

Research on the Supply Chain Partnership Relations with Association Rule Analysis

Huiying Zhang¹, Dezhi Zhu¹, Jin Du², Xin Chen²,

¹ School of management, Tianjin University, Tianjin 300072, China

² Tianjin Enterprises Survey Organization, Tianjin 300050, China
zhanghuiying@nankai.edu.cn

ABSTRACT

Due to that many fuzzy and uncertain association relations represented by some phenomenon in the process of enterprises cooperation in supply chain, data mining method was try out in this paper. At first, concept and deducing procedure was introduced, and then 300 samples from a survey was pretreatment, in succession cooperation relations of enterprises in supply chain was studied combined with association analysis method. Some acceptable conclusions were obtained, which can explain the relations between modality and measures of cooperation of enterprises reasonably.

Keywords: Association rule, Supply chain, Partnership

1. INTRODUCTION

Supply chain is a whole function net model that connects from supplier, manufacturer, distributor, retailer and the consumer. Purchase the material, manufacture the work-in-process and end product, and then transfer the product to the consumer by the distributed net, with the controlling flow of information, logistic and finance and surrounding the core enterprise [1]. Nearly every enterprise exists and develops in the supply chain. Many scholars, enterprisers and the consulers are researching the supply chain partnership. In this paper, we focus on the enterprise in Chinese Tianjin and sample of 300 enterprises. We studied the partnership from supplier and client at two aspects, one is how they exchange information and the other is how they solve the trade dispute. For the one aspect, we research the relationship of the cooperation range and time between information exchanges. For the other, we studied the relationship of the quantity of cooperator which has trade dispute and whether they sign the contract between the methods of solving the trade dispute. This research reflects the phenomena of Tianjin because basing on the data from the surveys, so it provides reference to the enterprises.

2. SUPPLY CHAIN PARTNERSHIP

Supply chain Partnership is the collaborative agreements of sharing information and profit, proportioning risk between enterprises. Supply chain can reduce transaction cost because of information exchange, mutual trust and fine partnership [2]. The effective partnership established by the collaborative enterprises will become the special resource and knowledge that can't be copied by others, so it is a private competitive advantage and protect the counterpart become opportunist. It's very important for all enterprise, not only manufacture but also service, to have stabile partnership. Establishing the useful partnership is the core competence for the enterprises to

enhance the satisfied level of the consumer and to achieve the win-win purpose between them by adding the product value [5].

3. DATA MINING AND ASSOCIATION ANALYSIS METHOD

3.1 Data Mining

Data mining is an integration research field across many subjects from the late middle of 90th. It includes database system, statistics, machine learning, visual and information science [4]. Its purpose is to find out the valuable knowledge and information from huge and complex data, by analyzing the data with machine learning and statistics and storing and managing the data with database and data warehouse. The statistics is old and has exact theory, includes described statistics and concluded statistics. But the machine learning method has fewer limits and provides more comprehended result than statistics [5].

3.2 Association Analysis Method

All methods of data mining are different and adapted different field [5]. Using the association analysis can find out the association rules in the data. These rules give the condition of the appearance of attribute-value in the dataset. More exactly, association rule describes the impact of the appearance of one item on the others, i.e. the interesting relation between items. It is popular in the marketing research and web analysis [5] [6]. There are also some scholars mining the text of the open question in survey.

The association rule can be formally stated as follows:

1) Let $I = \{i_1, i_2, \dots, i_m\}$ be a set of literals, called items.

2) Let D be a set of transactions, where each transaction T is a set of items such that $T \subseteq I$.

3) Associated with each transaction is an unique identifier, called its TID. 4) Let X is a set of items in I . A transaction T is said to contain X , if $X \subseteq T$.

An association rule is an implication of the form $X \Rightarrow Y$, where $X \subset I, Y \subset I$ and $X \cap Y \neq \Phi$. We must consider three value to be sure whether an association rule is strong:

- 1) Support s : probability that the transaction set D contains $X \cup Y$, $\text{support}(X \Rightarrow Y) = P(X \cup Y)$.
- 2) Confidence, c : conditional probability that the transaction set D having X also contains Y , $\text{confidence}(X \Rightarrow Y) = P(Y | X)$.
- 3) Lift [9]: the ratio of confidence to expected confidence of Y .

$$\text{cor}_{X,Y} = \frac{P(Y | X)}{P(Y)} \quad (1)$$

The rule is strong with the at least minimum support (called minsup) and minimum confidence (called minconf) and the lift is more than 1, where minsup and minconf are user-specified values.

Except for all above, a really interesting rule must have two conditions [9]:

- 1) It is exceptional and unusual.
- 2) It is potential useful.

These two conditions must be decided by the expert in the field, so the people must involve in the process of mining association rule.

Basing on the kind of the variable, the association rules can be divided to the qualitative and the quantitative. Basing on the kind of dimension in the data, they can be divided to the single-dimension the multi-dimension. In the survey, the qualitative and quantitative variable is the same, and a item correspond to a dimension, so the data is multi-dimension.

3.3 The Applicability of The Association Analysis Method for The Partnership

Because we use the data from survey, the reflected status of the enterprises is limited. We only use limited several items because the survey length. The relationship of the cooperation range and time between information exchanges is not casual, and the relationship of the quantity of cooperator which has trade dispute and whether they sign the contract between the methods of solving the trade dispute is not either. We just master the objective conditions. Even though there are some causal relationships between the items, these are not sufficient. We don't emphasize the

inevitable law but want express the rational correlation, so the association analysis method is suitable.

4. DATA DESCRIPTION

We gather data of the relation from the enterprise to the supplier or the client by the survey. The items are about the district and the contact method and how the enterprise solving dispute. These are:

- 1) The question "How the client (supplier) is distributed?" includes four answers, "in the same section", "in the same city but not the same section", "in the same country but not the same city" and "abroad". It expresses the distributing of the client and the supplier.
- 2) The question "How long the major client (supplier) has business with you?" includes five answers, "within 1 year", "from 1 year to 2 years", "from 2 years to 3 years", "from 3 years to 4 years" and "beyond 4 years". The time is longer, the relation is better.
- 2) The question "What's the method of how you connect with your client (supplier)?" includes five answers, "telephone or fax", "interview", "internet", "post" and "others". It's the information exchange method.
- 3) The question "What's the number of your major client (supplier)?" is a quantitative variable.
- 4) The question "What's the number of your client (supplier) dispute with you?" is a quantitative variable.
- 5) The question "Do you sign contract with your client (supplier)?" is a binary variable.
- 6) The question "Which method is to solve dispute?" include "go to court", "court of arbitration" and "negotiation".

There are 200 manufacture enterprises and 100 service enterprises in our sample which include big, middle and small company and different kinds.

5. PRE-PROCESS AND ANALYSIS

We deal with the data use the EM (Enterprise Miner) tool of the SAS (Statistic Analysis System) software after defining the question and gathering the survey data.

5.1 Data Pre-process

The original data is usual not designed special for the data mining, so we must revise or transform the data to the acceptable form for the mining tools. These are

clearing, integration, transforming and reducing. In this paper, we use the transforming method.

- 1) Because the original data is in the Excel table, we use the "import" function of SAS to import them in the SAS database first. There are 300 cases in the sample. Each case is a row and each item is a column in the database.
- 2) We use the ratio of the number of having dispute to the number of major client (supplier) to express whether the disputing partners are much. It's a quantitative variable. For the new qualitative variable, we must make the quantitative variable discrete before association analysis. In this paper we use the cluster to provide two new variables basing on the space partition. One is for the client number and the other is for the supplier number.
- 3) The question "How long the major client (supplier) has business with you?" has five answers. Considering the internet is widely used after 1997, so we transform to two variables using the dividing line, 1997.
- 4) The association analysis in the EM tool need transaction database, so we design a program in SAS to convert the relational database to transaction database.

5.2 Mining Analysis Rules In EM

The EM mining tool in SAS is a graphical operation, so it's easy to use. We just indicate the database, the variables and the analysis method. In the association analysis, the input variables are ID (transaction identify) and Item (transaction item). We identify the minsup is 5% and the minconf is 30% and the lift more than 1, so we get lot of rules. We filter the useful from all by controlling the antecedent set and the consequence.

5.3 Result

A perfect supply chain is involved the foundation of the collaboration between enterprise, supplier and client about the logistics and information exchange, and the foundation of the contract rule [4]. So we analyses method of information exchange and solving the trade dispute.

5.3.1 The association analysis for the information exchange

First, we analyses the relation between the tow items, distributing and collaboration time, to the information exchange method. Because the positive and correlative rules are interesting, so we only discuss the rules with big lift. The results are:

- 1) The companies with client on abroad and collaborated beyond 3 years mostly use the internet for information exchange. Support is 5.33%. Confidence is 34.04%. Lift is 2.55.
- 2) The companies with client on abroad have more relation with the internet information exchange method. Support is 6.67%. Confidence is 30.30%. Lift is 2.27.
- 3) The companies with domestic client have more relation with the traditional information exchange methods, "telephone and fax" and "interview". Though the lift is less, the confidence is much.
- 4) The companies with supplier on abroad and collaborated beyond 3 years mostly use the internet for information exchange. Support is 4.39%. Confidence is 34.21%. Lift is 3.89.
- 5) The companies with supplier on abroad have more relation with the internet information exchange method. Support is 5.41%. Confidence is 33.33%. Lift is 3.79.
- 6) The companies with domestic supplier have more relation with the traditional information exchange methods, "telephone and fax" and "interview". Though the lift is less, the confidence is much.

The result show that more and more companies use internet in the supply chain to contact the partner on abroad so they reduce the cost by the e-business. But the companies contact the domestic partner just use the traditional method, "telephone and fax" and "interview".

5.3.2 The association analysis for how they solve the trade dispute

We only consider the cases have trade dispute because most cases in the sample are no dispute.

First, we analysis the relation between the quantity of cooperator which has trade dispute and whether they sign the contract to the methods of solving the trade dispute. We only consider the rules with big lift. The results are:

- 1) The companies with more disputed client have more relation with the "negotiation" method. Support is 21.88%. Confidence is 100%. Lift is 1.42.
- 2) The companies with less disputed client have more relation with the "go to court" method. Support is 21.88%. Confidence is 37.84%. Lift is 1.27.
- 3) The companies with more disputed supplier have more relation with the "negotiation" method.

Support is 15.63%. Confidence is 100%. Lift is 2.21.

- 4) The companies with less disputed supplier have more relation with the "go to court" method. Support is 10.94%. Confidence is 46.67%. Lift is 3.32.
- 5) The companies with less disputed supplier have more relation with the "negotiation" method. Support is 28.13%. Confidence is 94.74%. Lift is 2.09.
- 6) There is no significant relation between whether they sign the contract to the methods of solving the trade dispute.

The result shows that the companies solve the dispute by tight collaboration and negotiation when they have more disputed client or supplier considering the long profit, and the companies with less disputed client or supplier tend to go to court to solve the occasional accident.

5. CONCLUSIONS

In this paper, we try to analyse the survey by the data mining and don't limit to the traditional statistics for the uncertain problem. We analyze the survey data from 300 samples of enterprise in Chinese Tianjin and find out a series of appropriate rules about the supply chain partnership. These rules can explain the relations between modality and measures of cooperation of enterprises reasonably and reflect the status of Chinese supply chain partnership and help the manager to make decision. So this method is suitable to the association

REFERENCES

- [1] Ma, Shihua, etc. *Supply Chain Management*, Mechanism Industry Press, 2000.
- [2] Xu, Shujun, etc. "A study on transaction cost of supply chain enterprises", *Industrial Engineering and Management*, Vol.6, No. 6, pp25-27, 2001.
- [3] Ma, Xinan, etc. "Supply Chain Partnership and Partner Selection", *Industrial Engineering and Management*, No. 4, pp33-36, 2000.
- [4] Jiawei Han, Micheline Kamber. *Concept and Technology of data mining*. Mechanism Industry Press, Aug, 2001.
- [5] Indranil Bose, Radha K. Mahapatra, "Business data mining — a machine learning perspective", *Information Management*, pp211-225, 2001
- [6] Qi-Yuan Lin, Yen-Liang Chen, Jiah-Shing Chen, Yu-Chen Chen, "Mining inter-organizational retailing knowledge for an alliance formed by competitive firms", *Information Management*, pp431-442, 2003.
- [7] Dai Wensheng, etc. "Association Rule in data mining". *Statistical Research*, Vol.20, No.8, pp40-42, 2002.
- [8] Tian Jinlan, Zhang Suqin, Huang Gang. "Investment risk rules mining in insurance business data with association rules method", *J Tsinghua Univ (Sci & Tech)*, Vol. 41, No. 1, pp 45-48, 2001.
- [9] Cai Weijie, etc. "Survey of Association Rule Generation". *Computer Engineering*, Vol.27, No.5, pp31-33, 2001
- [10] Zhao Lindu, "Supply chain management in the era of knowledge economics". *Journal Of Southeast University (Natural Science Edition)*, Vol.32, No.3, pp514-522, 2002