Marketing event optimization

D. A. Selby

We present an algorithm for performing multi-channel marketing event optimization. Previous related work, which made use of a complex mixed-integer linear program to generate a marketing plan, was capable of providing plans for only groups of individuals, referred to as micro-segments, and considered only the direct mail channel. Today, most firms use multiple channels, such as e-mail, call centers, and direct mail, to contact customers with marketing events. Our method successfully overcomes many past restrictions and reduces the run time for the marketing scenarios so that computation can take place on a daily basis. The algorithm used is an advanced form of a greedy heuristic which—given a set of marketing events by channel, a set of individuals, some constraints, and the concepts of saturation and cannibalization—determines the optimal set of marketing events to present to an individual customer. The algorithm is embedded in a solution that is designed to operate as an interactive what-if scenario planner, or as a batch-oriented job that can continually maintain each customer's future contact plans in an optimal fashion.

Introduction

Marketers, especially those in merchandizing and financial services environments, are responsible for developing and maintaining customers who will respond to particular marketing propositions. Let us first consider market traders, also known as stall-holders, who sell new or secondhand goods from outdoor or indoor stalls. Small traders such as these have a marketing proposition that involves how they lay out their market stall in order to entice a customer to purchase from them rather than someone else in the marketplace.

For larger traders who can afford the use of mass media, this kind of enticement is typically achieved by brand and mass marketing conducted through the use of TV, radio, billboards, and other media. In these cases, the marketer often chooses the recipient on the basis of the neighborhood demographics of the customers, and nothing is known or traceable with respect to individual customers. This form of marketing tends to be an extremely inexact science in contrast to direct marketing, which concerns itself with promotional efforts in which the receiver, an individual, is selected by the firm, and the response is measured. The information on these individuals, actions performed upon them, and the reaction to these actions is collected in a database. Therefore, this form of direct marketing is often also referred to as database marketing.

In the case of the majority of marketing departments, achieving the direct-marketing business goal requires a substantially large monetary investment. However, surprisingly, even today, a strong desire for better return on investment is not always accompanied by a desire to clearly understand how to achieve this return [1].

To place our discussion of marketing in perspective, consider that the implementation of even a modest direct-marketing plan may cost a firm tens of millions of dollars. In many firms, this expenditure may be as high as 15% of sales costs. This may appear to be a large percentage; however, the reason for such percentages becomes apparent when we consider the composition of the cost. For example, with the use of direct mail, a company incurs a delivery cost (e.g., sixty cents per item), a printing cost (e.g., \$1.20 for a 20-page catalog), a creative design cost, and the cost of running a marketing department. Here, the total delivery and printing cost is \$1.80. If the target audience comprises 100,000 customers, the company in our example is considering an expenditure of at least \$180,000. If we execute this marketing event once a month, expenditures would be at least \$2.1 million dollars for the year.

0018-8646/07/\$5.00 © 2007 IBM

Thus far, we have been considering a single marketing channel. When considering other typical channels used in database marketing, including call-center outbound calls and e-mail, the problem becomes more complex and a great deal more expensive. Additionally, the costs discussed so far do not reflect any promotional costs such as discounts and free packaging.

Mathematical optimization is based on the principle of producing the best allocation of finite resources. Actual marketing solutions obviously do not involve an infinite budget or channel capacity that enables the delivery of every event to every customer. This means that we need a mathematical approach in order to determine the most cost-effective set of customers to receive delivery of proposed events. Therefore, marketing departments are interested in testing the degree to which profitability is affected by hypothetical scenarios involving budget changes during a particular interval of time or the increase in call-center capacity that enables more outbound calls to be made.

In the course of introducing our solution and the algorithm which underpins it, we can mention the reasons as to why marketing optimization may be very effective. For example, marketing optimization may reduce computational time and cost, which in turn allows the analysts to create a better planning cycle and test more hypothetical marketing scenarios. Such optimization allows us to produce an "organic" marketing plan for each customer over time. We refer to our marketing plans as organic because we continually update our understanding of an individual from daily sales, and we adapt the plan for an individual through continuous optimization of the customer contact stream. Marketing expenditures are treated as an investment, and the optimization model balances a set of finite resources. Marketing optimization also addresses the critical issues of cannibalization and saturation. Cannibalization may occur when customers receive a marketing event to which they would respond favorably by making a purchase, but they are given insufficient time to respond before they receive a better offer for the same or similar goods or service. In this situation, the customers finally purchase from the latter offer, and therefore the cost of performing the first event is wasted. The term *saturation* refers to a scenario in which customers are contacted too frequently with marketing events, which causes them to reach a point at which they ignore communication from the firm. Both of these situations are critical for a marketing department to consider, because both represent wasted advertising dollars.

We begin our discussion by considering related marketing event optimization research. We then describe the existing algorithm for market optimization and follow this with a description of our new algorithm.

Related work

Marketing event optimization rose to prominence in the mid-to-late 1990s, a time during which the consulting services of IBM led much of the early work after several marketing departments requested assistance on topics that went beyond budget optimization and included the challenges of saturation and cannibalization of customers with marketing materials, particularly direct-mail materials. At the time, IBM provided a state-of-the-art solution that attracted media interest and was a finalist for the Franz Edelman operational research prize [2]. The complete solution remains a unique asset of the IBM services division.

This initial IBM solution was extremely computationally intense, which meant that it could run at most only one scenario per week, and it relied on the segmentation of customers through the use of a clustering algorithm [3, 4] that placed customers in "similar" groups. In such a case, market segments may consist of the accumulation of sparse demand observations (i.e., few individuals in a given segment), and this accumulation turned a problem that was infeasible with current mixed-integer linear programming (MILP) technology into a problem that was feasible to solve. Clustering methods do have certain shortcomings, and when such methods are applied to complex data sets such as these marketing sets, it is often easy for analysts to draw incorrect inferences and conclusions.

The setup and configuration for each installation of the solution was extremely labor-intensive. Hence, our goal in further developing the IBM Marketing Event Optimization Solution was to simplify the approach so that it could be deployed by personnel with little background in statistics or operations research in order to leverage this powerful technology. We also felt that the computational performance and IT infrastructure requirements of the solution had to be improved. We wanted to ensure that the runs and what-if scenario planning would be inexpensive, allowing the consumer of the technology to investigate many options quickly and therefore develop the most cost-effective customer contact strategies.

Factors of accuracy

If our technique is compared with a mixed-integer linear program, the overall error term may be higher using our approach; however, the more straightforward solution that makes use of an MILP is computationally too expensive to implement for the situations that we are studying. Note that from a business standpoint, this marketing problem has no completely infallible solution. Customers are complex entities who make purchases for many psychological and other reasons. At best, a marketer can make customers aware of a market proposition and hope that it will stimulate them on some level by generating desire, appealing to them as a result of a particular price aspect, or reminding them of something they had thought about purchasing and had not yet had time to do so.

We argue that the inaccuracies of past approaches result from the adoption of clustering techniques and apportioning a marketing plan to only groups of customers, referred to as *micro-segments*. Also, as we have indicated, the computational cost of solving the mixed-integer linear program model meant that the application was run less frequently than a business might desire and was more complex to both deploy and run on a practical operational basis. Our approach discards the use of a formal MILP and uses a less computationally intensive technique.

The major contributions of this solution are as follows. First, the solution can be implemented at the level of an individual customer: We no longer require micro-segments. Also, the solution can be run frequently, even daily, and can therefore take advantage of ongoing sales reports from individuals. The performance allows what-if scenarios to be run quickly, allowing for easy adoption of new business objectives. The solution does not require the skills of statistical and operational research professionals and can be executed by business analysts who are associated with a marketing department.

Underlying philosophy

The processes and results of *direct* marketing tend to be scientifically measured, whereas brand and mass marketing are often less well measured and tracked [5]. In the case of direct marketing, marketing events are typically executed chronologically for potential customers. That is, one offer follows another offer, each as an isolated process. Often, little regard is given to how one marketing event affects another, as would be the case when events with similar content occur within a short period of time. We note that it is possible to shift the emphasis from the event to the customer. This shift allows for better tracking of return on investment, better scheduling of events, improved customer experience, and higher customer satisfaction levels.

The underlying concepts of marketing event optimization are rooted in financial portfolio optimization techniques, where customers are considered as financial instruments and marketers invest in such customers using a balanced portfolio approach. Using this approach, an asset clustering of customers is normally undertaken before the optimizer is deployed. In order to better understand the phrase *asset clustering*, note that if we consider customers as an investment instrument in a stock market scenario, each customer has asset value to a business. Thus, the term *asset group* refers to a clustering that divides customers into groups on the basis of their contribution to the business and what we believe might be their future value. For example, we may have a cluster of shoppers who choose only premium-grade products, and they contribute a certain margin to the business compared with other clusters. From this information, we can determine an appropriate advertising investment for them. Thus, this asset clustering groups customers into asset groups, which may be considered as investment classes. Each group represents different customers who adopt a similar approach in the way they respond to the marketing proposition. This analysis can be considered in a number of dimensions, such as merchandise category, timing of purchases, and responses to contacts. The high-level planning tool helps to determine the investment strategies to be adopted by the marketing department. This part of the optimization process is germane to the new distribution of financial resources following the shift to customer-centric marketing. We note that business process change is an integral part of adopting the IBM Marketing Event Optimization Solution. However, experience shows that business process change is often overlooked in the deployment of a large proportion of analytical solutions, and this can lead to the failure of any firm to achieve desired results.

By adopting the IBM Marketing Event Optimization Solution, clients can form a better understanding of the amount of marketing that is applied at a customer level compared with a program level, where the term *program* refers to a set of marketing events such as mailings, phone calls, and the use of billboards. Overall returns can be maximized and, more importantly, the customer experience can be enhanced by avoiding saturation and by maintaining a pertinent contact stream.

Marketing budgets are under constant pressure

Marketing departments face a set of common issues or circumstances that must be addressed by effective optimization technologies [6]. The most common challenge is a reduction in the marketing budget. As we have discussed, the overall marketing budget tends to be a very large amount of money, in some cases up to 15% of sales revenue, and this makes the marketing budget an attractive target for short-term cost-cutting actions.

The typical approach used to address a budget reduction involves the removal of events from the marketing plan of equal or greater value than the required budget reduction. However, these cost-saving measures reduce the richness of the contact plan and therefore diminish the customer experience. In contrast, the proposed marketing optimization process and underlying algorithm allow an acceptable reduction in contacts because we carefully optimize each customer's

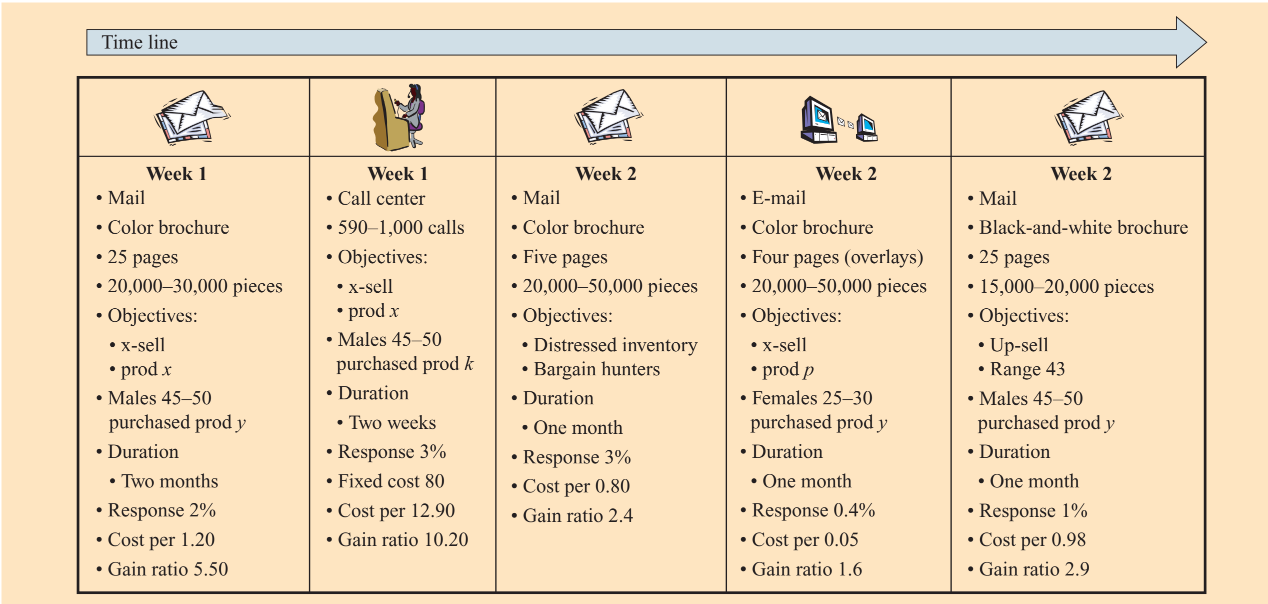

Figure 1

Example marketing event contact plan. The phrase "20,000–30,000 pieces" refers to printed brochures or catalogs. The term "x-sell" refers to the cross-selling of merchandise, as would be the case for trying to sell a customer a camping stove after he has purchased a tent. "Males 45–50" refers to males of ages 45–50. "Cost per 1.20" refers to a $1.20 cost for each mailed envelope. "Gain ratio" refers to the expected gain or profit from performing the marketing event. "Range 43" indicates a merchandise range category, such as "hot pink summer clothing" or "food mixers."

contact plan so that the overall marketing plan can continue to remain rich and dense while permitting the required budget reduction and without compromising the achievable sales plan.

Another common issue encountered by marketing departments requires the justification of the contact plan. Marketing event optimization methodology offers, for perhaps the first time in many marketing departments, a process and method that can be used to justify monetary investment in the marketing event contact plan (**Figure 1**), which specifies certain kinds of marketing activities over a period of time and is discussed in more detail in the section entitled "Preliminaries."

Current state

As we have been discussing, marketing departments invest extraordinary sums in marketing activities, often with little understanding of how these investments relate to an individual customer's net value to the enterprise. A trend exists for enterprises to use a customer lifetime value calculation, which provides a projection of the value of a customer to the enterprise. This estimation makes use of an indexation in order to rank a customer file and can therefore be used to establish an upper bound on marketing investments for that individual. This calculation of a customer value model tends to be difficult to derive, and for this reason is not often used in practical applications [7].

Marketing departments are organized around a marketing plan. This plan, normally prepared one year in advance, provides details of all of the events for the next twelve months across all of the possible channels in chronological order of execution. Each marketing event is selected in turn, and sophisticated predictive scoring models (e.g., regression models and other methods used to produce propensity-to-respond scores) are used to generate target lists of customers for each event. These lists are typically produced serially and in isolation, which tends to prevent marketing cannibalization and saturation from being addressed.

Figure 2 schematically represents an example of marketing cannibalization and saturation by showing how two marketing events may affect sales volume through time. These topics are discussed in detail in the following two sections.

Cannibalization

In Figure 2, marketing Event 1 occurs early in time. For example, Event 1 may correspond to the mailing of a product catalog. This catalog is sent to some group of potential customers. Sales begin to grow, as indicated by the red curve. Event 2, which may correspond to the mailing of another catalog that lists the same merchandise with a 10% discount, is scheduled to start when the red sales curve begins to drop. This catalog is sent to a group of potential customers that overlaps with the first one. In fact, no process exists to distinguish customers who will respond to Event 1 from those who are more likely to respond to Event 2. Thus, this second group of potential customers includes some who will have already received catalogs from Event 1. Note the drop in the expected returns from Event 1.

The ideal outcome, from a marketing and financial point of view, is to obtain a separate sale from each separate event. Cannibalization occurs when separate events generate only a single sale. Without cannibalization, profit may be considered equal to the sales margin minus the cost of Event 1. However, cannibalization changes this equation so that profit is now equal to the sales margin minus the cost of Event 1 and of Event 2.

The net effect of this cannibalization is that money spent on the initial contact is less effective, therefore increasing the cost of sales. This becomes a more significant issue as the density of the contact program increases; thus, the objective should be to achieve the maximum sales from Event 1 before issuing Event 2. The marketing event optimization model is designed to provide the best combination of relationship and affordability for each customer contact across time.

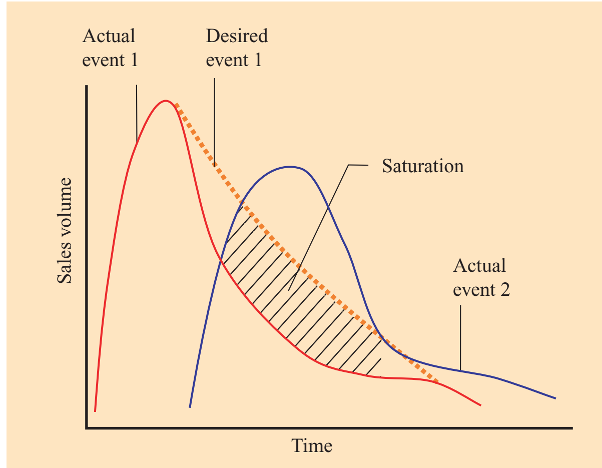

Figure 2

Example of cannibalization and saturation.

Saturation

Focusing on marketing events in isolation may cause even the best customers to quickly suffer from contact fatigue, or saturation. Receiving Event 1 and Event 2 may saturate this customer and generate no sales. Also, customers who might have responded to Event 1 are disgruntled at receiving Event 2, and therefore may make no purchase. One way of measuring this saturation is by surveying customers who appear to have entered this state in the marketing sequence. Without saturation, we would expect that the profit equals the sales margin minus the cost of Event 1 or that the profit equals the sales margin minus the cost of Event 2. Obviously, saturation changes these equations because without a sale there is no profit.

Even the best marketing departments experience the baffling circular maze of ineffectively spending increasing amounts of money on more and more media in order to reach customers. Regression models and other modeling techniques are often promoted as the solution to these kinds of problems, but experience demonstrates that such techniques promote the same high-RFM (recency, frequency, and monetary value) customer set repeatedly with very little variation. (The term *recency* refers to the amount of time that has passed since the customer last engaged in the activity that marketers are trying to encourage.) Standard approaches such as these saturate the best customers with promotions. More than almost all other factors, this saturation generates a loss of loyalty to a marketing proposition. Our algorithm, by examining customers across time and assigning to events a mathematical relationship, such as "cannibalization metrics," mitigates saturation. In our solution, mitigation is achieved indirectly by assuming a relationship between cannibalization and saturation. We use a cannibalization matrix, defined in [8–10], to address only lost sales that result from similarity in content and timing, not customer fatigue. Our hope is that by addressing cannibalization we are also addressing saturation (i.e., customer fatigue).

Our contributions

We want to maximize the expected return from our marketing spending. Understanding the amount of return for each dollar spent should clearly be a key performance indicator for a marketing department. We express this in terms of a return ratio, as depicted in **Figure 3**. The ratio is computed from the expense incurred for producing n events divided by saturated gain, which is defined as the probability of acceptance of a marketing event multiplied by the estimated gain, or profit. We have applied for a patent [11] for our methodologies involving return ratios, which are a key output of the optimization process. For a given marketing dollar investment, the return ratio allows marketers to understand the value of the expected gain. The IBM Marketing Event Optimization Solution also provides this return ratio at the level of a particular event, thereby identifying events that are producing low returns. As noted, the ratio is computed during the optimization process and is based on the anticipated saturated gain

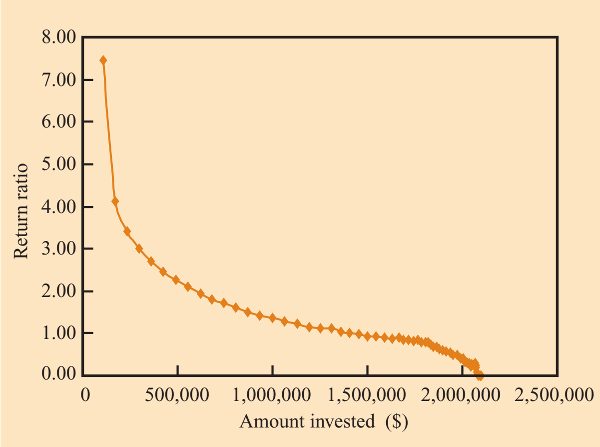

Figure 3

Marketing event return on investment ratio.

from a given marketing event. Sales cost (fixed and for a particular contact) is divided by gain and accumulated in a graph such as Figure 3. We compute the decay curve in Figure 3 during the optimization process so that the user can assess the optimal investment cutoff point.

Referring to Figure 3, if our targeted return ratio is 1.58 (100% investment + 58% return), our financial investment need not exceed approximately $800,000. Any investment over this amount does not achieve the required 1.58 return ratio. The phrase "100% investment" refers to the total amount we are willing to spend.

Preliminaries

In this section, we introduce some notation that is used throughout the remainder of the paper. We define some quantities and functions to help us describe the algorithms that follow.

A *contact plan*, sometimes referred to by marketing departments as a *storyboard*, is a common tool or method used by marketers to identify the set of marketing events to be performed, typically over a twelve-month horizon. Figure 1 is a pictorial representation that illustrates the first few weeks of such a contact plan. It is easy to see from this plan that the risk of cannibalization is quite high in the second week.

Let $E = \{e, c\}$ contain the indices for events being considered for selection, where e is the event and c is the customer. A *contact stream* is a specific subset of the contact plan that is generally associated with a customer. A contact stream can be thought of as a Boolean vector of contact flags in which "true" corresponds to "select this event for the customer," and false corresponds to "do not select this event." The stream is always sorted in an ascending date order.

A *contact stream expense* is the total advertising cost associated with all promotions in this stream. This consists of two components, *sc* and *expense per*, where *sc* represents a setup cost. For direct mail, this may represent printing and creative design costs. For direct-call offers from contact centers, *sc* may represent the training cost for the call-center operative. *Expense per* is the cost of each item. The contact stream expense is used to enforce event size and budgetary constraints on a per-customer basis.

A *contact stream reward* is used to measure the financial benefit of the contact stream and is determined as a result of sending an event stream to a customer. The *expected reward* is a value derived from historical knowledge of how this form of event has yielded responses. For new and unknown events, we evaluate the expected reward by approximation, on the basis of similarity to other events.

The cannibalization matrix is denoted by $\mathbf{S} = [s(p, q)]$. $s(p, q)$ represents the fraction by which the expected reward of promotion q is reduced by contact promotion p. As noted, the hope has always been that saturation is mitigated by treating cannibalization. However, this has not been directly measured by this work, although the subject is discussed in more detail in [10].

Algorithm overview

For a given customer and a contact plan, subject to promotion event quantity and individual customer budgets, the objective of the marketing event optimization model is to assign contact streams to customers that maximize the total financial reward, minimize contact stream expenses, and minimize the cannibalization effects.

Because the compute demands of the algorithm are small, it can be run every day. As sales are recorded, a customer's response scores to each future event will increase or decrease, causing the optimal contact plan to change over time. Technologies such as the relational database system DB2* and the Intelligent Miner* scoring service provide database-embedded predictive modeling [4]. This allows new scores to be produced on demand, and these scores are one of the primary inputs to the optimization algorithm. If the system is deployed on a parallel computation platform, the scoring takes place in parallel without an explicit change to the algorithm. The application of this approach for the computation of saturation and cannibalization is further described in [10]. Our solution incorporates a custom user-defined function that is employed to compute the cannibalization matrix indices dynamically on the basis of the attributes of each of the events. Cannibalization is expressed as a value from 0 to 1, and it represents three business elements: timing, similarity of events, and additional promotional activity. Details associated with these areas are individually customized for each installation of the solution. The cannibalization matrix is computed as an outer product join.

The importance of the three factors mentioned above can be understood in more detail as follows. The *timing* of two events is important, because if two events are close in time, this may lead to a marketing saturation effect. With respect to *similarity of event*, if the two events are very similar or identical (even if the events are distributed through different marketing channels), cannibalization may result. The nature of the *promotional activity* is important, as discussed previously in reference to Figure 2, with respect to the example of two catalogs with identical content. The second catalog comes with a 10% discount, and sending this one week after the first catalog to the same customer would produce a cannibalization effect.

The marketing event optimization algorithm scales linearly with respect to events and customers. Comparison of the previous approach to the new approach shows dramatic differences. The previous approach processed 5,000,000 customers with 36 events in 12 hours, using a dedicated IBM SP2 parallel computer with four CPUs and also using IBM OSL (Optimization Subroutine Library) software to solve the optimization problems. The system solved 10,000 integer-programming problems and one 20,000-variable linear-programming problem. The new algorithm can handle a problem of the same size in less than one hour using a single-CPU POWER5* computer.

Our ability to handle the same-sized problem results from the integration of the algorithm in DB2. In addition to online data storage, the database product provides the ability to use its computational engine in order to dynamically produce inputs to the algorithm on the basis of the data in the database. One of the steps that we complete inside the database is pre-saturation of the contact stream. This means that any event already processed for that customer must be integrated into the stream of proposed events. This is essential because we need the contact stream to be continuous. An actual event sent yesterday should have some impact on an event proposed for today.

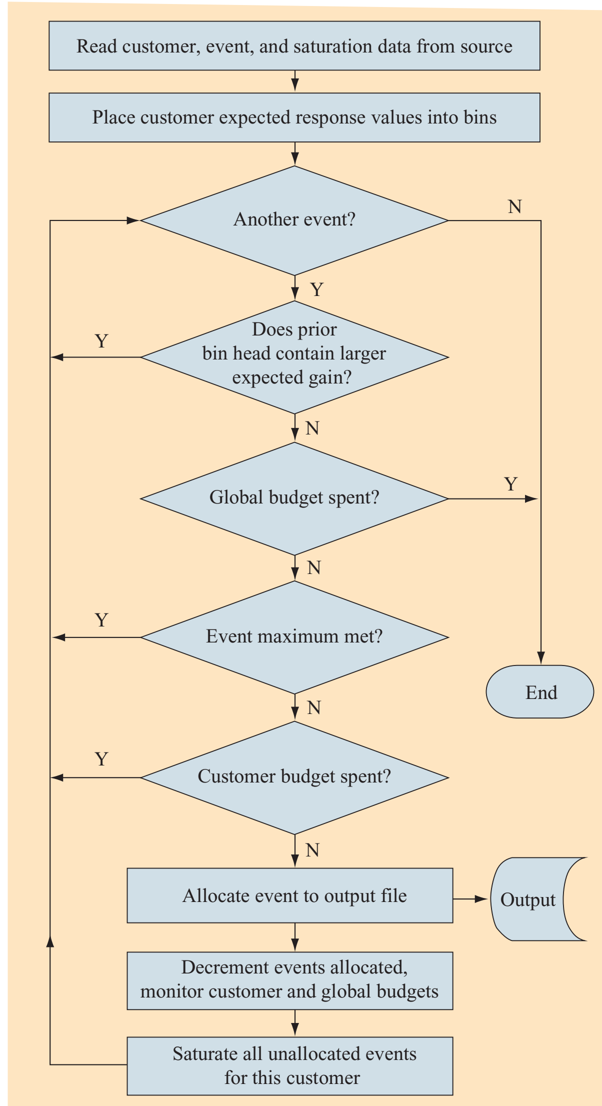

Figure 4

Flowchart for marketing event optimization. The expected gain in the second diamond from the top refers to the probability of response multiplied by the gain or profit. (Bin head: the bin that is currently first.)

Algorithm flowchart

We discuss the algorithm in two parts, beginning with a description of the computer science aspect of the algorithm. In the next section, we describe the mathematical aspects. The key to solving the problem efficiently is described in several patent applications [12–14], but here we provide a summary by way of a flowchart of the algorithm used for the IBM Marketing Event Optimization Solution (**Figure 4**).

We allocate events on the basis of the highest return for each event, and therefore meet the maximized or optimized budget return objective. In other words, each customer is allocated none, some, or all of the possible events based on the optimization run. This allocation of events is performed in the third rectangle from the bottom of the flowchart in Figure 4. The starting point for the data at entry to the flowchart is based upon the customer and the event with the highest potential monetary gain. In other words, the first item chosen is the one which is expected to result in the maximum return. More particularly, the starting point is based upon the probability of response times the expected gain. We then saturate all other non-allocated events for that customer (bottom of the figure) before moving to the next event that maximizes the return. Note that non-allocated events are those that are still available to be allocated to that customer from all possible events. In order to visualize the process, imagine that we have a lottery drawing in which we start with a number of balls. We draw a ball, leaving behind unallocated ones that may still be used in the process. It is possible that the next most profitable event may correspond to the same customer, but this is unlikely, so we would anticipate the next allocation to be associated with another customer in the input set. In other words, from a business context it is unlikely that one customer would generate the most profit from all of the events in the possible stream.

During the run, we monitor the other constraints of the global and customer budgets. We also ensure that the numbers of events do not exceed their maximums. Once we decide to allocate an event, we need to saturate all other events remaining in the stream for that customer.

Mathematical formulation

In this section, we first present the mathematical formulation as published in [8–10]. We then compare this prior work with the formulation as implemented in the marketing event optimization algorithm. Next, we describe the models in terms of the indices, parameters, variables, constraints, and objective.

Prior work

The goal of the mail stream generation (MSG) model is to generate candidate mail streams with the best possible net profit (i.e., gross profit minus expense) for each micro-class j (also referred to as a micro-segment), given the budget requirements for asset class k. An asset class is a group of individual customers who have been selected by a clustering algorithm because they respond similarly to the marketing proposition. An example of this class are customers who always and only purchase once per year up to a certain shopping-basket value in one merchandise category. We are willing to invest a certain amount of marketing dollars in this group. Micro-classes are a further subset of the asset class who will receive a particular stream of marketing mailings. The following are lists of key elements required for the model:

Indices:

$K = \{k\}$ The asset classes.

$J^k = \{j\}$ The micro-classes for asset class k.

$P = \{p, q\}$ The promotions to be mailed within a specified time horizon.

Parameters:

$R_p^{k,j}$ The expected "reward" (e.g., gross profit) from promotion p for customers in micro-class j from asset class k.

E_p The advertising expense for promotion p.

$S_{p,q}$ The cannibalization of "reward" (e.g., gross profit) from promotion q by promotion p.

$\underline{B}_k$ The lower bound on advertising expense to spend per customer from asset class k.

$\overline{B}_k$ The upper bound on advertising expense to spend per customer from asset class k.

Decision variable:

$y_p^{k,j}$ Equals 1 when promotion p is mailed to customers in micro-class j from asset class k, and equals 0 otherwise.

Constraint:

$$\underline{B}_k \le \sum E_p y_p^{k,j} \le \overline{B}_k \qquad \text{Mail stream budget}$$

Objective:

$$z = \sum_p (R_p^{k,j} - E_p) y_p^{k,j} - \sum_{p,q} R_p^{k,j} S_{p,q} y_p^{k,j} y_q^{k,j} \qquad \text{Maximize}$$

In order to generate n candidate mail streams for each micro-class, the interval from $\underline{B}_k$ to $\overline{B}_k$ is divided into n intervals, and MSG was solved n times with the appropriate corresponding bounds.

Furthermore, in order to convert the problem from being a 0–1 quadratic optimization problem to a 0–1 linear optimization model, we linearized the quadratic term by using the following decision variable, constraints, and substitutions. In particular, the new decision variable $w_{p,q}^{k,j}$ equals 1 when promotion p and promotion q are mailed to customers in micro-class j from asset class k, and equals 0 otherwise. As an additional constraint, we enforce the linear variable to represent valid quadratic solutions, $y_p^{k,j} + y_q^{k,j} - w_{p,q}^{k,j} \le 1$.The new objective is represented by the maximization of

$$z = \sum_p (R_p^{k,j} - E_p) y_p^{k,j} - \sum_{p,q} R_p^{k,j} S_{p,q} w_{p,q}^{k,j}.$$

Note that the MSG model is now linear, with integer 0–1 variables. The MSG is solved using the branch-and-bound algorithm provided by the IBM OSL [15].

We next address the mail stream selection (MSS) model. The goal of the MSS model is to select the best mail stream for each micro-class j, given the budget requirements for asset class k.

Indices:

$K = \{k\}$ The asset classes.

$J^k = \{j\}$ The micro-classes for asset class k.

$P = \{p, q\}$ The promotions to be mailed within a specified time horizon.

$M^{k,j} = \{m\}$ The candidate mail streams for asset class k, micro-class j.

Parameters:

$R_m^{k,j}$ The expected "reward" (e.g., net profit) from candidate mail stream m for customers in micro-class j from asset class k.

$F_m^{k,j}$ The advertising expense of sending candidate mail stream m for customers in micro-class j from asset class k.

$C^{k,j}$ The number of customers in micro-class j from asset class k.

$A_{p,m}^{k,j}$ Equals 1 if promotion p is in candidate mail stream m for micro-class j within asset class k.

$\underline{Z}_p$ The lower bound on the quantity of promotions p to send.

$\overline{Z}_p$ The upper bound on the quantity of promotions p to send.

$\underline{B}_k$ The lower bound on advertising expense to spend per customer from asset class k.

$\overline{B}_k$ The upper bound on advertising expense to spend per customer from asset class k.

Decision variable:

$x_m^{k,j}$ The number of customers in micro-class j from asset class k who receive candidate mail stream m.

Constraint:

$$\underline{Z}_p \le \sum_{k,j,m \in M^{k,j}} A_{p,m}^{k,j} x_m^{k,j} \le \overline{Z}_p, \ \forall p \in P \qquad \text{Promotion quantity}$$

$$C^{k,j}\underline{B}^k \le \sum_{k,j,m \in M^{k,j}} F_m^{k,j} x_m^{k,j} \le C^{k,j}\overline{B}^k, \ \forall k \in K \qquad \text{Asset class budgets}$$

$$\sum_{m \in M^{k,j}} x_m^{k,j} = C^{k,j}, \ \forall k \in K, \ j \in J^k \qquad \text{Micro-class mailing requirement}$$

Objective:

$$\sum_{k,j,m \in M^{k,j}} R_m^{k,j} x_m^{k,j} \qquad \text{Maximize}$$

MSS is a linear optimization problem that was solved with the simplex algorithm provided by the IBM OSL. We are optimizing the total "reward," which can be interpreted as gross profit minus advertising expense and cannibalization, summed across all of the streams selected for mailing. We can control the quantity of promotions (which allows the direct-marketing firm to meet postal requirements for quantity discounts) as well as the advertising spent on each asset class, and ensure that every customer receives a mail stream. Note that the solution of the MSS is continuous. In practice, however, we have noticed that the solution almost always has integer values. When this is not the case, we have developed heuristics that assign the "fractional" customer to the best mail stream of all the mail streams. The MSS selects the mail stream according to micro-class so that all customers in that class receive the same stream.

We note that this "two-models" approach is suboptimal in the sense that the use of groups that all receive the same stream is less optimal than actually choosing a stream for an individual, but we use this approach primarily because addressing the full problem at a single-customer level in a single straightforward model would generate an MILP model with billions of variables and equations, which is well beyond the capacity of current technologies to solve.

New approach implemented in the IBM Marketing Event Optimization Solution

In contrast to the approaches just discussed, our technique uses less-complex mathematics that simplifies the underlying model and does not require us to add linearization steps. Again we describe the model in terms of indices, parameters, variables, constraints, and objective.

First, we consider the contact stream generation model. The goal of market event optimization is to generate candidate contact streams (with the best possible net profit, i.e., gross profit minus expense) for each customer j given the customer budget.

Indices:

$J = \{j\}$ The (unclassified) set of customers.

$P = \{p, q\}$ The marketing events to be sent within a specified time horizon.

Parameters:

R_p^j The expected "reward" (e.g., gross profit) from event p for a customer.

E_p The advertising per-piece expense for event p.

F_p The fixed advertising event p, expressed as a one-time charge.

$S_{p,q}$ The cannibalization of "reward" (e.g., gross profit) from event q by event p.

$\overline{B}_j$ The upper bound on advertising expense to spend per customer.

G The global upper bound on spending.

Decision variable:

y_p Equals 1 when event p is delivered to customer j and equals 0 otherwise.

Constraint:

$$\sum_p E_p y_p^j \le \overline{B}_j \qquad \text{Overall event stream budget}$$

$$\sum_p F_p + \sum_{pj} E_p y_p^j \le G \qquad \text{Now incorporating fixed costs for events}$$

Objective:

$$z = \sum_{pj} (R_p^j - E_p) y_p^j \qquad \text{Maximize}$$

$$- \sum_p F_p - \sum_{pj}^{q} R_p^j S_{p,q} y_p^j y_q^j$$

In this simple model, $\sum_p F_p$ is a constant term and does not contribute to the optimization. The actual code (i.e., algorithm pseudocode) is a little more sophisticated, since it does take into account that no fixed cost is incurred where no instances of a given event are allocated. However, the current version does not optimize the decision to include or exclude a given event. Note also that the algorithm eliminates the MSS equation. We are optimizing the total "reward" and can control the quantity for marketing events, as discussed for the previous approach.

Assessing potential returns from marketing event optimization

Firms are likely to benefit from such marketing event optimization if they have the following business characteristics: a customer contact plan with more than ten events per year and one or more channels, media, or merchandise types; a mix of business units that contact an overlapping customer set; a concern about the number of events being received by certain customers; and/or a desire to better align the strategic perspective of the firm with that of its customers.

The goal of optimization is to produce the best allocation of finite resources, and in this case, the finite resource is the investment of a firm in database marketing. Benefits tend to vary depending on the density of the contact plan and type of merchandise being marketed, the overall quality of the data available, and the variance in profit margins across media and offers. The following formula can be used as a rough estimate of the benefits arising in one tactical area, namely changes in event response behavior:

$$rB - \{[(1-a)rB]Pm\} = \text{estimated pre-tax savings.}$$

Here, a refers to accuracy and denotes the percentage of advertising dollars not spent, given a reduction r in advertising investment that would otherwise have generated sales. B is the discretionary (or variable) advertising budget (advertising costs minus such items as allocations for systems or salaries used to produce the advertisements); m is the average profit margin on sales; and P is the productivity of advertising dollars (average dollars of sales per dollar of advertising). As noted, r is the recommended or desired percentage reduction in the discretionary advertising budget. This formula is only a rough estimate because other factors can contribute to the actual savings realized, such as the amount of advertising investment redirected into prospecting for new customers.

The process of migrating to this optimization approach varies from one firm to another, but the development of a customized plan that is tailored to the environment can be achieved after gaining an understanding of the existing contact plan, offer set, available data, data structures, predictive modeling practices, and business objectives. In all cases, strong senior management involvement for embracing business process change is essential [8, 9].

Conclusion

The IBM Marketing Event Optimization Solution expands upon prior related work in marketing optimization. Our solution, detailed in this paper, provides a new algorithm that scales linearly with respect to events and customers, producing as much as twelve times improvement in performance. This improved performance, coupled with the reduced computing resource requirements, provides an optimization system that can be run daily. Moreover, the analyst can undertake numerous what-if scenario plans to help a business fully model the return on investment aspects of any proposed changes to the marketing plan with respect to the entire customer file. The contact plan can now be produced at the customer level and no longer requires the use of micro-segments. The inherent skill level required to deploy and execute the solution has been changed from the skills of statistical and operational research professionals to those of business analysts concerned with marketing issues. We believe that the marketing optimization solution can significantly improve return on investment and customer satisfaction, and we are negotiating with several customers in order to deploy the solution.

Acknowledgments

I thank Stephen Todd for reviewing the mathematics used in this paper and for useful discussion, and I thank Jim Andersen for his diligent proofreading and provocative comments on the draft paper.

*Trademark, service mark, or registered trademark of International Business Machines Corporation in the United States, other countries, or both.

References

1. M. R. Greene, "Market Risk: An Analytical Framework," *J. Marketing* **32**, No. 2, 49–56 (1968).
2. "IBM Press Release: Fingerhut and IBM Named Finalists in Prestigious Franz Edelman Award for Innovative Business Intelligence Technology," 2000; see *http://www-03.ibm.com/press/us/en/pressrelease/1741.wss*.
3. P. Michaud, "Clustering Techniques," *Future Generation Computer Syst.* **13**, No. 2/3, 135–147 (1997).
4. IBM Corporation, Data Warehouse Edition Intelligent Miner Demographic Clustering, IBM Publication Order G507-1464-00; see *http://www-306.ibm.com/software/data/iminer/*.
5. J. Roska, *Ducks in the Henhouse: The Challenges of Integrating Direct and Brand Advertising*, Roska Direct Marketing, Montgomeryville, PA, 2004.
6. P. Kotler, *Marketing Management: Analysis, Planning, Implementation and Control*, 8th Ed., Prentice-Hall, Inc., Englewood Cliffs, NJ, 1994.
7. D. Raab, "Marketing Systems: Building a Customer Value Model," 2005; see *http://www.dmreview.com/article_sub.cfm?articleId=1037754*.
8. E. Bibelnieks and M. P. Haydock, "Horizontal Marketing: Optimized One-to-One Marketing," IBM Whitepaper, 1998, IBM Somers, New York. Inquiries for whitepapers may be sent to the IBM Center for Business Optimization at *http://www-935.ibm.com/services/us/bcs/html/bcs_centeroptimization.html*.
9. E. Bibelnieks and M. P. Haydock, "Optimizing Targeted Marketing Advertising Investments," *Operations Research in Industry*, T. Ciriani, S. Gliozzi, E. Johnson, and R. Tadei, Editors, Macmillan Press, London, 1999.
10. E. Bibelnieks, M. Bullock, M. Haydock, M. Schiller, W. Kugel, E. Pednault, N. Soderquist, and H. Crowder, "System and Method for Increasing the Effectiveness of Customer Contact Strategies," U.S. Patent 6,567,786, 2003.
11. D. Selby, M. Ramsey, and S. Todd, "Method, System and Program Product for Graphically Representing an Optimization Model," END920050159US1, IBM patent disclosure filed, 2005.
12. M. Ramsey and D. Selby, "Optimal Sequencing of Marketing Events," END920050031US1, IBM patent disclosure filed, 2005.
13. D. Selby and M. Ramsey, "Dynamic Selection of Groups of Outbound Marketing Events," END920050007US1, IBM patent disclosure filed, 2005.
14. M. Ramsey, D. Selby, and S. Todd, "Increasing of Customer Contact Strategies," END920050161, IBM patent disclosure filed, 2005.
15. D. G. Wilson and B. D. Rudin, "Introduction to the IBM Optimization Subroutine Library," *IBM Syst. J.* **31**, No. 1, 4–10 (1992).

Received September 15, 2006; accepted for publication October 17, 2006; Internet publication May 17, 2007

David A. Selby *IBM Global Business Services, Center for Business Optimization, Hursley Park, Winchester, Hampshire SO21 2JN, UK (David_Selby@uk.ibm.com).*

Mr. Selby received a Higher National Certificate (HNC) in computer science from Portsmouth Polytechnique in 1982. He is a Fellow of the British Computer Society. Since joining IBM in 1977 in the area of manufacturing, he has held many positions, including microcode engineer, software engineer, research scientist, and consultant. Mr. Selby has been awarded several patents in the field of advanced analytics, and he was named an IBM Master Inventor in 2004. He has received an IBM Outstanding Innovation Award for sustained innovation in real-time analytics. Mr. Selby is currently a solution architect for both the real-time fraud and abuse solutions and the IBM Marketing Event Optimization Solution.