A Framework for Targeting Banner Advertising On the Internet

Katherine Gallagher and Jeffrey Parsons
Faculty of Business Administration, Memorial University of Newfoundland
St. John's, NF, Canada A1B 3X5
{kgallagh, jeffreyp}@morgan.ucs.mun.ca

Abstract

Constraints that limit accurate targeting of advertising in traditional media may not hold in cyberspace. This paper presents a model for effectively and efficiently targeting hypermedia-based banner advertisements in an online information service. The model takes advantage of information technology to micro-target banner advertisements based on individual characteristics of users. A simple version of the model, which has the virtue of ease of development, is presented. Enhancements are also proposed. These require more effort to develop, but may lead to even more precise targeting of advertisements. Implementation of this framework may benefit both online advertisers and online consumers.

1. Introduction

Cyberspace is a rapidly growing new medium for commerce. To date, a great deal of industry attention has focused on electronic transactions over the Internet. Although rapid growth is predicted over the next few years [10, 17, 21], actual sales thus far have been only moderate: users appear to regard the Internet primarily as a source of product information--when it comes time to pay, they prefer to buy offline by more conventional means [12, 14].

Responding to consumers' desire for information, businesses in large numbers have developed sites on the World Wide Web (WWW or Web). Most commercial Web sites describe the firm and its products and/or services, and many offer opportunities for visitors to the Web site to provide feedback and ask for specific information. As well, some Web sites collect information from visitors in order to improve future offerings. Some sites also support ordering and payment. The interactive potential of Web sites is particularly exciting, as it facilitates relationship marketing and customer support, eliminating the obstacles of geography and time [14, 22]. Not surprisingly, then, industry and scholarly research has recently focused on making Web sites more appealing and useful to visitors [13]. However, a Web site can only be effective if current and prospective customers visit it. Attracting this audience is currently a major challenge.

In this paper, we address the challenge of attracting a defined target audience to a Web site via *banner advertising*. We propose a framework for effectively targeting banner advertising in an electronic marketplace in a manner that benefits both advertisers and consumers. It allows advertisers to reach consumers who are more likely to be interested in the products and/or services offered by the company, and exposes consumers to information about products and services that they are likely to be interested in purchasing. Although the framework is discussed in terms of the Internet, we believe it will be relevant to whatever form the "information superhighway" eventually assumes. The framework takes advantage of the capabilities afforded by information technology for collecting and processing data about users. The next section examines trends in the electronic marketplace. Subsequently, the current state of advertising in this medium is discussed. Thereafter, a framework for targeting banner advertising, supported by appropriate information technologies, is proposed. Finally, opportunities for further research are discussed.

2. Marketing and Advertising in an Evolving Electronic Marketplace

The Internet began in the early 1970s as a US government research project designed primarily for the needs of the military. It expanded in the 1980s to serve the international academic and research communities [19, 23]. In the 1990s, businesses began to appear on the Internet. Although accurate estimates are obsolete as soon as they are made, it is clear that today tens of millions of people have access to the Internet [16] through over 100,000 computer networks in 150 countries--and the numbers continue to increase [14]. Two types of developments are particularly noteworthy with regard to this growth.

First, a large and ever expanding number of affluent, educated consumers are using the Internet [11]. This concentration of very desirable consumers has led to a surge in commercial interest. Prior to 1990, nodes on the Internet were predominantly academic institutions. In 1990, about 1,000 businesses had Internet connections. By June 1995, over 21,000 businesses were online, and the growth in commercial connectivity shows no sign of slowing [8].

Second, the emergence of the hypermedia-based WWW,

1060-3425/97 $10.00 © 1997 IEEE

together with point-and-click multimedia interfaces such as Netscape, have greatly increased usability of the Internet for persons without extensive computer training. The development of "applet" technology, such as Java, which allows programs to run on a variety of platforms, increases the transparency of various Internet services. In other words, as technology continues to evolve, it is no longer an obstacle to, but an enabler of, electronic commerce.

In this environment, companies are seeking ways to use the Internet effectively [1, 3, 13, 22]. One active area in electronic commerce involves using the Internet as a medium to communicate persuasive product and service information via advertisements. These take various forms, the most common of which are corporate Web sites and banner advertising. We define a banner advertisement as:

- paid communication (via text, graphics, video and/or audio) of information about an organization and/or its products and services
- by an identified sponsor
- embedded within, and visually distinct from, information provided by an online service
- with hypermedia links to the sponsor's Web site.

We distinguish banner advertising from simple hypermedia links (paid or not) to commercial Web sites: banner advertising conveys a message even if the user does not follow the link; simple links can only convey a message if the user follows the link. Banner advertisements are also distinct from what [14] refer to as "flat ads," single page advertisements that do not contain hypermedia links. In this paper, we restrict our discussion to banner advertising that appears in the course of users' browsing and searching activities on information services, such as Yahoo! (http://www.yahoo.com) and Excite (http://www.excite.com), that provide an entry point to Internet resources. Appendix 1 shows a banner advertisements by the Saturn automobile company.

Scant attention has been paid to banner advertising by researchers. This may be because banners seem relatively insignificant, especially when compared with the interactive richness of Web sites. Technical specifications for banner advertisements severely limit creative options and preclude any consumer-firm interaction beyond the consumer's selection of the hypermedia link to the associated Web site (Excite, for instance, specifies that "all banners are 468x60 pixels, gif format only, maximum file size is 7k" [9]). Banner advertisements are, however, very important and interesting when viewed as part of a *system* that converts browsers and searchers into Web site visitors and, ultimately, customers. In their model of this conversion process, Berthon, Pitt and Watson [3] identify a sequence of tasks. First, users must be made *aware* of the Web site, then they must be *attracted to* and *locate* the site. Once users have found the Web site, the task is to turn that hit into a visit, ensuring there is some meaningful *contact* between the firm and the consumer; then to *convert* the visit into a purchase. The final task is to get purchasers to return to the Web site and *repurchase*. Each task in the sequence is dependent on the successful execution of the previous task.

Our view of the role of banner advertising in this system is as a mechanism to make target audience members aware of a firm's Web site and to attract those users to the site. We define two concepts critical to understanding this role. *Attraction effectiveness* is the number of target audience members who reach a company's Web site via a banner advertisement hypermedia link divided by the number of target audience members who use the information service on which the advertisement appears. *Attraction efficiency* is the advertising cost per target audience member attracted to a company's Web site via a banner advertisement.

There is some evidence that the attraction efficiency of banner advertising is low. A recent estimate indicates that only 1-2% of banner advertisements lead viewers to seek additional information (e.g., by selecting a hypermedia link to the company's Web site) [5]. Since information services charge advertisers based on number of exposures (e.g., [9, 24]), the cost of attracting a single target audience member to a Web site is at least 50 to 100 times what it would be if all users who were exposed to the advertisement selected the hypermedia link. (The cost is even higher if some users selecting the link are not target audience members.) Increasing attraction efficiency by reducing wasted exposures should therefore be a priority. (An additional motivation for improving performance of banner advertising in converting searchers and browsers into Web site visitors arises from recent events such as the agreement between Procter & Gamble and Yahoo! which provides for payment based on the number of people who actually seek additional information (by selecting a link from a banner advertisement) rather than those who are merely exposed to the advertisement [20]. Such arrangements are expected to pressure online services to eliminate wasted exposures [5].)

The estimate cited above does not provide evidence on the attraction effectiveness of banner advertising. The fact that only 1-2% of exposed users select a link to the advertiser's site is irrelevant to effectiveness if all target audience members using the information service are among this group. However, since banner advertisements on online information services are shown selectively to users, there will generally be the possibility that some target audience members who use the information service will not be exposed to the advertisement and, hence, will be unable to link to a company's Web site via it. Depending on the strategy used to select advertisements for users, a large number of target audience members may be missed.

We contend that both the attraction effectiveness and efficiency of banner advertising can be improved by

precisely *targeting* advertisements based on characteristics and behavior of individual users of information services. Moreover, such targeting can be more precise than the targeting possible in traditional media. For example, visitors to a "Travel" page on an information service may be good targets for an advertisement for discount airfares, as would readers of the Travel section of a newspaper. But the fact that the online visitors have made a series of decisions and taken a series of actions (i.e., selecting only a subset of highlighted links within a hierarchical menu of categories) to reach the Travel page, rather than some other page (e.g., the Home Decorating page) suggests they may have a greater interest in travel than, say, readers who unintentionally come upon the Travel section of a newspaper and decide to read it. Since these exposures are more likely to be target audience members, attraction effectiveness can be improved. Targeting individual users strategy should also lead to fewer wasted exposures, since the advertisement would not be shown to users who have not reached the Travel page, thereby improving attraction effectiveness. (See Appendix 2 for a similar example.)

At present, targeting of banner advertising does not always occur. For example, Appendix 3 shows an advertisement for Honda that appeared when Organic Gardening was selected from a hierarchical menu of categories. People interested in organic gardening may not be the best prospects for automobiles, as they are likely to be more environmentally sensitive than the general population and may feel that cars unnecessarily harm the environment.

Nevertheless, online information services do currently provide some targeting capability. As of August 1996, both Yahoo! [24] and Excite [9] offered advertisers three options: general rotation, geographic or content targeting, and keyword-based targeting. With "general rotation," banner advertisements rotate randomly through user searches and browsing on the site. The Honda advertisement that appeared on the Organic Gardening page in Appendix 3 was probably in general rotation. Restricted rotations permit advertisers to purchase space in specified content areas or by geographic region. For example, financial institutions can limit the exposure of their banner advertisements to users searching or browsing Business categories, and Canadian advertisers can choose to have their banner advertisements shown only to users who are searching or browsing in the Yahoo! Canada site. These two options are analogous to the targeting offered by traditional media such as newspapers, magazines, television, and radio [4].

The third option, keyword-based targeting, makes greater use of the targeting potential of information services. A company can buy keywords so that whenever a user enters one of those keywords during a search, s/he will be exposed to the company's banner advertisement. This ensures that the banner advertisement is presented only to people with a demonstrated interest in the area. For instance, a marketer of golf equipment might buy the keyword "golf." Every time a user enters "golf" in a search, a banner advertisement for the equipment would appear. This is analogous to the more precise targeting provided by magazines.

While these are useful strategies, they fail to take full advantage of the targeting potential of banner advertising. Current technology provides the capability to develop sophisticated and detailed profiles of individual users of information services based on individual characteristics and past patterns of behavior in using the information service. The next section proposes and describes informally two versions of a model for targeting banner advertising by using the information technology on which an online information service is built.

3. A Model for Targeted Advertising

In traditional media, the quality of the information available constrains an advertiser's ability to target advertising effectively and efficiently. For example, many media buying decisions are based on data provided by research bureaus such as the Audit Bureau of Circulations (ABC), Business Publication Audit of Circulation (BPA), Arbitron, and A.C. Nielsen, which collect data on the demographics and media habits of consumers, and sometimes on product usage and brands [4]. These survey data are cross-tabulated to develop a profile of the audience of each media vehicle. The audience profile is then compared to the target audience profile identified by the advertiser to determine where there is a good match. For instance, an automobile manufacturer might identify the target audience for an advertisement for a particular model of car as middle-income females, 18 to 34, with busy lifestyles. Based on research bureau data, as well as the experience and judgement of the media planner, media vehicles with good reach in that demographic group would be chosen. Realistically, though, this type of targeting is usually very approximate. For instance, no matter how well the media vehicle audience profile matches the target audience profile, it is likely that only a portion of the audience would be in the market for a new car.

Online banner advertising may be able to overcome this problem. It is possible to target users very precisely because data can remain associated with individuals, so advertisers can select exactly the users to whom they wish their advertising to be exposed. It may be possible, for example, to identify which users will be in the market for a new car in a particular year. The remainder of this section describes two versions of a model for targeting banner advertising by taking advantage of the technological capabilities of the online environment. The model is designed to be appropriate for use by information services which sell

advertising space.

3.1. Basic version

The basic version of the model requires that users be assigned unique identifiers (e.g., user accounts) when they first connect to the information service. Subsequently, they provide these identifiers each time they connect. Users also complete an online questionnaire the first time they use the information service. (Incentives to complete the questionnaire may be provided by informing users that the information will be used to filter out advertising for products in which they are likely not to be interested.) The questionnaire allows data to be collected on several dimensions, including: (1) demographic attributes such as geographic location, income, family lifecycle stage, occupation, and sex; (2) psychographic attributes such as travel patterns and hobbies; and (3) product and brand usage attributes. This element of the basic model permits a banner advertisement to be directed to users (and only those users) who fit certain criteria, assuming data were collected on relevant attributes. For instance, a banner advertisement for baby strollers could reach parents of children under five years old--and only individuals in that group.

In contrast, research bureau data uses demographic correlates (e.g., males and females, 18 to 34) to identify media vehicles that attract a relatively large proportion of the people in the identified demographic group [4]. The media vehicles thus chosen may miss members of the target group (e.g., older parents) and reach consumers not in the target group (e.g., people who are between 18 and 34 but do not have young children). Even audience data based on cross-tabulations, while they supply information on more variables, still cannot isolate individuals who are in the target audience. (For example, research bureau data may allow an advertiser to identify a magazine whose audience includes a large number of people between 18 and 34 who have young children, but there will still be some readers who are not in the target market.)

The second element of the basic model involves eliciting the target audience profile from advertisers. An advertiser can specify a target audience using any number of attributes about which data have been collected. These can be expressed conjunctively and/or disjunctively. For example, a specification may indicate that an advertisement is to be presented to all users who (1) have household incomes over $50,000, and (2) either work in a job that involves travel at least four times per year or have travelled on vacation in at least four of the past five years.

In this version of our model, the questionnaire determines the data collected about each user. The content of the questionnaire will vary depending on the nature of the information service, expected users, and expected advertisers. However, it is imperative to design the instrument carefully, in consultation with advertisers based on anticipated relevant target audience attributes.

The final element of the model consists of a mechanism to select banner advertisements to display to users. The target audience profiles supplied by advertisers provide a screening mechanism over users. Each time a user connects, his/her profile is compared to all target audience profiles from all advertisers. The user's profile will actually match some subset of those profiles. If the number of matches is small (and the session is long), it will be feasible to display all banner advertisements associated with the matched profiles during the user's session. However, if the number of matches is larger (or the session is short), presenting all advertisements associated with the matched profiles may overwhelm the user. In such a case, it will be necessary to present only a selection of the identified target advertisements. A rationing system would be needed so that users are not deluged with banner advertisements while advertisers are assured of access to users who match the target audience profile.

In summary, the basic model has three elements: individual user profiles, individual advertisement target audience profiles, and a selection mechanism for presenting advertisements to specific users who match the target audience profile. This framework potentially eliminates wasted exposures and provides the capability to reach every single user who matches the target audience profile (this may not be realized if a rationing system is used). Users also benefit, since they will see advertisements only for products likely to be of interest to them.

3.2. Enhanced Version

The basic version of the model relies on users completing a questionnaire when they initially use an information service. This is a straightforward mechanism to collect data about user characteristics for the purpose of targeting advertisements. A similar approach has been incorporated in a commercial product for use with online catalogs to direct shoppers to products in which they are interested [15]. However, the advantage of simplicity is offset by several potential limitations. First, such information may become outdated, sometimes quickly, as user preferences and characteristics change. To some extent, information can be kept up-to-date by either readministering the questionnaire periodically or giving the user the opportunity to update her/his information (e.g., by a menu option or hypertext link) each time s/he connects to the information service. However, each of these strategies is intrusive and may impose an unwarranted burden on users in order to maintain currency of information.

A second, and perhaps more serious limitation of the

questionnaire strategy is that it is subject to two potential types of bias. First, the questionnaire designer will want to identify as many user attributes relevant to potential advertisers as possible. As the number of attributes increases, so does the length of the questionnaire, creating the possibility of higher mortality in completing the questionnaire (especially since it may be more difficult to induce users to complete it because they are both physically and psychologically remote), thereby increasing the potential nonresponse bias [7]. Second, the questionnaire method is plagued with well-known problems, such as errors due to inaccurate recall, telescoping, social desirability concerns, and cognitive biases, as well as ambiguity, intimidation, confusion, and incomprehensibility [2].

In view of these potential problems, it is appropriate to enhance the model so that it does not rely on user self-reports, can accommodate changing user characteristics and preferences, and is less constrained by the choice of questions. Fortunately, information technology may provide assistance in each of these areas.

Current technology allows a considerable amount of data about user search activities (both deliberate search and browsing) to be collected unobtrusively and analyzed to determine patterns. (We are dealing here only with the capabilities of the technology, not with the ethical issues such capabilities raise. However, we recognize that ethical issues must be considered explicitly in the design of systems based on our model. For instance, we believe users should be aware that such information may be collected, and how it may be used, and consent to this activity before using an information service.) In the enhanced model, we propose that patterns of search and browsing behavior exhibited by users while using an information service determine which advertisements are shown to that user during current or future sessions. In the remainder of this section, we provide a general overview of this approach.

As before, this model relies on assigning a unique identifier to each user for recording her/his searching and browsing activities while using the information service. Each session constitutes a "record", consisting of data such as: sites visited in order; pattern of navigation through a hierarchical category structure (as in Yahoo!); choice of search terms in keyword-based searches; and reaction to previously exposed targeted banner advertisements (e.g., which linked Web sites are selected and visited by the user and which ones ignored). The aggregate of such records for each user provides a profile from which preferences can be implicitly generated. As a simple example, if a user has made several searches using keywords such as "Atlantic salmon" and "fly fishing", and has visited the site of the Angling Club Lax-a of Iceland (http://www.ismennt.is/fyr_stofn/lax-a/uk/angl_uk.html), s/he may be targeted for a banner advertisement for a fishing lodge in Alaska. However, if a user has previously been exposed to the same or similar banner advertisements but has not visited linked Web sites when there was an opportunity to do so, s/he may not be shown these banner advertisements in future.

This version of the model has the advantage of transparency. A user simply visits a service for whatever purpose s/he has in mind. Data are collected unobtrusively in the course of the visit. Moreover, the data reflect actual user behavior, rather than attitudes, intentions, or reported behavior captured through a questionnaire. Hence, the quality of data derived from user behavior should be superior to that of questionnaire data, for purposes of targeting advertisements.

A disadvantage of this model is the preparatory work involved on two fronts. First, it is not clear how to structure the data collected during visits so that useful information can easily be coded for storage and later extraction. Research is needed to develop useful and efficient coding mechanisms for storing such data as sequences of visits and search terms used. We expect this can be handled using conventional database structures such as relations (tables); however, the design of a relational database for this purpose is itself a distinct research issue. Second, the ability to store the required data does not necessarily mean useful information can be extracted from it. Further research is required to determine the types of analyses that yield insights into user characteristics and preferences hidden in the data.

The enhanced model should be used in conjunction with the basic model. A questionnaire may be very effective for identifying various demographic data relevant to advertisers but impossible to ascertain simply from users' online search and browsing behavior. However, since demographic data has limitations for effectively targeting consumers of most products, the enhanced model of data collection may yield complementary data on preferences from patterns of online search and browsing behavior.

The next section describes an implementation architecture for the basic version of the model. Extensions that support the enhanced version of the model remain as future research.

4. An Implementation Architecture

The architecture required to implement the basic version of the model consists of two parts: data structure to represent user profiles and target audience profiles, and an algorithm to select banner advertisements to display to a user. This section describes these components.

4.1. Data Structure

To target banner advertisements, two types of profiles are needed: profiles describing users of the information service; and profiles describing the target audience for advertisements, as defined by advertisers. Each profile can be modeled as a set of attributes.

We assume there is a finite "universe" of attributes, $A = <a_1,...,a_N>$, that may potentially characterize users or target audience members.

4.1.1. User Profile. Each user, u_i, of the service can be described by a record consisting of values of the universe of attributes, $R_i = <a_1(u_i),...,a_N(u_i)>$, where $a_n(u_i)$ (n=1,...,N) denotes the value of attribute a_n for user u_i. This may be implemented in a relational database in which a table is defined whose primary key is a user identifier, and remaining attributes are those in A. Each row in the table contains the profile of one user. (A more elaborate data structure is needed to support the enhanced model, since data must also be kept about the pattern of behavior of a user over one or more sessions.) All attributes need not be applicable or relevant to a particular user; hence, null values are permitted.

A simple example serves to illustrate this structure. Consider a universe consisting of three attributes: age, income, and number of dependents. Suppose there are two users of a service. When those users have completed a profile questionnaire, the resulting data may be stored in a relational table as:

USER

user_id	age	income	dependents
u_1	26	34000	0
u_2	45	54000	2

4.1.2. Target Audience Profile. A target audience profile is associated with each banner advertisement. A profile may be expressed as:

(1) *A characterization of an "ideal" target audience member.*
Such an ideal can be described by a record consisting of values of the universe of attributes, $T_i = <t_1,...,t_N>$, where t_n (n=1,...,N) is a specific value of attribute a_n. Some values may be null, indicating that any values of those attributes are permitted for the ideal; and/or a

(2) *A characterization of the "acceptable" target audience.*
Generally, an advertiser is interested in reaching those within specified ranges of the attributes of interest. Given N attributes of interest, acceptability can be thought of as a region in N-dimensional space. This region can be defined by specifying ranges of acceptable values for various attributes in the universe. Combinations of attributes may be expressed: conjunctively, indicating that users in the target region must satisfy all the conditions or restrictions; disjunctively, indicating that acceptable users must satisfy one of a set of conditions; or using a combination of disjunctions and conjunctions.

Note that "distance" from the ideal point may become relevant if an advertiser has to choose a subset of users whose profiles fall within the acceptable region.

Operationally, profiles for ideal or acceptable users can be maintained in a relational database structure. In the case of ideal profiles, a table can be defined in which each row describes the ideal target audience member for each advertisement. The primary key for this table consists of an identifier for the advertisement, while the remaining attributes are those of the universe of attributes of interest. Since not all attributes may be relevant in specifying an ideal, null values are permitted.

To illustrate, consider a simple example in which there are two advertisements, each with a different target audience profile, designated T_1 and T_2. The ideal target profile for T_1 is users aged 35 with incomes of $50,000 (no restrictions on number of dependents), while that for T_2 is users aged 25 with incomes of $25,000 and no dependents. These profiles are shown in the following relational table.

TARGET

ad_id	age	income	dependents
T_1	35	50000	-
T_2	25	25000	0

In this example, the space of profiles is three-dimensional. Since the "dependents" attribute is not relevant in describing the target audience of the first advertisement, the ideal profile can be depicted as a point in two-dimensional space, as shown in Figure 1. In the case of target audience profiles based on attribute values within a range, the necessary data structure can be provided by a table whose rows describe the acceptable ranges of specified attributes for each profile. Each row in this table provides a lower and an upper bound on a specified attribute for a specified profile. The primary key of this table consists of the identifier of the profile plus the name of the attribute. To illustrate, consider a variation of the example above. Suppose the profile T_1 is no longer age = 35, income = $50,000, but is relaxed: the advertiser will accept any user between 20 and 50 with an income between $40,000 and $60,000. Similarly, suppose the profile T_2 is relaxed to encompass ages from 20 to 30 and incomes from $20,000 to $30,000. Such profiles might be stored as follows:

RANGE

ad_id	attribute	lower	upper
T_1	age	20	50
T_1	income	40000	60000
T_2	age	20	30
T_2	income	20000	30000

In this case, the acceptable profile can be depicted as a region in two-dimensional space. Figure 2 shows the profile for T_1.

It is possible that both ideal and acceptable profiles could be generated by the same advertiser. By overlaying Figures 1 and 2, shown in Figure 3, we note that the ideal point need not lie at the geometric center of the acceptable region.

Figure 1
Ideal Profile

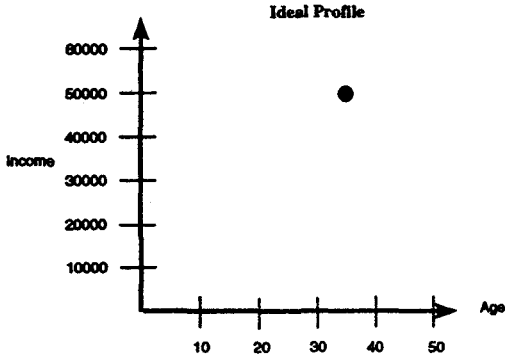


Figure 2
Acceptable Profile

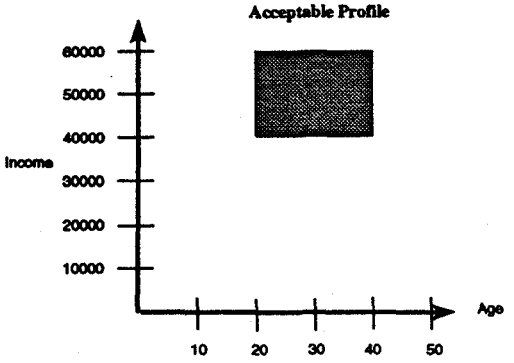
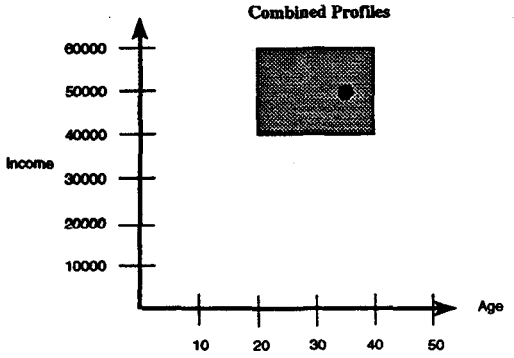


Figure 3
Combined Profiles



To handle measures of "distance" from an ideal, ranges of values on relevant attributes can be replaced with advertiser-specified information about the acceptable distributions of values over attributes. For instance, an advertiser may specify a mean (ideal) and standard deviation for an attribute if "acceptability" is normally distributed about a central value. Other measures of central tendency and dispersion may be appropriate for attributes in which the range of acceptability is quantified differently. The data structure of the RANGE table can be modified to accommodate this additional complexity.

4.2. Selecting Advertisements for Users

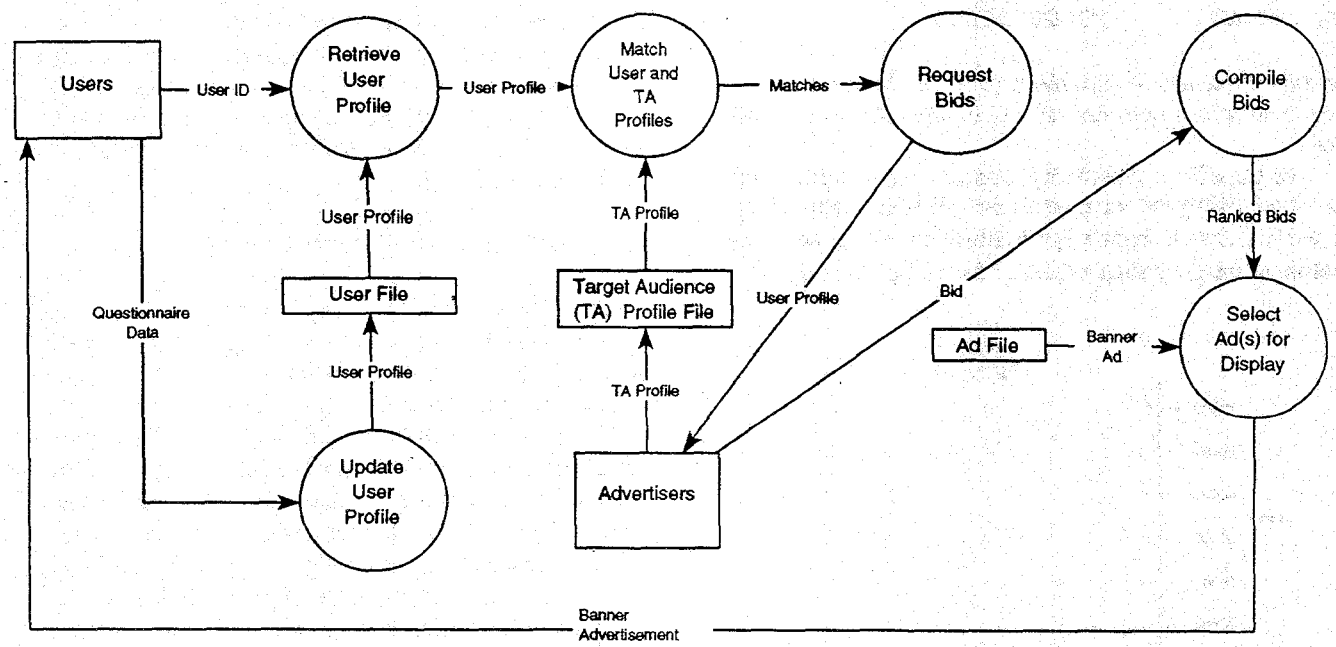
The primary challenge in effectively and efficiently targeting banner advertising is matching user profiles with target audience profiles. Figure 4 uses a data flow diagram to depict the matching process described below.

When a registered user visits an information service, his/her profile is retrieved. This profile is then compared with the target audience profiles of banner advertisements currently being run by the information service. Each target audience profile is associated with a banner advertisement. For each target audience profile, if there is no match with the user profile, the associated advertisement is dropped from further consideration.

After the comparisons are completed, a set of matched target audience profiles from a variety of advertisers remains. If this set is small, it may be feasible to show all the associated banner advertisements to the user during the session. In general, though, it will be necessary to select some subset of advertisements from the matched set to display to the user. We envision that the advertisers whose advertisements are in the matched set will compete for the opportunity to have their banner advertisements displayed to the user.

The concept of acceptable regions in target audience profiles provides a basis for competition. Profiles accommodate the possibility that some users within the region of acceptability may be more desirable to an advertiser than others. Hence, a distance metric capturing the relative desirability of a user with respect to an ideal profile is possible. It is not the purpose of this paper to propose or evaluate metrics. However, recognizing a notion of distance allows the possibility for advertisers to "bid" for the opportunity to display an advertisement to a user. Such bids would be determined by the advertiser, based on variables such as the user profile (to determine the distance from the ideal target audience profile) and advertising budget. It may be feasible to automate this by having software agents associated with each advertisement that would calculate the distance measure for the user and formulate a bid based on this, in addition to other

Figure 4
Selecting Advertisements for Users



information such as whether the user had seen this advertisement, or other advertisements for the same or similar products, in previous sessions (information which could be carried as part of the user profile).

When bids are received, they can be ranked. The banner advertisement corresponding to the winning bid is displayed to the user. Other advertisements may be displayed according to their ranking if there is an opportunity to display additional advertisements (e.g., if the user engages in several search or browse activities during a session).

This architecture provides guidance for implementing the basic version of the model. We present next a simple example showing how the architecture operates.

4.3. Example

Consider the relational database tables USER, TARGET, and RANGE presented earlier. Suppose first that the user with profile u_1 connects to the information service. This user's profile, consisting of the database record $<u_1, 26, 34000, 0>$ is retrieved from the USER table. Next, target audience profiles T_1 and T_2 are retrieved from the TARGET table. These identifiers determine the attributes whose profile ranges have to be selected from the RANGE table. Next, the age range for T_1, namely (20,30), is retrieved from RANGE. Since the age value of u_1 is 26, there is a match on this criterion. So, the salary range for T_1, (40000,60000) is retrieved. Since the user u_1 does not match this criterion of the target audience profile (salary is 34000), the advertisement corresponding to the profile T_1 will not be shown to u_1. Applying the same operations to the target audience profile T_2, u_1 would not be exposed to the advertisement corresponding to T_2 since the income of u_1 (34000) is greater than the upper bound of 30000 specified in the target audience profile T_2. Thus the user with profile u_1 would not be exposed to any banner advertisements when s/he used the information service. This is efficient, since showing either banner advertisement to the user with profile u_1 would entail a cost and constitute a wasted exposure.

Suppose now that the user with profile u_2 connects to the information service. This user's profile, $<u_2, 45, 54000, 2>$, is first retrieved from the user profile file. The target audience profiles T_1 and T_2 are then retrieved. Applying the matching algorithm, a match will be found between u_2 and the profile T_1. (Note that the target audience profiles in our example do not specify restrictions on number of dependents; hence, any values are permitted on this attribute.) However, there is no match between u_2 and T_2 since the income of u_2 (54000) is beyond the upper bound of income for T_2 (30000). Hence, the user with profile u_2 will be exposed to the advertisement corresponding to the target audience profile T_2.

This example does not show the full scope of the model, since there is no case where there are two or more target audience profiles that match a particular user profile. To illustrate this, consider an additional target audience profile having only the condition that user income must be at least 50000. This profile would require adding the following record to the RANGE table: <P_3, income, 50000, - >. Now if the user with profile u_2 connects to the information service, a match with both T_2 and T_3 will be found. In this case, the advertisers (or software agents) responsible for T_2 and T_3 will be contacted and provided with the profile u_2. Each advertiser (agent) will prepare a bid indicating how much it is willing to pay to have the banner advertisement corresponding to its profile exposed to the user with profile u_2. These bids are compiled and returned to the information service, where they are ranked. If we allow that the user will be shown only one advertisement, the one which placed the highest bid will be chosen.

In summary, this model makes use of rich, multiattribute data at the individual user level in determining whether each one will be exposed to a particular banner advertisement. This leads to more effective and efficient targeting than is possible using strategies such as general rotation, which does not use data at the individual level, and restricted rotation or keyword search, which rely on only a single data item about an individual in determining which banner advertisement(s) to present to the user. It is also more effective and efficient than targeting in the traditional media, which does not use any data at the individual level.

5. Future Research

This paper has presented a framework for leveraging information technology to target online banner advertising more effectively to benefit both users (who would be exposed only to advertising that is very probably of interest to them) and advertisers (whose advertisements would reach only those users who fit the target audience profile). This framework is, however, merely a starting point. Additional research on several fronts is needed before its potential can be realized. Several specific research concerns have already been noted. In addition, there are more general issues.

First, a system supporting the basic version of the model, based on the implementation architecture presented in this paper, should be implemented.

Second, an implementation supporting the enhanced version of the model is needed. This will require research to develop a more sophisticated database structure that can preserve users' searching and browsing behavior over time. In addition, techniques for detecting patterns of behavior are needed.

Third, both theoretical and empirical research is needed to explore agent bidding in the context of the framework proposed in this paper.

Fourth, empirical work needs to be done to evaluate the relative effectiveness and efficiency of this framework. A priority should be to compare the (1) the basic version of the model, (2) the enhanced version of the model, and (3) existing approaches to targeting advertisements. For instance, it would be interesting to test whether placing a banner advertisement on a relevant page (e.g., an advertisement for a new movie on the Entertainment page of an information service) would be more or less effective than directing the same advertisement to individual users selected on the basis of their answers to a questionnaire (i.e., the simple version of the model) or their search and browsing behavior (i.e., the enhanced version of the model).

Finally, the utility of this framework in other online contexts should be investigated. For instance, this approach could be used in developing Web sites that are more useful to visitors. Visitors with different profiles could automatically be shown different pages more likely to be of interest to them, eliminating the need for them to search the Web site for the information they desire.

6. Conclusion

Cyberspace is a new medium for advertising. In 1994, Edwin Artzt, chairman of Procter & Gamble, the largest advertiser in the United States, warned advertising agencies to "get their interactive act together" [6, p. 75]. As the advertisements in Appendices 1 and 3 show, even major advertisers and their agencies may not be taking full advantage of the opportunity to target their online banner advertising. The information technology that makes the WWW possible also allows the unobtrusive collection of detailed information about user interests based on their online searching and browsing. Advertisers should not assume that the same constraints that make media planning in traditional media a very inexact science also apply to online advertising.

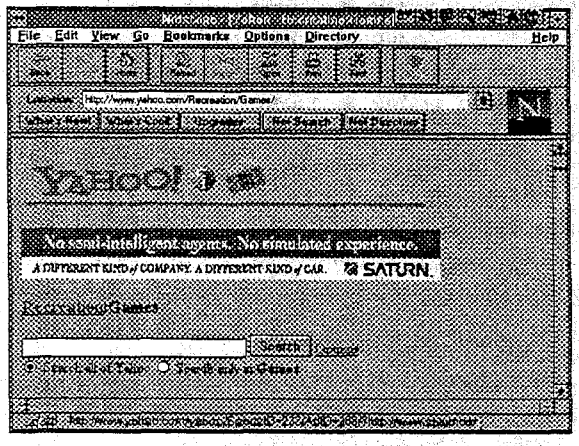
7. References

1. Armstrong, Arthur and John Hagel III (1996), "The Real Value of On-line Communities," *Harvard Business Review* 74(3), 134-141.
2. Barnes, James G. (1991), *Research for Marketing Decision Making*. Toronto, ON: McGraw-Hill Ryerson.
3. Berthon, Pierre, Leyland F. Pitt, and Richard T. Watson (1996), "The World Wide Web as an Advertising Medium: Toward an Understanding of Conversion Efficiency," *Journal of Advertising Research* 36(1), 43-54.
4. Bovée, Courtland L., John V. Thill, George P. Dovel, and Marian Burke Wood (1995), *Advertising Excellence*. New York, NY: McGraw-Hill.
5. *Business Week*, June 3, 1996, 44.
6. Castro, Janice (1995), "Just Click to Buy," *Time*, Spring, 74-

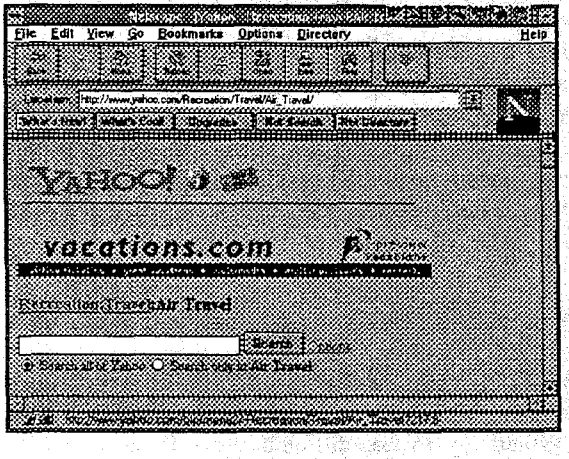
75.
7. Churchill, Gilbert A. (1991), Marketing Research: Methodological Foundations, Fifth Edition. Orlando, FL: Dryden.
8. Cortese, Amy (1995), "Cyberspace: Crafting Software that Will Let you Build a Business Out there," Business Week (February 27), 78-86.
9. Excite (1996), Excite Advertising Rate Card. [http://www.excite.com/info/advertising.html]
10. Greenwald, John (1995), "Battle for Remote Control" TIME Special Issue: Welcome to Cyberspace (Spring), 71-73.
11. Gupta, Sunil and Jim Pitkow (1995a), "Consumer Survey of WWW Users: Preliminary Results from the Third Survey," University of Michigan Business School. [http://www.umich.edu/~sgupta/hermes]'
12. Gupta, Sunil and Jim Pitkow (1995b), "Consumer Survey of WWW Users: Preliminary Results from the Fourth Survey," University of Michigan Business School. [http://www.umich.edu/~sgupta/hermes]
13. Hoffman, Donna L. and Thomas P. Novak (1996), "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations," Journal of Marketing 60(3), 50-68.
14. Hoffman, Donna L., Thomas P. Novak, and Patrali Chatterjee (1995), "Commercial Scenarios for the Web: Opportunities and Challenges," Journal of Computer-Mediated Communication, Special Issue on Electronic Commerce 1(December). [http://shum.huji.ac.il/jcmc/vol1/issue3/hoffman.html]
15. Information Week, "One-to-One Marketing Made Easy", Information Week, December 18, 1995, 110.
16. Kahn, Robert E. (1994), "The Role of Government in the Evolution of the Internet," Communications of the ACM 37(8), 15-19.
17. Karlgaard, Rich (1995), "Bozo Explosion," Forbes ASAP (December 4), 9.
18. Landau, Susan, Stephen Kent, Clint Brooks, Scott Charney, Dorothy Denning, Whitfield Diffe, Anthony Lauck, Douglas Miller, Peter Neumann, and David Sobel (1994), "Crypto Policy Perspectives," Communications of the ACM 37(8), 115-121.
19. Leiner, Barry M. (1994), "Internet Technology," Communications of the ACM 37(8), 32.
20. Ottawa Citizen, April 29, 1996, A10.
21. Press, Larry (1994), "Commercialization of the Internet," Communications of the ACM 37(11), 17-21.
22. Quelch, John A. and Lisa R. Klein (1996), "The Internet and International Marketing," Sloan Management Review 37(3), 60-75.
23. Tetzeli, Rick (1994), "The Internet and Your Business," Fortune 129 (March 7), 86-88+.
24. Yahoo! (1996), Sponsorship Program Opportunities [http://www.yahoo.com/docs/pr/adop.html]

8. Appendices

Appendix 1



Appendix 2



Appendix 3