Chapter 7

PERSONALIZATION AND TECHNOLOGY-ENABLED MARKETING

Lee G. Cooper

UCLA Anderson Graduate School of Management, 110 Westwood Plaza, PO Box 951481, Los Angeles, CA 90095-1481

INTRODUCTION

Despite a few very heralded failures in 2000, online retail is a vital and growing sector. While the rest of the economy was sinking further into recession, on-line sales grew 21% to $51.3 billion in 2001, jumped 48% to $76 billion in 2002, and are expected to increase to $96 in 2003. Approximately 70% of on-line retailers showed positive operating margins in 2002, up from 56% the year before.[1] As we come out of recession the U.S. Commerce Department reports that the increase in online retailing is five times as great as the increase in the rest of the retail sector.[2] Most online retailers (63%) updated their inventory management systems to better manage their supply chain.[3] On-line retailing leads in Forrester's industry-by-industry analysis of Website usability.[4] Personalization leads this trend.

Personalization, One-to-One Marketing, and Technology-Enabled Marketing all refer to the basic process of using computer-mediated environments to create an experience for the consumer that seems tailored to his or her particular needs and interests. Anticipated, relevant, and timely are the criteria that Godin (1999) offers in his discussion of personalization in permission-based email and Web marketing. Jupiter Research (Foster, 2000) efers to the personalization chain as an iterative five-step business cycle: storage, access, mining, tuning, and targeting. The consumer's response to a targeted offer initiates a new cycle. Thus, we should think of personalization and technology-enabled marketing as the demand end of a dynamic merchandising system.

---

[1] *The State of Retailing Online 6.0: Performance Benchmark Report*, June 2003, Shop.org.
[2] Wall Street Journal, August 22, 2003.
[3] *The State of Retailing Online 6.0: Performance Benchmark Report*, June 2003, Shop.org.
[4] Forrester, "First Look: Research Highlights for Forrester Clients," August 7, 2003.

This chapter considers personalization in computer-mediated environments, and the algorithms that underlie it. It looks at the basics of modern marketing—segmenting, targeting, positioning, and purchase-event feedback —and relates how these basics are applied in computer-mediated environments. I begin with several fictionalized illustrations of personalization in action, I then discuss personalization in terms of the marketing fundamentals (i.e., segmenting, targeting, positioning, and purchase-event feedback). I follow this with a discussion of the basic approaches to developing a personalized approach to customers (i.e., clustering, artificial intelligence systems, collaborative filtering, profiling, and segment-based learning). How to address the basic business questions comes next (i.e., "Who are our best customers?" "What products do these customers purchase?" "What other products and services do we have that these customers might like?" and "How can we acquire more customers like these?"). The chapter concludes with a real-market test of the improvement in customer response that comes from segment-based learning, and a few expectations for future developments in personalization.

1. AN ILLUSTRATION

Frank Stone points his browser to his favorite Web music store. Because he is a repeat customer, the site sees him coming. While the local bricks-and-mortar music store is stuck with the retail space that greets all customers and the low-wage personnel that makes customer memory almost impossible, this Web merchant, CD-Direct, recognizes Frank as a returning customer and designs the whole storefront on the fly to better appeal to his preferences. This is not science fiction. Such customization capability is off-the-shelf technology. What is new is the ability unobtrusively to connect learned customer preferences with this customization.

A Web merchant can have almost limitless inventory, but very limited visual opportunity to make the right offer. Display real estate is very scarce. Thus, taking advantage of what is known about the customer is essential. Frank is male, 25–34, in ZIP code 16611. Geographically, that is in western Pennsylvania (region: rural). Geo-demographically, that is the Lunch Pail Rural segment,[5] with average income $40,853 (from the U.S. Census), mostly blue-collar workers, with high-school education or less. So Frank is greeted with a screen that welcomes him back personally (an easy look-up). The home-page features the Bruce Springsteen "Greatest Hits" CD that is then the most popular in this segment, and the Nirvana "Nevermind" CD on which the store got a special deal. Both of these are filtered to ensure the store isn't offering Frank something he already bought there. If Frank has bought the Springsteen

---

[5] The segmentation scheme is discussed later.

CD, a Tom Petty CD that is second on the chart for Lunch Pail Rural replaces Springsteen. On the top right are two banner ads: the first for Penn Pizza, a regional restaurant tagged to his ZIP code and segment; the second a Rusty "Jeans for your life" banner, the most popular banner for Frank's segment and gender. At the bottom left is a cross-sell box that features Springsteen's Tour 2000 tickets available on HOTTIX.com that is tagged by a manual rule to both ZIP code and segment. If Frank has bought any CDs that he hasn't rated for CD-Direct, a box shows up on the middle right asking for his feedback. The Web merchant keeps a list of all unrated CD purchases, requesting feedback after enough days for evaluation have past.

Mary Brennan clicks her Favorites link to CD-Direct. Mary is female, 55+, in ZIP code 10024—the New York Metro area (region: highly urban). Geo-demographically, that is the Downtown Elite segment, with average income $96,158, mostly executive jobs, with bachelor degrees and up. For Mary, the homepage comes up featuring Miles Davis' "Kind of Blue," the most popular CD in the Downtown Elite segment, and the Nirvana album highlighted for all customers, again filtered to ensure that CD-Direct doesn't waste scarce feature space on CDs Mary has already bought there. The two banner ads at the top right feature "Fly JFK to London"—a banner tagged to her segment and ZIP code—and "Time for a change," a Risot.com ad showing a woman's watch, one of the most-clicked banners for Mary's gender-segment combination. The cross-sell box at the lower left features "Uptown: Quality Real Estates since 1922," tagged to her segment and ZIP code. A music-genre ID in Mary's cookie could identify her as having made enough classical music purchases to be a known classical-music fan. The Web storefront could then present the most popular classical CDs in the Downtown Elite segment, filtered to ensure the store isn't offering something already bought there. Tabs to the left of the banner ads direct Mary deeper into the site, to sections devoted specifically to each of the music genres—classical, jazz, pop, rock, and hip-hop—where screens are personalized with the offers most popular for her segment and gender in that particular genre, as well as standard search functions for finding what she wants. Banner ads rotate to the next-most popular in her segment with each successive screen requested.

If a new customer arrives, the Web music store can configure the screen without genre preferences, but with the cross-genre picks of the most profitable or otherwise most desired segment. Or the site can experiment to find the offers that most likely lead to new customer acquisition across segments—exposing only a small number of new customers to what might be failing combinations. If the new customer arrives by clicking a banner ad from another site, specific agreements could send an otherwise anonymous segment tag along with the new user, so that the offers could be tailored to segment preferences even for first-time visitors. Similarly, targeted email campaigns for new customers can

come with hot links that allow for traditional list-scoring as well as segment tags that help customize offers.

2. THE BASICS OF TECHNOLOGY-ENABLED MARKETING

These illustrations highlight the roles of Segmenting, Targeting, Positioning, and Purchase-Event Feedback—the fundamentals of modern marketing. The Web doesn't change the fundamental principles, but it does create new opportunities for using these fundamentals.

2.1. Segmenting

The foundations for modern approaches to market segmentation were codified in a special issue of the *Journal of Marketing Research* in 1978 (Wind, 1978). Behavioral segmentation dominated academic efforts at marketing-methods development for the subsequent 15 years. Following a traditional social-science, statistical approach, a sample of a firm's consumers would typically be intensely measured on attitudes, opinions, and interests; lifestyles and values; or other characteristics that the firm considered relevant to its marketing efforts. A multivariate statistical model (e.g., factor analysis, multidimensional scaling, cluster analysis, or discriminant analysis) would either classify or aid in classifying consumers into relatively homogeneous groups or *segments*. The firm would look for easily identifiable (i.e., actionable) keys that could tag the customers or potential customers in desired segments so that not every person had to be measured intensely. *Targeting* would involve a firm deciding which segments constituted its best market opportunity. Messages would be *positioned* to appeal to the known characteristics or preferences of the targeted segments.

The diffusion of checkout scanners in retail environments broadened the opportunity for behavioral segmentation. While scanner-based shopper panels started as small adjuncts to the retail tracking services, the rich record of frequent shopping choices enabled more powerful statistical approaches to segmentation (e.g., latent-class analysis, latent-mixture models, and choice-based individual differences models for multidimensional scaling). The ubiquitous loyalty-club cards for grocery shoppers enable point-of-purchase targeting at the segment or even individual level. But a major difference exists between the grocery shopper and the Internet shopper. While over 98% of customers who enter a grocery store end up purchasing, typically less than 2% of customers who show up at an e-commerce site end up purchasing.

Despite the paucity of *behavior* in the records of most e-commerce sites, behavioral approaches to segmentation still dominate. The analyticals favorite include:

- *Clustering* — Uses statistical techniques to group site visitors with similar characteristics into segments,
- *Artificial Intelligence*, or "AI" — A range of technologies, including natural language processing, expert systems, and neural networks,
- *Collaborative Filtering* — Uses algorithmic techniques to infer preferences based on similar behavior from others, and
- *Profiling* — Characterizes individual consumers based on their interaction with Web site elements.

Clustering methods are essentially an interchangeable collection of heuristic statistical techniques that work on rectangular arrays of site-visitor/customer data to classify groups or segments that are internally homogeneous, yet differ from segment to segment. The collection of variables used in behavioral approaches to clustering or profiling needs to be selected uniquely for each site. The analytical core of these methods considers all the selected variables as dependent measures—with the primary goal of grouping. The interpretation of what each group represents, is also a site-by-site heuristic process. The prediction of purchase (or some other like criterion) is typically external to the clustering or profiling system.

AI systems can work on irregular arrays, but still require site-by-site determination of the overall information space. AI systems are often criterion oriented—implicitly or explicitly attempting to predict a particular outcome, such as purchase. The "black-box" AI systems, such as neural nets, remove the need for site-by-site interpretation. The cost, however, comes potentially in not understanding the rules that determine management actions on one's Website. The expert-systems approaches often included in AI techniques are an exception to the black-box methods. Expert systems result in manual rules, which are discussed in Section 4.5.

Collaborative filtering is both least familiar and most representative of the first three approaches. There is a way of understanding collaborative filters that is very much like clustering—but with the criterion-related goal of determining a recommendation or offer for a customer. Think of a huge table of numbers that has a row for each individual who visits Amazon.com and a column for each book that has ever been bought on Amazon.com. A customer would have an entry of "1" in the column if that individual bought that particular book. A "0" would indicate no purchase. If you multiply the entries in the rows for any two customers, the resulting row for this pair of customers would have a "1" only where both individuals bought the same book. If you add the "1"s in that row, you would get a rough indication of how similar the book tastes of these two individuals were. Of course, you would also tend to have higher numbers for people who read more books, but there are ways of dealing with such issues. Now, think about two customers who have a relatively high similarity score. Most likely they have similar tastes, but have not

read exactly the same books. Why not recommend to Customer A the books that Customer B has bought that Customer A has not, and vice versa? This is conceptually what collaborative filters help managers do—take the choices of individuals and use these choices as a basis for making recommendations to similar individuals.

A visual analog of the Table 1 described above would have a dimension for each book and position each individual according to which of the books that customer bought. This is simple to visualize with two books. All of the customers who bought neither book would be at the origin of the space, all of the people who bought only Book 1 would be grouped at [1,0], all of the people who bought only Book 2 would be grouped at [0,1], and all of the people who bought both books would be grouped at [1,1]. With two books, we could have four groups of individuals. As the number of books increases, the number of possible groups grows exponentially, but the number of actual groups doesn't, since not all possible combinations of books are bought. Tastes emerge from the overlapping patterns of books that similar people tend to read. Further, we know how close various subgroups are to each other using similarity metrics such as that described above or simple distance measures between the centers of different groups. If we look closely, we might find groups at proximal locations that all liked modern mystery novels, books on cooking, or high-technology management books. By drawing approximate boundaries around such groups, we could develop recommendation schemes that capitalize on the taste similarities in those *neighborhoods*.

Many open questions surround collaborative filtering. *How do you draw boundaries around a neighborhood?* This is a heuristic procedure that requires close inspection and detailed knowledge of the business domain. *How often do you update the neighborhood?* Ideally you would update the information in real time, but the requirement to reinterpret the boundaries makes even monthly or quarterly updating rarely practical. Much useful information is lost in the interim. *Can you develop a scalable business model if each e-commerce site must be filtered anew?* The management personnel who possess the domain expertise that aids interpretation are not the same as the personnel who understand the analytical techniques. *How many purchases are required before you can classify a customer as belonging to a particular neighborhood?* The conventional wisdom is that collaborative filters are most beneficial in *high-touch* e-commerce sites (i.e., sites where repeat visitation and frequent purchase is the rule). DVD rental services are good examples of high-touch sites. After a customer has made a dozen or more rental choices, collaborative filters do a better job of reflecting consumer preferences. The exception here concerns when household DVD rentals reflect the preferences of different people in the household. How does the site know if the recommendation is for a child or an adult, a man or a woman? *How do you account for whether*

or not buyers liked a product (i.e., purchase event feedback)? Collaborative filters do not naturally accommodate customer satisfaction feedback. Recommending books or DVDs that similar customers chose, but did not like, is a bad idea. *What good are purchase records when visitors outnumber buyers 50-1?* The conventional wisdom also notes collaborative filters are relatively useless when dealing with new customers or when trying to convert visitors into customers.

Profiling is typified by Engage.com's approach. The company collected 800 supposedly anonymous pieces of information on each Internet user who received Engage.com ads. Just like wanting to know which of 800 books a person bought, you should realize that most of this huge table is empty. Knowing 800 things about an individual may be as useful as knowing nothing. The data are simply too sparse to reveal stable, interesting, and useable patterns in 800-dimensional space—particularly when behavioral response to a banner ad is so rare. Mining for patterns in 800-dimensional space is the kind of endeavor I would assiduously avoid. It reminds me of the full-page ad in the *Wall Street Journal* showing a very full-bodied mid-20s young man, arms folded, leaning up against a grocery-store wall, wearing only diapers:[6]

> DSS transaction no. 009511265: Loaded, Queried, Analyzed . . . At 6:32 PM Every Wednesday, Owen Bly Buys Diapers and Beer. Do Not Judge Owen. Accommodate Him.

I do not believe you can build a business around the 14 people for whom this ad rings true.

The other basic approach discussed in the *JMR Special Issue* on Segmentation involved geo-demographic segments. These approaches, which use U.S. Census data or geo-coded database information to segment neighborhoods, remain popular in the direct-mail niche of marketing. In direct mail, a 2% purchase rate can be quite profitable, unlike the grocery-shopping situation. This is more akin to what we see on the Internet.

In the not-too-distant past, a 2% click rate on a banner ad wasn't unusual, and converting 2% of the people who clicked through to your e-commerce site into purchasers was within normal expectation. Let's talk about purchases per million ads (PPMM). The old expectation was around 400 PPMM. With average click rates falling to .25% to .5%, the conversion rate is down to 50–100 PPMM. Further, while most grocery shoppers fill their market baskets with many items, the modal purchase on the Internet is one item. So, while behavioral segmentation has great potential in the behavior-rich grocery-shopping arena, behavior is rare in the e-commerce arena, and consequently the behavior-based approaches to segmentation are less applicable to the Internet.

---

[6] Wall Street Journal, Ad for Tandem Computers Incorporated, May 15, 1997, page B3.

Geo-demographic segmentation can be a real asset in e-commerce. The basic information typically used includes ZIP code, age, and gender—called ZAG in the direct-marketing world. For the Internet, ZIP code is the key. It is the *action* key, in that many sites acquire ZIP codes by offering customizable services such as weather, local films, or cultural events. It is the *conceptual* key, in that ZIP code ties you into the mother of all secondary sources—the U.S. Census. While the individual census records contain very little information, and are kept secret for many decades, a sample within each census block is questioned much more extensively. The aggregate data are available within a couple of years after each decennial census.

The clusters used in the illustration at the beginning of this chapter came from a hierarchical clustering of selected geo-demographic variables from the U.S. Census. Particularly interpretable clustering patterns appeared in solutions with 68 clusters, 45 clusters, 22 clusters and 11 clusters. The particular illustration came from the 22-cluster solution. This number of clusters provided the data density (given the amount of traffic on a mid-to-large size e-commerce site) that allows statistically stable patterns of preference to be learned in short order. At the 22-cluster level, these segments range in size from Ethnic Elite, which constitutes over 10% of U.S. households, to Native Experience, which makes up less than .3% of households.

The U.S. Census is updated every 10 years, while, of course, neighborhoods undergo continuous change. This inevitable change is more of a problem for local retailers dealing with narrow product mixes and a few ZIP codes. For national merchandisers, the coarser classification into 22 segments means the overall cluster averages change much less. The negative of that inevitable change is outweighed by the advantages of having a single segmentation scheme that can transcend product categories (enabling cross-selling), and multiple communications media (enhancing new customer acquisition).

2.2.   Targeting

A tremendous advantage of an exogenous segmentation scheme, based on the U.S. Census such as this one, is that we do not need to re-segment for each web site. Firms can see the segment mix of their current customers, find out if certain segments are over- or under-represented in comparison to the U.S. population, and, by matching margin data with this segmentation, find out which segments are the most or least profitable. The segmentation system provides the ability to keep score in a way that aids targeting decisions. This kind of a segmentation tool levels the playing field between big and small firms.

The customization capability of Internet sites gives Web merchants the ability to target multiple segments with different appearances and offerings. Obviously, bricks-and-mortar businesses cannot reset display windows for each

passing customer, but e-retailers can. They can try to attract their most desired segments and not appeal to the least-profitable ones. A single e-retailer can be a virtual mall, emphasizing service and high-end products in some segments while emphasizing costs and economy in others.

2.3.   Positioning

While I believe that the 68-cluster solution is most descriptive of the population, the 22-cluster system pushes the envelope for how articulated the system could be given the amount of traffic we could expect in practice. Data density is the important issue. Even ignoring age, which is often inaccurately reported, one needs to tract three levels of gender (i.e., male, female, and unknown). This implies tracking 66 groups for the 22 clusters. The more groups you have, the more data you need in order to be sure you know the most popular offering.

In each segment (or possibly each segment-gender-age combination), we have to position N offers. We need to know what are the most popular N offers and the order of their popularity, given only limited data from which we can learn this order. For a new offer, we need to know if it falls within the top N offers for a particular segment. We need to know this as quickly as possible.[7]

2.4.   Purchase-Event Feedback

We tend to forget this very important topic. Blockbuster clearly forgot purchase-event feedback in Take10—its initial experiment with recommendation systems (West et al., 1999). Take10 recommended 10 videos to customers based on their history of rentals. As pointed out above, a recommendation system may fail simply because it neglects to find out whether customers liked what it recommended. With profiling or collaborative filtering, one doesn't know how to incorporate customer satisfaction into recommendation systems. With a recommendation system that uses segment-based learning, coming up with weighting for recommendations is a straightforward, empirical matter. We accumulate some number of votes for each purchase by segment members and let positive feedback add to that count while negative feedback diminishes that count. Since most of the behavioral decision theory literature says that negatives have more impact than positives, we can give more weight to negative feedback than to positive feedback.

---

[7]   Eric Bradlow and I have worked on the statistical issues underlying these basic problems.

3.   THE TRADITIONAL SWAP BETWEEN UNIVERSITY FACULTY AND BUSINESSES

In the scanner-data era, beginning in the early 1980s, the basic swap between university faculty and businesses involved *data* for *methods*: Data intermediaries such as A.C. Nielsen and IRI, either directly or through the Marketing Science Institute, provided store-tracking data and/or scanner-panel data to academics in major research universities around the world. Empirical marketing scientists developed methods for addressing basic management issues, putting these methods into the public domain through articles in top-tier journals such as *Management Science*, *Marketing Science*, and *The Journal of Marketing Research*. Manufacturers, and the data intermediaries themselves, could develop proprietary versions of the methods they found useful. Academics got to work on interesting problems, advance the state of knowledge, and advance their careers. Not a great deal of consulting resulted from this arrangement, basically because grocery retailers work on notoriously slim margins, and they don't have the management expertise to readily use the intellectual property resulting from the publications. Doctoral education benefited because PhD students were equipped and eager to handle the advanced statistical work. That was enough to sustain the model. Not very much of it filtered into the MBA classroom, because teaching advanced statistics to MBA students is a painful experience. Even when simulators were developed to eliminate the need to understand the statistical side, MBA students were not very enamored with the mundane complexities of tactical promotion planning and retail category or brand management. This is partly why grocery retailers find it hard to attract good MBAs.

I had been involved from the very beginning in setting up the university infrastructure to deal with the onslaught of scanner data. Because of my early work in market-share analysis (Nakanishi and Cooper, 1974), UCLA was one of the very first to receive such data. Gerry Eskin, one of the founders of IRI, heard me speak to a Procter-and-Gamble-sponsored invitational conference in the fall of 1982 (Cooper, Nakanishi, and Hanssens, 1982), thought my work could be applied to scanner data, and sent a tape full of scanner data out to UCLA with Penny Baron, another IRI founder, who began a stint as a visiting faculty member the following January.

Marketing academics were just as interested in trying to tame this newest data source, Internet data. What they tended to encounter were overworked and overstressed staffs of Internet companies that were unlikely to carry through on promises of data, or huge, raw Web site logs that required horrendous amounts of processing before even a glimmer of market intelligence would show. None-the-less, marketing researchers have persevered, developing methods for modeling log-file data from online content sites to estimate advertis-

ing exposure (Chatterjee, Hoffman, and Novak, 2003), panel-data models of purchase conversion (Ansari and Mela, 2003; Bucklin et al., 2002; Sismeiro and Bucklin, 2003; Moe, 2003; Moe and Fader, 2001, 2004; Bucklin and Sismeiro, 2003; Moe and Fader, 2002; Park and Fader, 2002; Moe et al., 2002), and navigational methods seeking to understand the antecedents of purcha se (Montgomery et al., 2002).

Since Internet data rapidly becomes massive, academics would benefit from other agencies or intermediaries doing the programming and most of the data handling. The Marketing Science Institute played a facilitating role with retail scanner data, and an analogous role would be helpful with Internet data. Most of the faculty contribution would be formalized in equations. Since you don't patent equations, the intellectual-property (IP) issues do not seem insurmountable. IP concerns should not be an inhibiting force.

4. A BUSINESS-INTELLIGENCE SUITE

Harnessing the demand end of the dynamic merchandising chain, as a personalization/recommendation engine inherently should, makes the development of business-intelligence suites relatively straightforward. The basic questions are: "Who are our best customers?" "What products do these customers purchase?" "What other products and services do we have that these customers might like?" and "How can we acquire more customers like these?"

4.1. Who Are Our Best Customers?

The easiest way to track this basic question is by a plot such as that in Figure 1, based on an analysis of data provided by eHobbies.com.[8] The X-axis shows profits per customer and the Y-axis show the dollar share of sales for each segment. For eHobbies.com, Ethnic Elite, Small Town Success, Wealthy Commuter, and Affluent Elite are the segments with above-average profits per customer and above-average dollar shares. In the spring of 2000, these were eHobbies.com's most valuable segments. Finding out what products they like, and using this knowledge for create up-selling and cross-selling opportunities, are basic elements of retail strategy. The "Margin Squeezers" are segments with above-average dollar share of sales, but below-average profits per customer. Is there a way to structure benefits such as free shipping that can increase the margins on these high-dollar segments? The "Underdeveloped Segments" have high profit per customer, but low dollar share. These segments

---

[8] I would like to thank Seth Greenberg, CEO of eHobbies.com, for allowing me to use these data.

are secondary candidates for cross selling aimed at increasing the dollar volume. Finally, the "Low Value Segments" buy little and have small profit per customer at current margins.

4.2. What Products Do Our Best Customers Purchase?

To grow current customers, you must know their tastes. eHobbies.com has four major product groups: Models, Trains, Radio Controlled, and Die Cast. Models are most popular with ages 35–44, those in the Southeast region, and the Close-Knit Hispanic segment. The Close-Knit Hispanic segment contains households with average income and education, mostly Hispanics, and living in homes of average value. Many depend on public assistance. This is an Underdeveloped Segment in Figure 1. Trains are most popular with older (65–74) customers, male customers, and the Affluent-Elite segment—wealthy, highly educated, white-collar careerists who are mostly Caucasian and Asian-Americans who live in expensive homes in urban areas. Radio-Controlled products are most popular with the young (25–29) in the Urban-Challenge segment—low-income African Americans and Caucasians with higher unemployment rates, and who live in older, low-value, dense housing, mostly in industrial cities. Many tend to have below-average education, commute via mass transit, and receive public aid—a Low-Value Segment in Figure 1. Die-Cast products are favored by women, older age groups (55–64), those living in the Far-West region, and Diverse-City-Prime segment—above-average in-
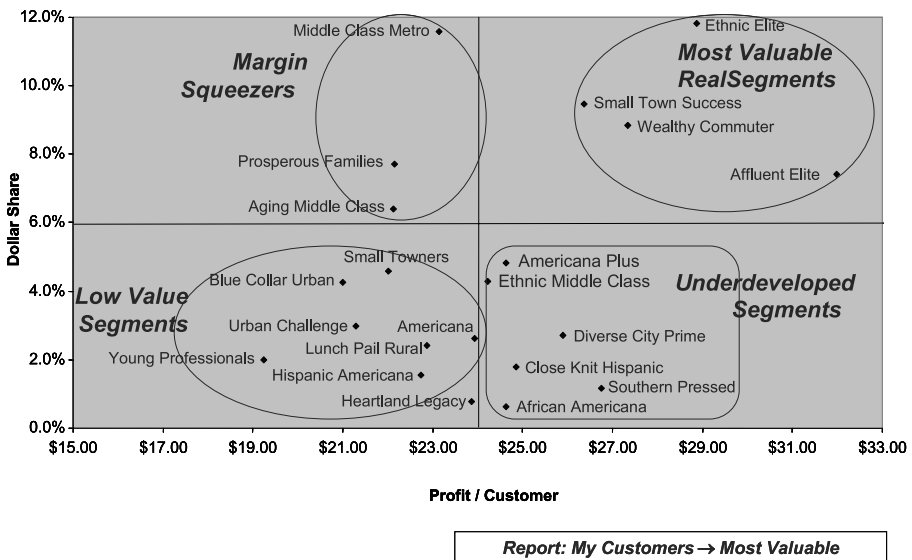
Figure 7-1. Crafting Segment-by Segment Marketing Strategies.

come, highly educated white-collar workers who are of diverse ethnicities including Hispanic, African American and Asian-American. Many live in urban areas, mostly in mid-Eastern states. This is an Underdeveloped segment for eHobbies.com.

When broken down by Most Valuable Segment, we find the Affluent Elite's favorite products are Traxxas Nitro 4-TEC RTR Car, Christmas Mixed Train Starter Set,

JRS XP652 Helicopter Radio System, O-27 Lionel Santa Fe Special Train Set, and the G 2-8-0 Rio Grande C-16 #268-Bumble Bee. The Ethnic Elite favorites are 1/18 Porsche 911 GT1 '96 Street—Blue, Futaba 8UHPS PCM Radio System, Traxxas Nitro 4-TEC RTR Car, TTR RAPTOR 49BB PRO, Kyosho Dodge Ram QRC Combo, and the LGB's 30th Birthday Starter Set. Small-Town-Success segment's favorites are O-27 Lionel Santa Fe Special Train Set, Traxxas Nitro Rustler Truck w/Radio, Losi XXX-T 2 WD R/C Truck, and Hobbico Skyrunner R/C RTF EP Airplane. And the favorites for the Wealthy-Commuter segment are Traxxas Nitro 4-TEC RTR Car, O-27 Lionel Santa Fe Special Train Set, O-27 Lionel NYC Flyer Train Set, Hangar 9 Skypack Pilot Package, and the JR F400 with 4-517 Radio System.

These are simple answers to basic questions—easily obtained when the information systems are organized by a marketing-driven logic.

4.3. What Other of Our Products and Services Might These Customers Like?

eHobbies.com had teams of veteran hobby-store owners coming up with manual rules for the complementary products to cross sell with each kit, but the empirical record provides an excellent substitute or complement for the hard-found expertise, and it's decomposable by segment. From the empirical record for the Affluent-Elite, Ethnic-Elite, and Wealthy-Commuter segments, we find the best cross sell for the Traxxas Nitro 4-TEC RTR Car is Dynamite BLUE THUNDER 20% high-performance fuel. For Small-Town Success it's the Futaba S148 Servo Precision J FM ("Genuine Futaba servos are the easiest and the most efficient way to upgrade your Futaba system").[9]

Looking at the cross-sell issue by product class gives you ideas that experts might not see. For dye-cast products, compliments are not popular. What sells in the same shopping cart are other dye-cast products. Multiple product discounts or other incentives to enhance the likelihood of larger shopping carts are indicated here.

For radio-controlled aircraft, the Hangar 9 Skypack Pilot Package is very popular at around $300 ("Hangar 9's Skypack package contains everything

---

[9]   This product description is from the eHobbies.com site.

Table 7-1. eHobbies.com's Experience Converting Browsers to Buyers

ZipSegment	Buyers	Browsers	Conversion
Ethnic Elite	1,680	2,07	37%
Small Town	1,580	698	69%
Wealthy Commuters	1,349	506	73%
Affluent Elite	992	428	70%
Diverse City Prime	488	292	63%

you need for takeoff. Convenient and comprehensive, the Skypack contains a JF F400 EX radio system, an MDS .40 FS Pro BB two stroke engine, one quart of Hangar 9's 'eroBlend fuel, a Master Airscrew 10 × 6 propeller and a Hangar 9 Start Up Field Pack which includes a glow igniter, fuel pump, chicken stick, four way wrench, two cycle glow plugs and starter tote.").[10] To an expert, putting this in your shopping cart is a sure sign that a beginning flyer is being set up. The empirical record and the expert would recommend a beginning trainer such as Hangar 9 Easy Fly 40 Trainer (Red or Blue).

The rules that generate these recommendations are simple *if–then* conditions. Rule-discovery algorithms, such as Cooper and Giuffrida (2000), Giuffrida et al. (1998), or specific ones as used here, easily capture the required evidence. They are readable and understandable by management with simple, empirical counts of the units, dollars, and margins involved. These are not the opaque results of collaborative filters or artificial-intelligence systems.

4.4.   How Can We Acquire More Customers Like These?

The first step in acquiring new customers is trying to convert current browsers into customers. Offering something of value to convert a stranger into a registered browser enables customization that can increase the odds of converting browsers to buyers. eHobbies.com's experience converting browsers to buyers in their best segments is very good, as can be seen in Table 1.

Armed with registration information, eHobbies.com can *push* emails to browsers to try to convert them to buyers. A prototypical effort is shown in Figure 2.

In the online world, acquiring completely new customers has proven to be an expensive and uncertain proposition. If the same exogenously determined segmentation scheme was used independently by a number of Websites, then eHobbies.com's media buyers could indicate a willingness to pay some premium if eHobbies.com's banner ads are exposed only to eHobbies.com's best segments; in this case, Affluent Elite, Ethnic Elite, Small-Town Success, and

---

[10] This product description is from the eHobbies.com site.

Prospecting efforts should recognize segment tastes

Figure 7-2. A prototypical effort.

Wealthy Commuter are the targets. This can be done anonymously, without the passing of customer identity, so that publisher sites don't violate their privacy policies.

Similar network benefits would accrue to email-list suppliers using the same exogenously determined segmentation scheme. Thus, emails could be crafted separately for each segment, with hot links that would cause recipients who chose to click through to be greeted at the site with segment-specific offerings. And finally, given that the action key is ZIP code, direct-mail lists can be bought for the ZIP codes in the desired segments. Thus, banner ads, emails, and direct mail can all be used to target the specific kinds of users that the site finds are its best customers. This is the benefit of a multi-channel segmentation system.

4.5. Improving Your e-Commerce Site

Putting the basics reported above into practice at an e-commerce site is straightforward, thanks to the standard capabilities of modern serve-side scripting languages. Figure 3 does this for eHobbies.com.

If management already knew all the things that the prior sections would help Websites learn, then the same benefits would accrue to Websites using

Figure 7-3. Increase profits by recommending profitable, targeted products and the right accessories.

manual-rules engines, such as BroadVision. Just as for the rules we learn with dataminers, manual rules are simple if-then statements that translate antecedent condition into actions (e.g., recommendations). If the Hangar 9 Skypack Pilot Package ends up in a shopping cart, then an expert would suspect a beginning R/C flyer is being set up, and consequently recommend a beginning training aircraft. As indicated above, eHobbies.com began by interviewing the veterans with years of hobby-shop experience to capture a set of such rules. While it is dangerous to assume you know all you need to know about your customers, many merchants moving from bricks-an-mortar operation to the web can use manual-rules engines to encapsulate the domain expertise they already have. I merely advocate that simple learning mechanisms, such as segment-based learning, be added to help the knowledge base grow. Manual rules and the rules from segment-based learning easily co-exist in personalization solutions.

5.   AD-OPTIMIZATION TEST

Personalization in computer-mediated environments will continue to grow. "Why? Because it works" (Foster, 2000).

A startup I worked with, which for the purposes of this chapter we will call Strategic Decision Corp., faced a compelling need to demonstrate that segment-based learning was a practical approach to Internet advertising optimizations problems. As with any Internet offer, we can learn how to target Internet banner ads more effectively using segment information. The special features of ad optimization do require fast learning, since the standard contract for a banner ad is only one month long, and budget constraints, since only a given number of exposures are paid for. As a condition of the C-Round venture funding the startup had to demonstrate that, for a consecutive, 30-day period, the segment-based learning and optimization produced at least 2X lift, including the time spent in learning, compared to a control group composed of these same campaigns. Some campaigns demanded so much traffic that optimization was nearly impossible. To satisfy the ad contract these banners had to be shown to too many segments for normal optimization. On the other hand, to insure that the optimization test was not just for cherry-picked campaigns, an additional condition require that, for the same 30-day period, the campaigns used to calculate the lift had to represent at least 50% of all paid traffic on the client site. The test criteria had to be satisfied before August 15, or the funding would be lost. This was a "bet the company" situation for the startup.

The startup began serving ads on the client site on May 19, 2000. Initially all ads were in a learning phase. But even by Day 2, some optimization was occurring. Figure 4 shows the cumulative performance. June 19th was the first day a full 30-day window of execution existed. By June 28th the criteria were satisfied. The startup had served approximately 153,000,000 banner ads on the client site without failure, and the Lift was 2.06.
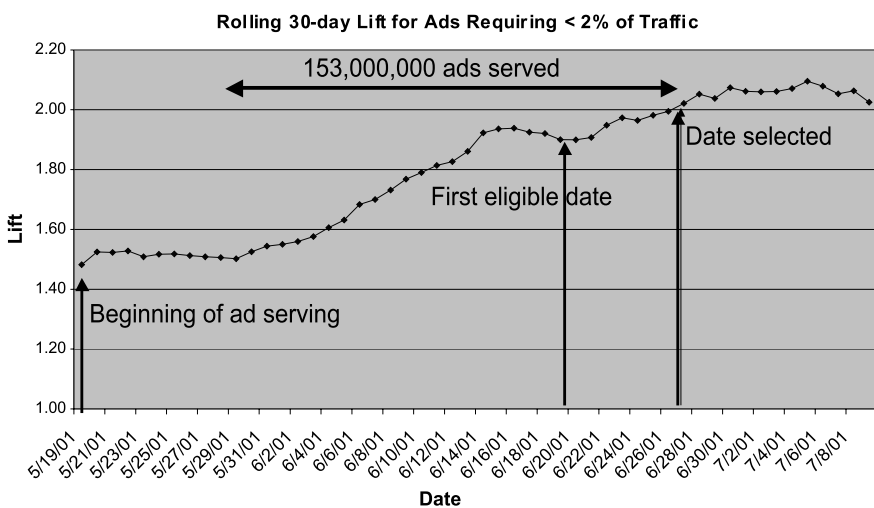

Figure 7-4. Ad Optimization Test Results.

The merchandizing side of personalization is much simpler to put into operation, than optimized ad serving. Product recommendations will rarely, if ever, require the 1,200 recommendations per second that can be demanded for banner ads. Products are typically not budgeted as to how many times they are to be shown in a month. And product life cycles are typically much longer that banner-ad life cycles, leaving much more time to benefit from what has been learned, before one item is replaced by another.

6. THE FUTURE OF PERSONALIZATION

Lately, managers are slowing down in their adoption of personalization techniques. Some have complained that, in the economic downswing that followed the Internet Bubble Burst, these techniques are too expensive and do not provide a big enough return on investment. One reason for this may be that we really do not know well enough consumers want from personalization. This is an area where marketing and behavioral scientists can add a lot of value—doing the research to uncover the *whys* so the techniques can work more efficiently at solving problems for which the consumers actually want answers. Lot of interesting research remains undone.

I am ever mindful of Alan Kay's famous slogan that the best way to predict the future is to invent it. Despite an alignment of agendas between marketing-science faculty at research universities and the most forward thinking Web merchants, I hesitate to forecast that faculty will invent the future of web merchandizing. I do, however, expect certain patterns to emerge:

1. Rule-based engines will grow at the expense of black-box systems. The basic tenet is that concerned management wants to know and understand the rules that are driving its marketing efforts. Whether rule-generating data-miners, manual rules, or some hybrid system will emerge is not clear. But I expect transparency to prevail. The exception to this lies in sensitive applications such as credit scoring, where not knowing the generating rules may have practical advantages for managers.
2. Despite the academic interest in behavioral segmentation, ZAG-based segmentation schemes will unite Web marketing with its natural brethren— direct marketing and database marketing. Multi-channel segmentation schemes that strengthen the ties between these allied fields will be favored.
3. Personalization engines will feed the demand component of integrated supply-chain systems.

As with any forecast, you are always on safer ground if the foundational elements are already positioned and operating. Such is the case with all three of these anticipated patterns.

REFERENCES

Ansari, Asim and Carl F. Mela (2003), "E-Customization," *Journal of Marketing Research*, 40, 2 (May), 131–145.

Bucklin, Randolph E., James M. Latin, Asim Ansari, Sunil Gupta, David Bell, Eloise Coupey, John D.C. Little, Carl Mela, Alan Montgomery, and Joel Steckel (2002), "Choice and the Internet: From Clickstream to Research Stream," *Marketing Letters*, 13, 3, 245–258.

Bucklin, Randolph E. and Catarina Sismeiro (2003), "A Model of Web Site Browsing Behavior Estimated on Clickstream Data," *Journal of Marketing Research*, (August), 249–267.

Chatterjee, Patrali, Donna L. Hoffman, and Thomas P. Novak (2003), "Modeling the Clickstream: Implications for Web-Based Advertising Efforts," *Marketing Science*, 22, 4, 520–541.

Cooper, Lee G. and Giovanni Giuffrida (2000), "Turning Datamining into a Management Science Tool," *Management Science*, 46, 2 (February), 249–264.

Cooper, Lee G., Masao Nakanishi, and Dominique M. Hanssens (1982), "Estimating Cross Competitive Influences on Market Share," The Procter and Gamble Invitational Psychometric Conference, October.

Foster, Cormac (2000), "The Personalization Chain: Demystifying Targeted Delivery," *Jupiter Vision Report, Site Operations Vol. 3*, New York: Jupiter Communications.

Giuffrida, Giovanni, Lee G. Cooper, and Wesley W. Chu (1998), "A Scalable Bottom-Up Data Mining Algorithm for Relational Databases." In *10th International Conference on Scientific and Statistical Database Management* (SSDBM '98), Capri, Italy, July 1998. IEEE Publisher.

Godin, Seth (1999), *Permission Marketing*, New York: Simon & Schuster.

Moe, Wendy W. (2003), "Buying, Searching, or Browsing: Differentiating between Online Shoppers Using In-Store Navigational Clickstream," *Journal of Consumer Psychology*, 13, 1&2, 29–40.

Moe, Wendy W. and Peter S. Fader (2001), "Uncovering Patterns in CyberShopping," *California Management Review*, 43, 4, 106–117.

Moe, Wendy W. and Peter S. Fader (2004), "Dynamic Conversion Behavior at e-Commerce Sites," *Management Science*, 50, 3, 326–335.

Montgomery, Alan L., Shibo Li, Kannan Srinivasan, and John C. Liechty (2002), "Predicting Online Purchase Conversion Using Web Path Analysis." Working Paper, Graduate School of Industrial Administration, Carnegie Mellon University.

Nakanishi, Masao and Cooper, Lee G. (1974), "Parameter Estimation for a Multiplicative Competitive Interaction Model-Least Squares Approach," *Journal of Marketing Research*, 11 (August), 303–311.

Sismeiro, Catarina and Randolph E. Bucklin (2003), "Modeling Purchase Behavior at an Ecommerce Website: A Conditional Probability Approach," *Journal of Marketing Research*, 40 (July), 249–267.

West, Patricia, Dan Ariely, Steve Bellman, Eric Bradlow, Joel Huber, Eric Johnson, Barbara Kahn, John Little, and David Schkade (1999), "Agents to the Rescue," *Marketing Letters*, 10, 3 (August), 285–300.

Wind, Yoram (1978), "Issues and Advances in Segmentation Research," *Journal of Marketing Research*, XV (August), 317–333.