Recommending from Content:
Preliminary Results from an E-Commerce Experiment

Mark Rosenstein and Carol Lochbaum
Telcordia Technologies, Inc.
445 South St.
Morristown, NJ 07960 USA
+1 973 829 2000
{mbr, ccl}@research.telcordia.com

ABSTRACT

We are conducting an ongoing experiment into the effects of various forms of recommendations on consumer behavior at a web site. In this paper, we report on measures of the usefulness and effectiveness of recommendations based on content. During a three month period, we provided recommendations on over 2000 products at an e-commerce web site. To evaluate the effectiveness of recommendations on customer behavior, we collected three sets of metrics. First, we measured the rate at which recommendations were actually viewed by visitors. Second, we analyzed the paths visitors took through the recommendations. Finally, we measured the impact of recommended items on number of items purchased and on revenue.

Keywords

recommendations, latent sematic indexing, experiment, electronic commerce

INTRODUCTION

We are exploring the potential of recommendations to effectively guide users to resources of interest for a variety of applications. We have previously [Hill, 1995] looked at using community-based recommendations in suggesting "taste" driven goods, such as video recommendations. In this paper we describe an ongoing experiment being conducted in a business-to-business context of recommending from content.

The Telcordia Information SuperStore is an e-commerce site which sells documents, self-paced training, and other Telcordia products and services. At the beginning of the experiment, the catalog contained approximately 3000 items. The Store has been available to Internet customers for over three years, with a mix of domestic and international visitors.

While our heart is in Community-Based Recommendation [Hill, 1995], the typical consumer behavior for this site, with its strong emphasis on task-based item selection, did not easily lead to a community-based solution. We turned instead to a content-based recommendation system. As the basis for this system, we used a Telcordia patented technology, Latent Semantic Indexing (LSI) [Deerwester, 1990]. LSI uses singular value decomposition to reduce the dimensionality of a document space. In the reduced space, semantically related documents are arrayed closer to one another than dissimilar documents. We exploited the closeness of related documents to provide recommendations.

THE EXPERIMENT

At the beginning of the experiment, the Store's catalog contained 3104 items for sale, over which the Store allows browsing and search of product information, such as product description and price. Of these, 2082 items were selected for indexing. Items eliminated from the index included free brochures, and items with abstracts of less than five words. Items that were available under different license terms were combined into a single item for indexing purposes.

LSI indexing was run on the abstracts of catalog entries using global entropy, after eliminating common words. This resulted in a reduced matrix of 5206 terms, and 106 factors. Using a cosine distance metric on these results, a list of the top 10 closest items was generated for each item. The goal was to provide four recommendations for each item. The decision to display four recommendations was somewhat arbitrarily made, and based mostly on screen real estate constraints.

Two additional filters were applied to derive the final recommendations from the 10 closest items. Some items were collections of other items in the catalog. We eliminated from the recommendations for these collections any individual items from the set, since these items were already displayed to customers via the existing interface. We also eliminated from recommendation any item with an LSI score of less than .6, since below .6, the relevancy of the items rapidly dropped off. From the remaining items, the four closest items were selected for recommendations. Under these constraints, we were unable to generate a full set of four recommendations for 285 of the catalog entries, but we maintained as many recommendations as were possible for those items.

There were a number of choice points for the experiment. One criterion was to make this first experiment as simple as possible, and then use the resulting data to explore more general environments. There are a number of places in a customer's visit where recommendations can be offered. We chose to offer recommendations when a customer added an item to a shopping basket. The recommendations were specific to the item added, and no notice was taken of any other items in the basket, except that if one of the four recommended items was already in the basket, it was eliminated from the recommended set. The titles of the recommended items were presented to the customer in a table, with each title also serving as a hyperlink to further product information, such as a detailed product description.

On August 9, 1999 recommendations went live at the Store. We report on data analyzed through Nov 18, 1999.

RESULTS

Since recommendations were only given after customers added an item to their shopping basket, we will focus on that subpopulation. We found it convenient to divide this data into sessions that resulted in an on-line order and sessions that did not. Of the total number of sessions in which a customer added an item to their basket, in 38% they completed the purchase on-line. We hypothesize that non-buying customers who put items into their shopping baskets are likely serious prospective buyers, who purchase through channels that are not tracked by the Store. Some customers place their orders via telephone. Other customers place orders through a centralized organization within their company, and use the shopping basket facility to communicate their orders. We have not yet determined what percentage of non-orders are accounted for by these two mechanisms.

We were able to give recommendations for 86% of the items that were ordered during the time period that we have analyzed. One catalog item, which is a service, rather than a product, accounted for 77% of the cases in which we were unable to provide a recommendation. Approximately a quarter of the time, when recommendations were presented, a customer viewed at least one recommendation by clicking on a recommended title's hyperlink. There was no difference between buyers and non-buyers in this clicking behavior.

In 9% of the cases, the customer replaced the item that generated a recommendation with one of the recommended items. We believe that these instances indicate that the customer found an item better suited to their task, though any conclusion must be tempered by the small sample size. We saw both a case where the customer replaced an item with one costing ten times the original and a case where the customer found a suitable document at one-seventh the price of the original. On the occasions when a customer purchased both the item producing the recommendation and a recommended item, on average, customers chose recommended items slightly more expensive than the item they originally placed in their basket.

The strongest case for the efficacy of recommendations would be made by comparing data from the period before recommendations to the period with recommendations. While that is our eventual goal, we do not yet have enough data for that comparison. What we have done in the interim is to look at customers' paths through the Store and their interaction with the recommended items. In all cases, the recommended item is added after its recommendation, and it is usually added immediately after the customer looks at the abstract for the recommended item. Using that as the baseline, 8% of the items added to baskets in orders were recommended items and this resulted in a 7% increase in revenue.

CONCLUSIONS

These preliminary results are very encouraging. The LSI technology combined with our methodology can generate recommendations in an automatic fashion, so the cost of providing recommendations is very low, compared to the 7% increase in revenue. This increase has given us the confidence to continue the experiment, so that we may gather sufficient data for more detailed statistical tests. Additionally, we are planning to inject recommender technology at other points in the customer experience at the web site.

FUTURE WORK

We see at least two future directions for this work. The first is to examine the effects of some of the choices we made in this experiment. We would like to see if there is a noticeable effect if we increase the number of recommendations provided for each item. We would also like to vary where in the customer's experience the recommendations are given. Currently, recommendations are presented when a customer adds an item to the shopping basket. The hypothesis was that customers would be focused on purchasing this class of item at that moment, and be more receptive to other items of that class. We would like to experiment with recommendations at two alternative points. The first is when the customer is looking at a description of an item, where alternative, but similar items may be of interest. The second is when the user starts the checkout process. LSI retrieval allows us to use the entire contents of the basket as a search term. If there is a coherent structure over the items in the basket, LSI is likely to retrieve items relevant to all items in the basket. From a nonscientific survey of customers' baskets, we believe this is more often the case than not, but an experiment is necessary to see if customers find this type of recommendation useful.

The second future direction for this work is to try other recommendation mechanisms and compare their effectiveness to LSI. We would like to integrate our knowledge of what items the customer viewed with what is recommended, as well as more community-based knowledge of what bundles of items other customers have found useful.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dennis Egan, who organized and leads an informal group to study data mining issues in the e-commerce world, and members of the group, who provided valuable feedback and insight into this work. In particular, Joel Remde did the required LSI analyses.

REFERENCES

1. Hill, Will, Stead, Larry, Furnas, George and Rosenstein, Mark. Recommending and Evaluating Choices in a Virtual Community of Use, in *Proceedings of CHI '95* (Denver CO, May 1995), ACM Press.

2. Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K. & Harshman, R.A. (1990). Indexing by Latent Semantic Analysis. Journal of the American Society for Information Science 41, 391-407.