Application of Association Rules in Analysis of Pavement Performance Attenuation Factor

Mingxin Nie[a], Yanlong Li[b]

School of Information Engineering, Wuhan University of Technology

Key Laboratory of Fiber Optic Sensing Technology and Information Processing, Ministry of Education, Wuhan, China

[a]624968304@qq.com, [b]1399477457@qq.com

Keywords: performance attenuation, association rules, data mining, factors.

Abstract: With the continuous growth of highway traffic mileage, the maintenance of highway has become increasingly important, especially the pavement performance attenuation has become the main disease influencing highway environment. In order to explore the factors of pavement performance attenuation, this paper used the multidimensional association rules of data mining technology to analyzed the interaction among various performance indexes like Pavement Condition Index, Riding Quality Index and so on, based on the technical data of pavement and traffic data provided by Highway Bureau of Hubei province; on other hand, we also analyzed the influence of traffic flow about the pavement performance through association rules technology. The result of association rules mining will be beneficial to preventive maintenance of highway.

1. Introduction

The pavement performances of highway response the technical condition of the highway and is the basic guarantee for the safe and normal driving of transportation vehicles. Therefore, studying the influencing factors of highway pavement performance attenuation and mastering the law of highway pavement performance fluctuation have important practical significance for highway preventive maintenance [1]. Taking current situation of Chinese preventive maintenance of highway into consideration, there are few studies on the factors of pavement performance attenuation. On the one hand, the research mainly focuses on the structure and material that require extremely professional knowledge and take a long time, but it is hard to have great breakthrough in short time [2]. On the other hand, the development of information technology has brought a new way for us to study the attenuation factors of highway pavement performance, especially the rapid development of data mining in recent years [3].

Multidimensional association rules in data mining are more suitable for the analysis of factors of highway pavement performance attenuation than traditional mathematical statistics methods. The concept of association rules was first proposed by R. Agrawal, then I. Bhandari used data mining technology to process NBA data and excavated valuable business intelligence successfully [4]. Kou Yu used association rules data mining to analyze customer data of telecommunications and discovered valuable information. This paper attempts to analyze the factors of highway pavement performance attenuation by using the data mining technology of multidimensional association rules[5].More concretely,this article will analyze the interrelationships among the internal index of pavement performance and external factors that the impact of heavy traffic to pavement performance attenuation,based on the technical status data of pavement performance and traffic flow from 2014—2016 provide by Highway Bureau of Hubei province.

2. Pavement performance evaluation and research of association rules mining

2.1 Pavement performance evaluation

Pavement performance evaluation is the judgement of highway to meet the requirement of

practical use, based on the collected data of pavement technical status and it can be divided into a single index evaluation and comprehensive evaluation. Single index evaluation is that evaluating highway status from every separate aspect of pavement performance, such as pavement damage evaluation, pavement driving quality evaluation, pavement rut depth evaluation, pavement skid resistance evaluation. The comprehensive evaluation of pavement performance is an overall evaluation of pavement performance for the sake of comparability of the status of each section. In general, we will refer to the comprehensive evaluation of pavement performance for short as pavement performance or pavement evaluation [6].

According to *The Evaluation Criteria for Highway Technical Status* issued by the Ministry of Transport of the People's Republic of China (hereinafter referred to as " *The Evaluation Criteria*"), the performance evaluation of asphalt pavement includes the following five aspects: pavement damage, roughness, rutting, skid resistance and structural strength. The pavement structure strength is the index of sampling evaluation, which is calculated and evaluated separately and the scope of evaluation is determined according to the requirements of pavement maintenance and geological conditions of roadbed; The performance evaluation of the cement concrete pavement includes three parts as follow: pavement damage, roughness and skid resistance; the performance evaluation of gravel pavement includes only a technical content of pavement damage. However, due to the factors that affect the performance of pavement are complicated, this paper will also make an analysis of the Subgrade Condition Index(SCI) and the Traffic-facility Condition Index (TCI) along the road. The current highway technical status evaluation system shown in Fig.1 [7], the range of each indicator is 0-100.

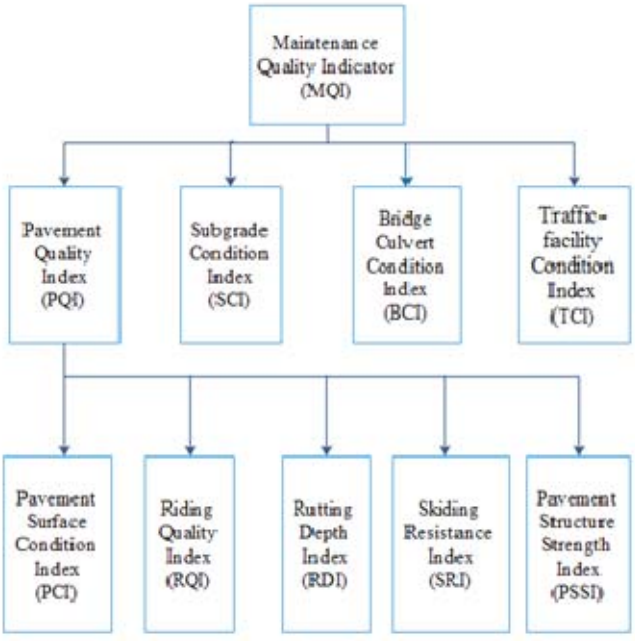

Fig.1 Highway technical status evaluation system

2.1.1 Pavement surface condition evaluation

Due to the differences in traffic flow and environment, different highways may have different types of damage in different extents and different scales. In order to compare the damage status quantitatively and synthesize various attributes of damages about each section, "*The Evaluation Criteria*" adopt the Pavement Surface Condition Index(called PCI for short) to evaluate pavement damage status and PCI was calculated by the Comprehensive Pavement Damage Rate (called DR for short).

DR can be calculated with (1) as follow:

$$DR = 100 \times \frac{\sum_{i=1}^{n} w_i A_i}{A}$$

where DR——the comprehensive pavement damage rate,the ratio of the equivalent area of various damages and the total highway area that surveyed (%);A_i——pavement damage area of category i ;A—— the total highway area that surveyed;w_i——highway damage weight parameter of category i,"The Evaluation Standards" for more details

The formula for calculating the PCI based on the DR of the road is as follows:

$$PCI = 100 - a_0 \times DR^{a_1}$$

where a_0——parameter of model, asphalt pavement uses 15.00, cement pavement uses 10.66,gravel pavement uses 10.10;a1——parameter of model, asphalt pavement uses 0.412, cement pavement uses 0.461,gravel pavement uses 0.487.

2.1.2 Pavement riding quality evaluation

The riding quality of the highway pavement is measured by the riding quality index (called RQI for short), which is calculated by the International Roughness Index(called IRI for short). The IRI can be measured by a professional reaction equipment. The relationship between ride quality index (RQI) and international roughness index (IRI) is (3):

$$RQI = \frac{100}{1 + a_0 e^{a_1 IRI}}$$

Where a_0——parameter of model, freeway and grade A highway use 0.026, the other roads use 0.0185;a_1——parameter of model, the highway and grade A highway using 0.65, the other roads use 0.58;

2.1.3 Pavement rutting depth evaluation

In recent years, due to the rapid increase of traffic flow, the problem of overloading are getting more serious every day,as a result that rutting has become a serious damage to highway surface in our country. In particular, the existence of rut on two-lane pavement shortens the service life of pavement seriously. Therefore, the detection and evaluation of the rutting depth of pavement is more and more important. According to the "T*he Evaluation Criteria",* we use the highway rutting depth index (called RDI for short) to measure the rutting technical status quantitatively, its formula is shown in (4).

$$RDI = \begin{cases} 100 - a_0 RD & (RD < RD_a) \\ 60 - a_1 (RD - RD_a) & (RD_a < RD < RD_b) \\ 0 & (RD > RD_b) \end{cases}$$

Where: RD——rutting depth, in millimeters (mm); RD_a——rutting depth limit, using 20mm; RD_b——rutting depth limit, using 35mm; a_0——parameter of model, using 2.0; a_1——parameter of model, using 4.0;

2.2 Association rules mining model

2.2.1 Definition of association rules

Association rules refer to the correlation between a large number of data item sets. Association rules's basic unit is transaction, and the transaction is composed of items, it is essentially the relationship between items and items. Generally, we use the term i_i, $I = \{i_1, i_2, ..., i_m\}$ represents the set of all items; T represents the transaction, and T is the set of some items; D represents the set of all the transactions, that is, the database.[8]

Let X and Y be a set of some items. An association rule is an implication of the form X => Y,

where X⊂I, Y⊂I, and X∩Y = Ø; X is called the premise or the left part, Y is called successor or right part. They are constrained by Support which represents the frequency of rules and Confidence which represents the confidence of rules.

The support of rule X => Y in database D is the ratio of transactions that contain both X and Y to all transactions in the transactions set, and it can be denoted as Support (X => Y) = P (X∪Y). Confidence of rule X => Y refers to the ratio of the number of transactions that contain both X and Y to the number of transactions that contain X, and it can be denoted as Confidence (X => Y)= P (Y | X).

In summary, the procedure of association rules mining can be divided into the following two sub-problems: 1) Find all items set with minimum support in transactions database D, they are called the maximum items set. 2) According to the maximum items set to generate the required association rules. For a maximum items set A, if the non-empty subset a of A satisfies the relationship Support (A) / Support (a)> = minconf(it is set by the situation). Then the association rule a => (A-a) is generated[9].

2.2.2 Analysis of apriori algorithm

Apriori algorithm is a traditional and effective method to obtain association rules, which uses the iterative method to find frequent items set. The Apriori algorithm takes advantage of the Apriori property — it means that all non-empty sub-sets of frequent items set must be frequent items set, and all the super-frequent items of non-frequent items set couldn't be frequent, Applying of this property can avoid blindly searching and improve efficiency of finding frequent items set.

The figure below illustrates the general flow of Apriori in multidimensional data mining.

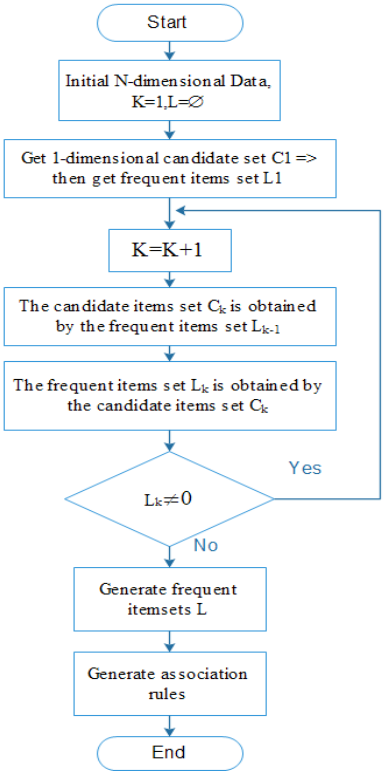

Fig.2 Flow diagram of Apriori

The basic process of Apriori algorithm is as follows [10]:

1) Finding all frequent 1-sets L1(it means the set contains only 1 item) according to the minimum support (called minS for short);

2) Starting the loop process until there is no maximum item set generated. The process of the loop is as follows: In step K, the k—dimensional candidate set C_k generated according to the (k-1)—dimensional frequent item set L_{k-1}, then geting the k—dimensional maximum item set L_k from C_k;

3) Generating association rules from the largest set of items.

3. Application of association rules in the attenuation of pavement performance

Based on the technical status data of pavement performance from 2014 to 2016 issued by Hubei Provincial Highway Bureau and the traffic flow data, the data model analyzed the internal correlation of pavement performance index and the data model analyzed the impact of traffic flow on the pavement performance index are established respectively, then mining the model using the association rule algorithm. However, since the performance indicators are continuous values and the association rules can not analyzed continuous values effectively, we discretized the performance indicators according to the criteria in Table 1.

Table 1 Performance indicators discretization standards

Performance index value	≥90	80~90	70~80	60~70	<60
grade	A	B	C	D	E

3.1 Internal impact of pavement performance indicators

Since ordinary highway are generally divided into cement pavement and asphalt pavement, the performance indexes of the two pavement are some differences. Asphalt pavement has five performance indicators and cement pavement has only three kinds of indicators. In order to having unified mining and analysis, we need to choose their common performance indicators for analysis. According to the data provided by Hubei Provincial Highway Bureau, we choose the following attributes of pavement performance indexes as attributes of internal influence data model: highway grade, PCI, RQI, RDI, Pavement Performance Composite Quality Index (PQI), Subgrade Status Index (SCI) and Technical Status Index of Facilities along the highway (TCI).

Using weka as the tool of association rule mining, we set up the minimum support with 10% and the minimum confidence with 60%, then use the internal data model of pavement surface performance index to mine the association rules,the results are shown in Table 2(inserted in next page).

Table 2 Pavement performance index internal influence association rules

Number	**Association Rules**	**Confidences**
1	technical level=B => RDI=E	1.00
2	RDI=E and SCI = E => TCI = E	1.00
3	PCI=E and RQI=E => PQI = E	0.99
4	PCI=A and RQI=A => PQI = A	0.97
5	RDI = E => RQI = E	0.96
6	RQI = D => PCI = D	0.83

At first, rule 1 shows that the probability of rut depth index(RDI) come to Grade E is to be about 100% from the current data if the highway technical level is Grade B, so we can know that the rutting depth index(RDI) has a strong correlation with the highway technical level; Secondly, rule 2 means that the probability of the technical condition index of the facilities along the highway being Grade E is 100% when the rutting depth index and the subgrade technical status index are both in Grade E, that implies the poor driving conditions of vehicles seriously affects the technical condition of the facilities along the highway and the technical conditions of the subgrade. Thirdly, rules 3 and 4 indicate that when the pavement damage index PCI and the driving quality index both are Grade E, the probability of the index PQI being Grade E is 99%. When the pavement damage index PCI and the driving quality index are both Grade A, the probability of PQI being Grade A is 97%. Therefore, it can be seen that PCI and PQI are two of the main factors affecting PQI; Fourthly, rule 5 shows that when the RDI is Grade E, there is a 96% probability that RQI is Grade E, we can see that rutting depth is the important impacting factor of leading to the attenuation ridding quality index; Fifthly, rule 6 shows that when the ridding quality index is E, the probability of pavement surface condition

index being E is about 83%,and because the ridding quality index is main measured by pavement roughness, we can infer that pavement roughness is an important factor leading to the attenuation of pavement condition index from rule 6.

3.2 Impact of traffic flow on pavement performance

The main function of highway is for the driving of vehicles. Therefore, traffic flow is the most direct and important external factors that affect the performance of highway pavement. According to the survey and analysis, it is found that the proportion of heavy vehicles in traffic flow has the greatest impact on the performance of pavement. Therefore, this paper mainly analyzes the impact of the proportion of heavy vehicles on the performance attenuation of pavement under different technical levels and different types of pavement. We choose the following attributes as the basis for the mining analysis: total traffic flow, proportion of heavy vehicles, PCI, RQI, RDI, PQI.

As above, weka is used as the mining tool of association rules and setting mining parameters minimum support with 10% and minimum confidence with 60%, we use technical data and traffic data provided by Hubei Provincial Highway Bureau to mine the association rules, the results are shown in Table 3.(inserted in the right column)

Rules 1, 2 and 3 indicate that the probability of PQI decay to Grade C is 99% when the proportion of heavy trucks in road traffic flow reaches 30%, The probability of the PQI level drops to Grade D is 99% when the proportion of heavy vehicles reaches 40% and the probability of the PQI level drops to Grade E is 97% when the proportion of heavy vehicles reaches 50%.

Table 3 Association rules about impact of traffic flow on pavement performance

Number	Association Rules	Confidences
1	Proportion of heavy vehicles = 30% => PQI = C	0.99
2	Proportion of heavy vehicles = 40% => PQI = D	0.99
3	Proportion of heavy vehicles = 50% => PQI = E	0.97
4	Proportion of heavy vehicles = 40% => PCI = D	0.96
5	Proportion of heavy vehicles = 50% => PCI = E	0.95
6	Proportion of heavy vehicles = 50% => RQI = D	0.90

From the first three rules we can know that the composite index PQI is negatively correlated with the proportion of heavy vehicles. Rules 4 and 5 indicate that when the proportion of heavily loaded vehicles reaches 40%, the probability of the pavement condition index PCI being Grade D is 0.96; when the proportion of heavy vehicles reaches 50%, the probability of PCI decreasing to Grade E is 95%. These two rules show that the proportion of heavy vehicles will also lead to more serious damage to the PCI. Rule 6 shows that when the proportion of heavy vehicles reaches 50% and the probability of RQI being Grade D is 90%, it can be seen that heavy vehicles is also an important factor that lead to the decrease of RQI.

4. Conclusions and prospects

This paper makes a preliminary mining of the highway maintenance technology status and traffic flow data provided by Hubei Provincial Highway Bureau using the association rules in data mining technology. The main analysis refer to the internal and external factors of highway pavement performance attenuation. The internal factor mainly include the interaction between the decay of performance index; the external factors mainly include the impact of the proportion of heavy vehicles on the attenuation of each index of pavement performance. Reached the following conclusions: 1) Rutting is easy to occur on roads with lower technical level and rutting is more serious; 2) rutting is deeper and the road sections with lower technical conditions in the subgrade is often being lower grade on technical status index of facilities along the highway (TCI); 3) The comprehensive index of performance PQI is mainly dependent on the pavement condition index PCI and the ridding quality index RQI; 4) The pavement condition index PCI and rutting depth index RDI also have a strong correlation; 5) The proportion of heavy vehicles will be serious affect the

performance of highway pavement indicators, and are all negatively correlated.

From the mining association rules, it can be seen that it is feasible to apply the association rule technology to the analysis of highway pavement performance attenuation. In the future, we can have deep analysis about the pavement performance and apply the results to highway preventive maintenance, then assisting the relevant departments to make decisions.

Acknowledgment

This work has been supported by Project of Hubei Provincial Highway Bureau under Grant Numbers 20141h0288.

References

[1] L.J. Sun, Asphalt pavement structure behavior theory, People's Publishing Press, 2010, pp.380-472.

[2] G. Wang. The applications of data mining in urban highway traffic, Chongqing: Chongqing University, 2016.

[3] Q. Ma, the research and application on Association rules algorithm, Taiyuan University of Technology, 2007.

[4] R.Agrawal,T.Imielinski,A.Swami,Mining association rule between sets of items in large database, Proc Acm Sigmod,New York,vol.22,1993,pp.207-216.

[5] I.Bhandari, Data mining and knowledge discovery, Boston: Kluwer Academic Publisher, 1991, pp.121-125.

[6] X.L.Liu, Highway preventive maintenance's optimization decision and research, Chang'an University, Xi'an, 2010, pp.11-28.

[7] H. Liu, S.Wu, Application of multi-dimentional association rule data mining in ship price influencing factor analysis, Journal of Shanghai maritime University,vol.34,December 2013, pp.31-37

[8] P.Sriratna ,P.Leesutthipornchai,Interesting-based association rules for highway traffic data, Computer Science & Engineering Conference,2015

[9] S.Y.He, Research and application on the technologies in mining association rules, Nanjing: Nanjing University of Aeronautics and Astronautics, 2010, pp.20-30.

[10] S.L. Hong, Y.H. Zhuang, K. Li, Data mining technology and engineering practice, Beijing: China Machine Press, 2014,pp.105-136.