13

Marketing models

Marketing decision processes are characterized by a high level of complexity due to the simultaneous presence of multiple objectives and countless alternative actions resulting from the combination of the major choice options available to decision makers. Therefore, it should come as no surprise that a large number of mathematical models for marketing have been successfully developed and applied in recent decades. The importance of mathematical models for marketing has been further strengthened by the availability of massive databases of sales transactions that provide accurate information on how customers make use of services or purchase products.

This chapter will primarily focus on two prominent topics in the field of *marketing intelligence*. The first theme is particularly broad and concerns the application of predictive models to support relational marketing strategies, whose purpose is to customize and strengthen the relationship between a company and its customers. After a brief introduction to relational marketing, we will describe the main streams of analysis that can be dealt with in this domain of application, indicating for each of them the classes of predictive models that are best suited to dealing with the problems considered. The subjects discussed in this context can be partly extended to the relationship between citizens and the public administration.

The second theme concerns *salesforce management*. First, we will provide an overview of the major decision-making processes emerging in the organization of a sales staff, highlighting also the role played by response functions. Then, we will illustrate some optimization models which aim to allocate a set of geographical territories to sales agents as well as planning the activities of sales agents. Finally, some business cases consisting of applied marketing models will be discussed.

Business Intelligence: Data Mining and Optimization for Decision Making *Carlo Vercellis*
 ISBN: 978-0-470-51138-1

13.1 Relational marketing

In order to fully understand the reasons why enterprises develop relational marketing initiatives, consider the following three examples: an insurance company that wishes to select the most promising market segment to target for a new type of policy; a mobile phone provider that wishes to identify those customers with the highest probability of churning, that is, of discontinuing their service and taking out a new contract with a competitor, in order to develop targeted retention initiatives; a bank issuing credit cards that needs to identify a group of customers to whom a new savings management service should be offered. These situations share some common features: a company owning a massive database which describes the purchasing behavior of its customers and the way they make use of services, wishes to extract from these data useful and accurate knowledge so as to develop targeted and effective marketing campaigns.

The aim of a *relational marketing* strategy is to initiate, strengthen, intensify and preserve over time the relationships between a company and its stakeholders, represented primarily by its customers, and involves the analysis, planning, execution and evaluation of the activities carried out to pursue these objectives.

Relational marketing became popular during the late 1990s as an approach to increasing customer satisfaction in order to achieve a sustainable competitive advantage. So far, most enterprises have taken at least the first steps in this direction, through a process of cultural change which directs greater attention toward customers, considering them as a formidable asset and one of the main sources of competitive advantage. A relational marketing approach has been followed in a first stage by service companies in the financial and telecommunications industries, and has later influenced industries such as consumer goods, finally reaching also manufacturing companies, from automotive and commercial vehicles to agricultural equipments, traditionally more prone to a vision characterized by the centrality of products with respect to customers.

13.1.1 Motivations and objectives

The reasons for the spread of relational marketing strategies are complex and interconnected. Some of them are listed below, although for additional information the reader is referred to the suggested references at the end of the chapter.

- The increasing concentration of companies in large enterprises and the resulting growth in the number of customers have led to greater complexity in the markets.
- Since the 1980s, the innovation–production–obsolescence cycle has progressively shortened, causing a growth in the number of customized

options on the part of customers, and an acceleration of marketing activities by enterprises.

- The increased flow of information and the introduction of e-commerce have enabled global comparisons. Customers can use the Internet to compare features, prices and opinions on products and services offered by the various competitors.

- Customer loyalty has become more uncertain, primarily in the service industries, where often filling out an on-line form is all one has to do to change service provider.

- In many industries a progressive commoditization of products and services is taking place, since their quality is perceived by consumers as equivalent, so that differentiation is mainly due to levels of service.

- The systematic gathering of sales transactions, largely automated in most businesses, has made available large amounts of data that can be transformed into knowledge and then into effective and targeted marketing actions.

- The number of competitors using advanced techniques for the analysis of marketing data has increased.

Relational marketing strategies revolve around the choices shown in Figure 13.1, which can be effectively summarized as formulating for each segment, ideally

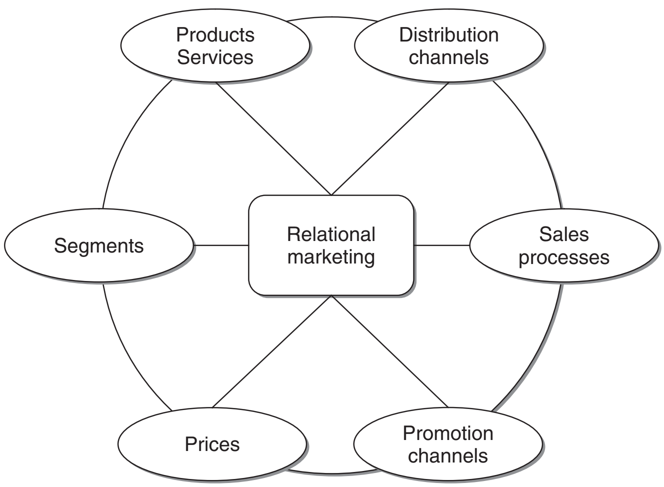

Figure 13.1 Decision-making options for a relational marketing strategy

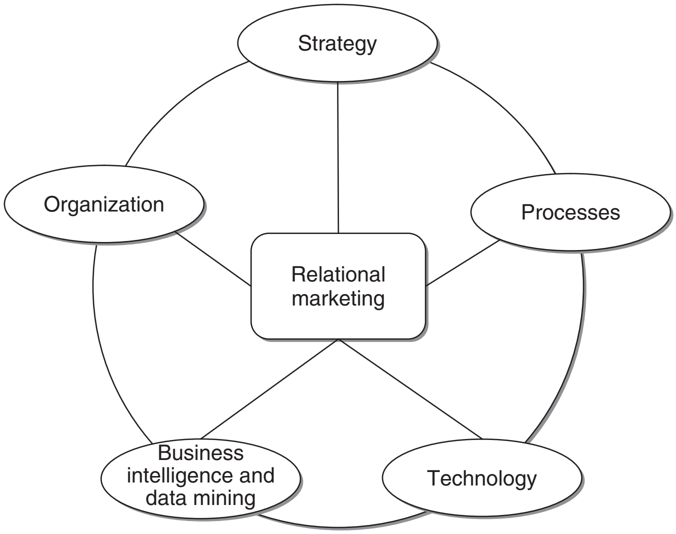

Figure 13.2 Components of a relational marketing strategy

for each customer, the appropriate offer through the most suitable channel, at the right time and at the best price.

The ability to effectively exploit the information gathered on customers' behavior represents today a powerful competitive weapon for an enterprise. A company capable of gathering, storing, analyzing and understanding the huge amount of data on its customers can base its marketing actions on the knowledge extracted and achieve sustainable competitive advantages. Enterprises may profitably adopt relational marketing strategies to transform occasional contacts with their customers into highly customized long-term relationships. In this way, it is possible to achieve increased customer satisfaction and at the same time increased profits for the company, attaining a win–win relationship.

To obtain the desired advantages, a company should turn to relational marketing strategies by following a correct and careful approach. In particular, it is advisable to stress the distinction between a relational marketing vision and the software tools usually referred to as *customer relationship management* (CRM). As shown in Figure 13.2, relational marketing is not merely a collection of software applications, but rather a coherent project where the various company departments are called upon to cooperate and integrate the managerial culture and human resources, with a high impact on the organizational structures. It is then necessary to create within a company a true *data culture*, with the awareness that customer-related information should be enhanced through the adoption of business intelligence and data mining analytical tools.

Based on the investigation of cases of excellence, it can be said that a successful relational marketing strategy can be achieved through the development of a company-wide vision that puts customers at the center of the whole organization. Of course, this goal cannot be attained by exclusively relying on innovative computer technologies, which at most can be considered a relevant enabling factor.

The overlap between relational marketing strategies and CRM software led to a misunderstanding with several negative consequences. On one hand, the notion that substantial investments in CRM software applications were in themselves sufficient to generate a relational marketing strategy represents a dangerous simplification, which caused many project failures. On the other hand, the high cost of software applications has led many to believe that a viable approach to relational marketing was only possible for large companies in the service industries. This is a deceitful misconception: as a matter of fact, the essential components of relational marketing are a well-designed and correctly fed marketing data mart, a collection of business intelligence and data mining analytical tools, and, most of all, the cultural education of the decision makers. These tools will enable companies to carry out the required analyses and translate the knowledge acquired into targeted marketing actions.

The relationship system of an enterprise is not limited to the dyadic relationship with its customers, represented by individuals and companies that purchase the products and services offered, but also includes other actors, such as the employees, the suppliers and the sales network. For most relationships shown

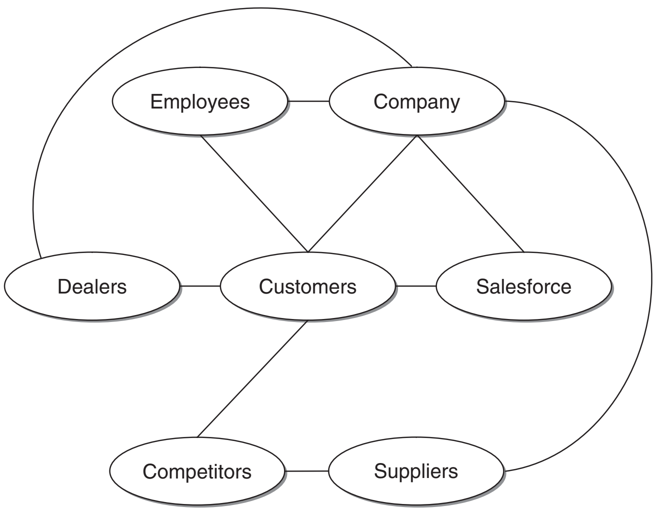

Figure 13.3 Network of relationships involved in a relational marketing strategy

in Figure 13.3, a mutually beneficial exchange occurs between the different subjects involved. More generally, we can widen the boundaries of relational marketing systems to include the stakeholders of an enterprise. The relationship between an enterprise and its customers is sometimes mediated by the sales network, which in some instances can partially obstruct the visibility of the end customers.

Let us take a look at a few examples to better understand the implications of this issue. The manufacturers of consumer goods, available at the points of sale of large and small retailers, do not have direct information on the consumers purchasing their products. The manufacturers of goods covered by guarantees, such as electrical appliances or motor vehicles, have access to personal information on purchasers, even if they rarely also have access to information on the contacts of and promotional actions carried out by the network of dealers. Likewise, a savings management company usually places shares in its investment funds through a network of intermediaries, such as banks or agents, and often knows only the personal data of the subscribers. A pharmaceutical enterprise producing prescription drugs usually ignores the identity of the patients that use its drugs and medicinal products, even though promotional activities to influence consumers are carried out in some countries where the law permits.

It is not always easy for a company to obtain information on its end customers from dealers in the sales network and even from their agents. These may be reluctant to share the wealth of information for fear, rightly or wrongly, of compromising their role. In a relational marketing project specific initiatives should be devised to overcome these cultural and organizational barriers, usually through incentives and training courses.

The number of customers and their characteristics strongly influence the nature and intensity of the relationship with an enterprise, as shown in Figure 13.4. The relationships that might actually be established in a specific economic domain tend to lie on the diagonal shown in the figure. At one extreme, there are highly intense relationships existing between the company and a small number of customers of high individual value. Relationships of this type occur more frequently in *business-to-business* (B2B) activities, although they can also be found in other domains, such as private banking. The high value of each customer justifies the use of dedicated resources, usually consisting of sales agents and key account managers, so as to maintain and strengthen these more intense relationships. In situations of this kind, careful organization and planning of the activities of sales agents is critical. Therefore, optimization models for *salesforce automation* (SFA), described in Section 13.2, can be useful in this context.

At the opposite extreme of the diagonal are the relationships typical of consumer goods and *business-to-consumer* (B2C) activities, for which a high number of low-value customers get in contact with the company in an impersonal

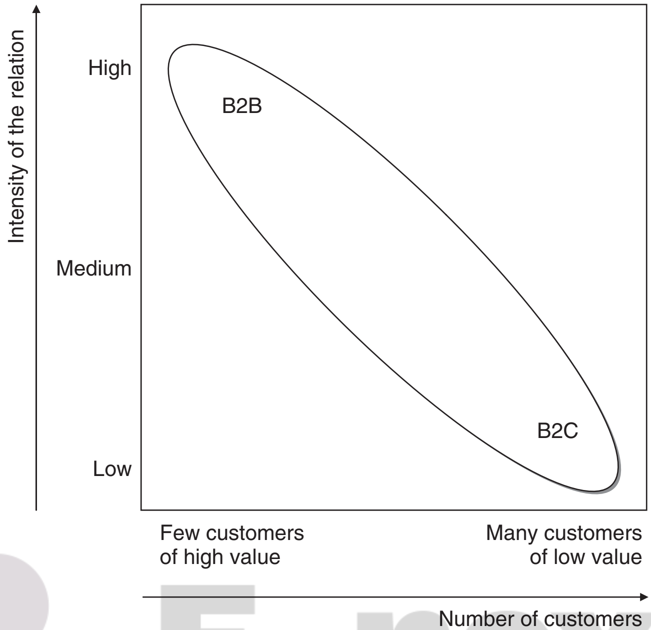

Figure 13.4 Intensity of customer relationships as a function of number of customers

way, through websites, call centers and points of sale. Data mining analyses for segmentation and profiling are particularly valuable especially in this context, characterized by a large number of fragmented contacts and transactions. Relational marketing strategies, which are based on the knowledge extracted through data mining models, enable companies develop a targeted customization and differentiation of their products and/or services, including companies more prone toward a mass-market approach.

Figure 13.5 contrasts the cost of sales actions and the corresponding revenues. Where transactions earn a low revenue per unit, it is necessary to implement low-cost actions, as in the case of mass-marketing activities. Moving down along the diagonal in the figure, more evolved and intense relationships with the customers can be found. The relationships at the end of the diagonal presuppose the action of a direct sales network and for the most part are typical of B2B relational contexts.

Figure 13.6 shows the ideal path that a company should follow so as to be able to offer customized products and services at low cost and in a short time. On the one hand, companies operating in a mass market, well acquainted with fast delivery at low costs, must evolve in the direction of increased customization, by introducing more options and variants of products and services offered

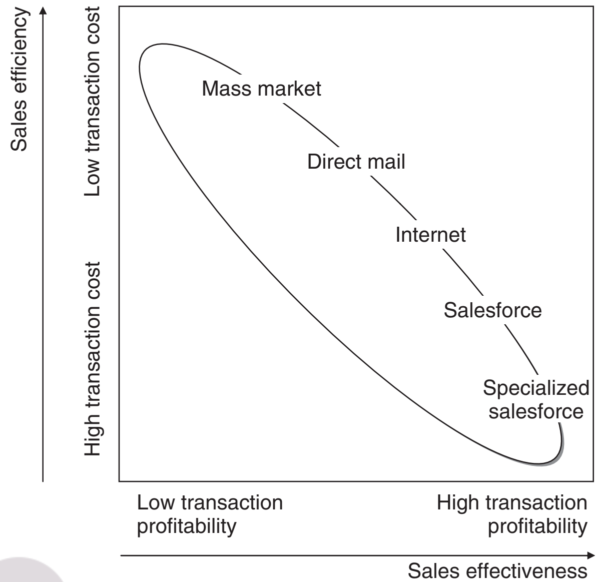

Figure 13.5 Efficiency of sales actions as a function of their effectiveness

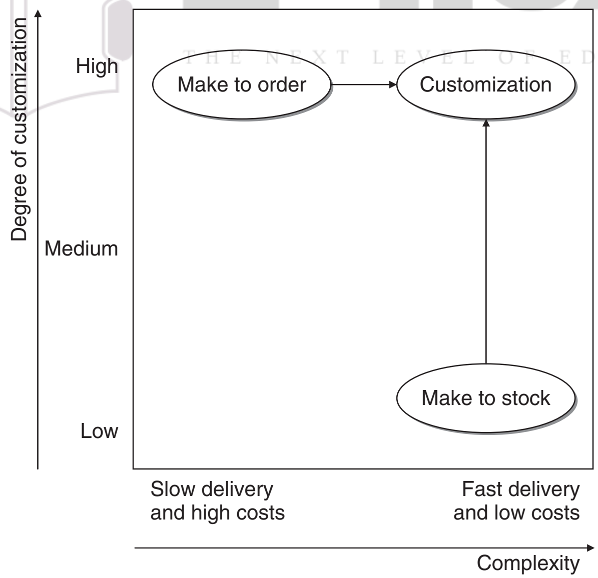

Figure 13.6 Level of customization as a function of complexity of products and services

to the various market segments. Data mining analyses for relational marketing purposes are a powerful tool for identifying the segments to be targeted with customized products. On the other hand, the companies oriented toward make-to-order production must evolve in a direction that fosters reductions in both costs and delivery times, but without reducing the variety and the range of their products.

13.1.2 An environment for relational marketing analysis

Figure 13.7 shows the main elements that make up an environment for relational marketing analysis. Information infrastructures include the company's data warehouse, obtained from the integration of the various internal and external data sources, and a marketing data mart that feeds business intelligence and data mining analyses for profiling potential and actual customers. Using pattern recognition and machine learning models as described in previous chapters, it is possible to derive different segmentations of the customer base, which are then used to design targeted and optimized marketing actions. A classification model can be used, for example, to generate a *scoring* system for customers according to their propensity to buy a service offered by a company, and to direct a cross-selling offer only toward those customers for whom a high probability of acceptance is predicted by the model, thus maximizing the overall *redemption* of the marketing actions.

Effective management of frequent marketing campaign cycles is certainly a complex task that requires planning, for each segment of customers, the content of the actions and the communication channels, using the available

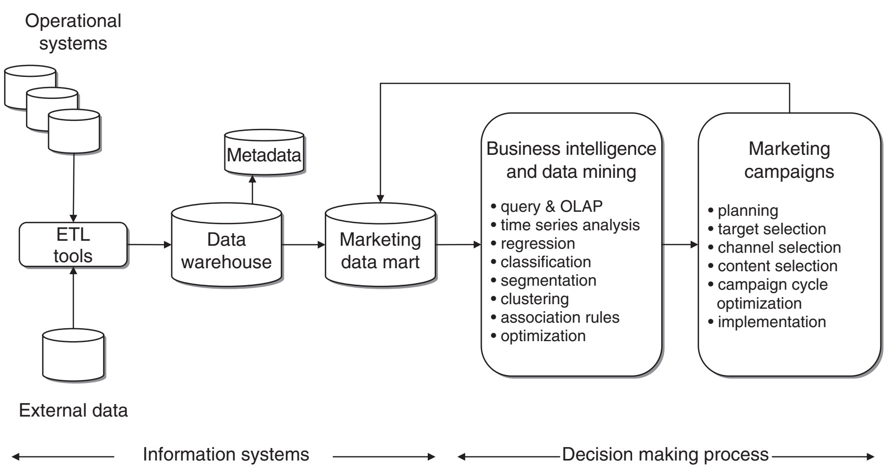

Figure 13.7 Components of an environment for relational marketing analysis

human and financial resources. The corresponding decision-making process can be formally expressed by appropriate optimization models. The cycle of marketing activities terminates with the execution of the planned campaign, with the subsequent gathering of information on the results and the redemption among the recipients. The data collected are then fed into the marketing data mart for use in future data mining analyses. During the execution of each campaign, it is important to set up procedures for controlling and analyzing the results obtained. In order to assess the overall effectiveness of a campaign, it would be advisable to select a *control group* of customers, with characteristics similar to those of the campaign recipients, toward whom no action should be undertaken.

Figure 13.8 describes the main types of data stored in a data mart for relational marketing analyses. A company data warehouse provides demographic and administrative information on each customer and the transactions carried out for purchasing products and using services. The marketing database contains data on initiatives carried out in the past, including previous campaigns and their results, promotions and advertising, and analyses of customer value. A further possible data source is the salesforce database, which provides information on established contacts, calls and applicable sales conditions. Finally, the *contact center* database provides access to data on customers' contacts with the call center, problems reported, sometimes called *trouble tickets*, and

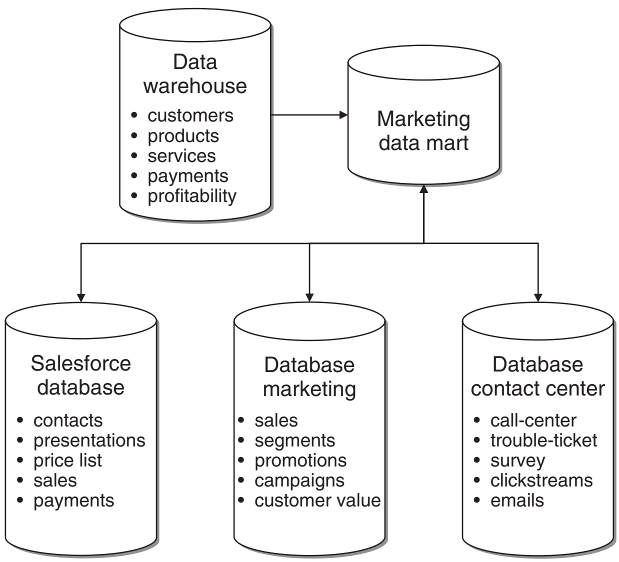

Figure 13.8 Types of data feeding a data mart for relational marketing analysis

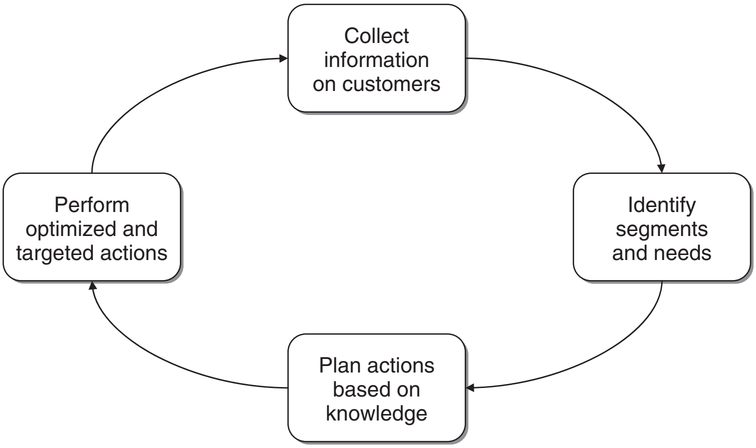

Figure 13.9 Cycle of relational marketing analysis

related outcomes, website navigation paths and forms filled out on-line, and emails exchanged between customers and the support center.

As shown in Figure 13.8, the available data are plentiful, providing an accurate representation of the behaviors and needs of the different customers, through the use of inductive learning models.

The main phases of a relational marketing analysis proceeds as shown in Figure 13.9. The first step is the exploration of the data available for each customer. At a later time, by using inductive learning models, it is possible to extract from those data the insights and the rules that allow market segments characterized by similar behaviors to be identified. Knowledge of customer profiles is used to design marketing actions which are then translated into promotional campaigns and generate in turn new information to be used in the course of subsequent analyses.

13.1.3 Lifetime value

Figure 13.10 shows the main stages during the customer *lifetime*, showing the cumulative value of a customer over time. The figure also illustrates the different actions that can be undertaken toward a customer by an enterprise. In the initial phase, an individual is a *prospect*, or potential customer, who has not yet begun to purchase the products or to use the services of the enterprise. Toward potential customers, *acquisition* actions are carried out, both directly (telephone contacts, emails, talks with sales agents) and indirectly (advertising, notices on the enterprise website). These actions incur a cost that can be assigned to each customer and determine an accumulated loss that lasts until a critical event

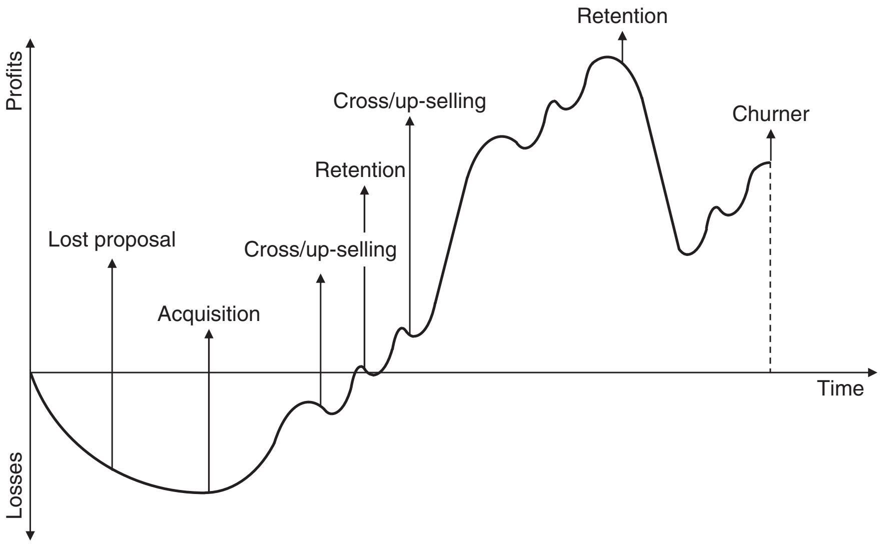

Figure 13.10 Lifetime of a customer

in the relationship with a customer occurs: a prospect becomes a customer. This event may take various forms in different situations: it may consist of a service subscription, the opening of a bank account, the first purchase at a retailer point of sale with the activation of a loyalty card. Before becoming a new customer, a prospect may receive from the enterprise repeated proposals aiming acquiring her custom, shown in the figure as *lost proposals*, which have a negative outcome. From the time of acquisition, each customer generates revenue, which produces a progressive rise along the curve of losses and cumulated profits. This phase, which corresponds to the maturity of the relationship with the enterprise, usually entails alternating *cross-selling*, *up-selling* and *retention* actions, in an effort to extend the duration and the profitability of the relationship so as to maximize the lifetime value of each customer. The last event in a customer lifetime is the interruption of the relationship. This may be *voluntary*, when a customer discontinues the services of an enterprise and switches to those of a competitor, *forced*, when for instance a customer does not comply with payment terms, or *unintentional*, when for example a customer changes her place of residence.

The progress of a customer lifetime highlights the main tasks of relational marketing. First, the purpose is to increase the ability to acquire new customers. Through the analysis of the available information for those customers who in the past have purchased products or services, such as personal socio-demographic characteristics, purchased products, usage of services, previous contacts, and the comparison with the characteristics of those who have not taken up the offers of the enterprise, it is possible to identify the segments with the highest potential. This in turn allows the enterprise to optimize marketing campaigns, to increase the effectiveness of acquisition initiatives and to reduce the waste of resources due to offers addressed to unpromising market segments.

Furthermore, relational marketing strategies can improve the loyalty of customers, extending the duration of their relationship with the enterprise, and thus increasing the profitability. In this case, too, the comparative analysis of the characteristics of those who have remained loyal over time with respect to those who have switched to a competitor leads to predictions of the likelihood of churning for each customer. Retention actions can therefore be directed to the most relevant segments, represented by high-value customers with the highest risk of churning.

Finally, relational marketing analyses can be used to identify customers who are more likely to take up the offer of additional services and products (*cross-selling*), or of alternative services and products of a higher level and with a greater profitability for the enterprise (*up-selling*).

The tasks of acquisition, retention, cross-selling and up-selling, shown in Figure 13.11, are at the heart of relational marketing strategies and their aim is to maximize the profitability of customers during their lifetime. These analysis tasks, which will be described in the next sections, are clearly amenable to classification problems with a binary target class. Notice that attribute selection plays a critical role in this context, since the number of available explanatory variables is usually quite large and it is advisable for learning models to use a

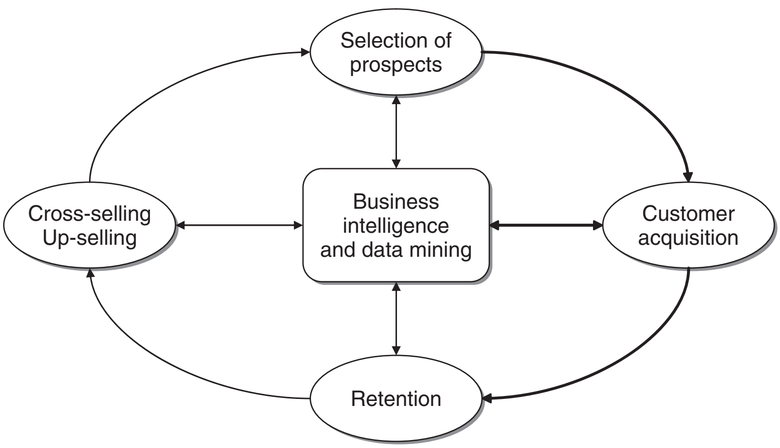

Figure 13.11 Main relational marketing tasks

limited subset of predictive features, in order to generate meaningful and useful classification rules for the accurate segmentation of customers.

13.1.4 The effect of latency in predictive models

Figure 13.12 illustrates the logic of development of a classification model for a relational marketing analysis, also taking into account the temporal dimension. Assume that t is the current time period, and that we wish to derive an inductive learning model for a classification problem. For example, at the beginning of October a mobile telephone provider might want to develop a classification model to predict the probability of churning for its customers. The data mart contains the data for past periods, updated as far as period $t-1$. In our example, it contains data up to and including September.

Furthermore, suppose that the company wishes to predict the probability of churning h months in advance, since in this way any retention action has a better chance of success. In our example, we wish to predict at the beginning of October the probability of churning in November, using data up to September. Notice that the data for period t cannot be used for the prediction, since they are clearly not available at the beginning of period t.

To develop a classification model we use the value of the target variable for the last known period $t-1$, corresponding to the customers who churned in the month of September. It should be clear that for training and testing the model the explanatory variables for period $t-2$ should not be used, since in the training phase it is necessary to reproduce the same situation as will be faced when using the model in the prediction stage. Actually, the target

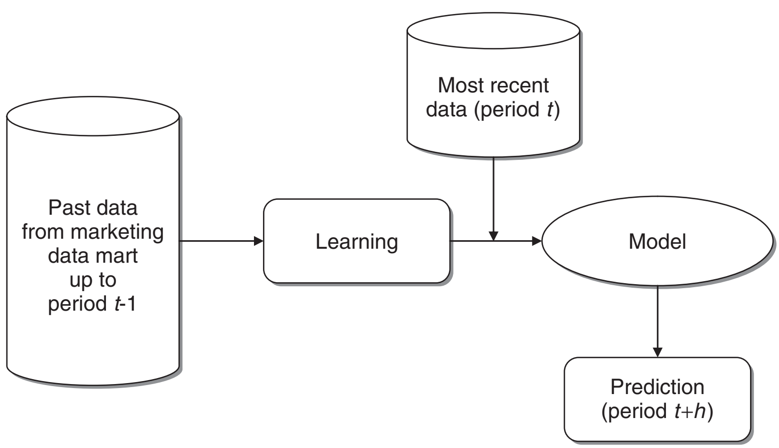

Figure 13.12 Development and application flowchart for a predictive model

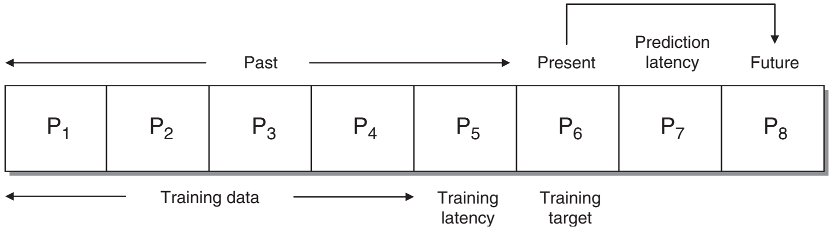

Figure 13.13 Latency of a predictive model

variable must be predicted $h = 2$ periods in advance, and therefore there is an intermediate period of future data that are still unknown at time t (the month of October in our example). To reflect these dynamics, the training phase should be carried out without using August data. In general, the $h - 1$ periods corresponding to data still unknown during the prediction phase, and not used during the training phase, are referred to as the model's *latency*, as shown in Figure 13.13.

13.1.5 Acquisition

Although retention plays a prominent role in relational marketing strategies, for many companies the *acquisition* of new customers also represents a critical factor for growth. The acquisition process requires the identification of new prospects, as they are potential customers who may be totally or partially unaware of the products and services offered by the company, or did not possess in the past the characteristics to become customers, or were customers of competitors. It may also happen that some of the prospects were former customers who switched their custom to competitors, in which case much more information is usually available on them.

Once prospects have been identified, the enterprise should address acquisition campaigns to segments with a high potential profitability and a high probability of acquisition, in order to optimize the marketing resources. Traditional marketing techniques identify interesting segments using predefined profiling criteria, based on market polls and socio-demographic analyses, according to a top-down perspective. This approach can be successfully integrated or even replaced by a top-down segmentation logic which analyzes the data available in the data mart, as shown in Figure 13.8 (demographic information, contacts with prospects, use of products and services of competitors), and derives classification rules that characterize the most promising profiles for acquisition purposes. Also in this case, we are faced with a binary classification problem, which can be analyzed with the techniques described in Chapter 10.

13.1.6 Retention

The maturity stage reached by most products and services, and the subsequent saturation of their markets, have caused more severe competitive conditions. As a consequence, the expansion of the customer base of an enterprise consists more and more of switch mechanisms – the acquisition of customers at the expense of other companies. This phenomenon is particularly apparent in service industries, such as telecommunications, banking, savings management and insurance, although it also occurs in manufacturing, both for consumer goods and industrial products. For this reason, many companies invest significant amounts of resources in analyzing and characterizing the phenomenon of *attrition*, whereby customers switch from their company to a competitor.

There are also economic reasons for devoting substantial efforts to customer *retention*: indeed, it has been empirically observed that the cost of acquiring a new customer, or winning back a lost customer, is usually much higher – of the order of 5 to 9 times higher – than the cost of the marketing actions aimed at retaining customers considered at risk of churning. Furthermore, an action to win back a lost customer runs the risk of being too late and not achieving the desired result. In many instances, winning back a customer requires investments that do not generate a return.

One of the main difficulties in loyalty analysis is actually recognizing a churn. For subscription services there are unmistakable signals, such as a formal notice of withdrawal, while in other cases it is necessary to define adequate indicators that are correlated, a few periods in advance, with the actual churning. A customer who reduces by more than a given percentage her purchases at a selected point of sale using a loyalty card, or a customer who reduces below a given threshold the amount held in her checking account and the number of transactions, represent two examples of disaffection indicators. They also highlight the difficulties involved in correctly defining the appropriate threshold values.

To optimize the marketing resources addressed to retention, it is therefore necessary to target efforts only toward high-value customers considered at risk of churning. To obtain a scoring system corresponding to the probability of churning for each customer, it is necessary to derive a segmentation based on the data on past instances of churning. Predicting the risk of churning requires analysis of records of transactions for each customer and identifying the attributes that are most relevant to accurately explaining the target variable. Again, we are faced with a binary classification problem. Once the customers with the highest risk of churning have been identified, a retention action can be directed toward them. The more accurately such action is targeted, the cheaper it is likely to be.

13.1.7 Cross-selling and up-selling

Data mining models can also be used to support a relational marketing analysis aimed at identifying market segments with a higher propensity to purchase additional services or other products of a company. For example, a bank also offering insurance services may identify among its customers segments interested in purchasing a life insurance policy. In this case, demographic information on customers and their past transactions recorded in a data mart can be used as explanatory attributes to derive a classification model for predicting the target class, consisting in this example of a binary variable that indicates whether the customer accepted the offer or not.

The term *cross-selling* refers to the attempt to sell an additional product or service to an active customer, already involved in a long-lasting commercial relationship with the enterprise. By means of classification models, it is possible to identify the customers characterized by a high probability of accepting a cross-selling offer, starting from the information contained in the available attributes.

In other instances, it is possible to develop an *up-selling* initiative, by persuading a customer to purchase an higher-level product or service, richer in functions for the user and more profitable for the company, and therefore able to increase the lifetime value curve of a customer. For example, a bank issuing credit cards may offer customers holding a standard card an upgrade to a gold card, which is more profitable for the company, but also able to offer a series of complementary services and advantages to interested customers. In this case too, we are dealing with a binary classification problem, which requires construction of a model based on the training data of customers' demographic and operational attributes. The purpose of the model is to identify the most interesting segments, corresponding to customers who have taken up the gold service in the past, and who appear therefore more appreciative of the additional services offered by the gold card. The segments identified in this way represent the target of up-selling actions.

13.1.8 Market basket analysis

The purpose of *market basket analysis* is to gain insight from the purchases made by customers in order to extract useful knowledge to plan marketing actions. It is mostly used to analyze purchases in the retail industry and in e-commerce activities, and is generally amenable to unsupervised learning problems. It may also be applied in other domains to analyze the purchases made using credit cards, the complementary services activated by mobile or fixed telephone customers, the policies or the checking accounts acquired by a same household.

The data used for this purpose mostly refer to purchase transactions, and can be associated with the time dimension if the purchaser can be tracked through a loyalty card or the issue of an invoice. Each transaction consists of a list of purchased items. This list is called a *basket*, just like the baskets available at retail points of sale.

If transactions cannot be connected to one another, say because the purchaser is unknown, one may then apply association rules, described in Chapter 11, to extract interesting correlations between the purchases of groups of items. The rules extracted in this way can then be used to support different decision-making processes, such as assigning the location of the items on the shelves, determining the layout of a point of sale, identifying which items should be included in promotional flyers, advertisements or coupons distributed to customers.

Clustering models, described in Chapter 12, are also useful in determining homogeneous groups of items, once an incidence matrix $\mathbf{X}$ has been created for the representation of the dataset, where the rows correspond to the transactions and the columns to the items.

If customers are individually identified and traced, besides the above techniques it is also possible to develop further analyses that take into account the time dimension of the purchases. For instance, one may generate sequential association rules, mentioned at the end of Chapter 11, or apply time series analysis, as described in Chapter 9.

13.1.9 Web mining

The web is a critical channel for the communication and promotion of a company's image. Moreover, e-commerce sites are important sales channels. Hence, it is natural to use *web mining* methods in order to analyze data on the activities carried out by the visitors to a website.

Web mining methods are mostly used for three main purposes, as shown in Figure 13.14: *content mining*, *structure mining* and *usage mining*.

Content mining. Content mining involves the analysis of the content of web pages to extract useful information. Search engines primarily perform content mining activities to provide the links deemed interesting in relation to keywords supplied by users. Content mining methods can be traced back to data mining problems for the analysis of texts, both in free format or HTML and XML formats, images and multimedia content. Each of these problems is in turn dealt with using the learning models described in previous chapters. For example, text mining analyses are usually handled as multicategory classification problems, where the target variable is the subject category to which the text refers, while explanatory variables correspond to the meaningful words contained in the text.

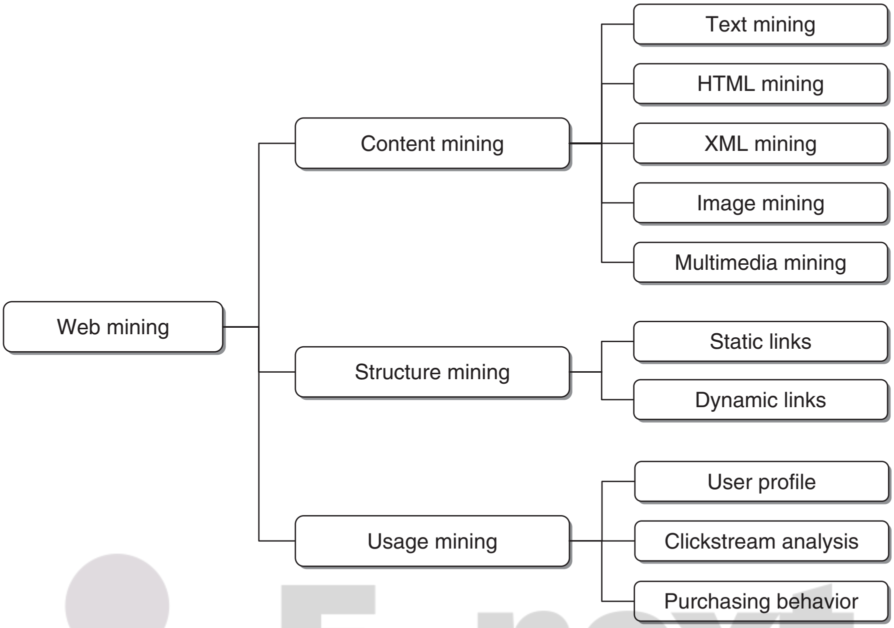

Figure 13.14 Taxonomy of web mining analyses

Once it has been converted into a classification problem, text mining can be approached using the methods described in Chapter 10. Text mining techniques are also useful for analyzing the emails received by a support center. Notice that the input data for content mining analyses are easily retrievable, at least in principle, since they consist of all the pages that can be visited on the Internet.

Structure mining. The aim of this type of analysis is to explore and understand the topological structure of the web. Using the links presented in the various pages, it is possible to create graphs where the nodes correspond to the web pages and the oriented arcs are associated with links to other pages. Results and algorithms from graph theory are used to characterize the structure of the web, that is, to identify areas with a higher density of connections, areas disconnected from others and maximal *cliques*, which are groups of pages with reciprocal links. In this way, it is possible to pinpoint the most popular sites, or to measure the distance between two sites, expressed in terms of the lowest number of arcs along the paths that connect them in the links graph. Besides analyses aimed at exploring the *global* structure of the web, it is also possible to carry out *local* investigations to study how a single website is articulated. In some investigations, the local structure of websites is associated with the time spent by the users on each page, to verify if the organization of the site suffers from inconsistencies that jeopardize its effectiveness. For example, a page whose purpose is to direct navigation on the site should be viewed by each user only briefly. Should this not be the case, the page has a problem due to a possible ambiguity in the articulation of the links offered.

Usage mining. Analyses aimed at *usage mining* are certainly the most relevant from a relational marketing standpoint, since they explore the paths followed by navigators and their behaviors during a visit to a company website. Methods for the extraction of association rules are useful in obtaining correlations between the different pages visited during a session. In some instances, it is possible to identify a visitor and recognize her during subsequent sessions. This happens if an identification key is required to access a web page, or if a cookie-enabling mechanism is used to keep track of the sequence of visits. Sequential association rules or time series models can be used to analyze the data on the use of a site according to a temporal dynamic. Usage mining analysis is mostly concerned with *clickstreams* – the sequences of pages visited during a given session. For e-commerce sites, information on the purchase behavior of a visitor is also available.

13.2 Salesforce management

Most companies have a sales network and therefore rely on a substantial number of people employed in sales activities, who play a critical role in the profitability of the enterprise and in the implementation of a relational marketing strategy. The term *salesforce* is generally taken to mean the whole set of people and roles that are involved, with different tasks and responsibilities, in the sales process. A preliminary taxonomy of salesforces is based on the type of activity carried out, as indicated below.

Residential. *Residential* sales activities take place at one or more sites managed by a company supplying some products or services, where customers go to make their purchases. This category includes sales at retail outlets as well as wholesale trading centers and *cash-and-carry* shops.

Mobile. In *mobile* sales, agents of the supplying company go to the customers' homes or offices to promote their products and services and collect orders. Sales in this category occur mostly within B2B relationships, even though they can also be found in B2C contexts.

Telephone. *Telephone* sales are carried out through a series of contacts by telephone with prospective customers.

There are various problems connected with managing a mobile salesforce management, which will be the main focus of this section. They can be subdivided into a few main categories:

- designing the sales network;
- planning the agents' activities;
- contact management;
- sales opportunity management;
- customer management;
- activity management;
- order management;
- area and territory management;
- support for the configuration of products and services;
- knowledge management with regard to products and services.

Designing the sales network and planning the agents' activities involve decision-making tasks that may take advantage of the use of optimization models, such as those that will be described in the next sections. The remaining activities are operational in nature and may benefit from the use of software tools for *salesforce automation* (SFA), today widely implemented.

13.2.1 Decision processes in salesforce management

The design and management of a salesforce raise several decision-making problems, as shown in Figure 13.15. When successfully solved, they confer multiple advantages: maximization of profitability, increased effectiveness of sales actions, increased efficiency in the use of resources, and greater professional rewards for sales agents.

The decision processes described in Figure 13.15 should take into account the strategic objectives of the company, with respect to other components of the marketing mix, and conform to the role assigned to the salesforce within the broader framework of a relational marketing strategy. The two-way connections indicated in the figure suggest that the different components of the decision-making process interact with each other and with the general objectives of the marketing department.

In particular, the decision-making processes relative to salesforce management can be grouped into three categories: *design*, *planning* and *assessment*.

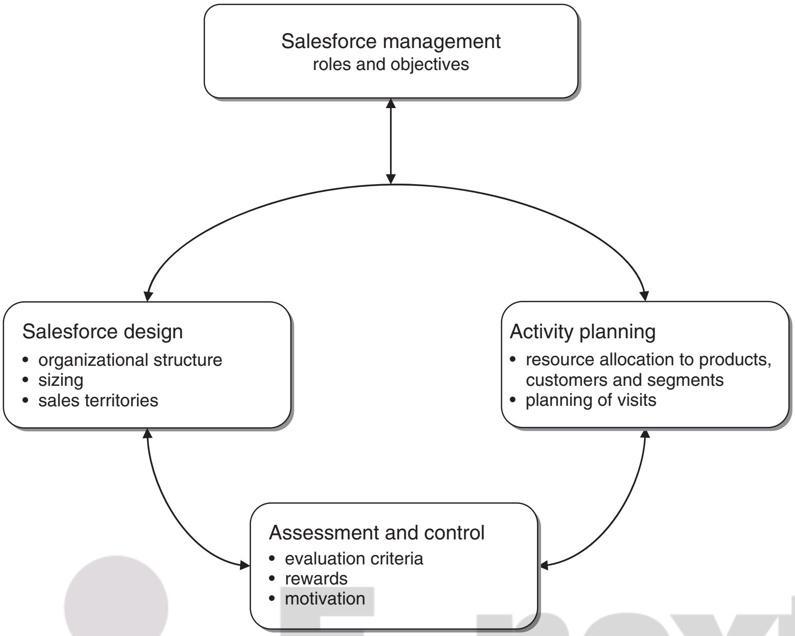

Figure 13.15 Decision processes in salesforce management

Design

Salesforce design is dealt with during the start-up phase of a commercial activity or during subsequent restructuring phases, for example following the merger or acquisition of a group of companies.

As shown in Figure 13.16, the design phase is usually preceded by the creation of market segments through the application of data mining methods and by the articulation of the offer of products and services, which are in turn subdivided into homogeneous classes. Salesforce design includes three types of decisions.

Organizational structure. The *organizational structure* may take different forms, corresponding to hierarchical agglomerations of the agents by group of products, brand or geographical area. In some situations the structure may also be differentiated by markets. In order to determine the organizational structure, it is necessary to analyze the complexity of customers, products and sales activities, and to decide whether and to what extent the agents should be specialized.

Sizing. Sales network *sizing* is a matter of working out the optimal number of agents that should operate within the selected structure, and depends on several factors, such as the number of customers and prospects, the desired level of

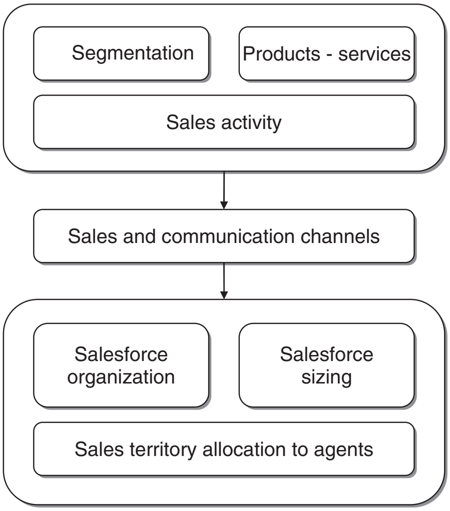

Figure 13.16 Salesforce design process

sales area coverage, the estimated time for each call and the agents' traveling time. One should bear in mind that a reduction in costs due to a decrease in the salesforce size is often followed by a reduction in sales and revenues. A better allocation of the existing salesforce, devised during the planning phase by means of optimization models, is usually more effective than a variation in size.

Sales territories. Designing a *sales territory* means grouping together the geographical areas into which a given region has been divided and assigning each territory to an agent. The design and assignment of sales territories should take into account several factors, such as the sales potential of each geographical area, the time required to travel from one area to another and the total time each agent has available. The purpose of the assignment consists of determining a balanced situation between sales opportunities embedded in each territory, in order to avoid disparities among agents. The assignment of the geographical areas should be periodically reviewed since the sales potential balance in the various territories tends to vary over time.

Decisions concerning the design of the salesforce should take into account decisions about salesforce planning, and this explains the two-way link between the two corresponding blocks in Figure 13.15.

Planning

Decision-making processes for *planning* purposes involve the assignment of sales resources, structured and sized during the design phase, to market entities.

Resources may correspond to the work time of agents or to the budget, while market entities consist of products, market segments, distribution channels and customers.

Allocation takes into account the time spent pitching the sale to each customer, the travel time and cost, and the effectiveness of the action for each product, service or market segment. It is also possible to consider further ancillary activities carried out at the customers' sites, such as making suggestions that are conducive to future sales or explaining the technical and functional features of products and services.

Salesforce planning can greatly benefit from the use of optimization models, as explained below.

Assessment

The purpose of *assessment* and *control* activities is to measure the effectiveness and efficiency of individuals employed in the sales network, in order to design appropriate remuneration and incentive schemes. To measure the efficiency of the sales agents it is necessary to define adequate evaluation criteria that take into account the actual personal contribution of each agent, having removed effects due to the characteristics that may make an area or product more or less advantageous than others. Data envelopment analysis, described in Chapter 15, provides useful models that can be applied to assess agents' performance.

13.2.2 Models for salesforce management

In what follows we will describe some classes of optimization models for designing and planning the salesforce. These models are primarily intended for educational purposes, to familiarize readers with the reasoning behind specific aspects of a sales network, through the formulation of optimization models. For the sake of clarity and conciseness, for each model we have limited the extensions to a single feature. Sales networks simultaneously possess more than one of the distinctive features previously described, and therefore the models developed in real-world applications, just like those described in the last section of the chapter, are more complex and result from a combination of different characteristics.

Before proceeding, it is useful to introduce some notions common to the different models that will be described. Assume that a region is divided into J geographical sales areas, also called *sales coverage units*, and let $\mathcal{J} = \{1, 2, \ldots, J\}$. Areas must be aggregated into disjoint clusters, called *territories*, so that each area belongs to one single territory and is also connected to all the areas belonging to the same territory. The connection property implies that from each area it is possible to reach any other area of the same territory. The time span is divided into T intervals of equal length, which usually correspond to weeks or months, indicated by the index $t \in \mathcal{T} = \{1, 2, \ldots, T\}$.

Each territory is associated with a sales agent, located in one of the areas belonging to the territory, henceforth considered as her area of residence. The choice of the area of residence determines the time and cost of traveling to any other area in the same territory. Let I be the number of territories and therefore the number of agents that form the sales network, and let $\mathcal{I} = \{1, 2, \ldots, I\}$.

In each area there are customers or prospects who can be visited by the agents as part of their promotions and sales activities. In some of the models that will be presented, customers or prospects are aggregated into segments, which are considered homogeneous with respect to the area of residence and possibly to other characteristics, such as value, potential for development and purchasing behaviors. Let H be the number of market entities, which in different models may represent either single customers or segments, and let $\mathcal{H} = \{1, 2, \ldots, H\}$. Let $\mathcal{D}_j$ be the set of customers, or segments of customers where necessary, located in area j.

Finally, assume that a given agent can promote and sell K products and services during the calls she makes on customers or prospects, and let $\mathcal{K} = \{1, 2, \ldots, K\}$.

13.2.3 Response functions

Response functions play a key role in the formulation of models for designing and planning a sales network. In general terms, a response function describes the elasticity of sales in terms of the intensity of the sales actions, and is a formal method to describe the complex relationship existing between sales actions and market reactions.

Sales to which the response function refers are expressed in product units or monetary units, such as revenues or margins. For the sake of uniformity, in the next sections response functions are assumed to be expressed as sales revenues. The intensity of a sales action can be related to different variables, such as the number of calls to a customer in each period, the number of mentions of a product in each period, and the time dedicated to each customer in each period.

In principle, it is possible to consider a response function in relation to each factor that is deemed critical to sales: the characteristics of customers and sales territories; the experience, education and personal skills of the agents; promotions, prices, markdown policies operated by the company and the corresponding features for one or several competitors.

Figures 13.17 and 13.18 show two possible shapes of the response function, obtained by placing the sales of a product or service on the vertical axis and the intensity of the sales action of interest on the horizontal axis. To fix ideas, we

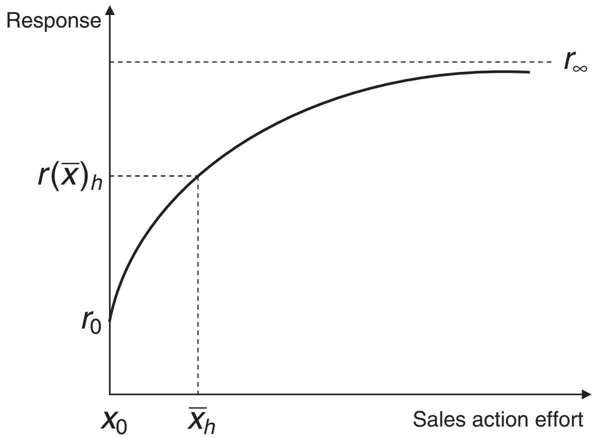

Figure 13.17 A concave response function

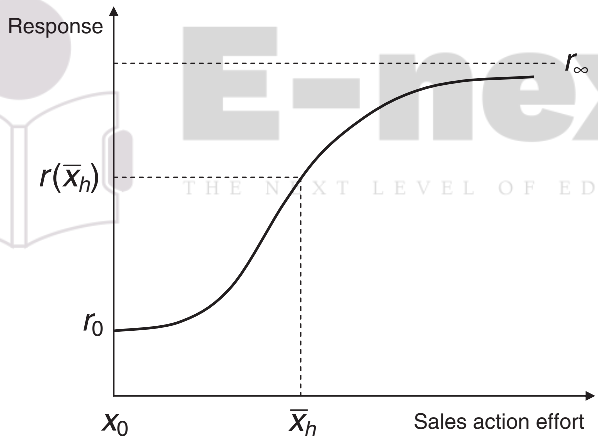

Figure 13.18 A sigmoidal response function

will assume that the number x_h of calls that a specific agent makes to customer h in each period of the planning horizon is placed on the horizontal axis.

The concave response function shown in Figure 13.17 can be interpreted in the following way: as the number of calls increases, revenues grow at a decreasing rate approaching 0, according to the principle of decreasing marginal revenues. In general, a lack of sales actions toward a given customer does not imply a lack of sales, at least for a certain number of periods. This is an effect of the actions executed in previous periods that lasts over time. For this reason the response function is greater than 0 at $x_h = 0$.

The sigmoidal response function in Figure 13.18 reflects a different hypothesis of sales growth as a function of the actions carried out. The assumption made in this case is that the central interval of values on the horizontal axis corresponds to a higher rate of sales growth, while outside that area the growth rate is lower.

It is worth noting that each decision concerning the allocation of sales resources is based on a response function hypothesis, which is implicit and unaware in intuitive decision-making processes, while it is explicit and rigorous in mathematical models such as those presented below.

Response functions can be estimated by considering two types of information. On the one hand, one can use past available data regarding the intensity of the actions carried out and the corresponding sales, to develop a parametric regression model through variants of regression methods. On the other hand, interviews are carried out with agents and sales managers to obtain subjective information which is then incorporated into the procedure for calculating the response function.

We will now show by means of an example how the procedure for estimating the response function works. Let $r_h(x_h)$ be the sales value for customer h associated with a number x_h of calls during a given period. More generally, the variable that determines the response function r expresses the intensity of the sales action that has been carried out. A parametric form should first be selected in order to express the functional dependence. The following function, which may assume both concave and sigmoidal shapes by varying the parameters, can be used:

$$r_h(x_h) = r_0 + (r_\infty - r_0)\frac{x_h^\sigma}{\gamma + x_h^\sigma}. \tag{13.1}$$

The parameters in the expression $r_h(x_h)$ have the following meaning: r_0 represents the sales level that would be obtained at a sales action intensity equal to 0, as a prolonged effect of previous actions; r_∞ represents the maximum sales level, irrespective of the intensity of the sales action; γ and σ are two parameters to be estimated.

To obtain an estimate of the four parameters appearing in the expression for $r_h(x_h)$ it is possible to proceed in two complementary ways. Past sales data can be used to set up a regression model and determine the values through the least squares method. In order to increase the value of the opinions of the sales agents, it is also possible to ask agents and sales managers to estimate the value of the parameters r_0 and r_∞, as well as the values of the expected sales at other three critical points of the response function: $r(\bar{x}_h)$, corresponding to the number of calls carried out at the time of the analysis, $r(\frac{1}{2}\bar{x}_h)$ and $r(\frac{3}{2}\bar{x}_h)$, associated respectively with increasing and decreasing the number of calls by 50% with respect to the current value. Based on a subjective evaluation of the five response values derived through the procedure described above, an estimate by interpolation of the scale parameters γ and σ is then obtained.

13.2.4 Sales territory design

Sales territory design involves allocating sales coverage units to individual agents so as to minimize a weighted sum of two terms, representing respectively the total distance between areas belonging to the same territory and the imbalance of sales opportunities for the agents.

Each region is subdivided into J geographical areas, which should then be clustered into I territories, whose total number has been determined beforehand. A sales agent will be associated with each territory, and she should be located in one of the sales coverage units, to be considered as her area of residence. It is further assumed that travel times within each area are negligible with respect to the corresponding travel times between a pair of distinct areas.

Each area will be identified by the geographical coordinates (e_j, f_j) of one of its points, considered as representative of the entire sales coverage unit. One might, for instance, choose the point whose coordinates are obtained as the average of the coordinates of all points belonging to that area. For each territory, let (e_i, f_i) denote the coordinates of the area where the agent associated with the territory resides. This area will be called *centroid* of territory i.

The parameters in the model are as follows: d_{ij} is the distance between centroid i and area j, given by

$$d_{ij} = \sqrt{(e_i - e_j)^2 + (f_i - f_j)^2}; \tag{13.2}$$

a_j is the opportunity for sales in area j; and β is a relative weight factor between total distance and sales imbalance.

Consider a set of binary decision variables Y_{ij} defined as

$$Y_{ij} = \begin{cases} 1 & \text{if area } j \text{ is assigned to territory } i \\ 0 & \text{otherwise.} \end{cases}$$

Define I additional continuous variables that express the deviations from the average sales opportunity value for each territory:

$$S_i = \text{ deviation from the average opportunity value } \frac{1}{I}\sum_{j \in \mathcal{J}} a_j \text{ for territory } i.$$

Hence, the corresponding optimization problem can be formulated as

$$\min \quad \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} a_j d_{ij}^2 Y_{ij} + \beta \sum_{i\in\mathcal{I}} S_i, \tag{13.3}$$

$$\text{s.to} \quad \sum_{j\in\mathcal{J}} a_j Y_{ij} - \frac{1}{I}\sum_{j\in\mathcal{J}} a_j \leq S_i, \qquad i \in \mathcal{I}, \tag{13.4}$$

$$\sum_{j\in\mathcal{J}} a_j Y_{ij} - \frac{1}{I}\sum_{j\in\mathcal{J}} a_j \geq -S_i, \qquad i \in \mathcal{I}, \tag{13.5}$$

$$\sum_{i\in\mathcal{I}} Y_{ij} = 1, \qquad j \in \mathcal{J}, \tag{13.6}$$

$$S_i \geq 0, \;\; Y_{ij} \in \{0, 1\}, \qquad i \in \mathcal{I}, j \in \mathcal{J}. \tag{13.7}$$

The purpose of constraints (13.4) and (13.5) is to bound by means of variable S_i the absolute deviation between each territory sales opportunity and the average sales opportunity, to make the assignment to territories more uniform with respect to sales opportunities, hence balancing the sales chances across the agents. Constraints (13.6) represent a multiple choice condition imposed to guarantee that each sales coverage unit is exclusively assigned to one territory, and hence to one and only one agent.

Model (13.3) is a mixed binary optimization problem, which can be solved by a branch-and-bound method, possibly truncated to limit the computing time and to achieve suboptimal solutions. Alternatively, an approximation algorithm can be devised for its *ad hoc* solution.

13.2.5 Calls and product presentations planning

Optimization models for calls and product presentations planning are intended to derive for each agent the optimal sales activity plan.

Calls planning

The aim of the first model described is to identify the optimal number of calls to each customer or prospect (taken together as *market entities* in what follows) located in the territory assigned to a specific agent. The objective function expresses the difference between revenues and transfer costs.

The decision variables are defined as

$$X_h = \text{ number of calls to market entity } h,$$

$$W_j = \text{ number of trips to market area } j,$$

while the parameters have the following meanings:

$$
\begin{aligned}
a_h &= \text{ strategic relevance of market entity } h, \\
c_j &= \text{ transfer cost to area } j, \\
v_j &= \text{ transfer time to area } j, \\
t_h &= \text{ time spent with market entity } h \text{ in each call}, \\
l_h &= \text{ minimum number of calls to market entity } h, \\
u_h &= \text{ maximum number of calls to market entity } h, \\
b &= \text{ total time available to the sales agent.}
\end{aligned}
$$

The corresponding optimization problem can be formulated as

$$\max \quad \sum_{h\in\mathcal{H}} a_h r\left(X_h\right) - \sum_{j\in\mathcal{J}} c_j W_j, \tag{13.8}$$

$$\text{s.to} \quad \sum_{h\in\mathcal{H}} t_h X_h + \sum_{j\in\mathcal{J}} v_j W_j \leq b, \tag{13.9}$$

$$X_h \leq u_h, \qquad X_h \geq l_h, \qquad h \in \mathcal{H}, \tag{13.10}$$

$$W_j \geq X_h, \qquad j, h \in \mathcal{D}, \tag{13.11}$$

$$X_h, W_j \geq 0 \text{ and integer}, \qquad h \in \mathcal{H},\ j \in \mathcal{J}. \tag{13.12}$$

Constraint (13.9) expresses a bound on the total time available to the sales agent within the planning horizon. Constraints (13.10) impose a lower and an upper bound, respectively, on the number of calls to each market entity. Finally, constraints (13.11) establish a logical consistency condition between the decision variables X_h and W_j.

Model (13.8) is a nonlinear mixed integer optimization problem. To obtain a solution, one may proceed as follows. First, the response function is approximated with a piecewise linear function, deriving a set of linear mixed integer optimization problems. These are then solved by using a branch-and-bound method, possibly truncated to limit the computing time and to achieve suboptimal solutions. Alternatively, again, an approximation algorithm can be devised for *ad hoc* solution.

Product presentations planning

The aim of this model is to determine for each period in the planning horizon the optimal number of mentions for each product belonging to the sales portfolio of a given agent. Through an index called *relative exposure* the model also incorporates the dynamic effects determined by the mentions of each product made in past periods.

The decision variables of the model are consequently defined as

$$X_{kt} = \text{ number of calls for product } k \text{ in period } t,$$

$$Z_{kt} = \text{ cumulated exposure level for product } k \text{ in period } t.$$

The parameters are

$$d_{kt} = \text{ number of units of product } k \text{ available in period } t,$$

$$p = \text{ maximum number of mentions for each product,}$$

$$\lambda = \text{ memoryless parameter.}$$

The quantity $\sigma(X_{kt})$ expresses the relative exposure of product k as a function of the number of times k has been mentioned in period t. The relative exposure formalizes the relationship between the level of cumulative exposure and the number of mentions made in period t through constraints (13.15) in the subsequent optimization model. The response function then depends on the level of cumulated exposure.

The resulting optimization model is formulated as

$$\max \quad \sum_{t \in \mathcal{T}} \sum_{k \in \mathcal{K}} d_{kt} r_k (Z_{kt}), \tag{13.13}$$

$$\text{s.to} \quad \sum_{k \in \mathcal{K}} X_{kt} \leq Kp, \qquad t \in \mathcal{T}, \tag{13.14}$$

$$Z_{kt} = \lambda\sigma(X_{kt}) + (1 - \lambda)\sigma(X_{kt-1}), \quad k \in \mathcal{K}, t \in \mathcal{T}, \tag{13.15}$$

$$X_{kt}, Z_{kt} \geq 0 \text{ and integer}, \qquad k \in \mathcal{K}, t \in \mathcal{T}. \tag{13.16}$$

Constraints (13.14) impose a limitation on the maximum number of mentions that can be made in each period. Constraints (13.15) express, through a recursive formula, the relationship between the cumulative exposure level and the number of mentions made in each period.

Model (13.13) is also a nonlinear mixed integer optimization problem, whose solution can be obtained analogously to model (13.8).

Calls and product presentations planning

The aim of the model described in this section is to determine the optimal number of calls to each market entity belonging to a given segment and, for each call, the number of mentions for each product in the sales portfolio. The aggregation of market entities into segments has the purpose of simplifying the estimation of the response function, by limiting its evaluation only to the segments identified within the scope of the analysis.

The decision variables are defined as

$$X_{kh} = \text{ number of mentions of product } k \text{ to a customer in segment } h,$$

$$W_h = \text{ number of calls to a customer in segment } h.$$

The parameters are

$$p_h = \text{ maximum number of products mentioned in each call to a customer in segment } h,$$

$$s_h = \text{ number of customers in segment } h,$$

$$b = \text{ maximum number of calls that can be made by each agent.}$$

The resulting optimization problem is formulated as

$$\max \quad \sum_{k\in\mathcal{K}}\sum_{h\in\mathcal{H}} r_{kh}(X_{kh}), \tag{13.17}$$

$$\text{s.to} \quad \sum_{k\in\mathcal{K}} X_{kh} \leq p_h W_h, \qquad h \in \mathcal{H}, \tag{13.18}$$

$$X_{kh} \leq W_h, \qquad k \in \mathcal{K}, h \in \mathcal{H}, \tag{13.19}$$

$$\sum_{h\in\mathcal{H}} s_h W_h \leq bI, \tag{13.20}$$

$$X_{kh}, W_h \geq 0 \text{ and integer}, \qquad k \in \mathcal{K}, h \in \mathcal{H}. \tag{13.21}$$

Constraints (13.18) express the limitation on the total number of mentions made to each segment. Constraints (13.19) represent a logical consistency condition between decision variables. Finally, constraint (13.20) establishes an overall upper bound on the number of calls that sales agents can make.

Model (13.17) is a nonlinear mixed integer optimization problem, to which remarks similar to those made regarding model (13.8) apply, in particular for its approximate solution.

A general model for sales resources planning

It is possible to provide a somewhat general formulation for salesforce planning problems by adopting a representation framework that involves listing, at least ideally, all tasks that can be assigned to each agent. The resulting model described in this section, like the one discussed in the previous section, derives the optimal plan for the sales agents across multiple time periods, taking into account different shared resources.

For each agent i and for each period t, the set of all possible sales actions, which represent the plan of calls and product presentations to different customers, is identified in advance and denoted by S_{it}. The required resources, denoted by the index $g \in \mathcal{G} = \{1, 2, \ldots, G\}$, represent the overall budget available to implement the sales actions, or other technical factors needed to adopt the different actions.

The required binary decision variables Y_{iut} are defined as

$$Y_{iut} = \begin{cases} 1 & \text{if action } u \in S_{it} \text{ is selected for agent } i \text{ in period } t, \\ 0 & \text{otherwise.} \end{cases}$$

The parameters are

$$\begin{aligned} S_{it} &= \text{ set of feasible actions for agent } i \text{ in period } t, \\ w_{giut} &= \text{ quantity of resource } g \text{ required to implement action } u \in S_{it} \\ &\qquad \text{by agent } i \text{ in period } t, \\ V_{gt} &= \text{ quantity of resource } g \text{ available in period } t, \\ v_{iut} &= \text{ profit value associated with action } u \in S_{it}. \end{aligned}$$

The resulting optimization problem is formulated as

$$\max \quad \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{I}} \sum_{u \in S_{it}} v_{iut} Y_{iut}, \tag{13.22}$$

$$\text{s.to} \quad \sum_{i \in \mathcal{I}} \sum_{u \in S_{it}} w_{giut} Y_{iut} \leq V_{gt}, \qquad g \in \mathcal{G}, t \in \mathcal{T}, \tag{13.23}$$

$$\sum_{u \in S_{it}} Y_{iut} = 1, \qquad i \in \mathcal{I}, t \in \mathcal{T}, \tag{13.24}$$

$$Y_{iut} \in \{0, 1\}, \qquad i \in \mathcal{I}, u \in S_{it}, t \in \mathcal{T}. \tag{13.25}$$

Constraints (13.23) express the upper limit on the amount available for each resource in each period. Constraints (13.24) represent a multiple choice condition imposed to guarantee that each agent in each period will perform exactly one action.

Model (13.22) is a binary optimization problem belonging to the class of generalized multiple choice knapsack problems, for whose solution remarks similar to those made for model (13.3) apply. Notice that it is usually advisable to reduce in advance the number of available actions for each agent, by means of a preprocessing phase aimed at discarding those actions that are regarded as less convenient or less profitable.

13.3 Business case studies

In this section we will briefly describe some business case studies that illustrate the application to real-world problems of the methods for marketing analysis presented above. For confidentiality reasons, numerical data for the examples presented will not be given. The purpose of these case studies is to offer readers some ideas on the possible fields of application of business intelligence systems in marketing-related decision-making processes.

13.3.1 Retention in telecommunications

Companies operating in the mobile telephone industry were among the first to use learning models and data mining methods to support relational marketing strategies. One of the main objectives has been customer retention, also known as churn analysis. The effect of market saturation and strong competition have combined to cause instability and disaffection among consumers, who can choose a company based on the rates, services and access methods that they deem most convenient. This phenomenon is particularly critical with regard to prepaid telephone cards, very popular in the mobile phone industry today, as they make changing a telephone service provider quite easy and of little cost. Due to the very nature of the services offered, telephone providers possess a vast array of data on their customers and are in the best position to achieve the maximum benefit from data mining in order to target marketing actions and to optimize the use of resources.

Company and objectives

A mobile phone company wishes to model its customers' *propensity* to churn, that is, a predictive model able to associate each customer with a numerical value (or *score*) that indicates their probability of discontinuing service, based on the value of the available explanatory variables. The model should be able to identify, based on customer characteristics, homogeneous segments relative to the probability of churning, in order to later concentrate on these groups the marketing actions to be carried out for retention, thus reducing attrition and increasing the overall effectiveness. Figure 13.19 shows the possible segments derived using a classification model, using only two predictive attributes in order to create a two-dimensional chart for illustration purposes. The segments with the highest density of churners allow to identify the recipients of the marketing actions.

After an initial exploratory data analysis, the decision is made to develop more than one predictive model, determining a priori some market macro-segments that appear heterogeneous. In this way it is possible to obtain several

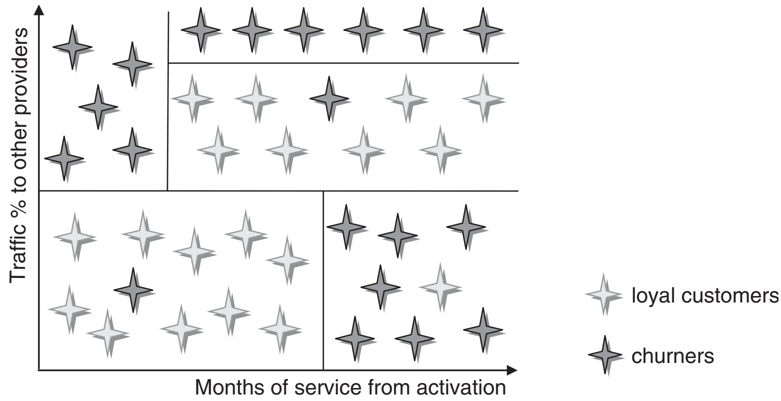

Figure 13.19 An example of segmentation for retention analysis in a mobile telephone company

accurate models instead of a single model related to the entire customer base. The analysis carried out using clustering methods confirms the appropriateness of the segments considered, and leads to the subdivision of customers into groups based on the following dimensions:

- customer type (business or private);
- telephone card type (subscription or prepaid);
- years of service provision, whether above or below a given threshold;
- area of residence.

The marketing data mart provides for each customer a large amount of data:

- personal information (socio-demographic);
- administrative and accounting information;
- incoming and outgoing telephone traffic, subdivided by period (weeks or months) and traffic direction;
- access to additional services, such as fax, voice mail, and special service numbers;
- calls to customer assistance centers;
- notifications of malfunctioning and disservice;
- emails and requests through web pages.

There are approximately 100 explanatory variables available for constructing predictive models.

Analysis and results

Once the dataset for developing the models has been extracted from the data mart, a detailed exploratory analysis can be carried out. On the one hand, it shows a certain number of anomalies, in the form of outliers and missing values, whose removal improves the quality of data. On the other hand, additional variables can be generated through appropriate transformations in order to highlight relevant trends and correlations identified by exploratory data analysis. After applying feature reduction and extraction, the new dataset contains about 150 predictive variables, after the addition of derived variables and removal of some original variables deemed uninfluential on the target. An indicator variable is defined to denote churning by a customer in cases where official notification of service discontinuation is not required, as is the case for prepaid cards. For the different macro-segments, the related *churning signal* is thus defined. For example, if a private customer makes fewer outgoing calls than a preset threshold and receives a number of incoming calls that is below a second threshold value, she is believed to be at a churning stage.

The retention analysis is therefore brought back to a binary classification problem, and models for each macro-segment are then constructed. Different methodologies and different parameterizations are used to obtain several alternative models, the most effective of which can be chosen later based on a comparison that takes into account the indicators of accuracy and the interpretability of the rules generated.

At the end of the development phase, two classes of predictive models are identified. The models based on support vector machines achieve a significantly higher accuracy than other methods, but the corresponding rules they derive are more cumbersome. Classification trees based on axis-parallel splitting rules lead to interpretable rules which are simple and intuitive, but achieve a lower accuracy. The former are preferable for generating the lists of optimal recipients to target marketing campaigns aimed at retention, while the latter are better for investigating loyalty and highlighting relevant market niches. Figure 13.20 shows the cumulative gain curve associated with a classifier based on discrete variants of support vector machines.

13.3.2 Acquisition in the automotive industry

Companies in the automotive industry are striving to develop initiatives aimed at strengthening competitiveness and increasing market share. In this scenario,

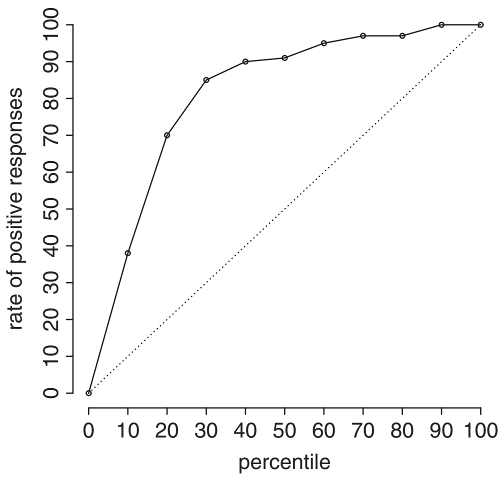

Figure 13.20 Cumulative gains chart for retention analysis in a mobile telephone company

relational marketing projects have been started with the purpose of optimizing the marketing actions and offering products in line with customers' needs, anticipating the evolution of markets and demand. A further element that leads to relational marketing initiatives targeting the sales network is the European directive called the *block exception rule* which introduced new scenarios in the relationships between manufacturing companies and partners in the sales channels, allowing dealers to carry out multi-brand sales activities, and promotion and sales actions across the entire EU territory.

In particular, the use of business intelligence methods is a key factor in strengthening the knowledge of prospects and customers and hence improving new customer acquisition processes and the loyalty of the customer base. It is worth observing that, for some markets in the automotive industry, it is possible to integrate the internal data contained in the marketing data mart with external data sources, which provide a thorough description of the purchases of automobiles and industrial vehicles made by private customers and by companies, easily found at motor vehicles registries. Such an opportunity can be used to enhance segmentation analyses.

Company and objectives

A company manufacturing industrial vehicles wishes to develop a predictive model that can assign to each prospect a score that indicates his propensity to positively respond to a marketing action aimed at acquisition. The main purpose of the model is to provide guidance for promotion actions carried out by the network of dealers. But also, in order to stay one step ahead of competitors, it is required to better understand the trends of the future demand, refining the knowledge of the customer base and of the market scenarios, and identifying the distinctive features that characterize current customers in order to design initiatives directed to stimulate new acquisitions.

Analysis and results

The data available for the acquisition analysis are subdivided into three categories.

Prospects and customers. The first group of data concerns current or potential customers of the company, including customers owning vehicles not necessarily produced by the company. Besides demographic information, the data include 15 explanatory variables that gather meaningful information, such as the top-rated type of vehicle owned by each prospect, the number of new and used vehicles produced by competitors and owned by each prospect, and the number of new and used vehicles bought in the past by each prospect.

Vehicles. The second group of data includes 16 attributes that enable the vehicles owned by each prospect to be identified, among them model, weight class, optional features, fuel type, first and last registration date and status (new or used vehicle). These attributes define for each vehicle the timing of sales transactions, or transfers of ownership.

Works. The third group of data refers to maintenance and repair work undergone by the vehicles during the warranty period and beyond. In particular, these data indicate the type of work carried out, a description of the problem, the vehicle mileage at the time of service, the car shop where the repair was carried out and the date of admission and release of the vehicle to and from the shop.

After exploratory data analysis, which led to the removal of a number of anomalies and outliers, selection of variables was carried out. The analysis led to the addition of new explanatory variables for the time interval that normally elapses between two subsequent purchases, the number of vehicles owned by each prospect, subdivided by weight class, the proportion of vehicles of other brands owned by each potential customer, and the proportion of vehicles of each class included in the vehicle portfolio of each prospect.

Classification models were then developed using different methods. Similar considerations to those in the previous section about retention analysis apply also in this case. Discrete support vector machines turned out to be the most effective method (see Figure 13.22), while classification trees generate rules that can be more easily interpreted (see Figure 13.21).

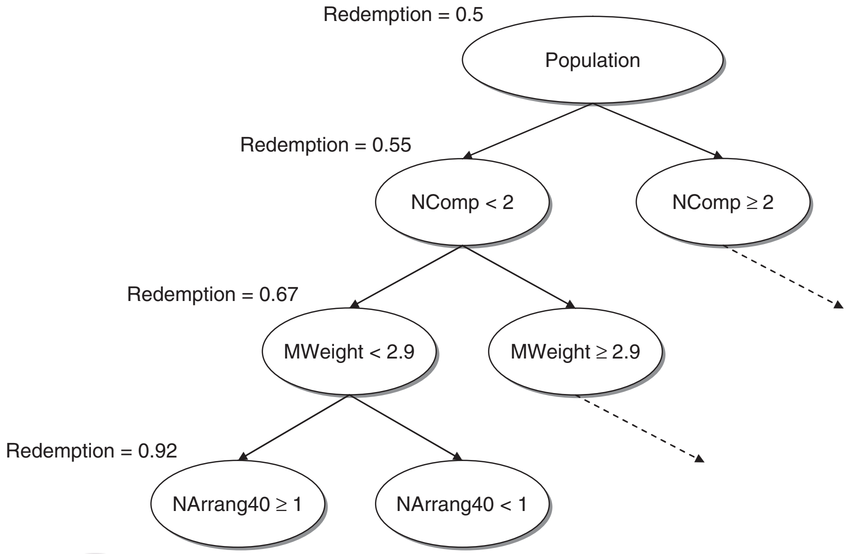

Figure 13.21 Classification tree for acquisition analysis in an industrial vehicle company

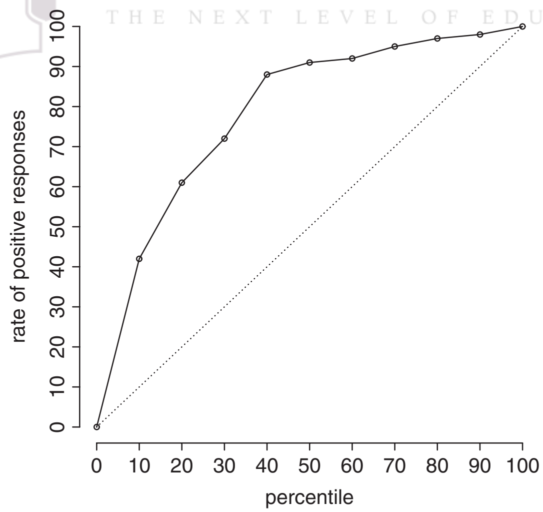

Figure 13.22 Cumulative gains chart for the acquisition analysis in an industrial vehicle company

13.3.3 Cross-selling in the retail industry

The company considered in this section operates in retail consumer electronics, and wishes to segment its customer base in order to optimize marketing actions aimed at promoting a specific product or group of products. The goal is therefore to develop a predictive model able to assign to each customer a score that indicates her propensity to respond positively to a cross-selling offer. Besides prediction purposes, the model should be used also to interpret explanatory factors that have a greater effect on the purchase of the product promoted. Finally, the model is to be used for assessing the existence of causal and temporal correlations between the purchase of the product promoted and the purchase of other items.

Analysis and results

The data available for cross-selling analysis are mainly transactional, referring to customers who have signed up for a loyalty card at one of the company's retail stores. These data include the following information:

- personal information (socio-demographic);
- date of signing up for the loyalty card, which can be regarded as the starting date for the relationship between customer and company;
- dates of first and last purchase, marking the boundary of the time interval within which purchases have been made by each individual customer;
- cash slips, indicating which items, and in what quantities, have been purchased by each customer;
- purchases of sale items made by each customer;
- participation in point-earning programs and related prizes won;
- consumer financing requested to make purchases.

Hence, a binary classification problem can be formulated, where the target variable corresponds to the purchase of the specific product to be promoted. Since the prediction should be available a month in advance, classification models should take into account the corresponding latency. Exploratory data analysis enabled the detection and removal of anomalies and missing data. Then a data preparation stage took place at which those variables were removed that showed a low correlation with the target. Finally, some new explanatory variables were generated through transformations of the original attributes, with the purpose of highlighting trends in the temporal sequence of the expenditure amounts.

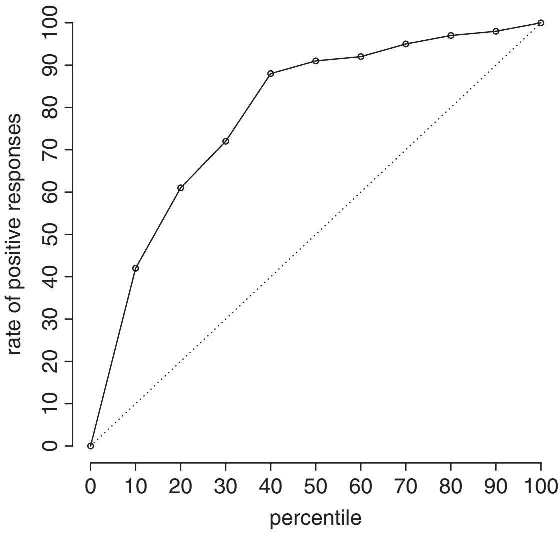

Figure 13.23 Cumulative gains chart for the cross-selling analysis in a retail company

At the end of the exploratory analysis, the dataset used to develop the classification models included approximately 120 explanatory attributes. Although the methods based on variants of support vector machines turned out to be more accurate, a model based on classification trees was deemed more appropriate, in consideration of the interpretability that was indicated among the primary objectives of the analysis. The cumulative gains curve shown in Figure 13.23 corresponds to a model based on support vector machines, able to reach a lift of 8 at a quantile of 0.05, corresponding to 5% of the customers considered by the model to have the highest propensity to buy the promoted product.

Attention then turned to the hypothesis of a causal correlation between the purchase of the product promoted and some other items, identified by marketing analysts on a subjective basis. We proceeded by evaluating the association rules which relate the product promoted to each of the other items, in a series of dyadic relations. The analysis showed that the presence of the product promoted in the head or body of the rule provides similar support and confidence values. This led to rejection of the hypothesis of a significant correlation between purchases. To further confirm this conclusion, sequential rules were analyzed to assess whether the purchase of the item promoted was preceded by the purchase of one of the other items considered, or whether the former was preceded by the latter. Also in this case, the analysis allowed the existence of causal relationships between purchases to be ruled out.

13.4 Notes and readings

There are several books devoted to relational marketing, among which we recommend Peppers and Rogers (1996, 2004), Dyche (2001), Egan (2001), and Bruhn (2002). In particular, for the role of data mining methods and learning models in relational marketing see Berson and Smith (1997), Berson *et al.* (1999), Parr Rud (2000), and Berry and Linoff (2004). Text mining methods are discussed in Drucker *et al.* (1999), Joachims (2002), and Nigam *et al.* (2000). For market basket analysis see Silverstein *et al.* (1998) and Lawrence *et al.* (2001). E-commerce applications are considered in Schafer *et al.* (2001), while Pei *et al.* (2000) and Cadez *et al.* (2003) analyze transactions on the web. The themes of salesforce automation are described in most books generally devoted to relational marketing that have been mentioned above. A more specific book on the subject is Zoltners *et al.* (2001). Optimization models for salesforce management are also considered in Eliashberg and Lilien (1993).