The Research Of Elevator Fault Diagnosis Method Based On Decision Tree Algorithm

Chang Liu[1], Xinzheng Zhang[1]*, Xindong Liu[1], Can Chen[2]

1. School of Electrical and Information Engineering, Jinan University, Zhuhai 519070, China

2. Guangdong Power Grid Company, Zhuhai Power Supply Bureau, Zhuhai Guangdong, China

* Corresponding author: tzhangxz@jnu.edu.cn

Abstract

Due to the complex structure of the elevator system and the complicated and diverse cause of the elevator operation faults, It is often difficult for maintenance personnel to comprehensively grasp the fault signs to determine the cause of the fault. Directing against this peculiarity, this paper proposes the elevator fault diagnosis method based on decision tree classification algorithm. We use decision tree classification algorithm to classify the fault diagnosis, and then we can establish a fault diagnosis model, which accords with the theory of elevator operation. Finally we use different kinds of fault data samples to test the method. Experimental results demonstrate the higher accuracy of our model. This paper provides a good assistant method by means of Fault diagnosis of elevator system at present.

1 Introduction

The elevator is a kind of conveyance that is widely applied in the building perpendicular transportation. Until the end of 2014, there have been 3.5 million elevators in China. Meanwhile, the fault diagnosis of elevator cannot be ignore.

It is difficult to diagnostic the fault of elevator with the complex diagnosis objects, limited testing methods and the inaccurate knowledge. At present, most of the literatures are about the fault diagnosis of simple system, but there are few reports about the fault diagnosis of elevator system at home and abroad.

Decision tree algorithm is one of the most commonly used methods of machine learning. It can find valuable one in a large number of information and help people to make the right classification[1]. It attempts to learn a mathematical model to describe a sample set, in which the samples are disorderly and unsystematic. During the operation of elevator, the monitoring data is generated continuously. Therefore, the fault diagnosis cannot be effectively analyzed by the common diagnosis method, while the decision tree can find out the internal relation between the data through a large number of data. So as to get the decision tree which accords with the operation mechanism of the system

On the above analysis, this paper proposed a fault diagnosis method of elevator based on decision tree. We research the elevator operation mechanism, signal monitoring and common fault, and then select the fault feature signal and fault signal of elevator. The decision tree in this paper is constructed by using MATLAB. The improve method is verified by an actual case of the Guangzhou subway company chebei station Thyssen elevator.

2 Decision tree classification algorithm

Decision tree classification algorithm is one of the important technology in data mining. It first select the optimal node as the root node by finding the most informative attributes of the sample data. Then it depends on different values of each node to determine the next node of the tree branch. So the cycle continues, we can from the root down and finally get the decision tree, whose construction like flow chart[2]. At present, there are two kinds of general decision tree classification algorithms: ID3 and C4.5. The C4.5 algorithm[3] is proposed by Quinlan in 1993 for the shortcomings of the ID3 algorithm, For instance, ID3 algorithm exist noise effects and discrete attributes. It uses information gain ration to replace the information gain to select each branch attributes. Compare with the characteristics of ID3 algorithm that easy to choose more attributes when selecting attributes by information gain, C4.5 algorithm has been optimized in decision efficiency and operating speed. In addition, C4.5 uses post-pruning to avoid the over-fitting and cross validation of tree. In this paper we use C4.5 algorithm.

2.1 Decision Tree Generation Algorithm

Before constructing decision tree, sample data need to be preprocessed. Discretization of continuous attribute' value is necessary if any attributes contain continuous value. Firstly, we sort the value of attribute A that contains continuous value according to the original data: a(1), a(2) … a(n). We make the average of two adjacent attribute value as split point of continuous attribute value temporarily. Then C4.5 algorithm calculate the information gain ration for each split point. The information gain ration of split point (a(1)+a(2))/2 is denotes as GainRatio(a). We select the split point with the maximum

information gain ration as the split point of the value of attribute A. The calculation formula of information gain and information gain ration:

The information gain of attribute A:

$$\text{Gain}(A) = I(A) - E(A) \qquad (1)$$

$$I(B) = -\sum_{i=1}^{m} P(B_i) \log_2 P(B_i) \qquad (2)$$

$$E(A) = -\sum_{j=1}^{n} P(A_j) \sum_{i=1}^{m} P(B_i|A_j) \log_2 P(B_i|A_j) \qquad (3)$$

$P(B_i)$ is the proportion of sample that belong to B_i in the total sample. $P(A_i)$ is the proportion of sample that belong to A_i in the total sample. A is the parent of B, $P(B_i|A_j)$ is the proportion of sample that belong to B_i in in the sample that belongs to A_i.

The information gain ration of node A:

$$\text{GainRatio}(A) = \frac{Gain(A)}{I(A)} \qquad (4)$$

C4.5 algorithm calculates the information gain ration of all attributes, and then chooses the attribute which has maximum information gain ration as the root node of the decision tree. Classify the attribute value of root node by the discrete method of attribute values. And then select the new root node of the subsample in the same way, until the sample number is 0.

2.2 Decision Tree Pruning Algorithm

After the decision tree finished, there might exist overfitting problem. C4.5 uses the pruning algorithm to remove noise and abnormal branching. The pruning algorithm is based on classification error rate. It the calculates the error rate before and after pruning according to the initial sample data, to decide whether to prune nodes[4].

It assume that we form the initial decision tree T. The sample number of some leaf node is n(t), The number of classification error is e(t). The classification error rate of the leaf node is r'(t).

$$r'(t) = [e(t) + 1/2]/n(t) \qquad (5)$$

T_t is subtree of T. The number of leaf node of T_t is $L(T_t)$, the classification error rate of T_t is $r'(T_t)$.

$$r'(T_t) = \sum_{T_t} \frac{[e(T_t) + 1/2]}{\sum_{T_t} n(T_t)} = \frac{\sum e(T_t) + \frac{L(T_t)}{2}}{\sum n(T_t)} \qquad (6)$$

In the quantitative analysis, this paper takes the number of total error instead of the error rate for simplicity.

$$e'(t) = e(t) + \frac{1}{2} \qquad (7)$$

$$e'(Tt) = \sum e(Tt) + \frac{L(s)}{2} \qquad (8)$$

In the PEP method of decision tree pruning. When the classification error rate of subtree T_t bigger than the classification error rate of node t after cutting into a leaf node, the pruning is proper. This node t is given the maximum attribute value.

$$e'(t) \le e'(Tt) + SE[e'(Tt)] \qquad (9)$$

$SE[e'(Tt)$ is standard error.

$$SE[e'(Tt)] = \sqrt{\frac{e'(Tt)[n(t) - e'(Tt)]}{n(t)}} \qquad (10)$$

Test the initial tree's all none-leaf nodes from root node to leaf nodes. Finally we can get the optimal decision tree.

After forming the finally decision tree, we can get the IF…THEN classification rules. And then new data sets can be classified and tested.

3 Example of Thyssen TE-evolution1 machine roomless elevator fault diagnosis model based on Decision Tree

3.1 The Elevator Fault Diagnosis Model based on Decision Tree.

The traction elevator is a kind of elevator that is widely applied in the building perpendicular transportation. The basic structure is shown in the figure 2[5]. Traction machine is the power source of elevator, it contains electric motor, reduction gearbox, brake and traction sheave and so on. Both ends of the Traction wire rope are connected with Car and Counterweight (or fixed in the Computer room), depending on the friction between Traction rope and Rope groove to drive the Car. The Guide wheel is separated from the Car and Counterweight spacing, using complex winding type can also increase tractive power. The Guide system is used to limit the movement freedom of Car and Counterweight, to make sure the Car moves up and down along the Guide rail only. The Guide rail is fixed on the Rail frame, the Guide rail frame bearing Guide rail and connected with well wall. The Guide shoe is arranged on the Car and the Counterweight frame, and is matched with the Guide rail. The Door system is mainly used to control the elevator door. Weight compensation device is used to compensate the effect that Car and the change of Counterweight side rope length to elevator balance. The Electric power drag system is used to control the real-time speed of elevator. Speed feedback device is implemented for speed regulation control. The Electric control system is used to control the operation of the elevator in real time. The Safety protection system includes all kinds of mechanical and electrical protection systems to protect the safety of the elevator[6].

It can be clearly seen from the above analysis that as a complex engineering system of mechanical and electrical equipment, there might occur many kinds of faults. Due to the fault symptom and fault cause is not one-to-one correspondence, but cross influence, the cause of the faults is more complicated. For example, if there exist noisy or

vibration in elevator operation[7], may be due to the traction steel wire rope tension uneven, may also be the wear of spring compression or both. On the other hand, if a component fails, it may cause multiple operational faults. It is found that there are many measurable state data in the process of elevator operation. To a certain extent, the data which is continuously changing reflects the process of elevator operation from no fault to fault. In this paper, these continuous state data are called the fault symptom.

In this section, the decision tree classification method is used to classify the fault symptom of the elevator, and establishes the fault diagnosis decision tree. So as to help maintenance personnel to fully grasp the causes of faults.
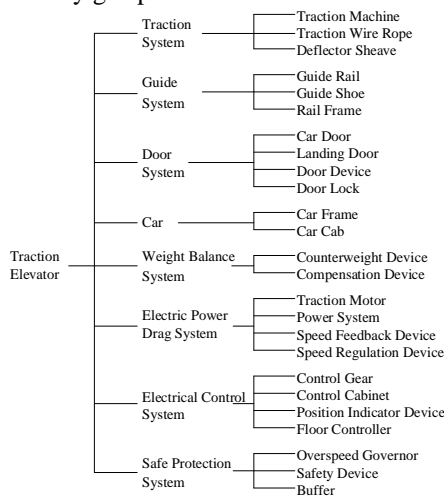

Figure 1: A system of Tractive Elevator

There are many fault symptoms and fault causes in elevator fault diagnosis system. Constructing decision tree directly will make the decision tree scale too large, and it will also affect the inference efficiency of decision tree. Therefore in this paper, the elevator is divide into several area, which based on the main operation system of elevator. We construct the corresponding fault diagnosis decision tree for each area. Then integrating the fault diagnosis results of each subsystem. Finally we can diagnose the fault of whole elevator system. This paper chooses the traction system as an example, screening all observable fault symptoms and fault causes[8,9], and establishing the elevator fault diagnosis model based on decision tree. The fault symptoms of Traction System listed in Table 1. The fault causes of Traction System listed in Table 2.

number	fault symptoms
X1	Elevator noise and Vibration
X2	The elevator is difficult to start or the running speed is low
X3	The host of Traction system is fever
X4	Fail of releasing electro-hydraulic shoe brake
X5	The wire rope and the traction wheel slip
X6	Double brake arm asynchronous when the gate opening and closing
X7	Braking force shortage after closing the gate

Table 1: The fault symptoms of Traction System

number	fault causes
Y1	The wear of motor bearing
Y2	The wear of stator and the rotor
Y3	Bearing wear of reduction gearbox
Y4	The oil starvation in Traction motor lubricating oil chamber
Y5	The gap is abnormal on both sides of electromagnetic brake
Y6	The electromagnetic coil of brake working abnormally
Y7	Wear of traction groove
Y8	Wear of Traction wire rope

Table 2: The fault causes of Traction System

Because there is no clear relationship between the various types of fault symptoms and fault causes, and the operation of elevator is a gradual process from run without fault to run with fault. So the number of fault symptoms collected is abnormal percentage. Fault symptoms are collected by mounting sensors. We collect 1000 sets of actual operating data according to the Thyssen elevator in Guangzhou Subway Company Chebei Station. The sample sets listed in Table 3. In which the Y is the fault cause, it has 9 attributes, Y1-Y8 are fault attributes, and fault-free attribute denote as No.

	fault symptoms						fault causes
	X1	X2	X3	…	X6	X7	Y
1	0.15	0.25	0.31	…	0.06	0.10	No
2	0.23	0.41	0.53	…	0.07	0.12	Y4
3	0.13	0.02	0.34	…	0.28	0.04	Y5
4	0.45	0.26	0.51	…	0.01	0.25	Y3
5	0.08	0.37	0.01	…	0.14	0.47	Y6
6	0.39	0.5	0.53	…	0.03	0.01	Y1
7	0.06	0.18	0.20	…	0.12	0.04	No
…	…	…	…	…	…	…	…
998	0.41	0.49	0.21	…	0.15	0.04	Y2
999	0.64	0.20	0.01	…	0.14	0.01	Y8
1000	0.04	0.12	0.09	…	0.02	0.02	No

Table 3: The fault sample set of Traction System

We input the samples into the decision tree classification program, and its software is programmed with MATLAB. Generating the initial decision tree according to the principle of maximum gain rate, and then pruning the initial decision tree, finally establishing the decision tree model of the traction system. As shown in Figure 2.
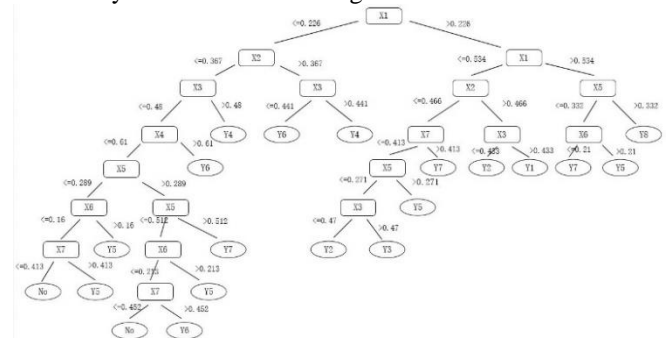

Figure 2: The decision tree of traction system

From the decision tree shown in Figure 2, we can get 20 IF-THEN rules as follows:
Rule No.1: IF (X1<=0.226 and X2<=0.367 and X3<=0.48 and X4<=0.61 and X5<=0.289 and X6<=0.16 and X7 <=0.413) THEN Y=No

Rule No.2: IF (X1<=0.226 and X2<=0.367 and X3>0.48) THEN Y=Y4

Rule No.3: IF (X1<=0.226 and X2<=0.367 and X3<=0.48 and X4>0.61) THEN Y=Y6

……

Rule No.20: IF (X1>0.534 and X5>0.332) THEN Y=Y8

3.2 Accuracy verification

In this section, after constructing the decision tree based on 1000 sets of elevator operation data, we collect 1000 sets of historical fault data again to verify the accuracy of decision tree. The historical fault data we collected contain fault types at different times. This paper uses the following data for quantitative analysis. The number of testing data is M, The number of testing data is M, N is the number of data for which the predicted results are in line with the actual situation. Diagnosis accuracy denote as P, P=N/M. The diagnosis results of the decision tree are shown in Table 4. The results of the test through calculations show the accuracy of elevator fault diagnosis model based on decision tree reached 76.3%, proving that the algorithm has high accuracy and practical value.

	fault symptoms					fault causes	diagnosis results
	X1	X2	X3	…	X7	Y	Y'
1	0.07	0.23	0.06	…	0.41	Y6	Y6
2	0.36	0.44	0.19	…	0.11	Y2	Y2
3	0.01	0.05	0.21	…	0.14	No	No
4	0.21	0.40	0.62	…	0.01	Y4	Y4
5	0.02	0.33	0.01	…	0.42	Y6	Y6
6	0.71	0.12	0.04	…	0.11	Y8	Y7
7	0.28	0.26	0.43	…	0.04	Y1	Y2
…	…	…	…	…	…	…	…
998	0.44	0.21	0.40	…	0.16	Y2	Y3
999	0.01	0.21	0.13	…	0.15	No	Y6
1000	0.07	0.10	0.14	…	0.26	No	No

Table 4: Comparison of decision tree diagnosis results

4 Conclusion

This paper mainly researches the elevator fault diagnosis method based on decision tree, through the analysis of elevator structure and the selection of fault symptoms, the huge amounts of elevator fault data into a clear fault diagnosis decision tree, compared to other complex fault diagnosis method, the elevator fault diagnosis model based on decision tree makes full use of the actual operation data of the elevator, and it more in line with the elevator operation mechanism. Furthermore, this method is intuitive and clear and has been proved that it can achieve high accuracy by the existing historical data. Therefore, the C4.5 decision tree classification method is very effective in elevator fault diagnosis, and has broad application prospects.

Acknowledgements

References

[1] Liu Tongming. "Data Mining Techniques and Its Applications", *National Defense Industry Press*. (2001)

[2] Liu Jiqing, Huang Jinhua. "Intelligent fault diagnosis method for equipment based on improved decision tree algorithm", *Manufacturing Automation*. 33(4), pp.30-33(2011)

[3] Quinlan J. "C4.5 Programs for Machine Learning", *Morgan Kauffman*. (1993)

[4] MANSOUR Y. "Pessimistic decision tree pruning based on tree size", *Machine Learning: Proceedings of the Fourteenth International Conference*. pp.195-201(1997)

[5] Yang Yang, Zhao Guojun, Liu Zheng. "Expert Systems for Elevator Controller Fault Diagnosis", *Techniques of Automation & Applications*. 21(5), pp.40-42(2002)

[6] Huo Rudong. "On the Elevator Machinery and Mechanical Structure", *Equipment Manufacturing Technology*. (07), pp.178-179(2010)

[7] Li Pin. "Vibration& Noise in Elevator Running", *Noise and Vibration Control*. (05):38-40(2004)

[8] Lin Jishu. "Frictional wear of Elevator's Tractive System", *Journal of Chengdu Textile College*. 24(04), pp.12-14(2007)

[9] Zhao Hongxing. "Analysis and design of the elevator system", *Science & Technology Information*. (04), pp.117(2010)