

A Framework for Knowledge – Driven CRM

Goon Tuck Choy^a and Ian Chai^b

^aFaculty of Information Science and Technology
Multimedia University, 75450 Ayer Keroh, Melaka
Tel: 03-89965168, E-mail: tcgoon@yahoo.com

^bFaculty of Engineering
Multimedia University, 63100 Cyberjaya, Selangor
Tel: 03-83125379, E-mail: ianchai@mmu.edu.my

ABSTRACT

In this paper we propose a framework to combine KDD (Knowledge Discovered in Databases) and CRM (Customer Relationship Management), with an emphasis on customer retention. The key aspect of the proposed framework is to enable adaptive use of knowledge discovered to predict customer buying patterns and capture interesting knowledge about customers.

Keywords

CRM (Customer Relationship Management), Data Mining, KDD (Knowledge Discovery in Databases), Knowledge – Driven CRM, Framework

1.0 INTRODUCTION

The need to better understand customer behaviour is of interest to many managers. However, the problem is that CRM means different things to different people. For some, it means direct e-mails but for others, it means mass customisation or development of products that fits individual customers' needs. For IT consultants, CRM translates into complicated technical jargon related to terms such as OLAP (online analytical processing) and CICs (customer interaction centres) (Winer, 2001).

If the definition itself remains ambiguous, what more the implementation. This is what motivated us to design a model to implement knowledge – driven CRM. So, in order to understand and build a model for knowledge – driven CRM, we must first define CRM and followed by defining what we mean by knowledge – driven CRM.

1.1 CRM

Customer Relationship Management, or CRM, is an information technology industry term used for methodologies, strategies, software, and other web-based capabilities that help an enterprise to coordinate, organise and manage customer relationships. A CRM comprises of a complete suite of customer and business intelligence applications. It enables the analysis of critical data for customer acquisition and retention costs, marketing campaign performance, product and customer profitability, distribution channel performance and customer service. The goal of CRM is to instil greater customer loyalty. It would be an advantage if we could respond faster to customer inquiries and have a deeper understanding of customers, thus increase marketing and selling opportunities.

1.2 Knowledge – Driven CRM

Knowledge - driven CRM involves the integration of an existing CRM system with knowledge discovered, being able to create a dynamic customer model that can evolve to accommodate customer behaviour while at the same time providing the basis for predictive management of customers' future needs.

It is very important to provide accurate and timely information to the sales team. The knowledge gathered by a particular person who has been interacting with the customer should lie with the organisation and not with the person alone. With sales representatives' mobility increasing by the day, it is difficult to expect the same person to come in contact with the customer all the time. When a new person from the organisation goes to the customer, he has to go about re-inventing the

wheel in order to establish rapport with the customer. This by itself is a risky process as the chances of either the customer getting frustrated trying to explain his needs to a new person each and every time, or the sales representative failing to establish a good rapport, is high. All this increases the organisation's risk of losing a customer. In such cases information of any sort pertaining to the customer is invaluable (Goon and Chai, 2003).

2.0 KNOWLEDGE – DRIVEN CRM MODEL

The knowledge – driven CRM process is organised into a number of phases; each phase consists of several generic tasks, some having second level tasks. These phases are called generic, because it is intended to be general enough to cover all possible knowledge – driven CRM situations. The generic tasks are intended to be as complete and stable as possible. Complete means covering both the whole process of knowledge – driven CRM, as well as all possible applications. Stable means that the model should be valid for yet unforeseen developments like new modelling techniques.

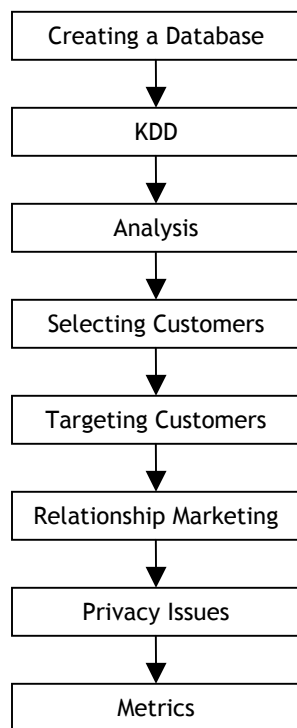


Figure 1: Knowledge - Driven CRM Model

Figure 1 shows a suggested eight phase model for implementing knowledge – driven CRM.

3.0 CREATING A DATABASE

An important phase to any complete CRM solution is the construction of a customer database or information files (Glazer, 1999). This is the foundation for any customer relationship management activity. For web-based businesses, constructing a database is a relatively easier task, as the customer transaction and contact information is accumulated as a natural part of the interaction with customers. For existing companies that have not previously collected much customer information, the task will involve seeking historical customer contact data from internal sources such as accounting and customer service.

3.1 Collecting Customer Data

Ideally, the database should contain:

- **Transactions** – This should include a complete purchase history with accompanying details (price paid, SKU (Stock-keeping Unit), delivery date).
- **Customer Contacts** – Today, there is an increasing number of customer contact points from multiple channels and contexts. This should not only include sales calls and service requests, but any customer - or company - initiated contact.
- **Descriptive Information** – This is for segmentation and other data analysis purposes.
- **Response to Marketing Stimuli** – This part of the information file should contain whether or not the customer responded to a direct marketing initiative, a sales contact, or any other direct contact.

The data should also be represented over time.

4.0 KDD

Knowledge Discovery in Databases, KDD is a broad process of finding knowledge in data. It involves the evaluation, and possibly, interpretation of the patterns to make the decision of what qualifies as knowledge. The goal is to

extract knowledge from data in the context of large databases and it does this by using data mining methods.

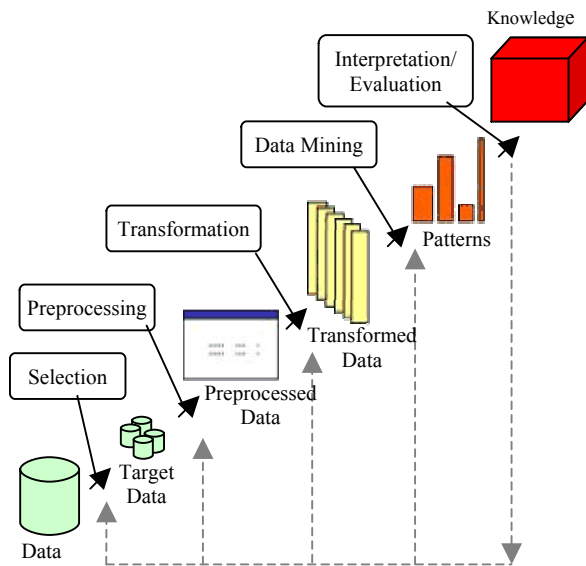


Figure 2: KDD Process

In KDD, an understanding of the data, such as the problem domain, relevant prior knowledge, and end-user goals needs to be developed. The process is then divided into five phases:

- **Selection** - Focus on subsets of variables, or data samples, on which discovery is to be performed
- **Pre-processing** - Cleaning and Pre-processing, e.g.:
 - Remove noise, or get more info to account for noise
 - Strategies to handle missing data
- **Transformation** - Reduction and Projection, e.g.:
 - Find useful features for data representation
 - Effectively reduce the number of variables
 - Find invariant representations of data
- **Data mining** - Choose data mining task:
 - Choose data mining algorithms e.g.: classification, clustering, regression, etc
 - Decide on appropriate models and parameters

- **Interpretation and evaluation** - Interpret and evaluate what classifies as knowledge

4.1 Data Mining

Data mining refers to the application of algorithms for extracting patterns from data without the additional steps of the KDD process. Databases only give us data about a particular customer, but data mining provides the intelligence needed to explore data, find patterns, rules and ideas. From the intelligence, we can explain what is going on and thus, predict what will happen.

In data mining, the goals are:

- **Prediction:** using some variables or fields in the database to predict unknown or future values of other variables of interest
- **Description:** focuses on finding human-interpretable patterns describing observed data

The goals are achieved through the following primary data mining tasks:

- **Classification** is learning a function that maps data items into one of several predefined classes
- **Regression** is learning a function which maps a data item to a real-valued prediction variable.
- **Clustering** is a common descriptive task where one seeks to identify a finite set of categories or clusters to describe the data.
- **Summarization** involves methods for finding a compact description for a subset of data.
- **Dependency Modelling** consists of finding a model which describes significant dependencies *between variables*.
- **Change and Deviation Detection** focuses on discovering the most significant changes in the data from previously measured or normative values

4.1.1 CRISP - DM Model (Chapman, Clinton, Kerber, Khabaza, Reinartz, Shearer and Wirth, 2000)

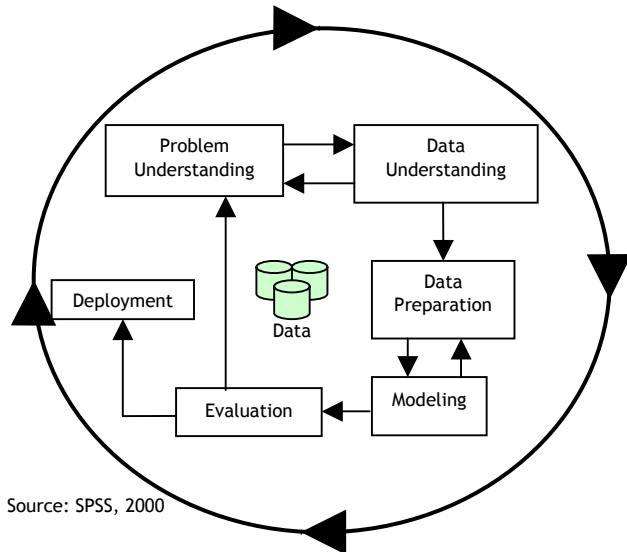


Figure 3: Phases of CRISP DM Reference Model

According to the CRISP DM reference model, the life cycle of a data mining project consists of six phases. Figure 3 shows the phases of a data mining process. The sequence of the phases is not fixed. Moving back and forth between different phases is always required. It very much depends on the results of each phase which phase or which particular task of a phase, has to be performed next. The arrows indicate the most important and frequent dependencies between phases.

The outer circle in Figure 3 symbolises the cyclical nature of data mining itself. Data mining is not over once a solution is deployed. The lessons learned during the process and from the deployed solution can trigger new, often more focused business questions. Subsequent data mining processes will benefit from the experiences of previous ones.

Here is a brief explanation of each phase:

- **Problem understanding** – This initial stage focuses on understanding the project objectives and requirements from a business perspective, then converting this knowledge into a data mining problem definition and a preliminary plan designed to achieve the objectives.

- **Data understanding** – The data understanding stage starts with an initial data collection and proceeds with activities in order to get familiar with the data, to identify data quality problems, to discover first insights into the data or to detect interesting subsets to form hypotheses for hidden information. Task:

- select source
- study data
- establish meta data
- variable type, e.g.:– quantitative or qualitative

- **Data preparation** – The data preparation stage covers all activities to construct the final dataset (data that will be fed into the modelling tool(s)) from the initial raw data. Data preparation tasks are likely to be performed multiple times and not in any prescribed order. Tasks include table, record and attribute selection as well as transformation and cleaning of data for modelling tools. This is often the hardest and takes up the most resources as raw data is inconsistent and often incomplete.

- **Modelling** – In this stage, various modelling techniques are selected and applied and their parameters are calibrated to optimal values. Typically, there are several techniques for the same data mining problem type. Some techniques have specific requirements on the form of data. Often a model is built and evaluation of that model is used to test techniques to see if suitable. Therefore, stepping back to the data preparation phase is often necessary.

- **Evaluation** – At this stage in the project you have built a model that appears to have high quality from a data analysis perspective. Before proceeding to final deployment of the model, it is important to more thoroughly evaluate the model and review the steps executed to construct the model to be certain it properly achieves the business objectives. One of the key objectives here is to determine if there is some important business issue that has not been sufficiently considered. At the end of this stage, a decision on the use of the data mining results should be reached.

- **Deployment** – Creation of the model is generally not the end of the project. Even

if the purpose of the model is to increase knowledge of the data, the knowledge gained will need to be organized and presented in a way that the customer can use it. It often involves applying “live” models within an organisation’s decision-making processes, for example in real-time personalisation of Web pages or repeated scoring of marketing databases. However, depending on the requirements, the deployment phase can be as simple as generating a report or as complex as implementing a repeatable data mining process across the enterprise. In many cases it is the customer, not the data analyst, who carries out the deployment steps. However, even if the analyst will not carry out the deployment effort it is important for the customer to understand up front what actions need to be carried out in order to actually make use of the created models.

5.0 ANALYSIS

The business world is a world of relationships, connecting people, places and things together. Relationships are everywhere and these relationships contain a wealth of information that most data mining techniques are not able to take direct advantage of (Berry and Linoff, 1997).

A variety of statistical methods such as cluster and discriminated analysis have been used to group together customers with similar behavioural patterns and descriptive data which are then used to develop different products or direct marketing campaigns. (Wedel and Kamakura, 1999) This phase may seem to be redundant as it is very similar to the evaluation phase during data mining. However, the objective of this analysis phase is to target the most profitable prospects for marketing campaigns.

By doing analysis, we hope to find:

- Items purchased on credit card, such as rental cars and hotel rooms, give insight into the next product that customer are likely to purchase
- Optional services purchased by telecommunications customers (call waiting, call forwarding, etc.) help

determine how to bundle these services together to maximise revenue.

- Banking services used by retail customers (accounts, hire purchases, mortgage, etc.) identify customers likely to want other services.
- Unusual combinations of insurance claims, indicating fraud and may lead to further investigation.
- Medical patient histories can give indications of complications based on certain combinations of treatments.

6.0 SELECTING CUSTOMERS

Given the construction and analysis of the customer information contained in the database, the next step is to consider which customers to target with the firm’s marketing programs. The results from the analysis can be of various types. If segmentation – type analyses are performed on purchasing or related behaviour, the customers in the most desired segments. (e.g., highest purchasing rates, greatest brand loyalty) would normally be selected first for retention programs. Other segments can also be chosen depending upon additional factors. For example, for promotions or other purchase including tactical decisions, if the customers in the heaviest purchasing segment already buy at rate that implies further purchasing is unlikely, a second tier with more potential would also be attractive. The descriptor variables for these segments (e.g., age, industry type) provide information for deploying the marketing tools. In addition, these variables can be matched with commercially available databases of names to find additional customers matching the profiles of those chosen from the database.

If additional customer – based profitability is also available through the data mining process, it would seem to be a simple task to determine on which customers to focus. Then, the marketing manager can simply choose those customers that are profitable (or are projected to be). The goal is to use the customer profitability analysis to separate customers that will provide the most long – term profits from those that are currently hurting profits. This allows the manager to “fire” customers that are too costly to serve relative to the revenues being produced. While this may seem contrary to being customer – oriented, the basis of the time – honoured “marketing concept,” in fact, there is nothing that says that marketing

and profits are contradictions in terms. The 80/20 rule often holds in approximation: most of a company's profits are derived from a small percentage of their customers. The point here is that, without understanding customer profitability, these kinds of decisions cannot be made.

7.0 TARGETING CUSTOMERS

Mass marketing approaches such as television, radio or advertising are useful for generating awareness and achieving other communications objectives, but they are poorly suited for knowledge – driven CRM due to their impersonal nature. More conventional approaches for targeting selected customers include a portfolio of direct marketing methods such as telemarketing, direct mail, and when suitable, direct sales. Companies should begin to discuss with their customers about these targeted approaches rather than talking “at” customers with the mass media.

Table 1: Targeting Customers

	Customer Acquisition			Customer Retention	
	Direct Mail to Rented List	Banner Advertising	E-mail to Rented List	Direct Mail to “House” List	E-mail to “House” List
Cost Per 1,000	\$850	\$16	\$200	\$686	\$5
Click-Through Rate	N/A	0.8%	3.5%	N/A	10%
Purchase Rate	1.2%	2.0%	2.0%	3.0%	2.5%
Cost Per Sale	\$71	\$100	\$286	\$18	\$8

Source: Forrester Research, 2000.

A recent study by Forrester Research, table 1 demonstrates that e-mail is a very cost-effective approach to consumer retention. Through lower cost per 1,000 names by using company's own database (the “house” list) and greater click through rates than those afforded by banner advertisements and e-mails sent to list rented by suppliers, companies can reduce cost per sale dramatically (Nail, 2000).

Your best customers are those customers who purchase frequently enough, the products or

service you provide, to be profitable. This type of customer generates regular income and is a very desirable customer for your business. Making an effort to target customers like these will rapidly increase your income while creating stability for your organisation.

8.0 RELATIONSHIP MARKETING

The objective of relationship marketing is to deliver a higher level of customer satisfaction as compared to the competitors.

Russell S. Winer in his article for the California Management Review, entitled “A Framework for Customer Relationship Management” gave a comprehensive set of relationship programs as shown in figure 4.

Source: California Management Review, 2001

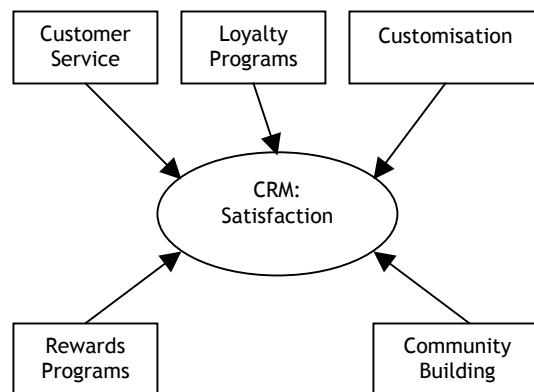


Figure 4: Customer Retention Programs

- **Customer Service** – Programs designed to gain repeat business. E.g.:
 - Reactive – where customer has a problem (product failure, question about bill), contacts the company to solve it.
 - Proactive – aggressive in establishing conversation with customers prior to complaints or other behaviour, suggesting a reactive solution.
- **Loyalty Programs** – provide rewards to customers for repeat purchasing. Patterns from spending behaviour can be generated.
- **Customisation** – create products and services for individual customers. Goes beyond 1-to-1 marketing. Can take

- advantage of the increased information available from customers to tailor products that at least give the appearance of being customised even if they are simply variations on a common base.
- **Community Building** – objective is to take prospective relationship with a product and turn it into something more personal. E.g.:
 - Devoting a section on web site to users and developers.

9.0 PRIVACY ISSUES

A CRM system depends on a database of customer information and knowledge – driven CRM depends on the analysis of that data for more precise targeting of marketing campaigns and relationship building activities. Obviously there exists a trade-off between the ability of companies to deliver better customised products and services and the amount of information needed to make this possible. During this Internet age, many consumers are very concerned about the amount of personal information that is contained in databases and how it is being used.

Almost every time a decision is made, there exists the possibility of legal repercussions. Therefore this is not a new issue at all. Direct marketers have mined databases for many years using analyses based on census tract data, magazine subscriptions, credit card transactions and many other sources of information (Winer, 2001).

Here we suggest that, before a customer signs up, a privacy statement should be attached, where the customer may choose whether or not to be included in future marketing campaigns. This may not be attractive to many, but most organisations tend to lure their customers to sign up by throwing in incentives. In marketing campaigns, do remind the customers that they had previously signed up for these campaigns and always give an option to de-register from them.

10.0 METRICS

The first six phases depend on performance targets and metrics to gauge their success, and enterprises must set measurable CRM objectives and monitor CRM indicators, gauge what is knowledge and what is not to successfully turn customers into assets. Without performance

management, a CRM strategy and associated program like the knowledge – driven CRM is destined to fail. These metrics have an internal and an external focus and link operations to strategy and corporate financial benefits. Each enterprise will have a unique set of metrics applicable to their situation (Kirkby, Thompson and Buytendijk, 2001).

The increased attention paid to CRM means that the traditional metrics used by managers to measure the success of their products and services in the marketplace have to be updated. Financial and market – based indicators such as profitability, market share, and profit margins have been and will continue to be important. However, in a CRM world, increased emphasis is being placed on developing measures that are customer – centric and give managers a better idea of how their CRM policies and programs are working.

Some of these CRM – based measures, both Web and non – Web – based, are: customer acquisition costs, conversion rates (from lookers to buyers), retention/churn rates, same customer sales rates, loyalty measures, and customer share or share of requirements (the share of a customer’s purchases in a category devoted to a brand) (Lehmann and Winer, 2001). All of these measures imply doing a better job acquiring and processing internal data to focus on how the company is performing at the customer level.

11.0 CONCLUSION

With the increased penetration of CRM philosophies in organisations, however not every CRM implementation has met, or will, meet expectations. The true essence of CRM is about acquiring, retaining, and developing your customers by using effective tools that will empower the people in your organisation. With knowledge – driven CRM, we are not only able to achieve that, but we may predict as well.

So, with this framework, we hope to achieve the following:

- Generic guidelines to implementing knowledge – driven CRM.
- Adaptive use of knowledge to combine KDD and CRM.

- Increased efficiency of future knowledge
– driven CRM implementation

12.0 REFERENCES

Berry M., Linoff G. (1997). *Data Mining Techniques: For Marketing, Sales, and Customer Support*, John Wiley & Sons, Inc.: 124-242

Chapman P., Clinton J., Kerber R., Khabaza T., Reinartz T., Shearer C., Wirth R. (2000). *CRISP - DM 1.0: Step by Step Data Mining Guide*, SPSS.

Glazer R. (1999). *Winning in Smart Markets*, Sloan Management Review, 40/4: 59-69

Goon T.C., Chai I. (2003). Knowledge – Driven CRM. In proceedings of MMU International Symposium on Information and Communications Technologies (M²USIC) 2003, [M019]. Petaling Jaya, Malaysia.

Kirkby J., Thompson E., Buytendijk F. (2001). *Getting the Best Out of CRM Performance Metrics*, Gartner Research, AV-14-9265.

Lehmann D., Winer R. (2001). *Product Management*, 3rd edition, Burr Ridge, IL, McGraw Hill.

Nail J. (2000). *The Email Marketing Dialogue*, Forrester Research Inc.

Wedel M., Kamakura W. (1999). *Market Segmentation: Conceptual and Methodological Foundations*, 2nd edition, New York, NY, Kluwer Academic Publishers.

Winer R. (2001). *A Framework for Customer Relationship Management*, California Management Review, 43/4: 89-105.

Zeithaml V., Rust R., Lemon K. (2001). elaborate on this section in their article, *The Customer Pyramid: Creating and Serving Profitable Customers*, California Management Review, 43/4.