INSULATION CONDITION EVALUATION TECHNIQUES OF CABLE BASED ON DECISION TREE

Zhang Tieyan Wang Chengmin and Jiang Chuanwen
Department of Electrical Engineering, Shenyang Institute of Engineering, Shenyang, 110006 China
Department of Electrical EngineeringShanghai Jiaotong UniversityShanghai, Huashan Road, No. 1954 China

Abstract: - The insulation conditions can be classed as batter, worse, failure etc. It is a important topic to diagnose the cable insulation condition based on corresponding daily repair data, test data and on-line monitoring data. The decision tree is used to determine the cable insulation condition while the subtree of every data is first performed and the subtree combination techonique is applied in this article. The case of a working cable is made at SPSS software and the fact are indicated that decision tree is a very effective classification method for insulation condition evaluation of cable.

Key-Words: - Decision tree; Classification; Data mining; Cable; Insulation condition;

1 Introduction

It is important to well and truly evaluate the insulation condision related with many factors for cable maintenance, operation and life management.

The classification decision model is established and the classification rule is performed based on large numbers of accumulated history data about cable maintenance and operation, thereby the insulation condition and failure can be judged by newly data acquisition.

The decision tree and decision rule are the tools to solve data classification problem in data minning. It is a process to classification by mapping the date iterms into a defined class. The classification can be gained by inductive learning algorithms with the classes and attribute vectors, i.e. the classification is the process, that discrete identity values(class) are distributed to unlabeled recorders. The classifier is the model(classification result) to forecast an attribute(class) under others are specified for the Test samples.

The more thechoniques can be regarded to classify, such as decision tree induce, Na ve-Bayes and neutral network etc. The decision tree[1] is adopted to classify the cable data for that cable insulation condition can be monitored in this article.

The decision tree which similar to flow chart is a tree structrue, where the inner node denotes the a attribute Test while the branch denotes the Test output and every leaf node denotes class or class distributing. The root node is located at topmost of a tree. The decision tree can be easily transformed to classification rule by Test the attribute at the decision tree with path from root node to the leaf node of tested sample for the classification of unknowned samples. The decision tree is generally used to the discrete and continuous attrbutes for establishing the forecast model with the simpleness and intuitions advantages compared with other data mining methods.

The decision tree is widely employed in power system, medical industry, finance and retail trade. The obvious advantages are exist ing in decision tree for the classification of incapable quantification attributes. The CART, CHAID and C4.5 are main arithmetics[2~7] in decision tree technique.

2 Insulation Condition Classification

The insulation condition can be classed as better, middling, worse and failure for on-line and outage cable as table 1.

There are three classes in other literatures[8], that number 2 and 3 in above table are incorporated to one.

3 Test Samples

The information reflected insulation condition comes mostly from cable Test data, examining and repairing data, on-line monitoring data. In this article, the assumptions are adopted, that all data are in same time clock and are checked due to delete the isolated point and vacancy value data. The date iterms are listed from table 2 to 4.

Tab. 1 Insulation Condition Classification

Number	Insulation Condition	Adopted Actions
1	better	To continue service.
2	middling	Be caution, to work under watch.
3	worse	The maintenance and repair are needed.
4	failure	To replace the cable.

Tab. 2 Daily Examining and Repairing Data

Number	Cable Types	Failure Types	Failure Degree	Repairing Effect	Service Time (Year)	Insulation Condition
1	Plastic Cable	Short Circuit	Worse	Middling	12.4	Worse
2	Oil and Paper Insulation Cable	Rupture	Middling	Better	15	Middling
3	Rubber Cable	Be Affected with Damp	Slight	Middling	16.7	Middling
4	…	Insulation Damage	Slight	Worse	17.8	Better
…		…	…	…	…	…

Tab. 3 Daily Test Data

Number	$\tan\delta$	Partial Discharge	…	AC Voltage Withstanding	Insulation Condition
1	<0.2%	Worse	…	12.4	Worse
2	>0.2%<0.5%	Middling	…	15	Middling
3	>0.5%	Slight	…	16.7	Middling
4	<0.2%	Slight	…	17.8	Better
…	…	…	…	…	…

Tab. 4 On-line Monitering Data

Number	$\tan\delta$	Leakage Current (μA)	…	Load (kA)	Environment Temperature (□)	Environment Humidity	Measure Time	Insulation Condition
1	<0.2%	>40<20	…	21	12.4	0.2	2000:3:20	Worse
2	>0.2%<0.5%	<20	…	12	15	0.4	2000:3:21	Middling
3	>0.5%	>40	…	23	16.7	0.3	2001:4:20	Middling
4	<0.2%	<5	…	31	17.8	0.6	2002:3:20	Better
…	…	…		…	…			…

The Daily examining and repairing data are listed in the table 2, here the historical failures are described. The failure degrees and repairing effect are classed as worse, middling and slight while failure types are various. The service time is counted before the failure. The insulation condition iterm is summarized by expert experiences in table 2.

The insulation condition evaluation criterion is certain because of less environment disturbances for cable test in laboratory. The data listed in table 2 and 3 can be intergrated due to the repaired effect may be proved by daily test data.

There are no criterion to judge cable insulation consition for on-line monitoring data, just the expert experiences and off-line test data can be used. The insulation condition criterion can be gained by choicing the test data as training samples, which are a great lot and sufficient though the test are not made.

4 Decision Tree Model

The sub-trees aiming at cable daily test data and on-line monitoring data are respectively established, and then they are integrated.

4.1 Sub-tree Model

The key problem is to choice the test attributes while the sub-trees are performed based on daily test data and on-line monitoring data. The minimum entropy is employed to choice the test atrributes as following:

$$\min \sum_{j=1}^{v} \frac{s_{1j} + \cdots + s_{mj}}{s} I(s_{1j}, \cdots, s_{mj}) \qquad (1)$$

where, the m is number of classes with value 3 or 4; The s is number of samples; The $j = 1, 2, \cdots, v$ denotes possible output value of chosen test attibutes if the sample set is marked as v subsets; The s_{ij} denotes sample number of class i in subset j; The $I(s_{1j}, \cdots, s_{mj})$ is information of the marking as:

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

where, the $p_{ij} = \frac{s_{ij}}{|S_j|}$ denotes probability of the samples in S_j, which belong to class i.

4.2 Sub-tree Integration

The sub-tree based on daily test data is showed is figure 1 while the sub-tree based on on-line monitoring data is showed in figure 2.

The same or similar attribute among two sub-trees is needed for integration, which is the basic request, otherwise two sub-trees are reciprocally independent tree.

There are no same attribute among the daily test data and on-line monitoring data, but the similar attrbute is in existence, for example, the service time attribute and measure time attribute are similar, so it is to change measure time attribute in on-line monitoring according to following formula:

Service Time= Measure Time – Working Moment (3)

therefore the measure time attribute can be replaced by service time attribute.

The association analysis is needed to choice the attribute with maximum comparability if there is no same or similar attribute in samples for sub-trees integration.

The approach can be also used in other sub-tree integration process, which is a method to simplify tree building. Firstly, the sub-trees are respectively performed by data classification in large numbers of samples. Secondly, the sub-trees are integrated as in figure 3.

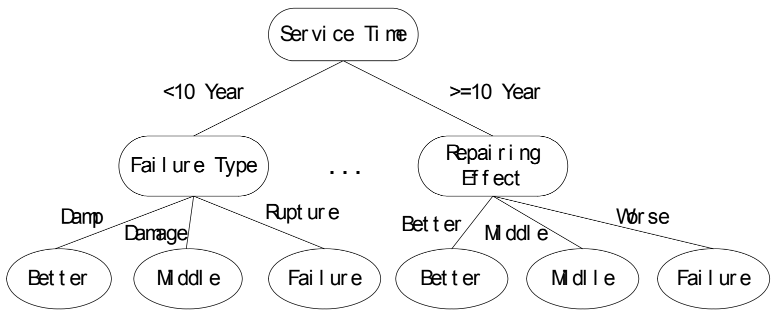

Fig 1 Decision tree for examine and repair data

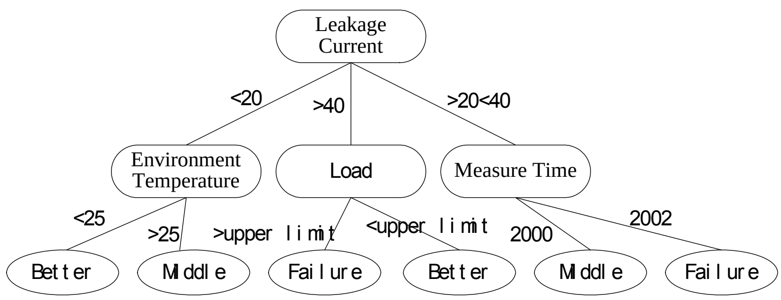

Fig 2 Decision tree for monitoring data

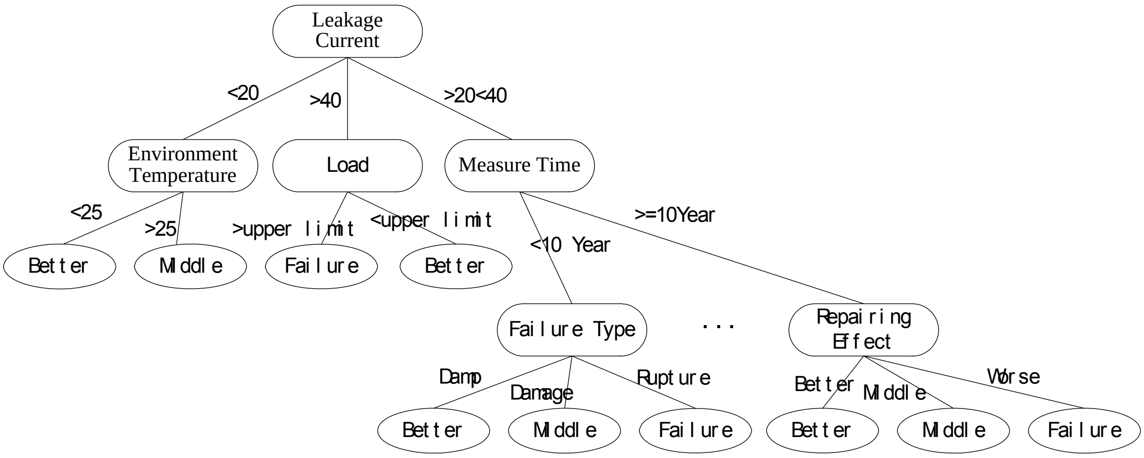

Fig 3 All decision tree

4.3 Classification Rules

The classification rules can be created by tree branch, which is similar to If-Then expression in C language due to be clear at a glance.

The classification rules can be regarded as the guide to judge cable failure and insulation condition and the knowledge to rightly use cable, avoid failure, in a word, it is propitious to cable life management.

5 Practices and Applications

There are great works to build management database in practice. The on-line monitoring data of a running cable are brought and the SPSS software is used to perform the decision tree with parameters $tg\delta$ and leakage current in this article. It is impossible that $tg\delta$ should observably change for single and centralized disadvantage, so the $tg\delta$ reflects universal shortcoming. The relation between above parameters and cable insulation condition can be gained by decision tree and collected data.

The leakage current and $tg\delta$ is considered as independent variables, and the cable insulation condition is considering as the result while decision tree is used to classification on SPSS software due to the classification rules are performed in "output" option.

It is to use SPSS software as showed in figure 4, and the results of final classification are listed in table 5.

The final three leafs as Node 2, Node 3 and Node 4 are determined by decision tree showed in figure 5, and it is indicated that leakage current and $tg\delta$ with value 98.78mA and 0.99% can be satisfied to test criterion. The parameter distribution is showed in figure 6.

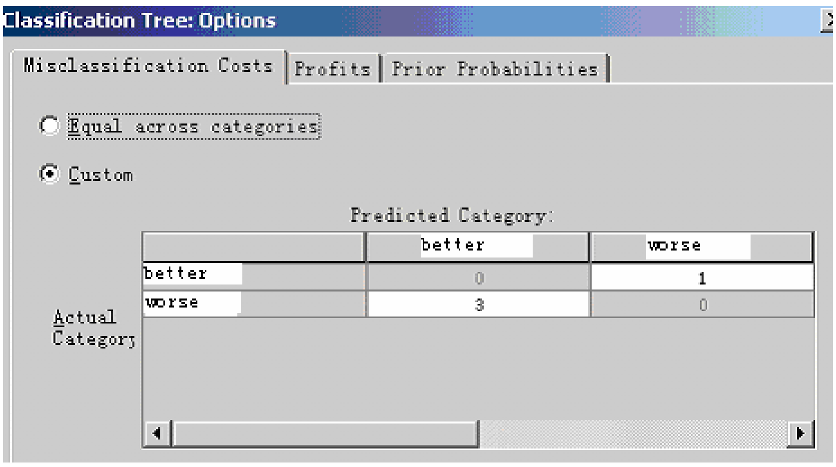

Fig. 4 Algorithm Correction

Tab. 5 Classification Results

Observed	Predicted		
	Better	Worse	Percent correct
Better	115	2	98.3%
Worse	0	283	100.0%
Overall Percentage	28.8%	71.3%	99.5%

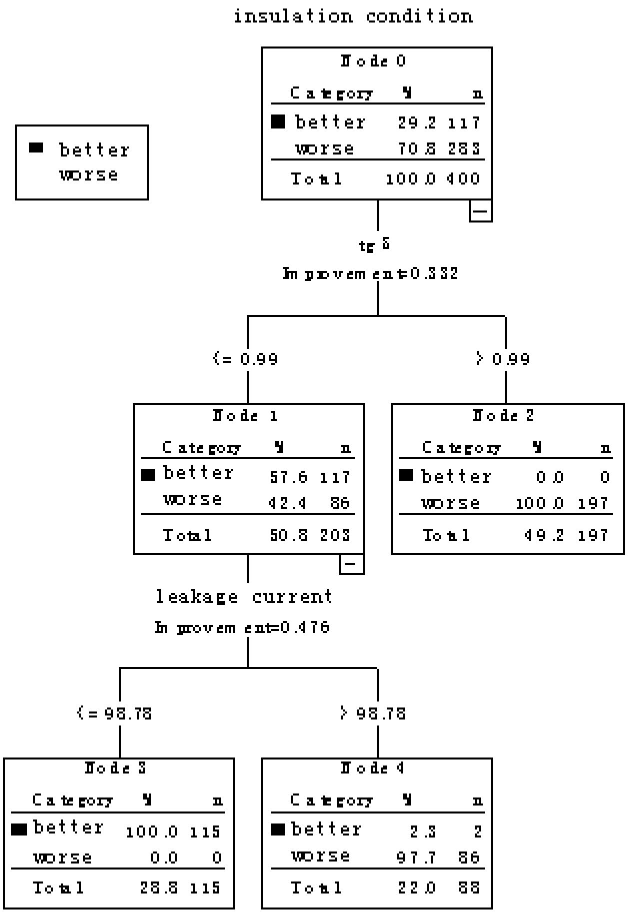

Fig. 5 The final Decision Tree

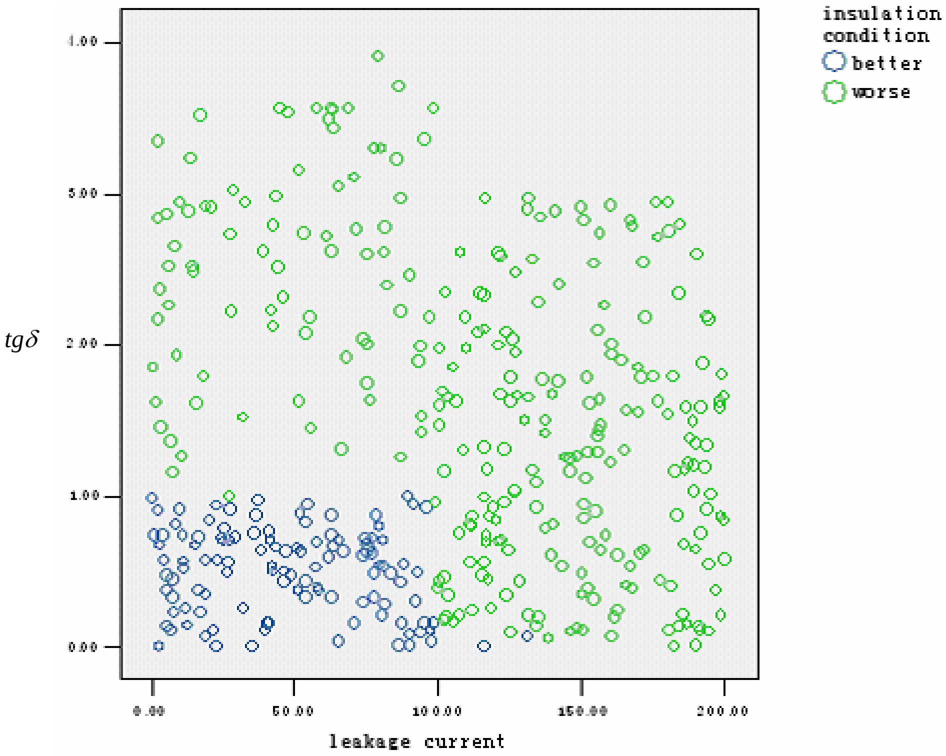

Fig 7 Insulation Condition Distribution

6 Conclusions

The decision tree is a typical classification technique, which is widely applied at each domain, but there is few reports in cable insulation condition monitoring. The better results and experiences are obtaind in this paper.

References:

1. Zhu Shaowen, Hu Hongyin, Wang Dequan, et al. Technique and Development of Decision Tree Mining, Computer Engineering, 2000Vol.26, No.10.
2. Quinlan J R. Induction of decision trees. MachineLearning, 1986,1: 81~106.
3. Michalski R S, Larson J B. Selection of the most representative training examples and incremental generation of VL1 hypotheses. Rept. No.78-867, Urbana-Champaign: Department of Computer Science, University of Illinois, 1978.
4. Cestnik B, Kononenko I, Bratko I. ASSISTANT 86: a knowledge elicitation tool for sophisticated users. In: Proceedings of EWSL-87, Bled, Yugoslavia, 1987.31~45.
5. Pagallo G, Haussler D. Boolean feature discovery in empirical learning. Machine Learning, 1990, 5:71~99.
6. Brodley C E, Utgoff P E. Multivariate decision trees. Machine Learning, 1995, 19:45~77.
7. liu Xiaohu, Li Sheng. An Optimized Algorithm of Decision Tree, Journal of Software, 1998, Vol.9, No. 10:797~800.
8. S. Kitai, K. Tokumaru, K. Hirotsu. Development of a Hot-Line Diagnostic method for XLPE Cables and the Measurement Results. IEEE Transactions on Power Delivery, Vol. 4, No. 2, 1989: 857~863.