

# Research on Customer Knowledge Acquisition Model based on Data Mining

Shen Nali

Southwest University of Political Science and Law, School of Management, Chongqing, China  
 Email:nali0518@163.com

**Abstract**—In the era of big data, with the acceleration of information and the intensification of market competition, customer knowledge, as an important resource of enterprise innovation, plays an increasingly important role in maintaining competitive advantage for enterprises. However, in the face of massive customer data, efficient discovery of useful customer information and knowledge becomes a difficult problem. On this issue, from the perspective of knowledge management, the definition and characteristics of customer knowledge is analyzed, the essence of customer knowledge acquisition based on data mining is revealed and a system model for customer knowledge acquisition is established in this paper.

**Index Terms**— Big data, customer knowledge, data mining, knowledge management, knowledge acquisition

Customer data, as one of the important big data resources in big data, contains great potential value. It is likened to a magical gold mine, and the customer knowledge excavated from it is "gold"[1]. Customer knowledge is the source of enterprise innovation, and is the key factor to gain and maintain customers, gain and maintain competitive advantage[2]. The effective use of customer information and knowledge can enable enterprises to find new market opportunities, develop new products, find potential customers, reduce costs and improve service, so as to create more value, make the enterprise keep a foothold and develop in an unprecedented fierce competition. Therefore, the acquisition and utilization of customer knowledge has become an important topic in the field of customer relationship and knowledge management. However, in the face of massive customer data, efficient discovery of useful customer information and knowledge has become a problem. Data mining technology can use pattern, association, correlation, classification, prediction, clustering and other methods to mine and discover customer knowledge from database, data warehouse or other database. Therefore, the use of data mining to obtain customer knowledge is an effective method, but the specific access mechanism and process need to be further studied. From the point of view of knowledge management, this paper analyzes the definition and characteristics of customer knowledge, reveals the essence of customer knowledge acquisition based on data mining, and establishes a system model of knowledge acquisition.

## I. DEFINITION AND CHARACTERISTICS OF CUSTOMER

## KNOWLEDGE

Customer knowledge is an important basis for the identification of customer needs, development of customer loyalty and marketing and innovation, including the knowledge about the customer, the knowledge that the customer needs and the tacit knowledge that the customer has<sup>[3,4]</sup>. The acquisition and utilization of these knowledge is one of the keys for enterprises to win customers and maintain competitive advantages.

Customer knowledge is closely related to knowledge, enterprise, customer and access technology. Therefore, customer knowledge has the following characteristics: ① Knowledge. ② Customer value. ③ Informatization.

## II. THE ESSENCE OF CUSTOMER KNOWLEDGE ACQUISITION BASED ON DATA MINING

Data mining is a key technology and link of customer knowledge acquisition. It mines and acquires customer knowledge from structured customer data (such as customer transaction data), semi-structured and unstructured customer data (such as customer reviews, complaints, logging data, etc.), and transforms the discrete and implicit knowledge in the process of transaction and situational communication into systematic, orderly and explicit knowledge. From the point of view of knowledge management, the acquisition process is essentially a knowledge creation spiral process based on data mining "Fig. 1."

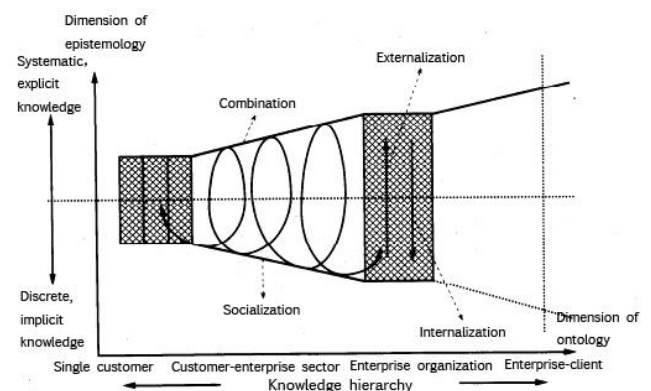


Figure 1. The essence of customer knowledge acquisition based on data mining

Customer knowledge acquisition based on data mining is a kind of intelligent technology and tool that uses data mining to acquire knowledge. The whole idea of the system construction still follows the principle of knowledge transformation and spiral, including socialization, externalization, combination, internalization and circulation of knowledge, as shown in "Fig. 1." Through the four modes of knowledge transformation, namely socialization, externalization, combination and internalization, tacit knowledge and explicit knowledge interact to create new knowledge. Socialization is the process of sharing experiences and creating tacit knowledge such as shared mental patterns and skills[5]. In the acquisition of customer knowledge, socialization is embodied in the process of communication between employees and customers, and sharing of information, experience and knowledge about customer knowledge among internal enterprise employees. Externalization is the process of representing tacit knowledge as explicit knowledge. After acquiring implicit customer knowledge, the employees will make it explicit by metaphor, analogy, concept, hypothesis and model, which can be recorded as information such as customer's behavior, psychology, preference, experience and so on by about employees, and these records are stored in related enterprise database for mining and analysis. Combination is the integration of various concepts into knowledge system, which combines different knowledge of enterprise information system and database to form new information and knowledge. For example, the customer transaction data, customer record data and product service data are collated and linked to form two new databases, namely, The relational database of transaction data and product services and the database of customer records and product services(as shown in "Fig. 2"). In addition, the data mining process itself contains explicit and combined methods to transform customer knowledge from discrete and implicit knowledge into systematic and explicit knowledge. Internalization transforms explicit knowledge into individual tacit knowledge. Customer knowledge is explicit knowledge obtained through data mining or knowledge acquisition systems. However, these knowledge need to be absorbed and utilized by employees and customers to reflect their value. This is a process of internalization of customer knowledge. Customer knowledge is acquired and deposited into customer knowledge base. In the enterprise, the related department, such as product research and development department, marketing department and so on, may call the knowledge according to own need, and apply it to R & D, production, marketing, sales and customer relationship management. In this way, customer knowledge is transferred from enterprise to customer, so as to realize the value of acquired customer knowledge and even create new value. Therefore, in the process of customer knowledge acquisition based on data mining, a process of socialization, externalization, combination, internalization and socialization has been carried out, as shown in "Fig. 1." Through this process, discrete and

implicit knowledge is transformed into systematic, explicit knowledge. The level of knowledge leaps from individual customers to the corporate sector, enterprise organizations, and even the entire network of firms and customers. At the same time, customer knowledge climbs and amplifies at the knowledge level.

### III. CUSTOMER KNOWLEDGE ACQUISITION SYSTEM MODEL BASED ON DATA MINING

According to the definition of customer knowledge, the customer knowledge to be obtained is roughly divided into three categories: Knowledge about customers, knowledge required by customers, and knowledge owned by customers. In view of these three different knowledge, its source and the form of expression are different, therefore the acquisition way also differs: For example, the knowledge about customers mainly comes from the observation of customer's consumption and transaction behavior, and this kind of knowledge is mainly obtained from customer transaction data; the knowledge required by customers and the knowledge they possess are mainly from customers' comments, suggestions and complaints about products and services, as well as marketing, service personnel and customers' conversations or interviews. The acquisition of this kind of knowledge depends on the mining of relevant text information. Therefore, when using data mining methods to acquire customer knowledge, we should not only use association rules and cluster analysis methods, but also combine text mining analysis to design a reasonable customer knowledge acquisition model, as shown in "Fig. 2." Figure 2 reveals the basic system model of customer knowledge acquisition, and the whole structure can be divided into three parts: data acquisition and collation, data mining process and customer knowledge acquisition and utilization.

#### A. Customer data acquisition and collation

The enterprise uses the existing information system and manual input method to input acquired customer data and information to the corresponding database through each data acquisition terminal, such as customer transaction database, product and service database, and text record database ("Fig. 2"). Among them, the product and service database is associated with the customer transaction database and the text record database, because the client transaction entities are mainly products and services. At the same time, transaction data and evaluation, recommendation, complaint and so on need to be recorded in this transaction. Therefore, the data of customer transaction database and product service database are collated, transformed and loaded to form "information database about customer" (represented by database A in "Fig. 2") in order to tap "knowledge about customers"; at the same time, the data of customer's text record database and product service database is collated and transformed to form customer database B, including text records of customer evaluations, recommendations, complaints related to

products, services and businesses, in order to tap "knowledge required by customers" and "knowledge owned by customers".

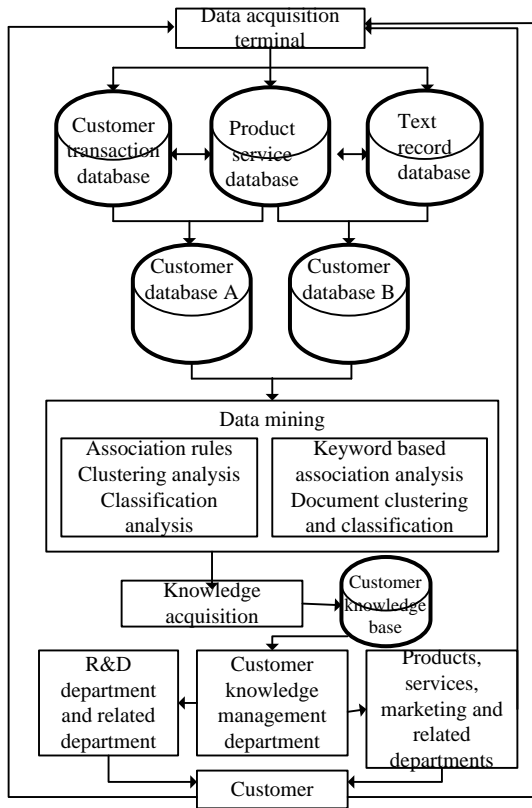


Figure 2. Customer knowledge acquisition model based on data mining

**B. Data mining process**

According to the essence and principle of customer knowledge acquisition based on data mining, we know that the acquisition process of customer knowledge is essentially a spiral process of customer knowledge creation using data mining technology. Obviously, this process can not be separated from the support of data mining technology. Therefore, it is necessary to analyze the core technology of data mining process. As shown in "Fig. 2", the key techniques of data mining are association rules, classification and clustering analysis and their algorithms. The object of analysis two new databases, A and B formed through the acquisition and collation of the initial customer data, that is, the association database of transaction data and product services, and the database of customer records and product services.

Association rules mine frequent patterns or knowledge from large amounts of data in databases A and B, discover interesting associations or related knowledge of data, to help enterprises to grasp customer's consumer behavior, habits, creativity and other knowledge in order to help managers make decisions. The basic idea of acquiring customer knowledge by association rules: First of all, according to the concept and definition of association rules, the management decision maker sets the minimum support threshold according to the prior knowledge to satisfy the customer's knowledge

acquisition condition (such as the probability of purchasing several goods at the same time); secondly, the association rules optimization algorithm (such as the famous Apriori algorithm or improved algorithm) is used to scan the database several times and find all the frequent itemsets, that is, a set of items whose support is greater than predefined minimum support; then, the association rules are generated from frequent large itemsets. The performance of association rules mining is mainly dependent on the determination of large itemsets. Once the large itemset is determined, the rules can be easily visualized, so it is the key to association mining. It is worth noting that in the study of customer knowledge acquisition, the database B is a text based database. Therefore, it is different from database A in terms of the method of association analysis. In text databases, keyword based association analysis is mainly used. That is to say, we should analyze the text data, process root, remove the dead word and so on, then, call association mining algorithms to find the association or correlation between them.

Classification is one of the most important methods in customer data mining. It treats customer data records in databases A and B as a training set (training, set) and analyzes them. Thus, an accurate description and model can be found for each predefined customer knowledge class. The training set consists of the database tuples and their associated class labels. Among them, class labels are usually defined according to some previous empirical data. For example, the class labels of tuple databases of bank customers can be expressed as "security" or "risk"; each customer database tuple in the training set contains several knowledge attributes forming an n-dimensional attribute vector  $X = (x_1, x_2, \dots, x_n)$  that describes the n metric of tuples on the n customer database attributes. Based on the analysis of the training data, classification algorithms such as decision tree, Bias classification, support vector machine, neural network, rule based classification are used to reveal customer data or knowledge classification rules or models. These rules and models are then used to categorize and predict customer data or knowledge without classification. Like association analysis, the classification of text records is different from the classification of transaction data. A text database is semi-structured or unstructured, and cannot be structured according to attribute - value pairs. That is, a keyword set associated with a document can not be organized into a set of fixed attributes or dimensions. Therefore, other efficient classification methods are needed, such as nearest neighbor classification, feature selection, Bias classification, support vector machines and association based classification.

Clustering is the process of dividing data objects into classes or clusters, so that the objects in the same cluster have very high similarity, while the objects of different clusters are highly different. In customer knowledge acquisition, clustering can analyze customer data, discover different customer groups and customer knowledge classes, and depict the characteristics of

customer groups and customer knowledge classes. Among them, clustering methods are mainly divided into method, hierarchical method, density-based method, grid-based method, model-based method, etc. These methods analyze the objects of the customer database, mining and discovering customer knowledge.

### C. Customer knowledge acquisition and application

Through data mining, the system extracts implicit and potentially valuable customer knowledge from the customer database and loads the customer knowledge into the customer knowledge base, as shown in "Fig. 2". In the customer's knowledge base, these customer knowledge has been explicit and owned by the whole enterprise organization. At the same time, some customer knowledge, such as the knowledge required by customers, will be shared by both customers and enterprises. When customer knowledge is acquired, it is shared and diffused among the enterprise organization or enterprise-customer network first. With the promotion and assistance of the customer knowledge management department, the R & D department draws and uses knowledge related to product and service innovation and applies them to the R & D process, while the production department extracts necessary knowledge, such as product, technology, process and other related knowledge, and combines the results of R & D department to improve technology and productivity. Similarly, other departments, such as financial departments, can obtain relevant knowledge and improve work efficiency. Marketing, sales, and other departments having more direct relationship with customers can apply customer knowledge to the marketing and sales activities of products and services, improve the quality of products and services, and create more value for customers. In this process, the enterprise organization and the customer constitute the knowledge interaction network. The enterprise applies customer knowledge to product innovation and service enhancement, and brings new experience to customers, resulting in new explicit or implicit customer knowledge. Similarly, marketing and sales departments use customer knowledge to targeted campaigns, promotions, sales campaigns, and interact with customers to obtain new first-hand customer data. In this way, the new data, information, experience and knowledge acquired by each department are then added into the knowledge acquisition system based on data mining to retrieve customer knowledge. In this way, we constantly create and apply customer knowledge, improve customer value, and maintain the competitive advantage of enterprises.

It is worth noting that in the process of acquiring and applying customer knowledge, the spiral process of knowledge creation is embodied again. Explicit customer knowledge acquired through data mining is transformed into new tacit knowledge by the various members of the enterprise-customer network, such as the team within the organization, individual employees or customer groups, customers by learning, absorbing, and internalizing. These new tacit knowledge form a new source of the spiral of customer knowledge creation. Then, through the

© ACADEMIC PUBLISHING HOUSE

knowledge communication and sharing within the enterprise client network, these implicit knowledge is externalized into explicit knowledge of data, text records and so on into the customer knowledge acquisition system, so as to realize the re-creation of new knowledge.

## IV CONCLUSION

The essence of customer knowledge acquisition based on data mining is the spiral of customer knowledge creation in enterprise customer network. Customer knowledge is constantly refined, absorbed, shared and diffused in this process, which benefits enterprises or customers from different levels. The data mining process follows the spiral principle of customer knowledge creation, analyzes and excavates useful customer knowledge from the vast amount of data and stores them into the customer knowledge base. Then, the enterprise's knowledge management department gives full play to its function of knowledge management and coordination, and distributes the customer knowledge to each required department through the knowledge acquisition system. In such a process, customer knowledge is acquired and applied and is constantly recycled. Finally, it achieves the goal of gaining, winning, maintaining competitive advantage and enhancing customer perceived value, satisfaction and loyalty.

## ACKNOWLEDGMENT

This paper is supported by Youth Fund Project of Humanities and Social Science Research Project, Ministry of Education of China "Research on Customer Knowledge Acquisition and Incentive Mechanism in Supply Chain Enterprises under Large Data Environment" (Project No.: 15YJC630108) and Project supported by academic research projects at the university level of Southwest University of Political Science and Law "Customer Knowledge Acquisition, Decision Making and Incentive of Manufacturers in E-commerce Environment" (Project No.: 2014XZQN-16).

## REFERENCES

- [1] V. M. Schonberger, K. Cukier, "Big Data: A Revolution that Will Transform How We Live," *Work and Think*, 2013.
- [2] S. VACHON, R.D.KLASSEN, "Environmental management and manufacturing performance: The role of collaboration in the supply chain," *International Journal of Production Economics*, 2008, 111(2): 299-315.
- [3] H. Gebert, M. Geib, "Knowledge-enabled customer relationship management: integrating customer relationship management and knowledge management concepts," *Journal of Knowledge Management*, 2003, 7(5): 107-123.
- [4] I. Nonaka, H. Takeuchi, "Enterprises creating knowledge -- the motive force of sustained innovation of Japanese and American Enterprises," *Beijing: Intellectual Property Publishing House*, 2006.
- [5] P. Simon, "Too big to ignore : the business case for big data," *Wiley*, 2013

**Shen Nali**, female, from Xianning, Hubei, Doctor, research director: Supply chain customer knowledge management and e-commerce, way of contact: 18983888365, E-mail :

nali0518@163.com, address: School of Management, No. 301 Baosheng Av., Yubei District, Chongqing