$$

$$\frac{du_{il}}{dt}=-\eta(t)\frac{\partial \boldsymbol{E}}{\partial u_{il}},\ l=1,2,\ldots,Q;\ i=1,2,\ldots,s. \tag{17}$$

According to gradient drop principles, each weight centre of l is updated by

$$u_{il}(t+1)=u_{il}(t)+\eta(t)(x_{ij}-u_{il})f_l^2(X_j). \tag{18}$$

Where $\eta(t)=\frac{1}{|t|}$ is learning decline factor.

Based on the above analysis, batch learning algorithm of classification process is generalized as follows:

i) In accordance with decision table, while output sort belongs to l, according to (18), calculate u_{il}.

ii) For l=1,2,…,Q, in turn update u_{il}, till $|u_{il}(t+1)-u_{il}|<\varepsilon$ where ε is a very small positive integer.

iii) A group of certain classification centre u_{il} can be achieved after batch learning training of networks, then according to (11), distance D_{jl} between X_j and $\boldsymbol{U}_l$ can be calculated, based on (12), function $f_l(X_j)=(1+D_{jl})^{-1}$ that expresses an extent that X_j belongs to l can be work out. Input data vector X_j is believed to belong to l because it is greater than gate limitation value $T(0\le T\le 1)$.

In this way, based on rough sets theory and fuzzy algorithm with Gauss subjection degree function as well as the above defined batch learning process, a fuzzy neural classifier is proposed shown in Fig.1.

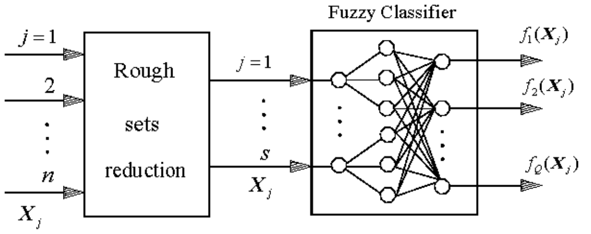

Fig.1 Diagnosis network model

IV. EXAMPLE ANALYSIS

There are 596 samples in all got from the data of DGA for transformer fault diagnosis knowledge library, in accordance with Rogers ratio code [11] of these samples as well as diagnosis results, consequently, decision table for transformer fault diagnosis is done as shown in Table I. Where condition attributes a, b, c and d respectively express the rates of capacities of five kinds of gas volumes such as CH_4/H_2, C_2H_6/CH_4, C_2H_4/C_2H_6 and C_2H_2/C_2H_4, decision attribute e is used to express diagnosis result, that is, 1–normal aging, 2–local discharge, 3–overheating (≤150 celsius degree), 4–overheating (150-200 celsius dugree), 5–overheating (200-300 celsius degree),6–naked metal overheating, 7–winding circumfluence, 8– iron-heart/hull circumfluence or tie-in overload, 9–electric arc discharge (non perfoliate discharge), 10–electric arc discharge (perfoliate discharge), 11–continuous discharge accident, 12–local discharge(creepage electricity trace). The sample number of history fault data in initial data base is represented by k, u express universe.

TABLE I
THE PRIMARY DECISION TABLE

u	k	a	b	c	d	e	u	k	a	b	c	d	e
1	60	0	0	0	0	1	10	3	1	0	1	0	6
2	38	3	0	0	0	2	11	42	1	0	1	0	7
3	35	1	0	0	0	3	12	35	1	0	2	0	8
4	30	2	0	0	0	3	13	50	0	0	0	1	9
5	35	1	1	0	0	4	14	35	0	0	1	2	10
6	40	2	1	0	0	4	15	35	0	0	2	2	11
7	43	0	1	0	0	5	16	5	1	0	2	0	11
8	5	0	1	1	0	5	17	45	3	0	0	1	12
9	35	0	0	1	0	6	18	25	3	0	0	2	12

After inspection we find that Table I has unharmonious decision information such as samples 10 and 11, 12 and 16 in u, then respectively calculating their rough subjection degree function values according to equation (3). For 10 and 11 in u, related to 6 in e, rough subjection degree function value is 3/45, and 42/45 related to 7 in e, we therefore can judge 10 in u is a kind of survey inference and omitted. Similarly, for samples 12 and 16 in u, with respect to 8 in e, rough subjection degree is 35/40, and 5/40 related to 11 in e, so we judge 16 in u is also a sort of error sentence, though, to keep statistic integrality, we add this five samples of 16 in u up to 12 in u and eliminate 16 in u, thus the revised decision information is shown in table II.

TABLE II
THE REVISED DECISION TABLE

u	k	a	b	c	d	e	u	k	a	b	c	d	e
1	60	0	0	0	0	1	9	38	0	0	1	0	6
2	38	3	0	0	0	2	10	42	1	0	1	0	7
3	35	1	0	0	0	3	11	40	1	0	2	0	8
4	30	2	0	0	0	3	12	50	0	0	0	1	9
5	35	1	1	0	0	4	13	35	0	0	1	2	10
6	40	2	1	0	0	4	14	35	0	0	2	2	11
7	43	0	1	0	0	5	15	45	3	0	0	1	12
8	5	0	1	1	0	5	16	25	3	0	0	2	12

After rough sets reduction Table II is still itself, reduction method is seen in [12]. Then based on Gauss subjection function (4), according to (5), order C=3 for a,c,d and C=2 for b, σ_i =0.75 for a and σ_i =0.5 for b,c and d, then we get fuzzy decision table shown in Table III. Each attribute importance factor $\alpha_{xi}(l)$ is calculated as follows:

for u=1, $\alpha_{xi}(l=1)=1$, $\alpha_{xi}(l\neq 1)=0$, $i=a,b,c,d$.

for u=2, $\alpha_{xi}(l=2)=1$, $\alpha_{xi}(l\neq 2)=0$, $i=a,b,c,d$.

for u=3, $\alpha_{xi}(l=3)=1$, $i=b,c,d$;$\alpha_{xi}(l=3)=35/65$,$i= a$, $\alpha_{xi}(l\neq 3)=0$ for $i=a,b,c,d$;

for u=4, $\alpha_{xi}(l=3)=1$, $i=b,c,d$; $\alpha_{xi}(l=3)=30/65$, $i= a$, and $\alpha_{xi}(l\neq 3)=0$, $i=a,b,c,d$.

Likewise, the other $\alpha_{xi}(l)$ (l=5-12) is also calculated by the above methods, here not to enumerate one by one.

According to equation (18), the centre of each output classification is worked out as follows:

$\boldsymbol{u}_1$=(0,0,0,0), $\boldsymbol{u}_2$=(3,0,0,0), $\boldsymbol{u}_3$=(1.54,0,0,0),
$\boldsymbol{u}_4$=(1.47,0,0,0), $\boldsymbol{u}_5$=(0,1,0.18,0), $\boldsymbol{u}_6$=(0,0,1,0),
$\boldsymbol{u}_7$=(1,0,1,0), $\boldsymbol{u}_8$=(1,0,2,0), $\boldsymbol{u}_9$=(0,0,0,1),
$\boldsymbol{u}_{10}$=(0,0,1,2), $\boldsymbol{u}_{11}$=(0,0,2,2), $\boldsymbol{u}_{12}$=(3,0,0,1.64).

According to Gauss subjection degree function, fuzzy variable vector of each output classification centre can be calculated as follows:

$\boldsymbol{u}_1$=(0,0.14,1,0.14,1,0,0.14,1,0,0.14,1).
$\boldsymbol{u}_2$=(1,0.14,0,0. 14,1,0,0.14,1,0,0.14,1).
$\boldsymbol{u}_3$=(0.15,0.99,0.12, 0.14,1,0,0.14,1,0,0.14,1).
$\boldsymbol{u}_4$= (0.12,0.99,0.15,1, 0.14,0,0.14,1,0,0.14,1).
$\boldsymbol{u}_5$= (0,0.14,1,1,0.14,0,0.26,0.87,0,0.14,1).
$\boldsymbol{u}_6$= (0,0.14,1,0.14,1,1,0.14,0,0,0.14,1).
$\boldsymbol{u}_7$= (0.03,0.8,0.41,0.14,1,1,0.14,0,0,0.14,1).
$\boldsymbol{u}_8$=(0.03,0.8,0.41,0.14,1,0.14,1,0.14,0,0.14,1).
$\boldsymbol{u}_9$= (0,0.14,1,0.14,1,0,0.14,1,0.14,1,0.14).
$\boldsymbol{u}_{10}$= (0,0.14,1,0.14,1,0.14,1,0.14,1,0.14,0).
$\boldsymbol{u}_{11}$= (0,0.14,1,0.14,1,1,0.14,0,1,0.14,0).
$\boldsymbol{u}_{12}$= (1,0.14,0,0.14,1,0,0.14,1,0.65,0.44,0).

TABLE III
FUZZY DECISION TABLE

u	$u_H(a)$	$u_M(a)$	$u_L(a)$	$u_H(b)$	$u_L(b)$	$u_H(c)$	$u_M(c)$	$u_L(c)$	$u_H(d)$	$u_M(d)$	$u_L(d)$	e
1	0.00	0.14	1.00	0.14	1.00	0.00	0.14	1.00	0.00	0.14	1.00	1
2	1.00	0.14	0.00	0.14	1.00	0.00	0.14	1.00	0 .00	0.14	1.00	2
3	0.03	0.80	0.41	0.14	1.00	0.00	0.14	1.00	0.00	0.14	1.00	3
4	0.41	0.80	0.03	0.14	1.00	0.00	0.14	1.00	0.00	0.14	1.00	3
5	0.03	0.80	0.41	1.00	0.14	0.00	0.14	1.00	0.00	0.14	1.00	4
6	0.41	0.80	0.03	1.00	0.14	0.00	0.14	1.00	0.00	0.14	1.00	4
7	0.00	0.14	1.00	1.00	0.14	0.00	0.14	1.00	0.00	0.14	1.00	5
8	0.00	0.14	1.00	1.00	0.14	0.14	1.00	0.14	0.00	0.14	1.00	5
9	0.00	0.14	1.00	0.14	1.00	0.14	1.00	0.14	0.00	0.14	1.00	6
10	0.03	0.80	0.41	0.14	1.00	0.14	1 .00	0.14	0.00	0.14	1.00	7
11	0.03	0.80	0.41	0.14	1.00	1.00	0.14	0.00	0.00	0.14	1.00	8
12	0.00	0.14	1.00	0.14	1.00	0.00	0.14	1.00	0.14	1.00	0.14	9
13	0.00	0.14	1.00	0.14	1.00	0.14	1.00	0.14	1.00	0.14	0.00	10
14	0.00	0.14	1.00	0.14	1.00	1.00	0.14	0.00	1.00	0.14	0.00	11
15	1.00	0.14	0 .00	0.14	1.00	0.00	0.14	1.00	0.14	1.00	0.14	12
16	1.00	0.14	0.00	0.14	1.00	0.00	0.14	1.00	1.00	0.14	0.00	12

The followings given out are two fault diagnosis practical examples.

Example 1: Gas volume capacities in transformer oil are tested as follows: $\varphi(CH_4)=21\times10^{-6}$, $\varphi(H_2)=26\times10^{-5}$, $\varphi(C_2H_2)=1\times10^{-6}$, $\varphi(C_2H_6)=15\times10^{-6}$, $\varphi(C_2H_4)=12\times10^{-6}$. Rogers ratio code is (3,0,0,0), after fuzzification, X=(1,0.14,0,0.14,1,0,0.14,1,0,0.14,1), the output results of fuzzy neural networks are shown as follows: $f_1(X)=0.76$,$f_2(X)=0.76$, $f_3(X)=1$, $f_4(X)=0.76$, $f_5(X)=0.67$, $f_7(X)=0.69$,$f_8(X)=0.74$,$f_9(X)=0.61$,$f_{10}(X)=0.48$,$f_{11}(X)=0.46$, $f_{12}(X)=0.61$. Set gate limit T=0.8, since $f_3(X)=1>T$, diagnosis result therefore should be 2(local discharge fault), field practical inspection proves the correctness of diagnosis result.

Example 2: All the tested gas volume capacities in transformer oil are shown as follows: $\varphi(CH_4)=74\times10^{-6}$, $\varphi(H_2)=60\times10^{-6}$, $\varphi(C_2H_2)=103\times10^{-6}$, $\varphi(C_2H_6)=58\times10^{-6}$, $\varphi(C_2H_4)=187\times10^{-6}$. Rogers ratio code is (1,0,2,1), after fuzzification, fuzzy characteristic vector is represented by X=(0.03,0.8,1,0.14,1,0.14,1,0.14,0.14,1,0.14), the output results of fuzzy neural classifier networks are shown as $f_1(X)=0.56$,$f_2(X)=0.52$,$f_3(X)=0.59$, $f_4(X)=0.52$, $f_5(X)=0.50$, $f_6(X)=0.57$,$f_7(X)=0.62$,$f_8(X)=0.77$,$f_9(X)=0.46$, $f_{10}(X)=0.57$, $f_{11}(X)=0.57$, $f_{12}(X)=0.56$. Set T=0.6, since $f_8(X)=0.77>T$, diagnosis result therefore is likely to be 8 in *e*, that is, iron-heart/hull circumfluence or tie-in overload, field inspection proves diagnosis result is correct.

V. Conclusions

i) Rough sets are used to harmonize inconsistent information in transformer fault diagnosis knowledge base and implement knowledge reduction, which can effectively drop problem-solving scale and system space dimensions, and simplify classifier networks architecture, improve real-time properties of malfunction diagnosis.

ii) After characteristic vector is fuzzed, fault-tolerance capacities of systems are dramatically improved, after input variables become fuzzy, user may random input characteristic attribute values not to care too more as ever. Fuzzification of outputs not can diagnose single or multi-fault and but can interpret them reasonably, and is a kind of flexible fault diagnosis method.

iii) The method owns excellent anti-interference capabilities and noise tolerance characteristics, because it considers the fuzzy relationship between output classes and input characteristic variables. In the conventional methods, the above relations are not fully considered, diagnosis results therefore are rigid, that is, flexibility is absent in process of diagnosis. Therefore, the proposed method in this text owns stronger robustness.

iv) The proposed neural classifier combines importance factors of system parameters into fuzzy reasoning methodology, and improves dependence degree of networks reasoning to input characteristics. Learning algorithm of networks is simple, training time is short, and is an ideal pattern classifier dut to its a full combination between input exactness values and fuzzy information. Which makes nonlinear characteristics of systems more rich, classification capabilities of systems therefore are dramatically improved.

References

[1]L. Yang, Y. Shang, X. F. Zhou and Z. Yan, "Probability reasoning and technique applied for identifying power transformer malfunction "(in Chinese), *Proceedings of the CSEE*. Vol. 20, no.7, pp.19-23, Jul. 2000.

[2]J. W. Zhang, D. G. Zhao and L. W. Dong, "An expert system for transform fault diagnosis based on fuzzy mathematics"(in Chinese), *High Voltage Engineering*. Vol.24, no.4, pp. 6-8, 1998

[3]H. C. Shu, X. F. Sun and D. J. Si, "A approach to founding and maintaining ES knowledge base for fault diagnosis of power trans former"(in Chinese), *Proceedings of the CSEE*, vol. 22, no.2, pp.31-35, Feb. 2002

[4]T. S. Bi, Y. X. Ni and F. L. Wu, "A novel neural network approach for fault section estimation"(in Chinese), *Proceedings of the CSEE*. vol. 22, no. 2, pp.73-78, Feb. 2002.

[5]Z. D. Zhang, H. B. Zhong and Y. H. Yang, "Substation alarm message dealing system based on ANN and ES"(in Chinese), *Automation of Electric Power Systems*. vol.25, no..9, pp.45-47,2001.

[6]Lee Heung-Jae, Ahn Bok-Skin and Park-Yong Moon, "A fault diag nosis expert system for distribution substations", *IEEE Trans. on Power Delivery*, vol.15, no.1,pp.92-97,2000

[7]Y. Du, P. C. Zhang and W. Y. Yu, "A substation fault diagnosis system based on case-based reasoning and rule-based reasoning"(in Chinese), *Power System Technology*. vol. 28, no.1, pp.34-37,2004

[8]L. F. Li, L. K. P. Liu and Y. H. Hu, "Architecture of multi-agent based distributed monitoring and control system for substations"(in Chinese). *Power System Technology*. Vol 27, no.6npp. 26-28,2003

[9] L. P. Zhao, S. f. Xie and Q. Z. Li, "Study on multi-agent based power quality monitoring and control system for electrified railway "(in Chinese), *Power System Technology*. Vol. 28, no.24,pp.69-73, 2004

[10]Pawlak, Z, "Rough sets", *International Journal of Information and Computer Science*. vol.11, no.5, pp.341-356,1982.

[11] Q. Zhang, Z. X. Han and F. C. Wen, "Analysis of rogers ratio diagnosis table based on rought set theory"(in Chinese), *The14th Proceedings CUS-EPSA*, VA, pp.742-747,1998.

[12] W. X. Zhang, Y. Liang and W. Z. Wu, *Information System and Knowledge Discover*(in Chinese), BeiJing: Science Press,2003.