A Decision Tree Approach for Steam Turbine-Generator Fault Diagnosis

Shaaban M. Shaaban[1] and Hossam A. Nabwey[1,2]

[1]*Department of Engineering Basic Science, Faculty of Engineering, Menofia University, Shebin Elkom, Egypt*
[2]*Department of Mathematics, Faculty of Science, Salman Bin Abdulaziz University, Al-Kharj, Saudi Arabia*
Shabaan27@gmail.com, eng_hossam21@yahoo.com

Abstract

Redundancy and inconsistency are universal features of the turbine vibration fault diagnosis. If we can provide a solution to the problem, it should be very meaningful that the fault diagnosis data included redundant and inconsistent information could be used to decision-making rules of fault diagnosis. A novel data mining approach for fault diagnosis of turbine generator unit is proposed based on a decision tree in this paper. In terms of history samples library of turbine generator faults, the method applies entropy-based information gain as heuristic information to select test attributes, and uses ID3 algorithm to generate the decision tree and distilling classification rules are handled. The research shows the method not only possesses rapid induction learning ability and classification speed, but also can effectively compress data and save memory, and is an effect turbine generator fault diagnosis method. In the end, a practical application indicates the validities of the method.

Keywords: *Turbine Generator; Fault diagnosis; Decision Tree; Data Mining*

1. Introduction

Turbine generator is a major device of the thermal plants to convert heat energy of steam into electrical energy through mechanical energy. The generator fault not only damages the generator itself, but also causes outages and loss of profits. Because of complexity and coupling of steam turbine-generator set vibration and structure It is difficult to establish relationship between fault cause and symptom with theoretical analysis way. In order to portray fault mode of steam turbine-generator set as possibly, improve the accuracy of fault diagnosis, and extract a lot of fault symptoms, the importance of fault symptoms is not identical and some are even redundant in fault diagnosis process [1]. At the same time, the information describing the fault mode may be inconsistent or contradictory, which was collected from steam turbine-generator set, in the actual fault diagnosis. For instance, the values of fault symptoms are identical and the fault types are different for tow fault examples. The redundant and inconsistent information exists, on the one hand, it is a waste of diagnosis resources; On the other hand, it directly impact on the making of simple, cost-effective diagnostic decision rules and the efficiency and real-time of fault diagnosis [2].

For the redundancy and inconsistency of the fault diagnosis of the problem, in order to get a better adaptability and ability to match the largest diagnostic decision-making rules. Handle the historical fault data of the steam turbine, and then establish diagnosis decision-making rules and format knowledge base. Using data mining methods of rough set to reason the fault diagnosis network model from decision-making table, and then extract diagnosis rules and provide a reliable basis for fault diagnosis [3]. Various artificial intelligent (AI) techniques have been proposed for fault diagnosis, such as the artificial neural networks (ANNs) [4-5], fuzzy logic (FL) theorem [6], fuzzy neural networks (FNN) [7] and expert systems [8]. In this research, the major fault diagnosis scheme is based on the vibration feature of rotary machines [9-10]. Combining the AI theorem and vibration features to improve the diagnostic accuracy, some theories were proposed including the ANNs based method [5], wavelet based techniques [11] and FL based scheme [12], and provide promising results. Wavelet neural network (WNN) has been applied on pattern recognition, such as classifying voices and images, with successful results [13]. The application of fault diagnosis on turbine generator appeared in [12], where the WNN must consider many parameters' initialization which is vital to the success of the complicated network. However, the learning time is too long to be applied on line. The fuzzy based method strongly depends on experienced experts. The inference rules and defuzzifier must be continuously revised and maintained. The ANNs are used the error back propagation to adjust weighted parameters, and achieve the nonlinear mapping relation to the desired diagnostic results. However, the local minimum problem, slower leaning speed and the weights interferences among different fault patterns are its major drawbacks.

In order to improve the instance, decision tree method was proposed to implement steam turbine generator fault diagnosis. The paper firstly presents the decision tree method and its generation process, and at last disposes the data of vibration steam turbine generator fault diagnosis according to the decision tree. The results show that the method can gain faster diagnosis speed and larger data-compression quantity as well as stronger fault-tolerant ability etc, and is an effective fault data disposal method for steam turbine generator fault diagnosis.

2. Decision Tree Analysis and Generation Method

A decision tree is a kind of tree structure similar to a flow chart, where each inner node expresses test or selection for an attribute, and every branch represents a tested output, but each leaf node all represents class or class distribution. The top node in tree is called root node. The structure of a simple decision tree is shown in Figure 1.

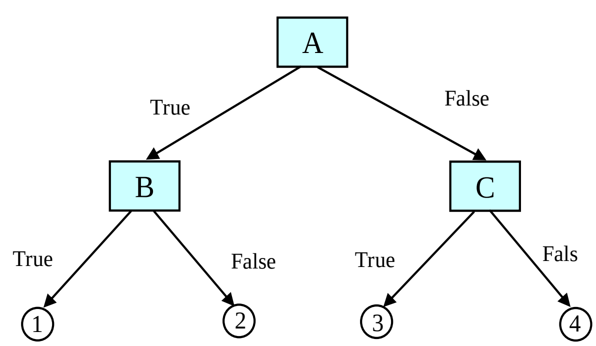

Figure 1. Structure of a Simple Decision Tree

In Figure 1, *A*, *B* and *C* represent test attributes, 1, 2, 3, 4 respectively express leaf nodes. Clearly, classification rules can be got as follows:

IF A=True and B= True THEN 1.

IF A=True and B= False THEN 2.

IF A=False and C= True THEN 3.

IF A=False and C= False THEN 4.

This shows that the decision tree may be used as a classifier, and also may be seen as a classification function, and it is very easy to generate decision rules from it.

2.1 Decision Tree Induction

The basic algorithm on decision tree induction is heuristic algorithm, by which the decision tree is recursively constructed from the top to the bottom of the tree. The algorithm is also an edition of ID3, based on disciplining samples, whose basic process is expressed as follows:

1) Establishing a root node N.

2) If all samples belong to a class, then N is seen as leaf node.

3) Otherwise, the algorithm applies entropy-based information gain as heuristic information to select the best attribute as class attribute, the attribute becomes a test attribute of the node N. In this algorithm, all attributes must be discrete, that is, discrete attribute values. Continuous values must be made discrete.

4) Based on every value of test attribute, a branch is constructed. In accordance with the branch, the samples are classified.

5) Like process, recursive decision trees are generated in each partition. Once an attribute occurs in a node, any offspring of the node may not be considered.

6) While satisfying one of the following conditions, recursive partition process stops.

(a) All samples are homogeneous.

(b) No residual attributes can be applied to further partition samples. Under the condition, multi-voting method is applied to yield class result.

(c) No sample is in one branch. In this time, leaf is constructed according to samples distribution in these classes.

The above process is a basic method for generation decision tree from the known training samples set. The followings presented are the selection or generation method of the test attributes.

2.2 Selection of Test Attributes

The selection of the best attribute is based on statistics characteristics of the samples. All the nodes in the tree apply information gain to weigh the choiceness of the selected test attributes. The method is usually called attributes measure or diverse choiceness scale. The attribute with the highest information gain (the largest entropy compression) is selected as test attribute, which makes the required information for samples partition results the fewest. The partition method can make the expected test attributes number down to the smallest, and ensure to find a simple tree.

Set S is a set with s data samples. Assume that data set yields m diverse classes, defined as C_i = (i=1, 2,…,m). order s_i is the samples class, the required expected information is given as follows:

$$I(s_1, s_2, ..., s_m) = -\sum_{i=1}^{m} P_i \log_2(P_i)$$

Where P_i is a probability that arbitrary sample belongs to class C_i, and estimated by s_i/s.

Set attribute A has v different values $\{a_1, a_2, \ldots, a_v\}$, based on A, S is divided into v subsets $\{S_1, S_2, \ldots, S_v\}$, where S_j comprises some samples in S, they have homogeneous value a_j in A. If A is selected as test attribute, then these subsets should correspond to the branches which are generated through node comprising set S Set s_{ij} is the samples number of class C_i in subsets S_j, the entropy of subsets generated by A is give out as follows:

$$E(A) = \sum_{j=1}^{v} \frac{s_{1j} + s_{2j} + ... + s_{mj}}{s} I(s_{1j} + s_{2j} + ... + s_{mj})$$

Set [$s_{1j}+s_{2j}+......+s_{mj}$] / s acts as weight of subject j, equals to the sum of the samples in subset (A=aj) divided by S. the smaller the value related to entropy is, the higher purity related to subsets partition is. For arbitrary given subsets Sj, we have

$$I(s_{1j}, s_{2j}, ..., s_{mj}) = -\sum_{i=1}^{m} P_{ij} \log_2(P_{ij})$$

Where $P_{ij}=s_{ij}/|s_j|$ is a probability, that is, samples S_j belong to class C_i .

The achieved information gain branched in A is

$$Gain(A) = I(s_1, s_2, ..., s_m) - E(A)$$

In other words, *Gain(A)* means a expected compression regarding entropy considering value of A. In this way, information gain of every attribute is worked out, and selecting an attribute with the largest information gain acts as test attribute in given set S. Based on this point, constructing a node and marking it, according to each attribute value, consequently, a branch is constructed

2.3 Distilling Classification Rules

In former Section 2 it has been expounded that a decision tree is a class function, it can be used as classifier and yield decision rules. The followings introduced are generation process on tree decision rules.

1) Yielding a rule for every route from root to leaf, the rule is given out by IF-THEN format.

2) The former part of rules is formed by attributes and values in given route, the latter part of rules is formed by leaf nodes, and is class prediction.

When the tree is very larger, IF-THEN rule is very effective and easy to understand. For one class, whose rules can be arranged in order according to their estimation precision As usual, a given sample can’t meet former part of any rules, default rules which appoint multitude classes are applied to dispose this status and add up to rules set of the result.

3. Fault Diagnosis Decision Table

A decision table must be established before using rough set for fault diagnosis of turbine generator unit. Seven typical frequency bands of the vibration signal based on the running frequency *f*, *i.e.*, (0.3-0.44)*f*, (0.45-0.6)*f*, *f*, 2*f*, 3*f*, 4*f*, $\geq 4f$, can be used to identify the machinery vibration fault [5]. The peak value of power spectrum in each band indicates a fault symptom at a particular frequency. A fault symptom attribute set is represented by C= {a, b, c, d, e, f, g }, V_{ci}= {0, 1}, i= 1,2,….,7. The decision attribute set is denoted $D = \{d\}$, $V_d = \{d_1, d_2, d_3, d_4\}$. In this paper, three typical vibration fault types are studied, oil-film oscillation *d*1, unbalance *d*2, asymmetry *d*3, and normal state *d*4.

After seven fault symptoms are normalized, the boundary values selected by experience are used to discretize the domains of the attributes into intervals according to different conditions [14].

If a, b, c $\in$ (0.35, 1], then a=1, b=1, c=1; otherwise equal to 0,

If d, e, f, g $\in$ (0.2, 1], then d=1, e=1, f=1, g=1; otherwise equal to 0.

Table 1 lists 300 cases of a turbine-generator unit vibration fault symptom after attribute discretization, where n represents the number of cases with the same attributes

Table 1. Fault Diagnosis Decision Table

U	n	Conditional Attributes							Decision
		a	b	c	d	e	f	g	D
1	10	1	1	1	0	0	1	0	1
2	20	1	1	1	0	0	0	0	1
3	10	0	1	1	0	0	0	0	1
4	30	1	1	0	1	1	1	1	1
5	10	1	1	0	0	1	0	0	1
6	10	1	1	0	0	0	0	0	1
7	10	0	1	0	0	0	0	0	1
8	10	0	0	1	1	1	1	0	2
9	10	0	0	1	0	1	0	0	2
10	10	0	0	1	1	1	0	1	2
11	10	0	0	1	0	1	1	0	2
12	10	0	0	1	0	0	0	1	2
13	10	1	1	1	0	0	1	0	2
14	20	1	1	1	0	0	0	0	2
15	10	0	1	1	0	0	0	0	2
16	10	0	0	0	1	1	1	0	3
17	10	0	0	0	1	0	0	0	3
18	10	0	0	1	1	1	1	0	3
19	10	0	0	1	0	1	0	0	3
20	10	0	0	1	0	1	1	0	3
21	10	0	0	1	1	1	0	0	3
22	30	1	1	0	1	1	1	0	3
23	10	1	1	0	0	1	0	0	3
24	10	0	0	0	0	0	0	0	4

4. Decision Tree Generation

According to equation (1) the overall expected information is

$$I(s_1, s_2, s_3, s_4) = -\sum_{i=1}^{4} P_i \log_2(P_i) = -\left(\left(100/300\right)\log 2\left(100/300\right) + ... + \left(10/300\right)\log 2\left(10/300\right)\right) = 1.7413$$

Based on (2) and (3), we can calculate

E(a) =1.5811, E(b) =1.5213, E(c) =1.4015, E(d) =1.5877, E(e) =1.5315, E(f) =1.6769, E(g) =1.6730

In accordance with (4)

Gain(a) =I-E(a) =1.7413-1.5811=0.1602, Gain(b) =I-E(b) =1.7413-1.5213=0.22,

Gain(c) =I-E(c) =1.7413-1.4015=0.3398, Gain(d) =I-E(d) =1.7413-1.5877=0.1536.

Gain(e) =I-E(e) =1.7413-1.5811=0.2098, Gain(f) =I-E(f) =1.7413-1.6769=0.0644,

Gain(g) =I-E(g) =1.7413-1.6730=0.0683.

Clearly, Gain(c)>Gain(b)>Gain(e)>Gain(a)> Gain(d)>Gain(g)>Gain(f). Therefore, we select c as test attribute. According to the attribute value v=0, 1 of a divides S into two incompatible subsets. According to Section 2.1, in each subset applying the above method respectively, unceasingly branching, in the end, a total tree is gained shown in Figure 2.

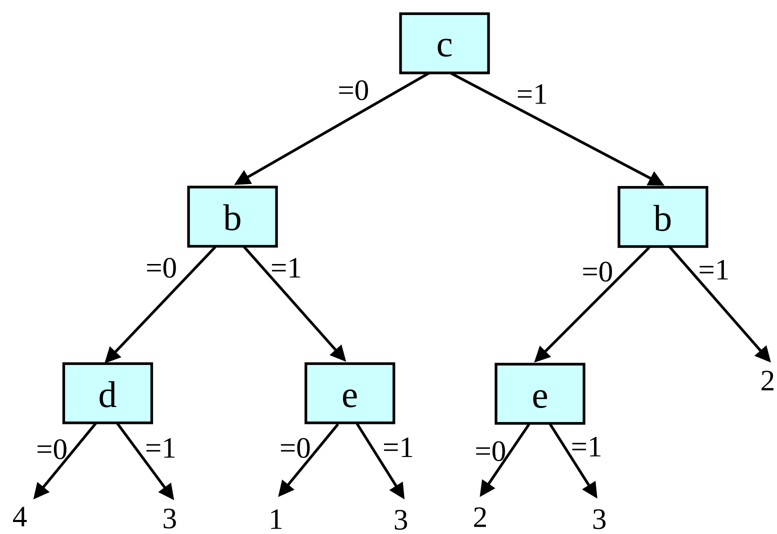

Figure 2. Decision Tree of Steam Turbine-generator Fault Diagnosis

5. Conclusion

Decision tree based steam turbine-generator fault diagnosis is a more flexible diagnosis method. It employs history data base or prior information of steam turbine-generator fault Diagnosis knowledge to establish a precise decision tree in terms of probabilistic statistic significance. In accordance with reasoning rules of the decision tree, the most important attributes are matched in the first place, and the attribute without any relationships to reasoning needn't compare, which consequently improves robustness of reasoning, at the same time, response time also is greatly reduced. In addition, for each classification rule, the best attribute is retained/selected based on information entropy, and this is very advantageous to data compression of large-scale data base. It may be seen that the decision tree can effectively mine latent knowledge from data sets and can discover law, and is an ideal data miner.

References

[1] P. Yang, J. Wu and Y. Feng, "Fuzzy Fault Diagnosis System for a 200MW Turbo-generator set", Automation of Electric Power Systems, vol. 25, no. 10, **(2001)**, pp. 45-49.

[2] W. Huang, X. Zhao, W. Wang and L. Dai, "Rough Set Model for Vibration Fault Diagnosis of Steam Turbine-generator Set", Automation of Electric Power Systems, vol. 28, no. 15, **(2004)**, pp. 80-84.

[3] Q. Guo and H. Wu, "A Novel Approach for Fault Diagnosis of Steam Turbine Based on Rough Set Data Mining Theory", Modern Electronic Technique, vol. 13, **(2006)**, pp. 074-078.

[4] R. Eslamloueyan, "Designing a hierarchical neural network based on fuzzy clustering for fault diagnosis of the Tennessee–Eastman process", Applied Soft Comput., vol. 11, **(2011)**, pp. 1407-1415.

[5] H. Li, C. X. Sun, R. J. Liao, W. G. Chen and X. S. Hu, "Improved BP Algorithm in Vibration Failure Diagnosis of Steam Turbine-Generator Set", Journal of Chongqing University, vol. 22, no. 5**, (1999).**

[6] S. M. Islam, T. Wu and G. Ledwich, "A Novel Fuzzy Logic Approach to Transformer Fault Diagnosis", IEEE Trans. on Dielectrics and Electrical Insulation, vol. 7, no. 2, **(2000)** April, pp. 177-186.

[7] C. Z. Wu, H. Yan and J. F. Ma, "Method Research of Noise Diagnosis Based on Fuzzy Neural Network", Proc. of 4th Int. Conf. on Signal Processing, Beijing, China, **(1998)** October 12–16, pp. 1370–1373.

[8] H. -J. Lee, B. -S. Ahn and Y. -M. Park, "A Fault Diagnosis Expert System for Distribution Substations", IEEE Trans. on Power System, vol. 15, no. 1, **(2000)** January, pp. 92-974.

[9] S. Umemara, *et al.*, "A New Health Monitoring System for Rotating Machinery", Mitsybishi Juke Giho, vol. 18, no. 6, **(1981)**, pp. 71-78.

[10] J. S. Mitcheell, "Machinery Analysis and Monitoring", Pennwell Publishing Company, **(1981).**

[11] H. Li, C. X. Sun, X. S. Hu, G. Yue, N. F. Tang and K. Wang, "Study of Method on Adaptive Wavelets for Vibration Fault Diagnosis of Steam Turbine-Generator Set", Journal of Electrical Engineering (China), vol. I5, no. 3, **(2000).**

[12] H. Li, C. X. Sun, X. S. Hu, G. Yue and K. Wang, "The Fuzzy Inputting and Outputting Method in Vibration Fault Diagnosis of Steam Turbine-Generator Set", Journal of Chongqing University, vol. 22, no. 6, **(1999).**

[13] H. Szu and T. Brian, "Neural Network Adaptive Wavelets for Signal Representation and Classification", Optical Engineering, **(1992)**, pp. 1907- 1916.

[14] W. Huang, X. Zhao, W. Wang and L. Dai, "Rough Set Model for Vibration Fault Diagnosis of Steam Turbine-Generator Set", Automation of Electric Power Systems, vol. 28, no. 15, **(2004)**, pp. 80-84.