

Rough Neuron network for Fault Diagnosis

Yueling ZHAO

College of Electrical Engineering , Liaoning University of Technology, Jinzhou, P.R.China
Email: zhao7267@163.com

Hui jin , Lihong Wang and Shuang WANG

College of Electrical Engineering , Liaoning University of Technology, Jinzhou, P.R.China
Email: { jinhui5868, zhdwanglihong , shuang_lg } @126.com

Abstract—Considering training time of traditional BP neural network is too long and it cannot solve the problems in the input vector with multiple-valued, a new method of BP neural network based on rough neuron is presented. A rough neuron can be viewed as a pair of neurons. One neuron corresponds to the upper boundary and the other corresponds to the lower boundary. Upper and lower neuron exchange information with each other during the calculation of their outputs. Firstly, the continuous attributes in diagnostic decision system are discretized with particle swarm optimization. Then, the reducts are found based on attribute dependence of rough set, and the optimal diagnostic decision is determined. Lastly, according to the optimal decision system, rough neuron network is designed for fault diagnosis. A practical example is given , the method is feasible and available.

Index Terms—rough Set, rough neuron, particle swarm, rough neuron BP neural network, fault diagnosis

I. INTRODUCTION

Intelligent fault diagnosis is an important developing direction of diagnostic technique. The traditional fault diagnosis method is based on the mathematical model of the system, mathematical model is dependent on the diagnosis system structure in [1], but many faults can cause changes in the structure of the system. Neural network technology has self-learning, nonlinear pattern recognition, associative memory and presumed ability, good fault tolerance and expansibility, so fault diagnosis is widely used in [2-4]. There are also many scholars improved neural network, for example, reference [5] optimized BP neural network with L-M method, there are many virtues such as fast convergence and hard into the local minimum values for the network. self-learning ability of BP neural network and interpreting ability of expert system is combined in [6], reliability and accuracy of system are increased.

But when the input information dimension becomes large, the above training time of neural network is too long. Rough Sets Theory was proposed in 1982 by Pawlak in paper [7], it is a new mathematical tool to deal

with vagueness and uncertainty. In reference [8-10], it has been applied in many areas such as pattern recognition, data mining and fault diagnosis. Its reduction operation can remove redundant attributes and redundancy samples, compression information space dimension, simplify the knowledge system, realize the pretreatment of information system. Literature[11,12] used Rough Sets Theory for pretreatment of input data of neural network, extracting key ingredients as input of network, although simplified structure of the neural networks, the operation mechanism of neural network was not improved.

This paper proposes a new neural network base on rough neuron, rough neuron network model by combining rough set theory with neuron concepts, which has broad meaning in real world. Rough neuron network can solve input vector with multiple-valued effectively. A new kind of Rough neuron network is proposed and applied in fault diagnosis.

. PRELIMINARIES AND NOTATIONS

A. Basic concepts of Rough Set

Definition 1 Let an information system is a 4-tuple $S = (U, A, V, f)$, where U is an objects collection (Domain); $A = C \cup D$ is an attribute set, subset C is called condition attribute and D is called decision attribute; A is an attribute set; V is a set of attributes,

$V = \bigcup_{a \in A} V_a$; $f: U \times A \rightarrow V$ is an information function, which designates each object $a \in A$, $x \in U$, $f(x, a) \in V_a$.

The information system that provided with condition attribute and decision attribute is called decision table.

Definition 2 Let an information system is a 4-tuple $S = (U, A, V, f)$, defined the set $\underline{R}(X) = \{x \in U \mid [x]_R \subseteq X\}$ as X 's lower approximation set, and $\overline{R}(X) = \{x \in U \mid [x]_R \cap X \neq \emptyset\}$ as X 's upper approximation set.

Definition 3 Let an information system is a 4-tuple $S = (U, A, V, f)$, defined the set $\bigcup_{x \in U / D} \{x \in U \mid [x]_c \subseteq X\}$

Footnotes: 8-point Times New Roman font;

Manuscript received January 1, 2008; revised June 1, 2008; accepted July 1, 2008.

Copyright credit, project number, corresponding author, etc.

as positive region of set D in set C , marks as $POS_{IND(C)}(IND(D))$ or $POS_c(D)$.

Definition 4 Let an information system is a 4-tuple $S = (U, A, V, f)$, $A = C \cup D$ and the dependability between D and C is defined: $k = \gamma_c(D) = |POS_c(D)|/|U|$. For the attribute of $c \in C$, if $\gamma_c(D) = \gamma_c - c(D)$, then attribute c is redundant attribute which is relative to decision-making attribute D . Or it is indispensable attribute. Under the importance of C relative to D , attribute c is defined: $SSGF(c, C, D) = \gamma_c(D) - \gamma_c - c(D)$. If any attribute is indispensable in C relative to D , then C and D are independent.

Definition 5 Let an information system is a 4-tuple $S = (U, A, V, f)$, $A = C \cup D$ and $P \in C$. If P and D are independent, and $\gamma_c(D) = \gamma_p(D)$, then P is called C 's relative reduction of D , and marks as $RED(C)$.

B. Rough Neuron

A rough neuron in reference [13] can be regarded as consist of two ordinary neurons, one of them is called the upper boundary \bar{r} , the other is called the lower boundary \underline{r} , as shown in figure 1.

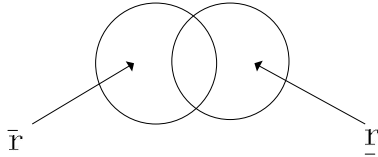


Figure.1 Rough neuron.

The input of a conventional, lower, or upper neuron is calculated using the weighted sum as:

$$input_i = \sum_j \omega_{ji} \times output_j \tag{1}$$

where i and j are either the conventional neurons or upper/lower neurons of a rough neuron. The outputs of a rough neuron r is calculated using a transfer function as:

$$output_{\bar{r}} = \max(f(input_{\bar{r}}), f(input_{\underline{r}})) \tag{2}$$

$$output_{\underline{r}} = \min(f(input_{\bar{r}}), f(input_{\underline{r}})) \tag{3}$$

The output of a conventional neuron i is simply calculated as:

$$output_i = f(input_i) \tag{4}$$

The connection between the upper boundary and the lower boundary represents information exchange, there are three kinds of connection between a rough neuron and another rough neuron, as shown in figure 2, Where figure 2-a shows fully connected between rough neuron r and s ; figure 2-b shows rough neuron r incents rough neuron s ; figure 2-c shows rough neuron r inhibits rough neuron s .

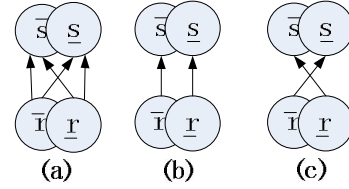


Figure 2. Rough neurons connection graph.

When the connection between neurons is existent, input formula of the upper boundary or the lower boundary neurons can be expressed as $I_j = \sum \omega_{ij} \times O_i$, neuron i and neuron j may be classic neuron, also can be the upper boundary or the lower boundary of rough neuron. Output formula of rough neuron can be expressed as $O_{\bar{r}} = \max\{f(I_{\bar{r}}), f(I_{\underline{r}})\}$, $O_{\underline{r}} = \min\{f(I_{\bar{r}}), f(I_{\underline{r}})\}$, function f using Sigmoid function.

C. Rough Neuron BP Neural Network

Rough neural network is the neural network which enabled the input value that considered an accidental error by having a revolving underwriting facility neuron into input with a rough set theory. It gets possible to handle uncertain data than a rough neuron. I can think about a rough neuron in rough neural network like a pair of a neuron. Therefore I can gain the favor by a rough set theory as a rough neuron and define two normal neuron as revolving underwriting facility neuron of one and take a lower limit of input as the other with an upper limit of input including an accidental error for one input as input of revolving underwriting facility neuron.

According to the traditional BP neural network, input value of each neuron is a precise value, but it may be two values or multiple values in some practical problems. Rough BP neural network can solve these problems. It is similar with traditional BP network, consists of three-layered feed forward network, namely: input layer, hidden layer and output layer, using fully connected between each layer. As shown in figure 3, input layer consists of rough neurons, each rough neuron is consist of two ordinary neurons, hidden layer or output layer is consist of ordinary neuron.

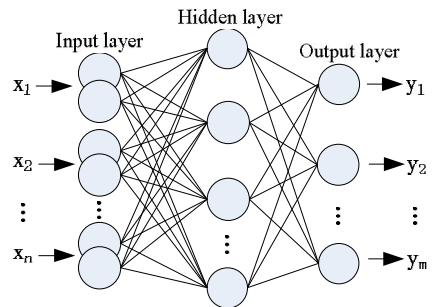


Figure 3. Rough neuron BP neural network.

. ATTRIBUTE DISCRETIZATION BASED PSO

Rough Sets Theory only deal with discretized data, so the continuous attribute should be discretized at first, then

reduction by using Rough Sets Theory. Meanwhile, it is very important to increase running speed of learning algorithm in next step, reduce actual space requirements and time consumption of algorithm, improve clustering ability.

A. Definition of Generalized Discretization

The main point of the existing discretization definition [14] is that condition attribute space is split using selected cuts. The existing discretization methods [14-17] always try to find the best cutting set. When the attributes are interval values, it is difficult to find the cutting set. So it is necessary to generalize the existing definition of discretization. The following description gives the generalized discretization definition.

Decision system $S = (U, CUD)$, where $U = \{x_1, x_2, \dots, x_n\}$ is a finite object set, and subsets $C = \{a_1, a_2, \dots, a_m\}$ and D are called as condition attribute set and decision attribute set respectively. V_a is value domain of attribute, $V_d = \{1, 2, \dots, r(d)\}$ and $r(d)$ are numbers of decision classes.

The point (a, c) can be defined as breakpoint c in a , where $a \in C$ and $c \in R$ (R is set of real numbers), $\forall a \in C$, set of breakpoint $\{(a, c_1^a), (a, c_2^a), \dots, (a, c_{m_a}^a)\}$ in $V_a = [l_a, r_a] \subset R$ is defined as class P_a of value domain V_a .

$$P_a = \{(c_0^a, c_1^a), (c_1^a, c_2^a), \dots, (c_{m_a}^a, c_{m_{a+1}}^a)\} \quad (5)$$

$$l_a = c_0^a < c_1^a < \dots < c_{m_a}^a < c_{m_{a+1}}^a = r_a \quad (6)$$

$$V_a = [(c_0^a, c_1^a) \cup (c_1^a, c_2^a) \cup \dots \cup (c_{m_a}^a, c_{m_{a+1}}^a)] \quad (7)$$

For a $\forall P = \cup P_a$, the new decision table is defined as $S^p = (U, C^p \cup \{d\})$ and $C^p = \{a^p : a^p(x) = i \Leftrightarrow a(x) \in [c_i^a, c_{i+1}^a)\}$, where $x \in U$ and $i \in \{0, \dots, m_n\}$.

The nature of discretization can be attributed to the choice of breakpoint, which it is space division for condition attributes. The m-dimensional space is divided into a finite regions. The object in each region is same value of decision-making.

B. Discretization based PSO

Particle swarm optimization is used to search the breakpoint of continuous attribute for achieving optimal cut sets. In reference [18-19], Particle swarm optimization algorithm is a computation intelligence technique, inspired by social behavior of birds flocking. It is advocated by Eberhart and Kennedy. All the particles are in search of the solution space, each particle has its own

position and velocity, both are expressed by vector. Each position vector of particle represents a solution of optimization problem, by substitution of it into the objective function can get an adaptive value. Decision attribute support degree of rough set theory is used as the optimization objective, preliminarily divide the condition attributes at first, then continuously adjust each divide point. If taking initial divide point as a particle, the process of continuous attributes discretization can be considered particles in search of optimal attribute values.

Speed and position formula below:

$$v[] = \omega \cdot v[] + c_1 \cdot rand() \cdot (G_{best}[] - present[]) + c_2 \cdot rand() \cdot (P_{best}[] - present[]) \quad (8)$$

$$present[] = present[] + v[] \quad (9)$$

Where $v[]$ is particle velocity, $n \times M$ matrix. n is number of particles, namely population size. M is solution space dimension, c_1, c_2 is accelerate constants respectively, take positive number. $rand [0,1]$ is random number in the range of $[0,1]$, $present[]$ is currently particle position, $n \times M$ matrix. ω is inertia weight.

Decision system $S = (U, CUD)$, where $U = \{x_1, x_2, \dots, x_n\}$ is the set of object, quality of classification $a_r(x) = \frac{card(R(x))}{card(U)}$ is a overall measure of the classification ability of decision table.

When particle swarm optimization algorithm is adopted, the initial division points are defined as the particles, where $C = \{c_1, c_2, \dots, c_n\}$ is condition attribute of decision table, $c_i (i = 1, 2, \dots, n)$ is optimization parameters. Quality of classification is a discrete indicator of optimization $f(x)$, the support of decision-making attributes keep without changing in Condition attributes. Every continuous attribute is discreted, the optimal solution of algorithm is particle final state.

Main algorithms are as follows:

Setting maximum iterations and iterative end conditions.

Initialize particle swarm, discretize decision table using each particle. Then, calculate the incompatibility degree and support degree of decision table, take them as fitness of each particle.

Memory individuals and groups corresponding to position of the best fitness, according to formula (8) and (9) in search of better breakpoint.

If it reaches the end condition, then it should jump step ; if it fails to reach the maximum iterations of presupposition, then it should repeat .

Judge the end condition, if it fails to reach it, then it should $n = n + 1$ and return step .

Take the optimal particle as a discrete point to divide decision table. The decision table of discretization is shown as TABLE .

. FAULT DIAGNOSIS BASED ON ROUGH NEURON

A. Rough Neuron Network System

Firstly, the continuous attributes in diagnostic decision system are discretized with particle swarm optimization. Then, the reducts are found based on attribute dependency degree of rough set theory, and the optimal decision system is determined and applied in fault recognition. The basic idea is shown in Figure 4.

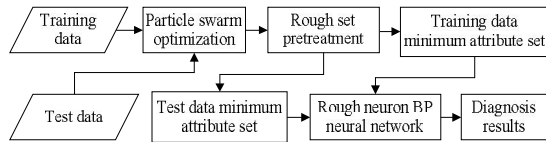


Figure 4. Rough neuron network system.

B. Data Discretization and Reduction

In large rotating machinery, rotor imbalance is a kind of common faults. A turbine fault diagnosis example is presented in [20], there are 11 symptoms, using “training-test” method, divides given data set into training data set and test data set (respectively random distribution according to 70% and 30%). According to section 4 sample data are discretized with particle swarm optimization. A decision table is obtained after discretization, as shown in TABLE 1, no fault expressed by 0 and fault expressed by 1.

TABLE I. DECISION TABLE OF DISCRETIZATION

U	Symptom													Fault
	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆	S ₇	S ₈	S ₉	S ₁₀	S ₁₁	S ₁₂	S ₁₃	
1	4	0	0	0	2	0	5	0	0	1	0	0	1	
2	4	0	0	0	2	0	3	0	0	1	0	0	1	
3	2	0	0	0	2	0	5	0	0	1	0	0	1	
4	4	0	0	1	2	0	5	0	0	1	0	0	1	
5	2	0	0	0	2	0	4	0	0	1	0	0	0	
6	2	4	0	4	0	2	2	2	0	0	2	0	0	
7	2	4	0	4	0	2	2	1	0	0	3	0	0	
8	3	4	0	4	0	1	2	2	0	0	2	0	0	
9	1	4	0	3	0	1	2	1	0	0	1	1	1	
10	1	4	0	4	0	4	0	2	0	0	3	0	0	
11	0	4	0	4	0	4	0	2	0	0	2	0	0	
12	0	4	0	2	0	4	0	1	0	0	1	0	0	
13	1	2	0	2	0	0	2	0	0	0	0	0	0	
14	1	3	0	3	0	0	0	0	0	0	0	0	0	
15	1	4	0	4	0	2	0	0	0	0	3	0	0	
16	1	4	0	4	0	4	0	0	0	0	3	0	0	
17	1	2	0	4	0	4	0	0	0	0	3	0	0	
18	3	1	0	1	1	0	0	0	0	0	1	0	0	
19	4	1	0	1	0	0	0	0	0	0	2	0	0	
20	4	4	0	2	0	0	2	1	0	0	3	1	0	
21	4	4	0	3	0	0	2	1	0	0	2	1	0	

Attribute reduction is deleting the unnecessary condition attribute. Thus after analyzing we get condition

attribute in reduction for decision rules of decision attribute. And people hope to get fewer conditions of reduction results if possible or get fewer decision rules. According to the definition of attribute dependability, in the information system the importance of any attribute $a \in C - P$ is $SIG(a, P, D) = \gamma(P \cup \{a\}, D) - \gamma(P, D)$. Visibly if $SIG(a, P, D)$ is greater and P is known, attribute a relative to decision attribute D is more important. Let an information system is $S = (U, C \cup D, V, f)$, core attribute set is $CORE$ and attribute reduction set is $RED(U)$, specific algorithm is as follows:

Get core attribute set $CORE$ of condition attribute and the dependability y_1 of attribute C relative to attribute D .

Get the dependability y_2 of core attribute set $CORE$ relative to attribute D . If $y_2 = y_1$, then jump , else let $C' = C - RED(U)$.

Get attribute importance of each attribute in C' and sort them by ascending order, maximum in them will be added to $RED(U)$.

Get the dependability y_3 of $RED(U)$ relative to attribute D . If $y_3 < y_2$, then delete the attribute which is just added to $RED(U)$ and return , else return .

Delete redundant condition attribute in $RED(U)$, at last a reduction can be obtained.

Reduct table 1 by using above method, core attribute set is $\{S_7\}$, the original decision table of condition attribute can be reduction for $\{S_2, S_4, S_7\}$.

In the process of discretization, cut sets of each attribute can be obtained, such as 0.1 is a breakpoint of $\{S_7\}$, according to the code rule $[0.06, 0.10)$ is class 1 and $[0.10, 0.34)$ is class 2. So the actual data should be 1 class 1, or class 2? According to the code rule, it should be for class 2, but if considering the error in data acquisition process, the breakpoint may actually be assigned to class 1. If the collected data meets the following formula: $|\lambda - \alpha| / |\beta - \alpha| < \epsilon$, where λ is actual numerical value, α and β denote respectively numerical value which based on discretization cut sets λ located in upper and lower classes of numerical interval, ϵ is error rate and $\epsilon = 0.05$ in this experiment. So after discretization the collected data can be considered two values, not a single value. If it does not meet that error formula, it can be changed into two equal values. Then it will meet the input form of rough neuron. So after discretization 0.1 should be “1,2”. According to the above method, discrete value of the minimum attribute set $\{S_2, S_4, S_7\}$ as shown in TABLE .

TABLE II.
DECISION TABLE OF REDUCING

U	Symptom			Fault
	S ₁	S ₂	S ₃	D
1	0	0	5	1
2	0	0	3	1
3	0	0	5	1
4	0	1	5	1
5	0	0	4	0
6	4	4	1,2	0
7	4	4	1,2	0
8	4	4	1,2	0
9	4	3	1,2	1
10	4	4	0	0
11	4	4	0	0
12	3,4	2	0	0
13	2,3	2	1,2	0
14	3	3	0	0
15	3	4	0	0
16	4	4	0	0
17	2,3	4	0	0
18	1	1	0	0
19	1	1	0	0
20	3,4	2	2	1
21	3,4	3	2	1

C. Diagnosis Analysis based Rough Neuron Network

(1) BP neural network

Number of neurons in neural network hidden layer directly relates to the network performance, BP neural network through the comparison of figure 5 before reduction, when number of neurons in hidden layer is 25, misclassification rate is almost 0;

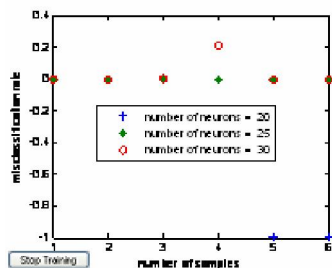


Figure 5. Misclassification rate before reduction.

Let BP neural network without Rough Sets reduces are model BP and network output value is Y_1 ;

TABLE III.
DIAGNOSIS RESULT OF MODEL BP NEURAL NETWORK

l	16	17	18	19	20	21
Y_1	0.0002	0.0000	0.0055	0.0055	0.9880	0.9892

Then Error curve of BP model(Figure 6).

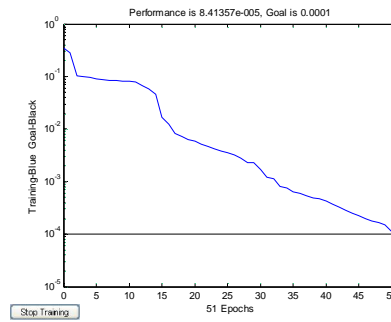


Figure 6. Error curve of BP model

(2) Rough set and BP neural network

let BP neural network with Rough Sets reduce is model RSBP1 and network output value is Y_2 ;

BP neural network through the comparison of figure 7 after reduction, number of neurons in hidden layer should be 6.

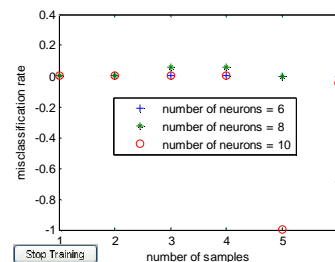


Figure 7. Misclassification rate after reduction

TABLE IV.
DIAGNOSIS RESULT OF MODEL RSBP1

l	16	17	18	19	20	21
Y_2	0.0000	0.0000	0.0038	0.0034	0.9889	0.9915

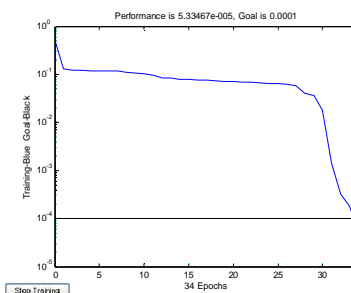


Figure 8. Error curve of model RSBP1

(3) Rough neuron BP neural network

Let rough BP neural network with rough set reduct is model RSBP2 (shown in Figure.9).

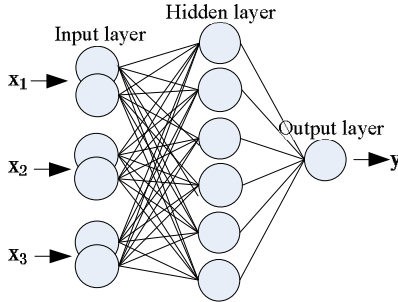


Figure 9. Rough neuron BP neural network

For training sample of No. p ($p = 1, 2, \dots, P$), Input of No. i rough neuron is property values which is conditional property of No. i in input state of No. p . The value is denoted as $x_{pi} = (\underline{x}_{pi}, \overline{x}_{pi})$. Then the output of No. i rough neuron :

$$\underline{z}_{pi} = \min(f(\underline{x}_{pi}), f(\overline{x}_{pi})) \quad (10)$$

$$\overline{z}_{pi} = \max(f(\underline{x}_{pi}), f(\overline{x}_{pi})) \quad (11)$$

Input layer of rough neurons is fully connected with the hidden layer of neurons. So input of hidden layer in No. j :

$$r_{pj} = \sum_{i=1}^n (\underline{\omega}_{ji} \cdot \underline{z}_{pi} + \overline{\omega}_{ji} \cdot \overline{z}_{pi}) + \theta_j \quad (12)$$

Where $\underline{\omega}_{ji}$ and $\overline{\omega}_{ji}$ is the connection weights of lower limit and upper limit, hidden layer of No. j and input layer of No. i respectively.

$$o_{pj} = g(r_{pj}) \quad (13)$$

Let v_{kj} is the connection weights in neuron of output layer NO. k and hidden layer No. j , θ_k is threshold, Then the output of output layer No. k :

$$y_{pk} = h(\sum_{j=1}^m v_{kj} o_{pj} + \theta_k) \quad (14)$$

Where $f(\cdot)$, $g(\cdot)$, $h(\cdot)$ is transfer function of Input layer, hidden layer and output layer respectively. Let

$f(\cdot)$ is 1, $g(\cdot)$ is function of Sigmoid, $h(\cdot)$ is function of purelin.

The network error is defined as:

$$E = \frac{1}{2} \sum_{p=1}^P \sum_{k=1}^l (d_{pk} - y_{pk})^2 \quad (15)$$

According to the steepest gradient descent method, the weights is varied along the negative gradient direction of the error function. The correction formula of weight value is derived as :

Output layer:

$$v_{kj}(t+1) = v_{kj}(t) + \eta \delta'_{pk} o_{pj} \quad (16)$$

$$\delta'_{pk} = c_k (d_{pk} - y_{pk}) \quad (17)$$

Hidden layer :

$$\underline{\omega}_{ji}(t+1) = \underline{\omega}_{ji}(t) + \eta' \delta'_{pj} \underline{z}_{pi} \quad (18)$$

$$\overline{\omega}_{ji}(t+1) = \overline{\omega}_{ji}(t) + \eta' \delta'_{pj} \overline{z}_{pi} \quad (19)$$

$$\delta'_{pj} = -\frac{e^{-r_{pj}}}{1 + e^{-r_{pj}}} \sum_{k=1}^n \delta'_{pk} v_{kj} \quad (20)$$

Where η , η' is learning factor respectively

Let network output value is Y_3 ; so model RSBP2 of actual output as shown in TABLE .

TABLE V.
DIAGNOSIS RESULT OF MODEL RSBP2

	16	17	18	19	20	21
Y_3	0.0000	0.0000	0.0000	0.0000	0.9998	0.9998

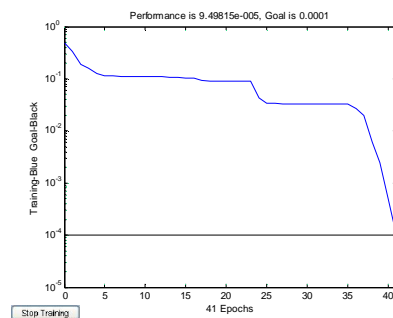


Figure10. Error curve of model RSBP2

Three models of error curve as shown in figure 7, 8 and 10, obviously model BP neural network reaches minimum error after 51 steps and model RSBP1 reaches minimum error after 34 steps. Both the result completely consistent, but the latter neural network has simple

structure, the shorter training. However, model RSBP₂ reaches minimum error after 41 steps, although training time between model BP neural network and model RSBP₂, its accuracy is higher, fully displays the feasibility of the model of the fault diagnosis.

. CONCLUSION

There are differences between traditional BP neural network and neuronal network that inputs data possessing. Neural network based on rough neurons is proposed in this paper, and given a rough neuron network structure. But due to rough set only can deal with discrete attribute, in the process of continuous attributes discretization, some characteristic data may be classified roughly, not an exact value. If these characteristic data are still taken as traditional BP neural network's input, it has to lose information. Rough neuron neural network compatible with multiple-valued input information by using rough neurons instead of traditional neurons, This also accord with the actual situation. The simulation results show that the complicated network structure leads to the running time of growth, the diagnostic accuracy is higher than the other model.

ACKNOWLEDGMENT

This work was supported by National Nature Science Foundation of Liaoning Province (NO.20071096), Educational Commission of Liaoning Province (NO. 2009A358)

REFERENCES

- [1] ZHENG X X, QIAN F. Research and development of fault diagnosis methods for dynamic system. *Control and Instruments in Chemical Industry*, 2005, 32(4): 1-7.
- [2] FENG J, LIAO Y, WANG J Q. Fault diagnosis based on ANN in water circulation system. *Journal of Hunan Institute of Engineering*, 2005.3, 15(1): 47-50.
- [3] SHI H F, LIU J R, MA Y F. The application of neural-network-based fault diagnosis expert system for sluice. *Chinese Journal of Scientific*, 2005.8, 26(8): 769-778.
- [4] LIAO J J. Research on data mining system for failure diagnosis based on neural network. *Information Technology*, 2009, 4: 31-33.
- [5] HUANG D, SONG X. Application of neural network to chemical fault diagnosis . *Control Engineering of China*, 2006.1, 13(1): 6-9.
- [6] GAO Y, YANG H ZH. Fault diagnosis method for chemical polymeric process based on neural networks. *Control Engineering of China*, 2005.5, 12: 52-54.
- [7] Pawlak Z, *Rough Sets*. *Informational Journal of Information and Computer Sciences*, 1982, 11(5): 341-356.
- [8] CHEN T, LUO J Q. Fuzzy pattern recognition for radar signals based on rough set fixed weight. *Computer Applications and Software*, 2009.1, 26(1): 25-27.
- [9] GUO Q L, ZHENG L. A novel approach for fault diagnosis of steam turbine unit based on fuzzy rough set data mining theory. *Proceedings of the CSEE*, 2007.3, 27(8): 81-87.
- [10] GUO X H, MA X P. Fault diagnosis feature subset selection using rough set. *Computer Engineering and Applications*, 2007, 43(1): 221-224.
- [11] GUO X H, MA X P. Fault diagnosis based on rough set and neural network ensemble. *Control Engineering of China*, 2007.1, 14(1): 53-56.
- [12] DONG J K, LI Y, GENG H. Research on airborne equipment fault diagnosis method based on rough set-neural network. *Avionics Technology*, 2008.3, 39(1): 37-41.
- [13] P.J. Lingras, *Rough neural networks*, *Proceedings of Sixth International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Granada, Spain, 1996, pp. 1445-1450.
- [14] Lin, T.Y.: Introduction to the special issue on rough sets. *International Journal of Ap-proximate Reasoning*. Vo.15. (1996) 287-289
- [15] Dai, J.H., Li Y.X.: Study on discretization based on rough set theory. *Proc. of the first Int. Conf. on Machine Learning and Cybernetics*. (2002) 1371-1373
- [16] Roy, A., Pal, S.K.: Fuzzy discretization of feature space for a rough set classifier. *Pattern Recognition Letter*. Vol.24. (2003) 895-902
- [17] usmaga, R.: Analyzing discretizations of continuous attributes given a monotonic dis-crimination function. *Intelligent Data Analysis*. Vol.1. (1997) 157-179
- [18] Eberhart R C, Kennedy J. *Particles swarm optimization [C]*. *IEEE International Conference on Neural Network*, Perth, Australia, 1995: 1942-1948.
- [19] Eberhart R C, Kennedy J. A new optimizer using particles swarm theory [C]. *Proc. of Sixth International Symposium on Micro Machine and Human Science*, Nagoya, Japan, 1995: 39-43.
- [20] YANG SH Z, DING H, SHI T L. *Diagnose reasoning based on knowledge [M]*. Beijing: Tsinghua University Press, 1993: 30-31.



Zhao Yueling, born in Jinzhou, China, in June 7th, 1972, and received his Ph.D degree in Northeastern University, Shenyang, China in 2006.

He is currently a associate professor with College of Electrical Engineering, Liao Ning University of Technology ,China. His current research interests include Intelligent Control , and Rough Set Theory Applications, et al.



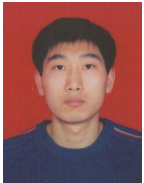
Jin Hui, born in Jinzhou, China, in August 6, 1963, and received his Ph.D degree in Northeastern University, Shenyang, China in 2009.

She is currently a associate professor with College of Automobile and transportation engineering, Liaoning University of Technology, China. Her current research interests include Intelligent Control, Logistics, supply chain management, and Production Planning, et al.



Wang Lihong, born in liaoyang, China, in October 10th, 1971, and received her Master degree in Dalian Maritime affairs university, Dalian, China in 2009.

She is currently a associate professor with College of Electrical Engineering, Liao Ning University of Technology ,China. Her current research interests include intelligent Control.



Wang Shuang, born in Panjin, China, in February 15th, 1984, received bachelor degree in Shenyang Institute of Chemical Technology, Shenyang, China in 2006.

He is a graduate student in Liao Ning University of Technology, China. His current research directions is intelligent control.