A Case for Analytical Customer Relationship Management

Jaideep Srivastava[1], Jau-Hwang Wang[2], Ee-Peng Lim[3], and San-Yih Hwang[4]

[1]Computer Science & Engineering
University of Minnesota, Minneapolis, MN 55455, USA
srivasta@cs.umn.edu
[2]Information Management
Central Police University, Taoyuan, ROC
jwang@sun4.cpu.edu.tw
[3]Chinese University of Hong Kong
Hong Kong, PRC
aseplim@ntu.edu.sg
[4]National Sun-Yat Sen University
Kaoshiung, ROC
syhwang@misserv.mis.nsysu.edu.tw

Abstract. The Internet has emerged as a low cost, low latency and high bandwidth customer communication channel. Its interactive nature provides an organization the ability to enter into a close, personalized dialog with individual customers. The simultaneous maturation of data management technologies like data warehousing, and data mining, have created the ideal environment for making customer relationship management (CRM) a much more systematic effort than it has been in the past. In this paper we described how data analytics can be used to make various CRM functions like customer segmentation, communication targeting, retention, and loyalty much more effective. We briefly describe the key technologies needed to implement analytical CRM, and the organizational issues that must be carefully handled to make CRM a reality. Our goal is to illustrate problems that exist with current CRM efforts, and how using data analytics techniques can address them. Our hope is to get the data mining community interested in this important application domain.

1   Introduction

As bandwidth continues to grow, and newer information appliances become available, marketing departments everywhere see this as an opportunity to get in closer touch with potential customers. In addition, with organizations constantly developing more cost-effective means of customer contact, the amount of customer solicitation has been on a steady rise. Today, with Internet as the ultimate low latency, high bandwidth, customer contact channel with practically zero cost, customer solicitation has reached unprecedented levels.

Armed with such tools, every organization has ramped up its marketing effort, and we are witnessing a barrage of solicitations targeted at the ever-shrinking attention span of the same set of customers. Once we consider the fact that potentially

good customers, i.e. 'those likely to buy a product', are much more likely to get a solicitation than those who are not so good, the situation for the good customers is even more dire. This is really testing the patience of many customers, and thus we have witnessed a spate of customers signing up to be on 'no solicitation' lists, to avoid being bombarded with unwanted solicitations..

From the viewpoint of the organizations, the situation is no better. Even though the cost of unit customer communication has dropped dramatically; the impact of unit communication has dropped even faster. For example, after a lot of initial enthusiasm, it is now widely accepted that the impact of web page banner advertisements in affecting customer opinion is practically negligible. On the other hand, the impact of targeted e-mails, especially with financial offers, is quite high. In essence, each organization is spinning its wheels in tying to target the same set of good customers, while paying insufficient attention to understanding the needs of the 'not so good customers' of today, and converting them into good customers of tomorrow. A clear example of this mutual cannibalism of customers is the cellular phone industry, where each service provider is constantly trying to outdo the others. "Customer churn" is a well-accepted problem in this industry.

A well-accepted wisdom in the industry is that it costs five to seven times as much to acquire a new customer than to retain an existing one. The reason is that the organization already has the loyalty of existing customers, and all that is required for retention is to meet the customer's expectations. For customer acquisition however, the customer must be weaned away from another organization, which is a much harder task. Given this, it is crucial that the selection of customers to target is done with care, and the right message be sent to each one. Given these needs, it becomes important for an organization to understand its customers well. Thus, one can consider customer relationship management to consist of two parts as follows:

CRM = customer understanding + relationship management

This equation is not new, since in the classical 'neighborhood store' model of doing business, the store had a highly localized audience, and the store owner knew practically everyone in the neighborhood – making it easy for him to meet the needs of his customers. It is the big corporations, serving a mass customer base, that have difficulty in understanding the needs of individual customers. The realization of this gap of knowledge has been one of the driving factors for the rapid adoption of CRM software by many corporations. However, the initial deployment of CRM software has been for the second part of the CRM equation, namely 'relationship management'. As described above, relationship management efforts without an understanding of the customer can be marginally effective at best, and sometimes even counter productive.

The approach that resolves this dilemma is the use of data analytics in CRM, with the goal of obtaining a better understanding of the needs of individual customers. Improved customer understanding drives better customer relationship efforts, which leads to better and more frequent customer response; which in turn leads to more data collection about the customer – from which a more refined customer understanding can be gained. This positive feedback cycle – or 'virtuous loop' as it is often called – is shown in Figure 1.
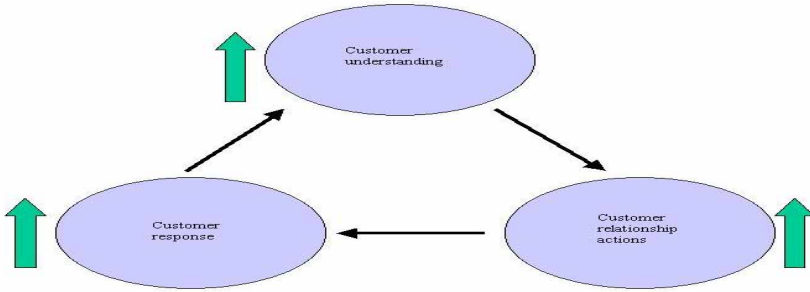
Figure 1. 'Virtuous circle' of CRM.

While this picture is very desirable, unfortunately there are a number of technical and organizational challenges that must be overcome to achieve it. First, much of customer data is collected for operational purposes, and is not organized for ease of analysis. With the advance of data analysis techniques, it is becoming feasible to exploit this data for business management, such as to find existing trends and discover new opportunities. Second, it is critical that this knowledge cover all channels and customer touch points - so that the information base is complete, and delivers a holistic and integrated view of each customer. This includes customer transactions, interactions, customer denials, service history, characteristics and profiles, interactive survey data, click-stream/browsing behavior, references, demographics, psychographics, and all available and useful data surrounding that customer. This may also include data from outside the business as well, for example from third party data providers such as Experian or Axciom. Third, organizational thinking must be changed from the current focus on products to include both customers and products, as illustrated in Figure 2. Successful adoption of CRM requires a change in focus by marketing from "who I can sell this products to?" to "what does this customer need?" It transforms marketing from "tactical considerations, i.e. "how do I get this campaign out of the door" to strategic focus, i.e. "what campaigns will maximize customer value?"
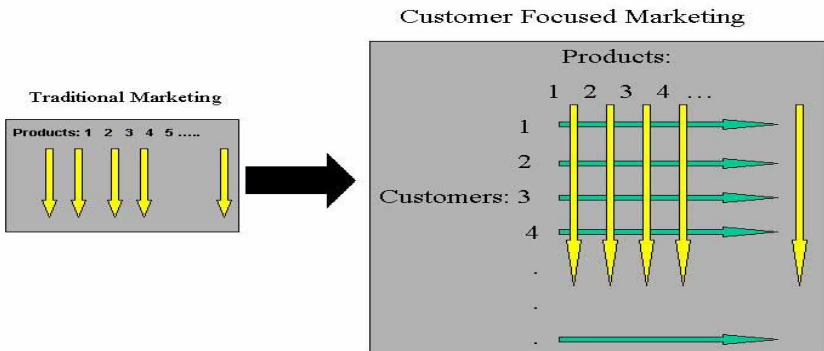

Figure 2. Change of focus from product only to customer+product.

The goal of this paper is to introduce the data mining community to the data analytics opportunities that exist in customer relationship management, especially in the area of customer understanding. As the data collected about customers is becoming more complete, the time is ripe for the application of sophisticated data mining techniques towards better customer understanding. The rest of this paper is organized as follows: in Section 2 we introduce the concept of analytical customer relationship management. Section 3 briefly describes the underlying technologies and tools that are needed, namely data warehousing and data mining. Section 4 describes a number of organizational issues that are critical to successful deployment of CRM in an organization, and Section 5 concludes the paper.

2   Analytical Customer Relationship Management

Significant resources have been spent on CRM, leading to the success of CRM software vendors such as Seibel, Oracle, and Epiphany. However, in the initial stages sufficient attention was not paid to analyzing customer data to target the CRM efforts. Simple heuristics and 'gut-feel' approaches led to profitable customers being bombarded with offers (often turning them off), while there being little attempt to develop today's the 'less valuable' customers into tomorrow's valuable ones. This lack of attention to customer needs is the cause of decreasing customer satisfaction across a wide variety of industries, as illustrated in Figure 3 [Heyg2001].[1]
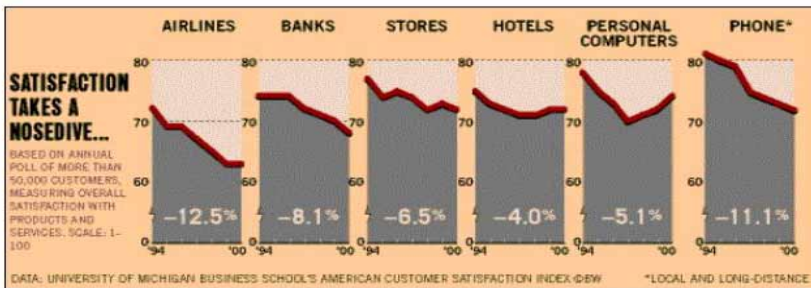

Figure 3. Declining trend in customer satisfaction index.

Fortunately, however, the tremendous advancement in data management and analysis technologies is providing the opportunity to develop fine-grained customer understanding on a mass scale, and use it to better manage the relationship with each customer. It is this approach to developing customer understanding through data analysis, for the purpose of more effective relationship management, that we call "*analytical customer relationship management(ACRM)*". ACRM can make the customer interaction functions of a company much more effective than they are presently.

---

[1] Of course, customer expectation keeps rising over time, and the source of dissatisfaction today is very different that that of a few years ago. However, this is a battle that all organizations must constantly fight.

2.1 Customer Segmentation

Customer segmentation is the division of the entire customer population into smaller groups, called *customer segments*. The key idea is that each segment is fairly homogeneous from a certain perspective – though not necessarily from other perspectives. Thus, the customer base is first segmented by the value they represent to an organization, and then by the needs they may have for specified products and services.

The purpose of segmentation is to identify groups of customers with similar needs and behavior patterns, so that they can be offered more tightly focused products, services, and communications. Segments should be identifiable, quantifiable, addressable, and of sufficient size to be worth addressing. For example, a vision products company may segment the customer population into those whose eyesight is perfect and those whose eyesight is not perfect. As far as the company is concerned, everyone whose eyesight is not perfect falls in the same segment, i.e. of potential customers, and hence they are all the same. This segment is certainly not homogeneous from the perspective of a clothing manufacturer, who will perhaps segment on attributes like gender and age.

A company's customer data is organized into *customer profiles*. A customer's profile consists of three categories of data, namely (i) identity, (ii) characteristics, and (iii) behavior. These categories correspond to the questions *Who the person is?, What attributes do they have?,* and *How do they behave?* Two types of segmentation can be performed based on the profile, namely

- group customers based on common characteristics, and identify their common patterns of behavior, and
- group customers based on common patterns of behavior, and identify their common characteristics.
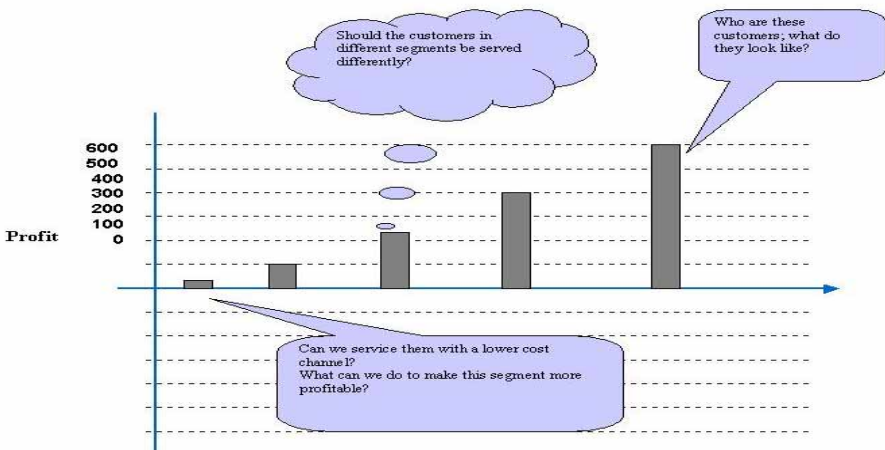

Figure 4. Segmentation of customers by profitability.

As shown in Figure 4, each customer segment represents a different amount of profit per customer; the treatment of each segment can be different. The figure shows examples of the type of questions the company can ask about segments. Also included

are some overall strategic questions about which segments to focus on, and how much.

2.2 Customer Communication

A key element of customer relationship management is communicating with the customer. This consists of two components, namely (i) deciding what message to send to each customer segment, and (ii) selecting the channel through which the message must be sent. Message selection for each customer segment depends on the strategy being followed for that segment, as shown in Figure 4. The selection of the communication channel depends on a number of characteristics of each channel, including cost, focus, attention, impact, etc.

Typical communication channels include television, radio, print media, direct mail, and e-mail. Television is a broadcast channel, which is very good at sending a common message to a very large population. While it is very effective in building brand recognition, it is difficult to target a specific segment, as well as to measure response at the individual customer level. Radio, like television is a broadcast medium, and hence difficult to use for targeted communication to individual customers. Some television and radio stations, e.g. public radio and public television, develop a fairly accurate sample of their listener/viewer base through periodic fundraisers. Print media like newspapers and magazines can be used for much more focused communication, since the subscriber's profile is known. However, the readership of print media is usually much larger than the subscription base – a ratio of 1:3 in the US – and hence for a large part of the readership base, no profile is available. Direct mail is a communication channel that enables communicating with individual customers through personalized messages. In addition, it provides the ability of measuring response rates of customers at the individual level, since it enables the contacted customer to immediately respond to the message – if so desired. Finally, given its negligible cost, e-mail is becoming the medium of choice for customer contact for many organizations.
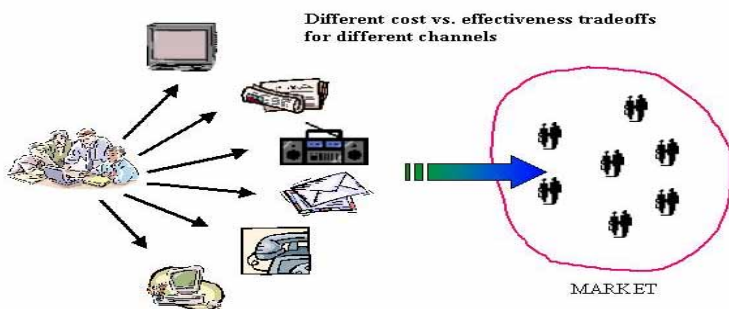

Figure 5. Formulating the optimal customer communication strategy.

Figure 5, courtesy of [Stev1998], illustrates the problem of formulating the customer communication strategy. Each communication channel has its own

characteristics in terms of cost, response rate, attention, etc. The goal of communication strategy optimization is to determine the (set of) communication channel(s) for each customer that minimizes cost or maximizes sale, profit, etc. While communication channel optimization has been a well-studied problem in the quantitative marketing literature, characteristics of new channels such as e-mail and the Web are not well understood. Thus, there is a need to revisit these problems.
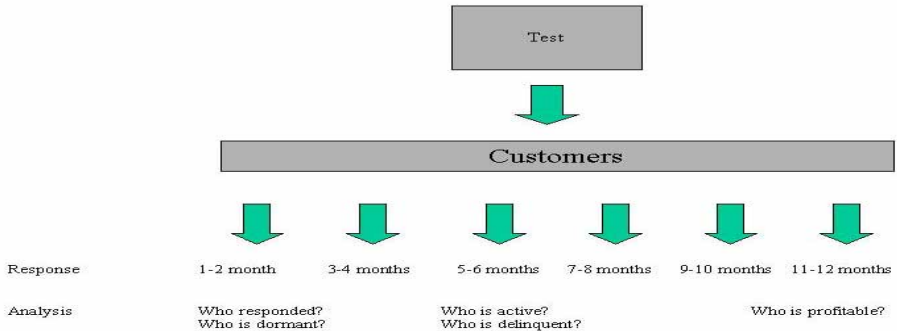

Figure 6. Analyzing the response to customer communications.

Sending the message to each customer through the chosen communication channel is not enough. It is crucial to measure the impact of the communication. This is done by using an approach called *response analysis*. As shown in Figure 6, response analysis metrics, e.g. number of respondents, acquired customers; number of active customers, number of profitable customers, etc. can be calculated. These are analyzed to (i) determine how effective the overall customer communication campaign has been, (ii) validate the goodness of customer segmentation, and (iii) calibrate and refine the models of the various communication channels used. While response analysis for traditional communication channels is fairly well understood, for new channels like e-mail and the Web, hardly anything is known. Understanding how customers relate to these new medium, which aspects they like and which they don't, and what are the right set of metrics to measure the usage of the medium, are all open questions.

2.3 Customer Retention

Customer retention is the effort carried out by a company to ensure that its customers do not switch over to the competition's products and services. A commonly accepted wisdom, acquired through substantial experience, is that it is 5 to 7 times more expensive to acquire a new customer than to retain an existing one. Given this, it is of paramount importance to retain customers, especially highly profitable ones. A good loyal customer base that persists for a long time is one of the best advertisements for a business, creating an image of high quality. This helps in attracting other customers who value long term relationships and high quality products and services.
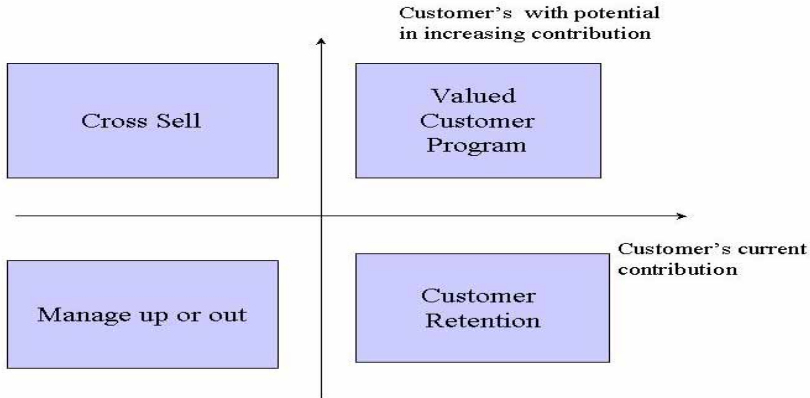
Figure 7. Treatment of various customer segments.

Figure 7 shows how a company thinks of its various customer segments, from a current and future profitability perspective. Clearly, the quadrants on the right bottom and the right top should be targeted for retention. In addition, the right top customer quadrant must be targeted for strengthening the relationship, as there is significant unrealized potential.

A successful customer retention strategy for a company is to identify opportunities to meet the needs of the customer in a timely manner. A specific example is of a bank that used the event "ATM request for cash" is rejected due to lack of funds" to offer unsecured personal loans to credit-worthy customers the next day. This offer was found to have a very high success rate, with the additional advantage of building customer loyalty. Classically, this analysis has been done at an aggregate level, namely for customer segments. Given present day analytic tools, it should be possible to do it at the level of individual customers.

2.4 Customer Loyalty

From a company's perspective, a loyal customer is one who prefers the company's products and services to those of its competition. Loyalty can range from having a mild preference all the way to being a strong advocate for the company. It is well accepted in consumer marketing that an average customer who feels closer to a company (high loyalty) is significantly more profitable than one who feels less close (low loyalty). Thus, ideally a company would like all its customers to become loyal, and then to quickly advance up the loyalty chain.

Figure 8, courtesy of [Heyg2001], illustrates the concept of tracking a customer to identify events in his/her life. Many of these events offer opportunities for strengthening the relationship the company has with this customer. For example, sending a greeting card on a customer's birthday is a valuable relationship building action – with low cost and high effectiveness.
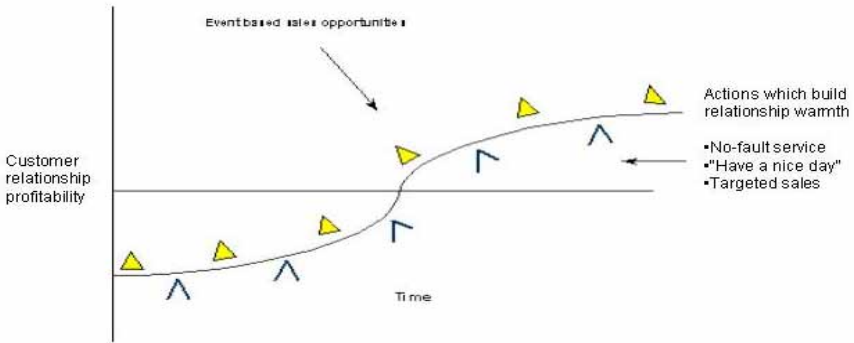
Figure 8. Lifetime impact of customer loyalty.

In marketing language this is called 'event marketing', where the idea is to use the occurrence of events as marketing opportunities. Sometimes even negative events can be used to drive sales. For example, a bank adopted the policy of offering a personal loan to every customer whose check bounced or there were insufficient funds for ATM withdrawal. This program was very successful, and also enhanced the reputation of the bank as being really caring about its customers.

The data mining community has developed many techniques for event and episode identification from sequential data. There is a great opportunity for applying those techniques here, since recognizing a potential marketing event is the biggest problem here.

3 Data Analytics Support for Analytical CRM

In this section we describe the backend support needed for analytical CRM. Specifically, we first outline a generic architecture, and then focus on the two key components, namely data warehousing and data mining.

3.1 Data Analytics Architecture

Figure 9 shows an example architecture needed to support the data analytics needs of analytical CRM. The key components are the data warehouse and the data analysis tools and processes.

3.2 Data Warehouse

Building a data warehouse is a key stepping stone in getting started with analytical CRM. Data sources for the warehouse are often the operational systems, providing the lowest level of data. Data sources are designed for operational use, not for decision support, and the data reflect this fact. Multiple data sources are often from different systems, running on a wide range of hardware, and much of this software is built in-house or highly customized. This causes data from multiple sources to be mismatched. It is important to clean warehouse data since critical CRM decisions will be based on it. The three classes of data extraction tools commonly used are - data migration which allows simple data transformation, data scrubbing which uses domain-specific knowledge to scrub data, and data auditing which discovers rules and relationships by scanning data and detects outliers.
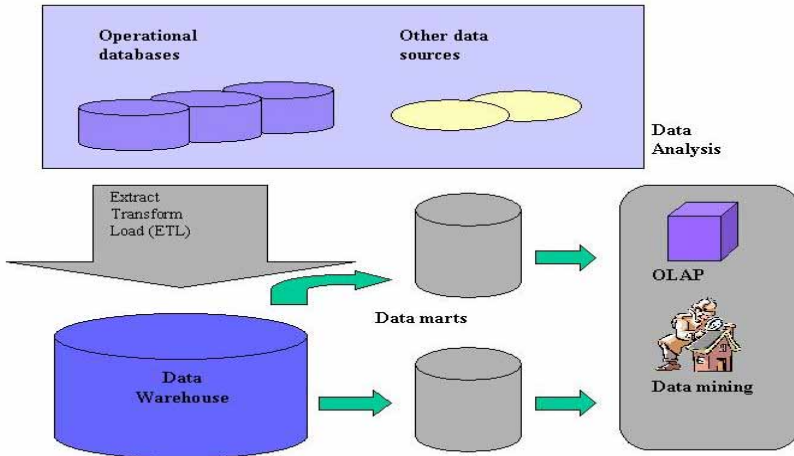
Figure 9. Data analytics architecture.

Loading the warehouse includes some other processing tasks, such as checking integrity constraints, sorting, summarizing, and build indexes, etc. Refreshing a warehouse requires propagating updates on source data to the data stored in the warehouse. The time and frequency to refresh a warehouse is determined by usage, types of data source, etc. The ways to refresh the warehouse includes data shipping, which uses triggers to update snapshot log table and propagate the updated data to the warehouse, and transaction shipping, which ships the updates in the transaction log.

The key entities required for CRM include *Customer, Product, Channel*, etc. Usually information about each of these is scattered across multiple operational databases. In the warehouse these are consolidated into complete entities. For example, the Customer entity in the warehouse provides a full picture of who a customer is from the entire organization's perspective, including all possible interactions, as well as their histories. For smaller organizations the analysis may be done directly on the warehouse, while for larger organizations separate data marts may be created for various CRM functions like customer segmentation, customer communication, customer retention, etc.

3.3 Data Mining

The next generation of analytic CRM requires companies to span the analytical spectrum and focus more effort on looking forward. The 'what has happened' world of report writers and the 'why has it happened' OLAP worlds are not sufficient. Time-to-market pressures, combined with data explosion, are forcing many organizations to struggle to stay competitive in the 'less time, more data' scenario. Coupled with the need to be more proactive, organizations are focusing their analytical efforts to determine what will happen, what they can do to make it happen, and ultimately to automate the entire process. Data mining is now viewed today as an analytical necessity. The primary focus of data mining is to discover knowledge, previously unknown, predict future events and automate the analysis of very large data sets.

The data mining process consist of a number of steps. First the data collected must be processed to make it mine-able. This requires a number of steps to clean the

data, handle mismatches in format, structure, as well as semantics, and normalization and integration. A very good book on the subject is [Pyle99]. Once the data has been cleaned up, various data mining algorithms can be applied to extract models from it. A number of data mining techniques have been developed, and the one to be applied depends on the specific purpose at hand. [HMS00] provides and excellent introduction to various data mining algorithms, while [Rud00] shows how they can be applied in the context of marketing.

Once a model has been developed, it can be used for two kinds of purposes. First is to gain an understanding of the present behavior of the customers. A model used for this purpose is called a *descriptive model.* Second is to use the model to make predictions about future behavior of the customers. A model used for this purpose is called a *predictive model.* The descriptive model, extracted from past behavior, is used as a starting point from which a predictive model can be built. Such an approach has been found to be quite successful, as is based on the assumption that past behavior is a good predictor of the future behavior – with appropriate adjustments. This holds quite well in practice.

4   Organizational Issues in Analytical CRM Adoption

While the promise of analytical CRM, both for cost reduction and revenue increase, is significant, this cannot be achieved unless there is successful adoption of it within an organization. In this section we describe some of the key organizational issues in CRM adoption.

4.1 Customer First' Orientation

Companies that offer a number of products and services have traditionally organized their customer facing teams, e.g. sales, marketing, customer service, etc. along product lines, called "Lines of Business (LOB)". The goal of any such product marketing team is to build the next product in this line; the goal of the sales team is to identify the customers who would be likely to buy this product, etc. This product line focus causes customer needs to be treated as secondary.

The customer focusing teams of an organization must be re-oriented to make them focus on customers in addition to product lines. These teams can be organized around well-defined customer segments, e.g. infants, children, teenagers, young professionals, etc., and each given the charter of mapping our product design, marketing, sales, and service strategies that are geared to satisfying the needs of their customer segment. As part of this, some of the activities might be targeted to individual customers.

4.2 Attention to Data Aspects of Analytical CRM

The most sophisticated analytical tool can be rendered ineffective if the appropriate data is not available. To truly excel at CRM, an organization needs detailed information about the needs, values, and wants of its customers. Leading organizations gather data from many customer touch points and external sources, and bring it together in a common, centralized repository; in a form that is available and ready to be analyzed when needed. This helps ensure that the business has a

consistent and accurate picture of every customer; and can align its resources according to the highest priorities. Given this observation, it is critical that sufficient attention be paid to the data aspects of the CRM project, in addition to the software.

4.3 Organizational 'Buy In'

While data mining and data warehousing are very powerful technologies, with a proven track record, there are also enough examples of failures when technology is deployed without sufficient organizational 'buy in'. As described earlier, the parts of the organization that will benefit the most from analytical CRM are the business units, i.e. marketing, sales, etc., and not the IT department. Thus, it is crucial to have 'buy in' from the business units to ensure that the results will be used appropriately. A number of steps must be taken to ensure this happens.

First, there needs to be a cross-functional team involved in implementing a CRM project in the organization. While the technical members on the team play an important role, an executive on the business side, who should also be the project owner and sponsor, should head the team. Second, processes need to be adopted, with an appropriate set of measurable metrics, to ensure that all steps for project success are being taken. Finally, incentives for performing well on the project should be included as part of the reward structure to ensure motivation.

4.4 Incremental Introduction of CRM

Introducing CRM into an organization must be managed carefully. Given its high initial cost, and significant change on the organization's processes, it is quite possible that insufficient care in its introduction leads to high expense, seemingly small early benefits, which can lead to low morale and excessive finger-pointing.
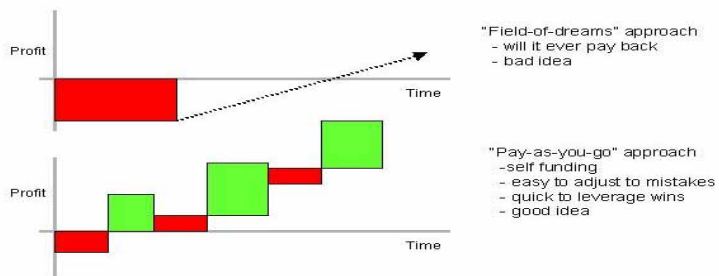

Figure 10. Incremental approach to CRM adoption

As shown in Figure 10, courtesy of [Fors2000], it is better to have an incremental 'pay-as-you-go' approach rather than a 'field-of-dreams' approach. The benefits accrued from the first stage become evident, and act as a catalyst for accelerating the subsequent stages. This makes the choice of the first project and its team very critical. Ideally, the project must be in a potentially high-impact area, where the current process is very ineffective. The ideal (cross-functional) team should have

enthusiastic members, who are committed, and are also seen as leaders in their respective parts of the organization. This will make the dissemination of the successes much easier.

5   Conclusion

The Internet has emerged as a low cost, low latency and high bandwidth customer communication channel. In addition, its interactive nature provides an organization the ability to enter into a close, personalized dialog with its individual customers. The simultaneous maturation of data management technologies like data warehousing, and analysis technologies like data mining, have created the ideal environment for making customer relationship management a much more systematic effort than it has been in the past. While there has been a significant growth of software vendors providing CRM software, and of using them, the focus so far has largely been on the 'relationship management' part of CRM rather than on the 'customer understanding' part. Thus, CRM functions such as e-mail based campaigns management; on-line ads, etc. are being adopted quickly. However, ensuring that the right message is being delivered to the right person, that multiple messages being delivered at different times and through different channels are consistent, is still in a nascent stage. This is often leading to a situation where the best customers are being over communicated to, while insufficient attention is being paid to develop new ones into the best customers of the future.

In this paper we have described how Analytical CRM can fill the gap. Specifically, we described how data analytics can be used to make various CRM functions like customer segmentation, communication targeting, retention, and loyalty much more effective. Our hope is that the data mining community will address the analytics problems in this important and interesting application domain.

6   References

[Cabe1998] Peter Cabena, Pablo Hadjinian, Rolf Stadler, Jaap Verhees, Alessandro Zanasi, *Discovering Data Mining: From Concept to Implementation*, Prentice Hall, 1998.

[Fors1998] Richard Forsyth, "Customer Relationship Marketing Requirements Definition Workshop," http://www.crm-forum.com/library/pre/pre-014/brandframe.html

[HMS00] Hand, David J., Heikki Mannila, Padhraic Smythe, "Principles of Data Mining" MIT Press, 2000.

[Heyg2001] Richard Heygate, "How to Build Valuable Customer Relationships", http://www.crm-forum.com/library/sophron/sophron-022/brandframe.html

[Fors2000] Richard Forsyth, "Avoiding Post-Implementation Blues Managing the Skills", http://www.crm-forum.com/library/pre/pre-025/brandframe.html

[Pyle99] Pyle, Dorian, "Data Preparation for Data Mining", Morgan Kaufmann Publishers, 1999, ISBN No. 1558605290

[Rud00] Rud, Olivia C., "Data Mining, Cookbook: Modeling Data for Marketing, Risk and Customer Relationship Management", John Wiley and Sons, 2000.

[Silvon] Silvon Software, "The Bottom Line of CRM: Know your Customer", http://www.crm2day.com/library/wp/wp0032.shtml

[Stev1998] Peter Stevens, "Analysis and Communication - Two Sides of the Same Coin", http://www.crm-forum.com/library/sophron/sophron-003/brandframe.html

[Stev1999] Peter Stevens, and John Hegarty, "CRM and Brand Management - do they fit together?," http://www.crm-forum.com/library/sophron/sophron-002/brandframe.html

[Swif] Ronald S. Swift, "Analytical CRM Powers Profitable Relationships – Creating Success by Letting Customers Guide You", http://www.crm-forum.com/library/ncr/ncr-073/ncr-073.html.

[Think] thinkAnalytics, "The Hidden World of Data Mining", http://www.crm-forum.com/library/ven/ven-051/ven-051.html

[Michigan] University of Michigan Business School Study, "American Customer Satisfaction Index", 2000.