

Research on Financial Risk Management of Automobile Supply Chain Based on Data Mining Technology

Zi-Gui CHEN

School of Management
Guangxi University of Science and Technology
Liuzhou, China
E-mail:754573221@qq.com

Shi-Ping GUAN

School of Management
Guangxi University of Science and Technology
Liuzhou, China
E-mail:477689835@qq.com

Abstract—With the rapid development of supply chain finance, the risk management problem ensues. In this paper, the data mining technology is used to study the large amount of data of the spare parts suppliers in the automobile supply chain, and to identify the characteristic attributes of the default suppliers, and try to help the core enterprises to evaluate and select the partners more efficiently. The results show that the six attributes such as "supplier locations" are highly correlated with the default risk of the supplier, and can help the core enterprises to identify the default suppliers. In addition, "registered capital" and "auto parts properties" are weakly related to default risk, we can delete them.

Keywords—data mining; automobile supply chain; risk management

I. INTRODUCTION

With the popularization of information technology and the globalization of economy, enterprises are faced with a more complex and competitive environment than ever before. The uncertainty of the external environment, prompting enterprises with sensitive response mechanism; And in today's market, has gradually formed a competition between supply chains, competition is the comprehensive strength of the supply chain, rather than the enterprise alone, and enterprises need to mutual cooperation and common development. Based on this, companies need to integrate themselves into the supply chain, the establishment of a good supply chain partnership to achieve the minimum cost or maximize the benefits. Therefore, how to evaluate and choose supply chain partners has become the main problems in the enterprise, is related to the development and efficiency of the whole supply chain. In practice, the common method is one by one on the target enterprise evaluation and selection; this should not be operating and inefficient. To this end, this paper uses data mining technology to identify the characteristics of the supply chain upstream default suppliers to improve the efficiency of enterprises to choose partners and reduce the risk of supply chain finance.

II. DATA MINING TECHNOLOGY

Data mining is a new commercial information processing technology. Its main feature is a large number of business data extraction, conversion, analysis and other model processing, from which to extract the key business decision-making

information [1]. Supply chain financial business will produce a lot of data at all times, and the advent of the era of big data and cloud computing, to strengthen the needs of analyzing and using these data, making data mining technology has come into play. Domestic scholars have done a lot of research on the application of data mining in the supply chain, Chen Chouyong used clustering and classification in data mining to analyze the selection of supply chain partners [2]; Wang Peng Wang summed up the problems in the construction of the supply chain, and through the data mining technology to help the hospital to allocate resources, transfer information, and promote the daily operation of the hospital [3]; Ju Yanbing proposed a supply chain simulation model to analyze the impact of random factors on the supply chain system and the degree of impact [4]; Liu Qing took Chinese pork as the research object, and used web data mining technology to reveal the risk of pork safety in China and put forward the corresponding countermeasures [5]; Chu Dianhui put forward a relational model of customer preference and customer attributes, and used the data mining technology to optimize supply chain product configuration [6]; Wang Jiao used data mining to study the attributes of appliance dealers and proposed financial risk prevention measures [7]; Dai Xin, He Ning constructed RFID data set, and based on this proposed a supply chain workflow model [8]; Li Ming, Yao Yuanyao studied the logistics supply chain, and through the BP neural network model to help core enterprises to choose partners [9]; Zhang Junmo made use of data mining to study the supply cycle and inventory, and improved the supply chain system [10]. The development of supply chain finance in China is not mature enough, and the related financial business is still in the primary stage of development. This paper uses data mining technology to identify the characteristics of auto parts suppliers with high risk of default, and provides reference for the core enterprises in the automobile supply chain when choosing partners.

III. IDENTIFICATION OF DEFAULT SUPPLIER ATTRIBUTES

Data mining is generally divided into predictive mining and descriptive mining. Predictive mining is on the basis of the current data to make prediction; descriptive mining is a brief summary of the data, which summarizes the general characteristics of the data [1]. The object of this study is the parts supplier in the upstream of the automobile supply chain; the purpose is to identify and describe the characteristic

attributes of the default suppliers and therefore belongs to descriptive data mining. The choice of suppliers as a starting point for research is because suppliers are one of the main sources of supply chain risk. Spare parts suppliers are in a weak position in the supply chain and are often required to be responsible for delivery, postponement, deferred payments, etc., which will affect the daily operation of suppliers, and then through the supply chain to influence the other enterprises. And the cost of the vehicle is mainly composed of the price of spare parts, the quality of the vehicle is also determined by the quality of the parts, whether the suppliers can produce spare parts on time and high quality is related to the normal production and operation of all the enterprises in the supply chain. It is assumed that a complete automobile partsacturers, distributors, and consumers. Obviously automobile manufacturers are the core companies in the supply chain, as shown in Figure 1.

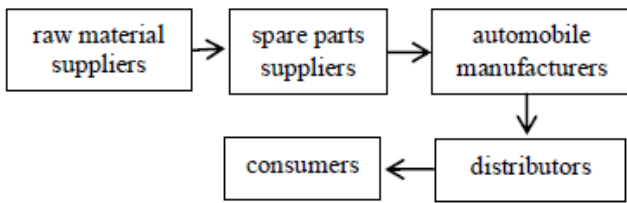


Figure 1. Automobile supply chain.

Then analyzing the automobile manufacturers how to apply data mining techniques to identify and select the parts supplier. In this paper, taking An Automobile Company (hidden the real name of the company) as an example and collecting the data of the upstream 25 suppliers to introduce the basic process of descriptive data mining. It mainly includes: collect and collate data, analyze attribute relevance, summarize characteristic attribute and output mining result.

A. Collect and Collate Date

In order to identify a supplier of spare parts with a high risk of default, it is necessary to put all the suppliers as a group and then dig out their common attributes. The collected automobile suppliers are located in 7 different regions, collect and establish a database, the main field names and values are shown in table 1.

TABLE I. THE MAIN FIELD NAME AND RANGE OF THE AUTO PARTS SUPPLIER DATABASE

Field Name	Value
supplier location	East China, Southern China, North China, central China, northeast, southwest, northwest
registered capital	10 million or less, more than 100 million of the following 10 million, more than 100 million
size	Large, medium and small
enterprise type	State-owned, joint venture, foreign-funded, private, others
whether it is the main customer of the supplier	major customers, general customers, small customers
auto parts properties	key accessories, main accessories, general accessories

Field Name	Value
distance between supplier and delivery point	0 ~ 500 kilometers, 500 ~ 1000 kilometers, 1000 ~ 2000 kilometers, more than 2000 kilometers
delivery season	spring, summer, autumn, winter
default risk	high, low
...	...

After establishing a database, has entered into the phase of data sorting, analysis and delete fields that are not relevant to the data mining topic. The theme of this study is using data mining techniques to identify the characteristics of parts suppliers with high default risk, and the supplier name has nothing to do with the attributes, therefore, is not given in table 1. Depending on the region where the supplier is located, the climate and geographical conditions, the average educational level and basic quality of the population, people's ideas, the logistics and social environment will be different, which are likely to be the reasons for the difference in corporate performance. As for the delivery season, it is basically fixed for every supplier, and Study the influence of the season to supplier delivery performance has certain significance. In addition, the distance between the place of delivery and the supplier is too long to delay the delivery of the supplier and make its performance ability becomes poor. When a field has more different values, the field can be classified, and count the number of each class, such as "registered capital".

B. Analyze Attribute Relevance

The attributes that are collected include the area of the supplier, the registered capital, the enterprise scale, etc., but not all of the attributes will result in a higher risk of default. Therefore, this paper adopts the method of information gain to identify the characteristic attributes which are weakly related or irrelevant to the default risk.

It is assumed that Y is a set of sample data with the number of y and that the class label of each sample is known. For example, "default risk" can be defined as the class label of each sample is high or low. Suppose there are n categories, class labels are X_1, X_2, \dots, X_n , Y contains y_i samples of X_i class, $i=1, 2, \dots, n$, so the probability that an arbitrary sample belongs to class X_i is y_i/y . The expected information for a given sample classification can be expressed as follows:

$$I(y_1, y_2, \dots, y_n) = -\sum_{i=1}^n \frac{y_i}{y} \log_2 \frac{y_i}{y} \quad (1)$$

Let the value of attribute Z be $\{z_1, z_2, \dots, z_m\}$, so the set Y is divided into $\{Y_1, Y_2, \dots, Y_m\}$, y_j contains those samples with a value of z_j . Let y_j contain y_{ij} samples of X_i , then the expected information divided by attribute Z is the entropy of Z.

$$E(Z) = \sum_{j=1}^m \frac{y_{1j} + y_{2j} + \dots + y_{nj}}{y} I(y_1, y_2, \dots, y_n) \quad (2)$$

The information gain divided on the field attribute Z can be expressed as:

$$\text{Gain}(Z) = I(y_1, y_2, \dots, y_n) - E(Z) \quad (3)$$

According to the above formula to obtain the information gain of each characteristic attribute, if the information gain value is large, this attribute is highly correlated with the supplier default risk. The threshold for setting the information gain is 0.05, and calculates the information gain for each field, the information gain of the “registered capital” and “auto parts properties” is found to be 0.041 and 0.027, so can delete these two attributes.

C. Summarize Characteristic Attribute and Output Mining Result

After analyzing the attribute relevance, we can group a set of parts supplier attribute with high default risk. The specific approach is to merge the records in which the value is completely equal, and then output the results. The results can be output through the form of charts, such as a bar chart, crosstab, etc., this will be more intuitive. In addition, the attributes can be taken out separately for analysis: According to supplier location, for example, the risk of default in the eastern region is significantly lower than that in the western region, which is basically consistent with the case of supply chain finance; Another example is the size of suppliers, the default risk of medium-sized suppliers is lower than the large and small suppliers, this may be due to the fact that medium-sized enterprises in the automotive industry are more resistant to risk than small enterprises and are easier to operate than large enterprises.

IV. THE SUPPLIER RISK CONTROL MEASURE

There are several measures for the core enterprise to control the risk of its suppliers: The selection of high quality suppliers, which mainly include the development of new suppliers and re-evaluation of existing suppliers, and eliminate unsuitable suppliers, these is a dynamic process of reciprocating; Sign quality agreement with supplier, the spare parts need to meet the quality standards, and the supplier shall not arbitrarily change the production process, liquidated damages can also be set up in the agreement to increase the default costs for supplier; To strengthen the monitoring of supplier behavior, so that suppliers will do their best to avoid default; Of course, also need to improve cooperation with suppliers, the establishment of a coalition of interests, to achieve win-win situation.

Supply chain finance in China started later than abroad, but is developing very fast. Supply chain financial risk is increasingly prominent, how to manage the risk of becomes a major problem. This paper establishes a database of auto spare parts suppliers, based on the use of descriptive data mining to identify the characteristics of the auto parts suppliers. In order

to ensure the stability and development of supply chain, the core enterprise can take measures to strengthen the supervision and management or even cancel the cooperation relationship with suppliers that have these characteristics. The traditional supplier evaluation method is more easily influenced by subjective factors, and the data mining technology is based on objective data, the identification of the supplier attributes as a standard of core enterprises to select partners, this not only improves the core enterprises to select partners is scientific, and more efficient. Different enterprises, in different stages of development, different industries, and different regions, with different nature and sizes, and its own suppliers also have different characteristics; Data mining technology has a good interpretation of this. The application of data mining technology to the supply chain management of enterprises, supply chain operation will produce massive information data, analyzing the data can help the medium and small enterprises and core enterprises in the supply chain improve the level of management decision-making and help the enterprises to develop better, so as to promote the sustainable and healthy development of the whole supply chain, and also has a certain reference for the financial institutions to carry out the supply chain finance business.

REFERENCES

- [1] Xie Jianglin, He Yiqing, Chen Tao. The Application of Data Mining in the Control of Supply Chain Finance Risk [J]. Journal of Nanchang University (Natural Science), 2008, pp.278-281.
- [2] Chen Chouyong, Lu Jinhong. Supply Chain Partner Selection Based on Data Mining Method [J]. The Journal of Quantitative & Technical Economics, 2002, pp.79-84.
- [3] Wang Peng. Construction of Hospital Supply Chain System based on Public Cloud and Data Mining [J]. Sci-tech Innovation and Productivity, 2015, pp.43-45.
- [4] Ju Yanbing, Wang Aihua, Wang Mingjin. Research on Mining Data from One Supply Chain Simulation System [J]. Microcomputer Information, 2006, pp.138-140.
- [5] Liu Qing, Zhou Jiehong, Yan Zhen. The Risk Identification of Pork's Quality and Safety of China from the Supply Chain Perspective and Policy Implications: Based on the Empirical Analysis of 1624 Food Quality Safety Cases [J]. Chinese Journal of Animal Science, 2016, pp.60-65.
- [6] Chu Dianhui, Zheng Hongzhen. Optimal Configurations of Supply Chain Products Based on Data Mining [J]. Computer Engineering, 2010, pp.276-278.
- [7] Wang Jiao. Application of Data Mining in Prevention of Supply Chain Financial Risks [J]. Logistics Technology, 2013, pp.380-382.
- [8] Dai Xin, He Ning. Study on Data Mining Problem of Supply Chain Oriented toward RFID Terminal Data Collection [J]. Logistics Technology, 2014, pp.434-436.
- [9] Li Ming, Yao Yuanyao. Research on Application of Data Mining Technology in Partner Selection in Logistics Supply Chains [J]. Logistics Technology, 2015, pp.152-154.
- [10] Zhang Junmo. Application of Data Mining Technology in Fine Management of Supply Chain [J]. Electronic Technology & Software Engineering, 2015, pp.199-201.