# Construction of Customer Relationship Management

Li Ju, Yu Xiaoting, Xu Wenbin, and Zhou Bei

*Abstract*—**Data mining is the process of discovering connotative, previously unknown, potentially useful and understandable information from large datasets. Data from the database of decision domain is often uncertain and incomplete, and resulting in difficulty of knowledge discovery from this kind of data. The concept of CRM and data mining are discussed in this thesis. Based on that discussion the application of data mining in CRM is demonstrated.**

*Index Terms*—**CRM, rough set, data mining, customer classification.**

## I. INTRODUCTION

In today's society, information explodes, enterprises through the establishment of the customer information database realize customer information management. However, after collecting and storing information on customers, suppliers and business partners, because of lack of useful information found hidden in the data, Business cannot convert these data into knowledge, cannot achieve the ultimate goal of at the right time through the right channels providing appropriate services to the right customers [1].

And data mining is in accordance with the specific goal, finding trends or patterns, to extract the potential, effective model.

In CRM, data mining is from a lot of data about customers, collecting customer information, using various analytical methods, through the disorder, the inner surface information and digging up the knowledge and laws, business can based on these laws or use this information design a mathematical model to make the results forecast for behavior which did not occur, and provide the basis for market planning, integrated management decision-making of the enterprise [2].

Now, a new business culture is developed. The face of these changes, the company urgently needs to implement new solutions, develop new strategies. Since the industrial revolution, it is the first time when mass production and mass marketing concept and customer relationship management by gradually this new concept of alternative. The company is now analyzing the life cycle of customers to participate in the assessment of customer value [3]. The traditional product-oriented model of customer-oriented gradually is replaced by the model. The traditional mass marketing approach by the new replaced one to one marketing. In the traditional approach, the marketing objective is to obtain more customers, expand customer base [4]. However, if it

access to new customer needs to spend more than to contact customers now [5]. To do so, the focus of marketing, customer base shift from expansion to the depth of excavation on customer needs. Showed changes from the so-called market share to wallet share.

## II. THE DEFINITION OF DATA MINING

Data Mining is from a large number of incomplete, noisy, fuzzy, random data, extracting implicit in them, people do not know in advance, but is potentially useful information and knowledge process [6]. In CRM, data mining is a large number of the customer data, collect customer information, using various analytical methods, through the disorder, the inner surface of the information dug up the knowledge and rules, companies can follow these rules or use the information to design a mathematical model to make the results of the predicted behavior not occurred, the comprehensive enterprise management decision, the market provide the basis for planning [7].

### A. Data Mining Applications

From process-oriented perspective, data mining can be divided into three broad categories, which is discovery, predictive model, statutory test method [8].

1. Found- to access the database to find hidden knowledge, it does not require any a priori knowledge and assumptions. It is divided into conditional logic, close relationships and connections, trends and changes [9].

2. The predictive model-get data from the database, and use them to predict the future. It is divided into model predictions, the results prediction.

3. The statutory test method - apply to obtain the data to find distinctive principles.

Data mining is used to build six kinds of models to solve business problems: classification, recession, time series, clustering, joint analysis and sequence discovery [3]. The first two, classification and regression are used to make forecasts, but the joint discovery and sequence discovery is used to describe the behavior. Clustering is used to predict or describe [10], [11].

#### 1) Statistical Mining Techniques

Statistics mining technology is a data mining technology which is currently the most widely used. And it can compress customer data, if customer data is abnormal, so it can find a large number of business customers hidden in the data.

Visualization. Visualization is a calculation way converting abstract symbols into geometry, so that researchers can observe the simulation and calculation process and results

#### 2) Neural networks

From structure, it simulated biological neural networks,

based on model and rule learning, establish three kinds of neural network models: feed-forward network, feedback network and self-organizing network. Through this approach you can complete classification, clustering and feature mining and other data mining tasks.

*3) GA*

GA is simulating biological evolution algorithm, it is composed of breeding (selection), crossover (recombinant), variation (mutation) Algorithms act on a group of possible solutions of a particular issue. They attempt to combine or "breeding" the best existing solution to produce a better solution. It makes the solution less abandoned according to "survival of the fittest" concept. And then improve breeding results.

*4) Decision Tree*

These decisions set by the classification of data sets generated rules. The world's most influential and oldest method is ID3 decision tree, and later other decision tree methods are developed. These rules can be used to classify new cases.

*5) K-Nearest Neighbor Algorithm*

K-nearest neighbor algorithm is based on analogy to learn. Training samples are described with n-dimensional numerical attribute. Each sample represents a point in n-dimensional space, so that all the training samples are stored in the n-dimensional pattern space.

*6) Visualization*

Including the understanding and synthesis of image, generate images according to a complex multi-dimensional data. Data visualization is including data, models and processes. Process visualization is to use data flow diagram to describe the knowledge discovery.

*7) Association Rules*

Data mining association rules is one of the classic technique, which initially started in the super shopping basket studies, focusing on the link between different attributes of the data. The process of knowledge discovery of the association rule can be roughly divided into the following steps: first the problem definition. Be familiar to background knowledge of applications, clear the task demands. Then, data extraction extracts relevant data from the database according to the demand. Second, the data pre-processing; Thirdly, data mining. According to the task of knowledge discovery in the form of pre-database, select the appropriate discovery algorithm, and then extract rules. Finally, knowledge assessment express the model which are found, become a knowledge; Assess the knowledge gained based on the assessment rules to determine the effectiveness of its novelty; according to requirements optimize the processing stage of some knowledge discovery process, until meeting the requirements.

*8) Rough Set Theory*

Rough set theory has the ability that uses incomplete information or knowledge to deal with some uncertain phenomena or disaggregate data according to some uncertainty results.

*B. Rough Set Basic Concept*

Definition 1：For a given decision-making system S= (U, $C \cup D$, V, f), reduction of condition attribute set C is a non-empty subset of C ——P. It meets:

① $\forall a \in P$, Can not be omitted by D

② POSp(D) = POSc(D)

Claimed: P is a reduction of C, the set of all reduction of C denoted RED(C).

By the reduction of the definition, every decision-making system reduction may have several, but reduction is equivalent, that is say they have the same classification ability. The reduction of nuclear is the most important attribute set, which includes all of the reduction.

Definition 2 ： In decision-making system $S = (U, C \cup D, V, f)$, $a \in C$ is defined as

$$\sigma_{(C,D)}(a) = \frac{\gamma_C(D) - \gamma_{C-\{a\}}(D)}{\gamma_C(D)} = 1 - \frac{\gamma_{C-\{a\}}(D)}{\gamma_C(D)}$$

In decision-making system $S = (U, C \cup D, V, f)$, the attribute importance of $P \subseteq C$ is defined as

$$\sigma_{(C,D)}(P) = \frac{\gamma_C(D) - \gamma_{C-\{P\}}(D)}{\gamma_C(D)} = 1 - \frac{\gamma_{C-\{P\}}(D)}{\gamma_C(D)}$$

Definition 3: the domain and dependence properties. With decision-making system $S = (U, C \cup D, V, f)$, C, D denote condition attributes and decision attribute respectively, positive region of the decision attribute under the condition attribute can be defined as $POS_C(D) = \bigcup_{X \in U/D} \underline{C}(X)$, POSC(D) denoted the division according to C's knowledge U/C, can be classified exactly U/D.

The dependence of decision attribute D on condition attribute C is defined as

$$k = \gamma_c(D) = \frac{|POS_C(D)|}{|U|}$$

dependence $\gamma_c(D)$ denoted the target objects under the condition attribute C that can be classified exactly U/D accounted for the ratio of the total object in the domain, expressed dependence of decision attribute on the condition attribute, denoted $C \Rightarrow_k D$.

*C. Rough Set Features*

1. Rough Sets does not require a priori knowledge. Fuzzy sets and evidence theory based on probabilistic methods are a common method of dealing with uncertain information, but these methods require some additional information or a priori knowledge, such as fuzzy membership function, the information sometimes is not easy to get. Rough set analysis methods use only information provided by the data itself, without any a priori knowledge.

2. Rough set is a powerful data analysis tool. It can express and deal with incomplete information: simplify the data and obtain smallest expression of knowledge in the premise of keeping critical information; identify and assess the dependencies between data, reveal the simple concept model; obtained the rule knowledge from empirical data.

3. Rough sets and fuzzy sets, respectively, depict the two sides of incomplete information: rough set based on the indiscernibility relation, focusing on classification, fuzzy set stress ambiguity of collection itself, (vagueness).

From the view of rough sets, rough sets can not be clearly

defined because of the lack of adequate knowledge of the domain, but you can approach with a pair of clear sets. Rough set and evidence theory based on probabilistic methods also have some overlaps, in practical applications they can complement each other.

4. Rough set theory has some unique perspective. These views make rough set particularly suitable for data analysis. The members of the rough set is an objective calculation, and only relevant to data, thus avoiding the subjective factors.

### III. CUSTOMER RELATIONSHIP MANAGEMENT (CRM)

Customer relationship management in the Internet age should be the use of modern means of information technology in the enterprise and its customers to create a digital, timely, interactive communication management system. Through sales, marketing and service departments and personnel to provide comprehensive personalized customer information, and strengthen the tracking service, and information analysis. To enable them to establish and maintain cooperative with customers and business partners, a series of fruitful "one to one relationship", so that enterprises can provide more efficient and considerate services, improve customer satisfaction, attract and retain more customers so as to increase sales, and through information sharing and optimizing business processes effectively reduce the cost of doing business.

Customer relationship management to better understand customers and enhance the effectiveness of customer contact channels to achieve interaction with customers, improve customer sales and reduce customer service costs. Customer Relationship Management building the first step [1] is to solve: how to find potential customers, identify key customers, how to measure customer value and other issues. Customer relationship management systems to data mining and OLAP based, to carry out customer classification, Ke for Ying Ye income, risks and costs related to factors such as analysis, prediction and optimization, transactions can be classified on the client to determine what customers most likely to buy and use a product.

#### A. The Characteristics of Customer Knowledge

Customer knowledge refers to knowledge which is benefit to the decision-making, from customer knowledge pattern, value and other terms, customer knowledge has the following characteristics [2]:

1. Value. Not all customer information is customer knowledge, and only those information which can effectively service for decision-making goal that is customer knowledge

2. Dynamic. Range of enterprise customers are constantly changing, these constant changes have determined the dynamics of customer knowledge.

3. Predictability. Customer buying patterns excavated using KDD technologies need to keep the amendment, and the customer knowledge must have predictable characteristics.,

4. Transformative. Customer knowledge identify customer relationships which have the different value, in the enterprise through a certain process can be converted to capital which reflects relationship between the business and customer.

#### B. Customer Classification Model

In CRM, knowledge discovery based on rough set can be applied in the following areas:

##### 1) Customer Profitability Analysis

Customers to buy directly related to profits of enterprise. Using data mining techniques to build customer profitability return prediction model. We can classify customers, high-value customers, low-value customers and worthless customer. for low-value and worthless customers, we can take certain measures to make low-value or worthless customers into profitable customers; Enterprises can place the limited energy and resources on the most profitable place. For high-value business customers, Enterprises can take different promotions and incentives to maintain customer loyalty to the enterprise. We can identify the trend of the customer return size from the prediction model. Accordingly, the company can take appropriate measures to promote good conversion or avoid the bad conversion.

##### 2) One to One Marketing

One to one marketing is to understand each client, and build lasting relationships. Through the characterization and classification, knowledge discovery technology based on rough set can classify a lot of customers into different classes, customers in each class have similar attributes, rather than similar properties in the different customers. Through data mining to understand the different customer preferences, to provide targeted products and services, which can greatly improve the satisfaction to enterprise and products.

##### 3) Cross-Selling

Cross-selling means a marketing process that you provide new products and services to your existing customers. Through relevant analysis, data mining can help you get the best reasonable sales match. Relevant analysis results can be used in two aspects of cross-selling: For the higher purchase frequency of the merchandise mix, identify those who purchased most commodities of a portfolio, sell them "missing" goods; Another aspect is that for each customer to find more suitable relevant laws, sell them the corresponding goods series.

##### 4) Obtain New Customers

In most commercial areas, the main indicators of business development, including the ability to obtain new customers. Including the discovery of new customers who do not understand your products, and also the customers who received services of competitors before.By the breakdown of these users, it can help companies to filter potential customers. The aims of customer classification is not only to achieve effective identification of customers, often used to guide the strategic resource allocation of enterprise customers manage.

##### 5) Customer Classification

Decision trees, clustering is a common tool for customer segmentation, in accordance with different criteria, such as: according to the customer's spending habits, consumer psychology, purchase frequency, demand for the product or the product profit contribution, which is divided into different user groups, in order to achieve the targeted customer service and develop targeted product to improve customer satisfaction, maximum digging customers lifetime value to enterprise. through the breakdown, it helps the

companies to provide different services according to the different characteristics of the customers, enables companies with minimum investment for maximum return.

### C. The Significance of Customer Relationship Management

An information and a decision will affect the company's existing customers and potential customers fate, can determine the success of the enterprise. How to retain old customers and develop new customers for any business is crucial. Customer Relationship Management (Customer Relationship Management, abbreviated CRM) enterprise through meaningful communication, understanding and influence customer behavior, and ultimately increase customer acquisition, customer retention, customer loyalty and repeat customers to the purpose of earning profits in the process cycle, is a purpose improve the relationship between business and customers a new management mechanism [1]. It is implemented in the corporate marketing, sales, service and technical support and customer-related business areas. Its core idea is based on customer-centric and improves customer satisfaction, improve customer relationships, thereby enhancing the competitiveness of enterprises.

### IV. DATA MINING IN CUSTOMER RELATIONSHIP MANAGEMENT

### A. Customer Life Cycle

The Customer life cycle is the stage of the relationship between customers and enterprises. Understanding of the customer life cycle is very important because it and the customer's income and directly related to customer returns. Market traders said that there are three ways to improve customer value. (1), increase the use value of their existing products; (2) to sell them more products or higher profits; (3) more time to maintain its customers (4). However, over time, customer relations have been changing. Gradually become enterprises and consumers constantly learning from each other. So, this is why the customer life cycle is important? It provides an understanding of customer behavior and the overall framework.

1. Potential customers - not the customers, but customers in target markets;

2. Respondents - for products and services generated interest in potential customers;

3. Active customers - now and in the use of products and services;

4. Previous customers - may lead to high-cost debts, or "bad" customers; they are no longer the target market part; or they have bought a competitor's product.

Customer life cycle for data mining in customer relationship management application provides a good framework. In data mining the input plane, customer life cycle to tell us what information can be used; in the output of the time, customer relationship management to tell us what information should draw attention. Data mining to predict when they become active when the customer returns, how long they will become active customers, how they will leave the [5]. In short: the difficulties China's retail presence and

focus of the present is how to integrate customer data, integration of customer contact channels to large number of customers information, used various channels to enhance customer Guanxi transaction data, mining customer value, Zhangwo Yewuguilv. Now it seems the retail industry in the management of customer life cycle stages can be used in data mining: data mining could help in determining customer characteristics, which can provide customers with targeted services; through data mining can be found to purchase certain products, customer characteristics, which can expand the customer base; also can find the characteristics of the loss of customers, and similar features in the event of loss of customers have not adopted before the targeted features. Therefore, customer maintenance, customer care and maintaining, the data mining tools are playing an active role.

### B. Data Mining in Customer Relationship Management in Specific Application

Customer segmentation, decision tree, clustering is a common tool for customer segmentation, in accordance with different standards, such as: customer's spending habits, consumer psychology, purchase frequency, demand for the product or profit contribution of products classified in different user groups to achieve specific customer service and development of targeted products, to improve customer satisfaction, maximum digging customer lifetime value to enterprises. This will help companies through the breakdown of the different characteristics according to the customer to provide differentiated services will enable enterprises to maximize the minimum return on investment.

#### 1) Customer Loyalty Analysis

Differences in customer demand and the scarcity of retail lead to the generation of customer value analysis. Retail customers based on customer contribution will divide customer value into high value, low value, medium value and so on. In order to provide a retention strategy to different customers, system will build customer value models, and do customer classification according to the customer's contribution on corporate profits, take different promotional tools and incentives for customers who have the different profitability.

#### 2) Obtain New Customers Businesses

Obtain new customers business, not only to maintain existing customers, more importantly, to discover the potential value of new user groups. Get new customers is to convince potential customers use your products or services. Based on historical data, use data mining techniques to establish a "customer behavior response" forecast model, to predict the future behavior of customers. Forecast their sales efforts on your reactions, can be divided into "negative reaction", "no response", "positive reaction." Negative response is not interested you say so, no response is demand for the product is dispensable state; positive reaction expressed interested in your product (but does not necessarily mean to buy), whereby positive response categories selected groups marketing.

#### 3) Customer Retention

Under the 80/20 principle, 80% from 20% of business profits back customers. Customer retention is to retain the possible loss of customers. The cost to acquire new

customers increased year by year, the old customers have a high turnover rate, customers get a new customer costs is to retain an old customer costs 5 times. According to Laikehaide and Sachenko school statistics, if the enterprise customer churn rate has dropped 5%, profits will increase by 25% -50%. Therefore, customers generally contribute to the business much more than new customers. We use data mining technology to keep abreast of customer satisfaction with the services and requirements, and timely analysis of the reasons for changes in the customer as soon as possible amendments to the existing service quality, improve performance, prevent the loss of a large number of customer base and reduce losses. To guide the company focus, the rational allocation resources provide clients with "one on one" personal service. Help enterprises firmly grasp the potential of existing customers and attract more customers, greater competitive advantage. For example, in a mall, for every customer to do a membership card for each customer to purchase a monthly amount or number of purchases to set a value, the daily behavior of customers in this value should be among, if a customer in a month the purchase amount or number of purchases less than the value that the customer appeared abnormal behavior, and probably a lost customer, at the moment, market management should analyze the causes in time, to live a certain strategy to retain customers.

*4) Customer Satisfaction*

Customer satisfaction is psychological experience after end of the consumer process. Therefore, improving customer satisfaction analysis of the retail business to retain customers has a positive effect. It is appropriate to establish customer satisfaction model. Analyze differences of different customer satisfaction, we should provide necessary customer care to the different requirements of different customers;

*5) Customer Profitability Analysis of Customers to Buy Directly Related to Profits*

Using data mining technology to build customer return forecasting model, we can classify customers, high-value customers, low-value customers and worthless customer. For low-value customers and worthless customer, can take certain measures to make low-value customers or no value into profitable customers; enterprises limited energy and resources on profitable enterprises the most useful place. For high-value business customers, can take different promotional tools and incentives to maintain customer loyalty enterprise [5]. From the prediction model can identify the trend of the size of customer returns. Return may be high returns from low return to change or return to the low returns from high fall trend. Accordingly, the enterprise can be targeted to take corresponding measures to promote good conversion, retention, or to avoid bad conversion.

*6) Analysis of the Corresponding Customer Responsiveness*

Build customer response model, looking for customer feature interested in a particular product, and then predict a response on customer to new product recommended.

Business activities are generated based on customer data analysis, So the system through the analysis of customer analysis, play the advantage of the customer analysis, under this analysis, we increase corporate profits and achieve the ultimate tasks of customer relationship management.

## V. CONCLUSION

In today's global economy, the tide of integration, CRM has become an essential business survival business strategy, It can help companies achieve profit maximization, and improve enterprise competitive Li. Data mining technology in the field, throughout the customer life cycle, the various stages of mining enterprise customers the greatest value of Tongshi You Zhuyu enterprise proper allocation of existing resource for businesses Juece support and Shangye function plays a central role. We use appropriate data mining tools for empirical analysis of customer information in order to find the actual needs of the customer and do the customer classification, find out how to make better product sales to increase product sales. By paying a smaller price increase related to customer satisfaction factor scores, which greatly increase customer satisfaction, strengthen customer relationships.

## REFERENCES

[1] A. Berson, S. Smith, K. Thearling, and Q. Z. He made such a translation, building data mining application for CRM (U.S.) [M] Posts & Telecom Press, 2001.

[2] Pilot Software. An introduction to data mining: discovering hidden value in your dataware- house. http://www3.shore.net/~kht/text/dmwhite/dmwhite.htm

[3] Edelstein H. Data mining：exploiting the hidden trends in your data. DB2 Online Magazine. http://www.db2mag.com/9701edel.htm

[4] Freeman M. The customer lifecycles. Intelligent Enterprise 1999；2 (16)：9.

[5] Chen Mingliang. Based on full life cycle of profitable customers segmentation [J]. Economic management, new management,2002,(20):42246.

[6] Pawlak Z. Rough sets: Theoretical aspects of reasoning about data. Boston: Kluwer Academic Publishers,1991

[7] Chmielewski, M.R., Grzymala-Busse, J.W. Global discretization of continuous attributesas preprocessing for machine learning, International Journal of Approximate Reasoning, 1996,15:319-331

[8] A.G Jackson, Z. Pawlak, S.R. LeClair. Rough sets applied to the discovery of materials knowledge, Journal of Alloys and Compounds, 1998,279:14-21

[9] Chan C C. A rough set approach to attribute generalization in data mining. Journal of Information Sciences, 1998,107:169-176

[10] W. Ziarko. Variable precision rough sets model. Journal of Computer System Science, 1993, 46:39~59.

[11] Katzberg JD, Ziarko W.Variable precision rough sets with asymmetric bounds. Ziarko W P. Proceeding of RSKD'93, Springer-Verlag, 1994, 167~177.

[12] Lenarcik A., Piasta Z. Discretization of Condition Attributes Space Intelligent Decision Support. Edited by Roman Slowinski Kluwer, 1992, 373~389.