OPTIMIZING OFFER SETS BASED ON USER PROFILES

Pelin Atahan, Monica Johar, and Sumit Sarkar
Özyeğin University, University of North Carolina at Charlotte, University of Texas at Dallas
pelin.atahan@ozyegin.edu.tr, msjohar@uncc.edu, sumit@utdallas.edu

Abstract

Personalization and recommendation systems are being increasingly utilized by ecommerce firms to provide personalized product offerings to visitors at the firms' web sites. These systems often recommend, at each interaction, multiple items (referred to as an offer set) that might be of interest to a visitor. When making recommendations firms typically attempt to maximize their expected payoffs from the offer set. This paper examines how a firm can maximize its expected payoffs by leveraging the knowledge of the profiles of visitors to their site. We provide a methodology that accounts for the interactions among items in an offer set in order to determine the expected payoff. Identifying the optimal offer set is a difficult problem when the number of candidate items to recommend is large. We develop an efficient heuristic for this problem, and show that it performs well for both small and large problem instances.

Keywords: Personalization, recommendation, e-commerce, probability theory

1. Introduction

Effective personalization can help firms reduce their customers' search costs and enhance customer loyalty. This, in turn, translates into increased cash inflows and enhanced profitability (Ansari and Mela 2003). Extant research has shown that in electronic shopping environments, personalized product recommendation enable customers to identify superior products with less effort (Häubl and Trifts 2000). These works have demonstrated that personalization can be an effective tool for firms.

The personalization process consists of two important activities, *learning* and *matching*. Learning involves collecting data from a customer's interactions with the firm and then making inferences from the data about the customer's profile. For instance, the relevant profile for a customer may be her membership in one of several possible demographic or psychographic segments, which could be based on age, gender, zip code, income, political beliefs, etc. (Montgomery et al. 2001, Wall Street Journal October 17 2007). Matching is the process of identifying products to recommend based on what is known about the customer's profile. Naturally, the quality of a customer's profile should impact the ability of the firm to provide high quality recommendations targeted towards sales (viz., the matching ability).

In this research, we examine how a firm can maximize its expected payoffs when making recommendations to users by leveraging the knowledge of the profiles of visitors to the site. In order to identify the best set of items to offer (e.g., links to a set of recommended items on a page that we call the *offer set*), a firm would first need a methodology to evaluate the expected payoff given an offer set. Then, the optimal offer set can be determined by selecting the set of items that maximizes the expected payoff for each page requested by the visitor based on what the firm knows about the visitor's item history (denoted by *IH*) and the profile. To evaluate the expected payoffs from an offer set the firm would need to evaluate the likelihood of each offered item being viewed and eventually purchased. The probability that an item will be viewed when provided in an offer set depends not only on the probability parameters associated with the item itself, but also on the other items in the offer set. Therefore, the interaction among items in an offer set should be accounted for when evaluating the expected payoffs from that offer set.

Extant literature has not formally analyzed the impact of the composition of an offer set on the resultant expected payoffs. Existing approaches that consider multiple recommendations typically sort association rules by some criteria like confidence or lift and simply take the top n items to recommend (Huang et al.

2004, Zaïane 2002). A novelty of the proposed approach is that it explicitly studies the impact of an item in the offer set on the probability of other items in the offer set being viewed and ultimately purchased when calculating the expected payoffs from that offer set.

In the next section, we present the framework to evaluate the expected payoffs from an offer set. A firm can evaluate all feasible offer sets using this framework and select the optimal one. We present in Section 3 an efficient heuristic approach to determine the offer sets quickly when the number of sets to evaluate is large. Section 4 discusses experiments to evaluate the performance of the proposed approach. Concluding remarks are provided in Section 5.

2. Evaluating the Expected Payoff from an Offer Set

The interactions between a visitor and the site are iterative in nature, with the firm providing a new offer set at each interaction (i.e., each time the visitor makes a page request). Figure 1 shows the choices faced by the visitor when provided with an offer set.

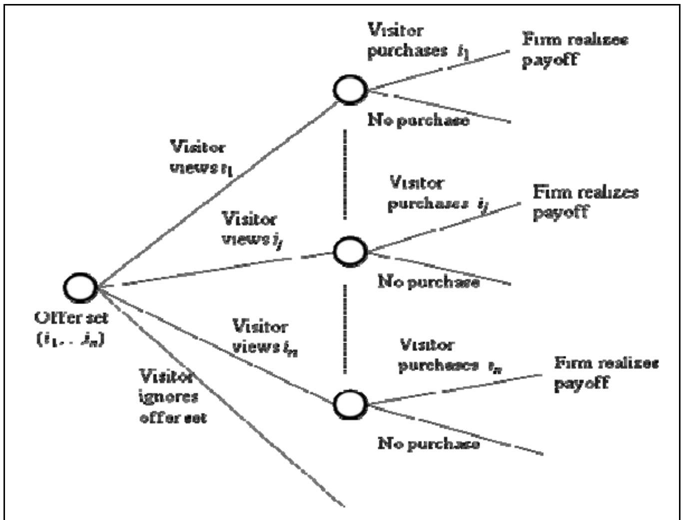

Figure 1. Interactions between a Visitor and the Site

Given an offer set, the visitor may either view detailed information on one of the offered items or ignore the offer set. When the visitor views information on one of the items (say i_j) by clicking on the appropriate link, the site provides detailed product information for item i_j, along with a new offer set (i.e., a new set of recommendations) in case the visitor does not like the product. If, on viewing the information on item i_j, the visitor decides to purchase that item, it results in a payoff to the firm. If the visitor does not purchase that item, then the visitor has the option of selecting an item from the new offer set for further evaluation, and the process repeats.

A visitor's decisions are driven by the visitor's profile and the items previously viewed by the visitor. A visitor's profile is represented by the set of possible classes (a_i) the visitor may be a member of, accompanied by the probability associated with each class. At any point in time, the visitor's item history is known to the site; and the site can drive a probability distribution of the visitor's profile information given the visitor's item history, i.e., the probability $P(a_i|IH)$ for each a_i (details of the belief revision process are suppressed for lack of space). To estimate the probability that a given visitor purchases an offered item i_j, the site needs to estimate the joint probability distribution of the visitor viewing the item (v_j), purchasing the item (s_j) and the visitor's profile, i.e., the site needs the joint probability $P(s_j,v_j,a_i|IH)$ for each a_i. This probability can be expressed as:

$$P(s_j,v_j,a_i|IH)= P(s_j|IH,v_j,a_i)P(v_j|IH,a_i)P(a_i|IH).$$

Given an offer set (O) and the knowledge about the visitor's profile, the firm can calculate the expected payoff from that offer set ($EP(O)$) in the following manner:

$$EP(O) = \sum_{a_i} \sum_{i_j \in O} P(s_j \,|\, IH, a_i, v_j, O) P(v_j \,|\, IH, a_i, O) P(a_i \,|\, IH) \omega_j \,.$$

ω_j is the profit realized from sales of item i_j. To simplify the exposition the profit from each item is assumed to be the same and equal to 1; we should point out that our approach can accommodate differentiated values for ω_j.

To operationalize this framework, the firm would need to estimate the following probability parameters associated with the choices made by the visitor.

- The probability that a visitor associated with a given profile and item history will view item i_j when presented with an offer set $O=\{i_1,\ldots,i_n\}$, i.e., $P(v_j|IH,a_i,O)$.
- The probability that such a visitor will purchase item i_j after viewing information on that item, i.e., $P(s_j|IH,v_j,a_i,O)$.

One approach to obtain the necessary parameters is by directly estimating them based on the historical data on customer interactions at that site. While that could be feasible for some of the above parameters, it would be very difficult for others because the number of feasible item histories and offer sets would be typically very large.

To help estimate the probability that a visitor will view information about an item that is part of the offer set we consider the use of association rules. For example, if a user has viewed items i_1 and i_2 and there exists a rule of the form $\{v_1,v_2\}\Rightarrow v_j$ this rule would provide $P(v_j|v_1,v_2)$.

To leverage the profile information of its visitors when making recommendations, a firm would need profile specific probabilities associated with user actions. For instance, to make gender specific recommendations a firm would need probabilities associated with male (m) and female (f) visitors' decisions to view and to purchase each item. For example, for the aforementioned rule, the firm would need the probabilities item i_j will be viewed by male and female visitors who have previously viewed items i_1 and i_2, i.e., $P(v_j|v_1,v_2,m)$ and $P(v_j|v_1,v_2,f)$. [1] Using the data from site's log files, the firm can also estimate the probability associated with item i_j being purchased by male and female visitors who have viewed the items in the rule antecedent, i.e., $P(s_j|v_1,v_2,m)$ and $P(s_j|v_1,v_2,f)$. The probability of purchasing item i_j is assumed to be independent of the other offered items conditioned on the visitor's class, visitor item history and the fact that the item has been viewed, i.e., $P(s_j|IH,v_j,a_i,O)=P(s_j|IH,v_j,a_i)$.

We next illustrate using an example how a site can estimate the probability that a visitor with a specific profile (e.g., gender) will view an item that is part of the offer set. The firm chooses a set of items to offer to a visitor based on the visitor's item history; the item history can be used to identify the *eligible* rules and the current belief about the visitor's gender. A rule is considered to be *eligible*, if its antecedent is a subset of the visitor's item history and the consequent is not a subset of the visitor's item history. We first discuss the methodology where there exist several rules with antecedents that match the visitor's item history. The situation where antecedents of rules are proper subsets of the visitor's item history is similar and discussed later.

Imagine that the firm has two eligible rules R_1: $IH\Rightarrow v_1$ and R_2: $IH\Rightarrow v_2$, and is considering offering i_1 and i_2. The site would need to determine the probability the visitor will view either of the offered items or ignore the offer set. The class specific probabilities associated with these rules are $P(v_1|IH,m)$, $P(v_1|IH,f)$, $P(v_2|IH,m)$, and $P(v_2|IH,f)$. A male visitor's likelihood of viewing information on item i_1 when presented in offer set O is the probability $P(v_1|IH,O,m)$. Assuming that the visitor views one of the two items (event V), the probability that the user will view item i_1 is

$$P(v_1|IH,m,O,V)=\frac{P(v_1,IH,m)}{P(v_1,IH,m)+P(v_2,IH,m)}=\frac{P(v_1|IH,m)}{P(v_1|IH,m)+P(v_2|IH,m)}.$$

[1] A firm can estimate profile specific probabilities from the profile information of a subset of its users. Such data is available from market research agencies like comScore or AC Nielsen which collect personal information from a large panel of users and track their online behavior.

Consequently, $P(v_2|IH,m,O,V) = 1-P(v_1|IH,m,O,V)$. The corresponding probabilities for female visitors are obtained analogously.

We next consider the situation where the visitor does not view either item. Based on the available rules, the firm knows the probability that each item is of interest to a male user with item history *IH*, i.e., $P(v_1|IH,m)$ and $P(v_2|IH,m)$. Then, the probability that a male visitor with item history *IH* is not interested in either of the items offered, denoted $P(\phi|IH,m,O)$, can be estimated as follows $P(\phi|IH,m,O) = 1 - P(v_1|IH,m) - P(v_2|IH,m) + P(v_1,v_2|IH,m)$.

Using the chain rule, the term $P(v_1,v_2|IH,m)$ can be written as $P(v_1|v_2,IH,m)*P(v_2|IH,m)$ or as $P(v_2|v_1,IH,m)*P(v_1|IH,m)$. The joint probability $P(v_1,v_2|IH,m)$ can be calculated directly if the firm has rules of the form $\{IH, v_1\}\Rightarrow v_2$ or $\{IH,v_2\}\Rightarrow v_1$. If neither of these rules is available, this implies a weak dependency between the two items given the profile and the item history. In that case, it is reasonable to assume the probability of viewing item i_1 is independent of the probability of viewing item i_2 conditioned on the profile and the item history, i.e., $P(v_1|v_2,IH,m) = P(v_1|IH,m)$. We then have $P(v_1,v_2|IH,m) = P(v_1|IH,m)* P(v_2|IH,m)$.

The probability that a male user with item history *IH* will view an offered link is then $P(V|IH,m,O) =1 - P(\phi|IH,m,O)$. The unconditional probability that a male visitor with item history *IH* will view item i_1 (i.e., without assuming the visitor must view an offered item), $P(v_1|IH,m,O)$, is then obtained as $P(v_1|IH,m,O) =P(v_1|IH,m,O,V)*P(V|IH,m,O)$. The probability $P(v_1|IH,f,O)$ can be estimated similarly. The above analysis easily extends to offer sets comprising of any number of items, and for profile attributes that can take any number of values.

As the size of the item history increases it will be difficult to find association rules with antecedents that perfectly match the entire item history. However, there will usually exist many eligible rules when the item history is large. The firm can then consider for inclusion in the offer set the consequents of eligible rules which have maximal antecedents (an antecedent is maximal if there does not exist another eligible rule whose antecedent is a superset of the target rule's antecedent). The rest of the procedure will remain unchanged.

3. Determining the Optimal Offer Set

The firm's objective is to select the offer set (including a predetermined number of items *n*) that maximizes its expected payoff. The items are chosen from consequents of eligible rules at each interaction. An obvious way to identify the offer set that maximizes the firm's expected payoffs would be to evaluate all feasible offer sets and then provide the offer set that leads to the highest expected payoff. However, when the number of items for consideration is large it may not be feasible to evaluate all possible offer sets in real time. We develop an efficient heuristic approach to determine the offer set in such situations.

3.1 Algorithm to Determine Offer Sets

Our approach selects items to include in the offer set in an iterative manner. It identifies items that have high probability of being viewed and purchased by visitors of each class, so that they contribute highly to the expected payoff from the corresponding class. It creates as many lists as the number of classes, where each list includes items more likely to be viewed by members of that class, i.e., items for which $P(a_i|IH,v_j)>P(a_i|IH)$. Then, items in each of the lists are sorted by their *item value.* An item's value for a given class is calculated by the product of an item's likelihood of being viewed and purchased by members of that class, i.e., item value of i_j in the list associated with class a_i is calculated as $P(s_j|IH,v_j,a_i)P(v_j|IH,a_i)$. The algorithm then compares the expected payoffs from offer sets that are created by adding the highest contributing item from each class. When comparing expected payoffs it disregards the likelihood of a user ignoring the entire offer set. Otherwise, the algorithm would be overly biased in earlier iterations to select links that have a high probability of being viewed.

4. Experiments

To validate our approach we have performed simulated experiments (we do not have access to real world data). We use expected payoffs from the identified offer sets as a measure of performance. We compare the performance of the proposed approach with that of the optimal offer set for many problem instances.

In our experiments, we used a binary class attribute for a visitor's profile. To generate the probabilities associated with a member of a class viewing an item, we generated the distribution of the profile of visitors $P(a_i|v_j)$ who view each item and each item's overall popularity $P(v_j)$ based on uniform distributions. We then obtained probabilities associated with each item being viewed by members of a specific class assuming a population prior $P(a_i)=P(a_{-i})=0.5$.

We expect the profiles of visitors who view an item to be correlated with the profile of visitors who purchase that item. The probabilities associated with purchasing the item $P(s_j|a_i,v_j)$ were generated by mixing a uniform distribution with the distribution of $P(v_j|a_i)$ associated with members of that class viewing the item where specific levels of correlation were created between the two probabilities. Then, the purchase probabilities were normalized to be between 0 and 0.3. We performed experiments on different datasets which had correlation levels of 0.6, 0.8 and 1.

To determine the optimal solution, we evaluated the expected payoffs from all possible offer sets and select the one that provides the highest expected payoff. In these experiments, the cardinality of the offer set is 8 and there are 40 candidate items. This leads to 76,904,685 possible offer sets to evaluate. We randomly generated 5 different datasets for each correlation level considered. On each dataset, for a given profile distribution, the proposed approach was implemented first to determine the offer set. Then each possible offer set was enumerated. The expected payoff from the offer set identified by the proposed approach was compared with the expected payoff from each of the other offer sets. We recorded the rank of the expected payoff from this offer set compared to all other offer sets and the percentage difference of the expected payoff from this offer set from that of the optimal offer set. We repeated the experiments on the same dataset for 11 different user profile distributions (profile probability for one class ranging from 0 to 1 in increments of 0.1). Then we conducted the same set of experiments on the datasets for each correlation level.

Table 1. Comparison with Optimal

User Profile Probability	Rank			Difference in Expected Payoff		
	Correlation Level			Correlation Level		
	0.6	0.8	1	0.6	0.8	1
0	25,652	22,326	203	-4.20%	-3.12%	-1.78%
0.1	3,186	3,572	30	-2.59%	-2.45%	-0.61%
0.2	1,028	1,391	6	-1.84%	-2.11%	-0.27%
0.3	295	284	5	-1.67%	-1.43%	-0.33%
0.4	351	388	35	-1.48%	-1.67%	-0.96%
0.5	11	716	19	-0.74%	-1.64%	-0.68%
0.6	37	975	8	-1.38%	-1.79%	-0.61%
0.7	71	134	5	-1.45%	-1.09%	-0.24%
0.8	627	641	83	-2.33%	-1.75%	-0.86%
0.9	1,446	1,373	235	-2.52%	-1.77%	-1.42%
1	5,745	2,617	237	-3.04%	-2.85%	-2.02%
Average	3,495	3,129	79	-2.11%	-1.97%	-0.89%

The results of these five sets of experiments are averaged for each correlation level in Table 1. Each row reports, for a given profile distribution, the rank of the solution in terms of expected payoff provided by the proposed approach among all possible offer sets and the average percentage difference in expected payoffs between the proposed offer set and the optimal offer set. The last row provides the results averaged over the different correlations considered.

The proposed approach performs well in all the experiments. In the worst case, the rank of the solution provided by the proposed approach is 25,652, which is in the top 1% of all possible offer sets. The overall performance of the proposed approach is even better in the experiments where the correlation level is higher, e.g., when the correlation is 1, the rank of the solution is within the top ten (out of more than 76 million) in several of the experiments. The performance can be explained as follows. Our approach considers the potential value of an item for one class and ignores the potential value of the item for the other class. In some cases, instead of including an item with the highest potential value for one class, it may be more profitable to include an item that has slightly lower potential value for that class, but much higher potential value for the other class. In such situations, because the proposed approach will fail to identify and include such items, the expected payoffs for the solution provided by the proposed approach may deviate more from the optimal solution. When the correlation level is high having items valuable for both classes is less likely. Therefore, the items identified by the proposed approach are more likely to be the most valuable items to include and the expected payoffs for the solution from the proposed approach will be much closer to the optimal expected payoff.

The percentage difference in expected payoff between the solution provided by the proposed approach and the optimal solution is quite small in general. It is around 2% or less on average. The performance of the proposed approach degrades slightly compared to the optimal approach at more extreme user profiles.

5. Conclusion and Discussions

Firms typically make multiple recommendations to visitors traversing their sites. However, extant research has not addressed how the multiple items in an offer set impact each other's view and purchase probabilities and hence a firm's expected payoffs from an offer set. We study how a firm should compose the offer set to maximize its payoffs from the recommendations. The framework presented would allow the firm to select the offer set that maximizes its expected payoffs based on the visitor's item history and the current beliefs regarding the visitor's profile. We propose an efficient heuristic algorithm to determine the offer sets quickly when there are a large number of items that are considered for inclusion in the offer set. Simulated experiments demonstrate that the heuristic performs well compared to the optimal approach. Ongoing experiments (not reported here) show that the performance of the proposed approach can be markedly better compared to that of a benchmark approach.

References

Ansari, A., and Mela, C. F. "E-Customization," *Journal of Marketing Research*, Vol. XL, May 2003, pp. 131-145.

Häubl, G., and Trifts, V. "Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids," *Marketing Science* (19: 1), 2000, pp 4-21.

Huang, Z., Chung, W., and Chen, H. "A Graph Model for E-Commerce Recommender Systems", *Journal of the American Society for Information Science and Technology*, 55(3) 2004, pp. 259–274.

Montgomery, A.L. "Applying Quantitative Marketing Techniques to the Internet," *Interface* (30:2), 2001, pp. 90-108.

The Wall Street Journal. "Firm Mines Offline Data to Target Online Ads," October 17, 2007.

Zaïane, O.R. "Building a recommender Agent for e-Learning Systems", in *Proceedings of the International Conference on Computers in Education(ICCE '02)*, 2002.