Fault Diagnosis Using Rough Set Theory

Yann-Chang Huang, Huo-Ching Sun, Wei-Chi Su

Department of Electrical Engineering, Cheng Shiu University
Kaohsiung, TAIWAN
huangyc@csu.edu.tw

***Abstract*—This paper has presented an effective and efficient approach to extract diagnosis rules from inconsistent and redundant data set of power transformers using rough set theory. The extracted diagnosis rules can effectively reduce space of input attributes and simplify knowledge representation for fault diagnosis. The fault diagnosis decision table is first built through discretized attributes. Next, the GA based optimization process is used to obtain the minimal reduct of symptom attributes. Finally, the rule simplification process is adapted to achieve the maximal generalized decision rules, which can be derived from inconsistent and redundant information. Experimental results demonstrate that the proposed approach has remarkable diagnosis accuracy than the existing method.**

***Keywords*— Fault Diagnosis, Dissolved Gas Analysis, Rough Set**

1. INTRODUCTION

Fault diagnosis of power transformers must strictly be periodically examined to find incipient faults and to protect them from further deterioration as early as possible [1-4]. Dissolved gas analysis (DGA) methods have been widely used [5-11] to detect incipient faults in transformers, which DGA identifies faults by considering the ratios of specific dissolved gas concentrations, their generation rates, and the total combustible gas detected by sampling and testing of the transformer insulation oil.

Fuzzy expert system [12] has been suggested to diagnose incipient faults of transformers. The crisp boundaries of the gas attributes to classify the fault types were fuzzified in this fuzzy expert system to handle the imprecision and incompleteness nature of the transformer fault diagnosis problem. The diagnosis results were promising; however, the fuzzy expert system could not learn from previous diagnosis results because the membership functions and the diagnostic rules were determined by practical experience or trial-and-error tests.

Fuzzy diagnosis systems [13,14] were developed for transformer fault diagnosis to acquire knowledge directly from training data and thus circumvent the disadvantages of the fuzzy expert system [12]. However, the numbers of classification attributes and fuzzy partitions were limited to reduce the number of decision variables to be determined, due to simultaneous determination of the membership functions and the inference rules in the diagnostic systems.

Artificial neural networks (ANNs) [15-17] have been presented to deal with the transformer fault diagnosis, due to their accurate and efficient performance in numerical modeling problems, built-in fault tolerance and real-time response in practical applications. The ANNs can acquire new experiences by incremental training from newly obtained samples. Moreover, they can interpolate and extrapolate from their experiences, yielding at least a best guess of the fault. The ANNs trained by an error back-propagation algorithm have great diagnostic capabilities. However, certain issues, such as local convergence and determination of the network configuration and control parameters (learning rate and momentum constant), must be resolved before the ANNs can become a practical tool.

This paper presents rough set theory (RST) to handle vagueness and uncertainty inherent in making decisions. RST has been applied to some branches of artificial intelligence and cognitive sciences, such as machine learning, knowledge discovery from databases, expert systems, inductive reasoning, pattern recognition and learning [18-20].

2. ROUGH SET THEORY

The basic definitions of the RST are briefly stated as follows.

An inform system S is an ordered pair $S = (U,A)$, where U is a nonempty, finite set called the universe, A is a nonempty, finite set of attributes. Each attribute $a \in A$ is a total function a: $U—V_a$, where V_a is the set of values of a, called the domain of a. In the RST, the elements of the universe are referred to as objects which are characterized through their attribute values.

An indiscernibility relation is a binary relation that identifies objects which have the same descriptions with respect to a set of attributes of objects. Let $S = (U,A)$ be an information system. Each subset of attributes $B \subseteq A$ defines a equivalence relation $IND(B)$, called an B-indiscernibility relation;

$$B = \{(x,y) \in U \times U : \forall a \in B, a(x) = a(y)\}. \quad (1)$$

Obviously, $IND(B)$ is an equivalence relation and

$$IND(B) = \bigcap_{a \in B} IND(a). \quad (2)$$

Given $S = (U,A)$, let $X \subseteq U$ be a set of objects, and $B \subseteq A$ be a set of attributes. The B-lower and B-upper approximation of X is defined respectively as follows:

$$\underline{B}X = \{x \in U : [x]_B \subseteq X\} \quad (3)$$

$$\overline{B}X = \{x \in U : [x]_B \cap X \neq \emptyset\} \quad (4)$$

where $[x]_B$ denotes an equivalence class of B containing $x \in U$.

The set $\underline{B}X$ is the set of all elements of U, which can be with certainty classified as members of X, with respective to the values of attributes from B; and the set $\overline{B}X$ is those elements of U, which can be possibly classified as members of X, with respective to the values of attributes from B.

Let $S = (U,A)$ be an information system with k objects. The discernibility matrix of S is a $k \times k$ matrix with entries c_{ij} consisting of the set of attributes from A on which objects x_i and x_j differ, i.e.,

$$c_{ij} = \{a \in A : a(x_i) \neq a(x_j)\}, \text{ for } i, j = 1, 2, \ldots, k. \quad (5)$$

A discernibility function f_S for S is a propositional formula of n Boolean variables, $a_1^*, ..., a_n^*$, with respective to the attributes $a_1, ..., a_n$, defined as bellows.

$$f_S(a_1^*, ..., a_n^*) = \bigwedge_{1 \le j < i \le n} \bigvee_{c \in c_{ij}^*, c_{ij} \neq 0} c \quad (6)$$

where $c_{ij}^* = \{a^* : a \in c_{ij}\}$.

The discernibility function f_S describes constraints which must hold to preserve discernibility between all pair of discernible objects from S. It requires keeping at least one attribute from each non-empty element of the discernibility matrix corresponding to any pair of discernible objects. The set of all prime implicants of f_S determines the set of all reducts of any information system $S = (U,A)$.

The B-reduct of A is the minimal subset of A, which provides the same classification of objects into elementary classes of B as the whole attributes A. The B-core of A is the essential part of A, which can not be eliminated without disturbing the ability to classify objects into elementary classes of B. A attribute $a \in B$ is superfluous in B, if $IND(B) = IND(B\text{-}\{a\})$; otherwise, a is indispensable in B. If all attributes $a \in B$ are indispensable in B, then B will be called orthogonal.

Subset $B' \subseteq B$ is a reduct of B, denoted $red(B)$, iff B' is orthogonal and $IND(B) = IND(B')$. The set of all indispensable attributes in B will be called the core of B, denoted $core(B)$, and defined as:

$$core(B) = \bigcap red(B) \quad (7)$$

A decision system $A = (U, C\{d\})$ is an information system for which the attributes are separated into disjoint sets of condition attributes C and decision attributes D, $C \cap D = \emptyset$.

A decision system is represented in the form of a decision table, in which its rows contain some objects and columns contain the values of attributes describing the objects, and the decision table contains rules specifying what decisions should be made when certain conditions are satisfied. Note that some redundant and inconsistent attributes may exist in the decision rules, and the reduction of the decision table must be further processed to eliminate these attributes.

3. FAULT DECISION ATTRIBUTES REDUCTION

A decision table must be established before using rough set for fault diagnosis of turbine-generator unit. The relationship between the ratio of gases and fault are complicated. In [21], 15 gas ratios are selected as the input features, including C_2H_2/C_2H_4, CH_4/H_2, C_2H_4/C_2H_6, C_2H_6/CH_4, CH_4/C_2H_4, C_2H_2/CH_4, C_2H_2/C_2H_6, C_2H_6/H_2, C_2H_4/H_2, C_2H_2/H_2, CH_4/total hydrocarbon, C_2H_4/ total hydrocarbon, C_2H_6/total hydrocarbon, C_2H_2/ total hydrocarbon, H_2/(total hydrocarbon plus H_2). Five faults were considered in this paper:

(1) no fault,
(2) lower energy discharge,
(3) high energy discharge,
(4) low temperature overheating,
(5) high temperature overheating.

A decision table [21] is considered in this paper, as shown in Table 1. As listed in this table, the number of input feature is from 15 down to 5, the number of rules is from 29 down to 20. Therefore, the dimension of database is greatly reduced. However, inconsistent cases exist in the decision rules. Table 2 shows that rules 7, 12 are inconsistent rules, and each rule has the same values of symptom attributes, but possess different decision attributes.

TABLE 1

DECISION TABLE [21]

U	c_1	c_2	c_3	c_4	c_5	*d*
1	1	0	1	0	2	1
2	0	0	2	2	2	2
3	0	0	2	2	1	4
4	1	1	0	1	0	5
5	2	2	1	0	1	3
6	2	2	0	0	1	3
7	1	1	0	0	0	1 or 5
8	1	2	1	1	0	5
9	1	1	1	1	0	4
10	2	2	0	0	2	3
11	1	2	0	0	0	5
12	0	1	1	2	1	4 or 5
13	0	0	2	1	2	2
14	0	1	0	0	0	5
15	1	1	1	2	1	4
16	2	1	2	1	2	2
17	2	2	1	1	1	3
18	2	0	2	1	1	1
19	0	0	1	1	0	4
20	1	2	0	1	0	5

In this paper, attribute reduction is a process to obtain a subset from the original set of symptoms of the given fault patterns. The attribute reduction is solved by an optimal process guided by the proposed genetic algorithm (GA) [22], and the proposed GA is used to compute the minimal reduct.

As shown in Table 2, there are 14 decision rules can be generated by the proposed GA, However, inconsistent cases exist in the decision rules. Table 2 shows that rules 3 and 13 are inconsistent rules, and each rule has the same values of symptom attributes, but possess different decision attributes.

A set of decision rules can be generated by the minimal reduct obtained from the GA; however, the rules may be inconsistent. This paper develops a process for maximum generalized decision rules from imprecise, incomplete and inconsistent diagnosis rules. In the process of rule simplification, the maximal number of symptom attribute values is removed without losing essential information, and the maximum generalized rules can be achieved. The process for obtaining the maximum generalized rules is presented in [22]. As listed in Table 3, 14 maximal generalized decision rules can be obtained from Table 2.

TABLE 2

REDUCED DECISION TABLE

No.	c_1	c_2	c_3	c_4	c_5	*d*
1	1	0	*	*	*	1
2	2	0	*	*	*	1
3	0	1	*	0	*	1or 5
4	2	1	*	*	*	2
5	0	*	*	*	2	2
6	2	2	*	*	*	3
7	0	0	*	*	1	4
8	1	1	1	*	*	4
9	1	*	*	2	*	4
10	*	0	*	*	0	4
11	*	*	0	1	*	5
12	0	*	0	*	*	5
13	0	1	1	*	*	4 or 5
14	1	2	*	*	*	5

*: don’t care

TABLE 3

MAXIMAL GENERALIZED DECISION RULES

No.	c_1	c_2	c_3	c_4	c_5	*d*
1	1	0	*	*	*	1
2	2	0	*	*	*	1
3	*	1	*	0	*	1
4	2	1	*	*	*	2
5	0	*	*	*	2	2
6	2	*	0	*	*	3
7	*	*	*	*	1	4
8	1	1	*	*	*	4
9	*	*	*	2	*	4
10	*	*	*	*	0	4
11	0	*	1	*	*	5
12	*	*	0	*	*	5
13	0	*	0	*	*	5
14	1	2	*	*	*	5

4. TEST RESULTS

The 12 samples listed in the Table 4 are used to test the diagnosis accuracies of the proposed approach and the conventional approach.

Experimental results demonstrate that the proposed approach has remarkable diagnosis accuracy (100%) than the existing method (83.3%) presented in [21].

TABLE 4

TEST DATA

U	c_1	c_2	c_3	c_4	c_5	fault	RS-ANN	RST
1	0	1	1	2	1	5	5	5
2	0	1	2	2	1	4	4	4
3	2	1	2	1	2	2	2	2
4	1	1	0	0	1	5	3	5
5	1	0	2	2	1	4	4	4
6	1	1	2	2	2	4	2	4
7	1	1	0	1	1	2	2	2
8	0	2	1	2	2	2	2	2
9	1	2	2	1	0	5	5	5
10	1	2	0	0	0	5	5	5
11	2	1	0	0	1	3	3	3
12	2	2	0	0	1	3	3	3
Diagnosis accuracy							83.3%	100%

5. CONCLUSIONS

This paper has presented an effective and efficient approach to extract diagnosis rules from inconsistent and redundant data set of power transformers using rough set theory. The extracted diagnosis rules can effectively reduce input features, simplify knowledge representation and fault diagnosis task. After the fault diagnosis decision table is built using discretized attributes, the GA based optimization process is further used to achieve the minimal reduct of input attributes. Then, the process of rule simplification is used to obtain the maximal generalized decision rules, which can be derived from inconsistent and redundant information. The test results confirm that the proposed approach is much more diagnostically accurate than the existing methods.

REFERENCES

[1] R. J. Heywood, A. M. Emsley, and M. Ali, “Degradation of cellulosic insulation in power transformers. Part 1: Factors affecting the measurement of the average viscometric degree of polymerisation of new and aged electrical papers,” IEE Proceedings—Science, Measurement and Technology, vol. 147, no. 2, pp. 86-90, 2000.

[2] A. M. Emsley, X. Xiao, R. J. Heywood, and M. Ali, “Degradation of cellulosic insulation in power transformers. Part 2: Formation of furan products in insulating oil,” IEE Proceedings—Science, Measurement and Technology, vol. 147, no. 3, pp. 110-114, 2000.

[3] A. M. Emsley, X. Xiao, R. J. Heywood, and M. Ali, “Degradation of cellulosic insulation in power transformers. Part 3: Effects of oxygen and water on ageing in oil,” IEE Proceedings—Science, Measurement and Technology, vol. 147, no. 3, pp. 115-120, 2000.

[4] A. M. Emsley, R. J. Heywood, M. Ali, and X. Xiao, “Degradation of cellulosic insulation in power transformers. Part 4: Effects of ageing on the tensile strength of paper,” IEE Proceedings—Science, Measurement and Technology, vol. 147, no. 6, pp. 285-290, 2000.

[5] P. S. Pugh, and H. H. Wagner, “Detection of incipient faults in transformer by gas analysis,” AIEE Transaction, vol. 80, pp. 189-195, 1961.

[6] J. J. Kelly, “Transformer fault diagnosis by dissolved-gas analysis,” IEEE Trans. on Industry Applications, vol. 16, no. 6, pp. 777-782, 1980.

[7] E. Dornenburg, and W. Strittmater, “Monitoring oil cooling transformers by gas analysis,” Brown Boveri Review, vol. 61, pp. 238-274, 1974.

[8] IEC Publication 599, Interpretation of the analysis of gases in transformers and other oil-filled electrical equipment in service, First Edition, 1978.

[9] R. R. Rogers, “IEEE and IEC codes to interpret incipient faults in transformers, using gas in oil analysis,” IEEE Trans. on Electrical Insulation, vol. 13, no. 5, pp. 349-354, 1978.

[10] M. Duval, “Dissolved gas analysis: It can save your transformer,” IEEE Electrical Insulation Magazine, vol. 5, no. 6, pp. 22-27, 1989.

[11] ANSI/IEEE Std C57.104-1991, IEEE guide for the interpretation of gases generated in oil-immersed transformers, IEEE Power Engineering Society, 1992.

[12] C. E. Lin, J. M. Ling, and C. L. Huang, “An expert system for transformer fault diagnosis and maintenance using dissolved gas analysis,” IEEE Trans. on Power Delivery, vol. 8, no. 1, pp. 231-238, 1993.

[13] Y. C. Huang, H. T. Yang, and C. L Huang, "Developing a new transformer fault diagnosis system through evolutionary fuzzy logic," IEEE Trans. on Power Delivery, vol. 12, no. 2, pp. 761-767, 1997.
[14] H. T. Yang and C. C. Liao, "Adaptive fuzzy diagnosis system for dissolved gas analysis of power transformers," IEEE Trans. on Power Delivery, vol. 14, no. 4, pp. 1342-1350, 1999.
[15] Y. Zhang, X. Ding, Y. Liu, and P. J. Griffin, "An artificial neural network approach to transformer fault diagnosis," IEEE Trans. on Power Delivery, vol. 11, no. 4, pp. 1836-1841, 1997.
[16] W. Xu, D. Wang, Z. Zhou, and H. Chen, "Fault diagnosis of power transformers: application of fuzzy set theory, expert systems and artificial neural networks," IEE Proceedings—Science, Measurement and Technology, vol. 144, no. 1, pp. 39-44, 1997.
[17] Z. Wang, Y. Liu, and P. J. Griffin, "A combined ANN and expert system tool for transformer fault diagnosis," IEEE Trans. Power Delivery, vol. 13, no. 4, pp. 1224-1229, 1998.
[18] Z. Pawlak, Rough Sets, International Journal of Information and Computer Science, Vol. 11, No. 5, pp. 341-356, 1982.
[19] Z. Pawlak, J. Grzymala-Busse, R. Slowinski and W. Ziarko, Rough Sets, Communications of the ACM, Vol. 38, No. 11, pp. 89-95, 1995.
[20] Z. Pawlak, Rough Sets and Intelligent Dada Analysis, Information Sciences, Vol. 147, No. 1, pp. 1-12, 2002.
[21] X. Yu and H. Zang, "Transformer Fault Diagnosis Based on Rough Sets Theory and Artificial Neural Networks," 2008 *International Conference on Condition Monitoring and Diagnosis*, Beijing, China, April 21-24, 2008.
[22] Y. C. Huang, H. C. Sun, and H. J. Lu, "Decision Rule Generation Using Data Mining Approach," paper submitted to *2009 International Conference on Advanced Information Technologies (AIT)*, Taichung, Taiwan, April 24-25, 2009.