An implemented e-commerce shopping system which makes personal recommendations

TONG SUN & ANDRÉ TRUDEL
Jodrey School of Computer Science
Acadia University
Wolfville, Nova Scotia, B0P 1X0
Canada
email: andre.trudel@acadiau.ca

ABSTRACT

We present an implemented customer oriented online shopping system called HOPE (Helpful Online Purchase Environment). Data mining methods are used to automatically generate suggestions, which vary according to the customer and section being browsed. Customers receive personalized suggestions that match their own purchasing habits and preferences.

KEY WORDS

E-commerce, online shopping, data mining, kNN.

1 Introduction

"Marketing" is a worn out term linked with unsolicited phone calls, emails, or even strangers knocking on our door. We can eliminate these unsolicited contacts by using the latest AI technologies.

In the online business world, the consumer-vendor relationship is no longer merely buy and sell. Consumer survey forms are not the only choice for vendors to know and understand their consumers. Instead of asking survey questions like "What products do you like? What products do you often buy?", vendors can watch their consumers' purchasing habits. In the digital online world, monitoring consumer web browsing alone can reveal where their purchasing intentions are. Consumer buying tendencies can also be predicted from their purchasing history. Instead of asking the question "Are you planning to buy product A in 3 months?", the vendor can see what pages customers like to browse but have not made any purchases yet. For example, I liked to go to the camera section whenever I was in a bricks and mortar shop that sold cameras. Before making a purchase, I did an extensive survey of what modules and prices were available at various shops. Most people will look before they buy, especially for expensive products. Suppose my shopping research was not done in the stores, but all on the web. The vendors can monitor my browsing and anticipate my purchase. I would be glad if I received a voucher saying "We know you've been interested in our product recently. Here is a one-hundred-dollar

voucher if you buy it within a month".

In summary, in the digital online world, the consumer-vendor relationship is not confined to "buy and sell". There is an on-going consumer-vendor interaction, and valuable information can be derived from this "extra" interaction.

The new marketing concept is implemented in the HOPE (Helpful Online Purchase Environment) system, which is a customer oriented online shopping system that does not need surveys. Marketing is aided by data mining methods, which are used to gain knowledge of customers' purchasing behaviors. HOPE can automatically give purchasing suggestions, which vary according to the customer and the section being browsed. A feature of the HOPE system is that it treats each customer as an individual. Customers receive personalized suggestions that match their own purchasing habits and preferences. Market researchers can also benefit from the underlying data structure which allows easier data aggregation and analysis.
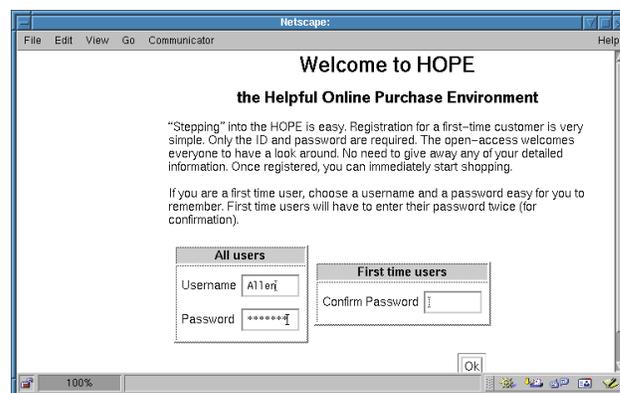
2 User Interface

Figure 1. Customer entrance

HOPE has two entrances, one for customers and the

other for the owner. The customer interface provides on-line registration and shopping. Registration for a first-time customer is simple. Only a customer generated ID and password are required (see Figure 1). The open-access concept welcomes potential customers who are browsing, since they do not need to provide any detailed information such as their name, email address, mailing address or credit card number. Once customers have registered, they can immediately start shopping. Unlike some other sites, no waiting for verification is necessary.
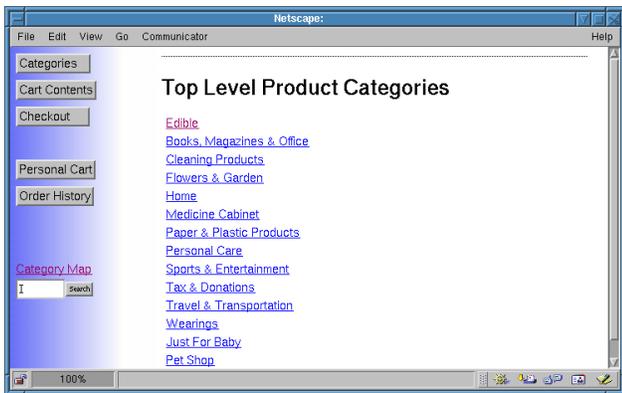

Figure 2. Top level product categories

Figure 2 is the first page shown to the user. The products are grouped into categories and sub-categories for easy navigation. For example, "sandal" belongs to the "shoes" category, which then belongs to the "wearing" category. Only the top level categories are initially shown. The customer can then drill down from the top categories into sub-categories, and finally down to the products where the actual shopping is done.

"Today's recommendation" (the personal flyers section), which appears after the categories (see Figure 3) is the advertisements. They inform the customer of discounts, sales or promotions for the goods that they tend to buy, with the prices shown in the square brackets. The recommendation system makes suggestions based on customers' own purchasing habits, and also on similar-interest customers. The concept is similar to word-of-mouth promotion in real life.

For example, after logging in (Figure 1), Allen is greeted with the top level product categories (Figure 2), along with an empty shopping cart (Figure 4). Allen likes baking, so the first thing he wanted to do was drill down the product category tree along the path "edible" → "Baking" → "Ready To Bake". Half way along the path, in the baking section, his favorite baking material (US ready to bake, large package) appeared at the top of the personal flyers section ("Today's Recommendation"). He bought two packages by typing "2" into the quantity entry box and checked it into the shopping cart with the check box with-
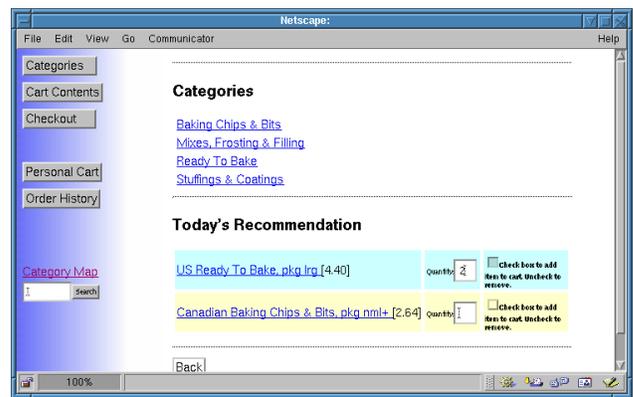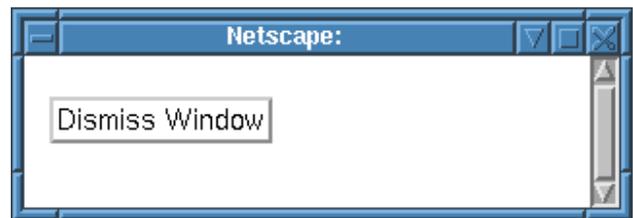

Figure 3. Bake category

Figure 4. Empty shopping cart

out going further down the product tree (Figure 3). The shopping cart window will immediately reflect the change (Figure 5).
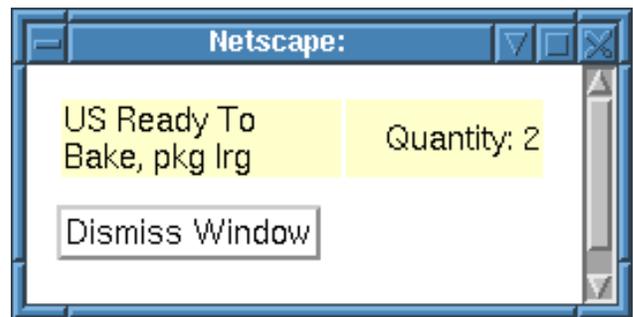

Figure 5. Shopping cart

Only relevant products are shown in the personal flyers section. When Allen was shopping in the baking section, he was not shown advertisements from other sections. When he changed to the "Snacks and Desserts" section, all advertisements were updated accordingly, reflecting the products that belonged to the section (Figure 6).

When Allen was traveling along the product category tree, his interest was in the product categories. The personal flyers were shown after the product categories. The personal flyers came as an aid, covering the products in the whole section while not interfering with Allen's browsing. The deeper he was down the product category tree, the
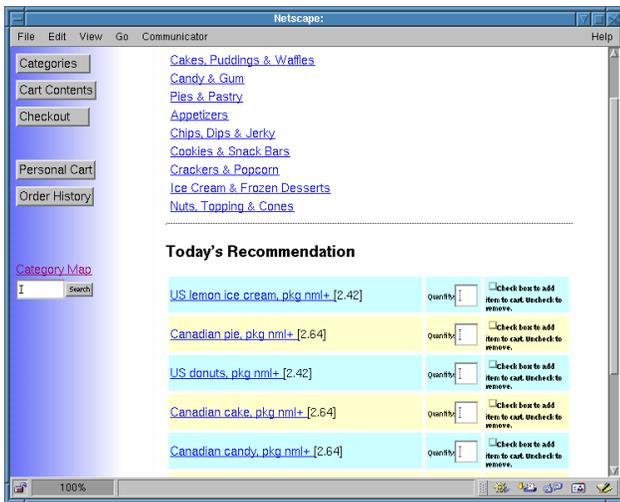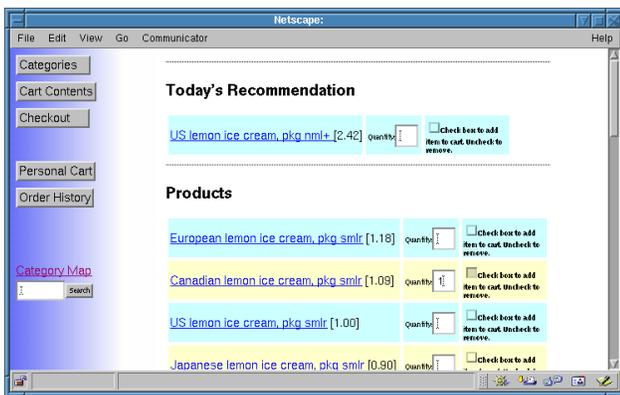
Figure 6. Adaptive personal recommendation

Figure 7. Product purchase page

more specific the advertisements became. When he entered an actual product page (e.g. lemon ice cream), the personal flyers were shown at the top of the page (Figure 7). The reason is that they are the products he is most likely to buy. The number of recommended products is small relative to the total number of products shown on the page (Figure 7). In this example, Allen did not choose his favorite but tried something new (one smaller package of Canadian lemon ice cream). This was the first time he tried it. His new choice will only affect the recommendation system if he repeatedly shows interest in this new product.

Note that all users are shown the list of categories at the top of Figure 3 and products at the bottom of Figure 7. The list of recommended products (either in the categories window (Figure 3) or products window (Figure 7)) is tailored to individuals and varies between users.

The "shopping cart" is used during the shopping experience by the customer to store chosen products. The shopping cart window is immediately updated whenever

products are chosen (Figures 4 and 5). After shopping, customers can "check out"(review and buy) their purchases. First-time shoppers are asked to enter detailed personal information. Returning customers however, can go ahead and purchase with their normal credit card, or have an option to use another card if they wish.

There is also a "regular shopping cart" (Figure 7, fourth menu button on the left hand side labelled "Personal Cart"), where customers can keep their frequently purchased products. The benefit is that customers can purchase their regular products directly from it. Some shopping sites on the web use this approach. The more items the regular shopping cart holds, the less helpful it is. Customers will find it harder to manage and locate goods in them, once the regular shopping cart contains too many items from different categories.
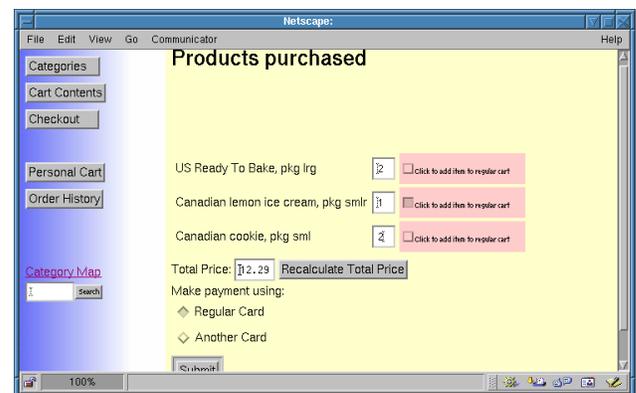

Figure 8. Check out page

Allen has finished his shopping. On the check out page (Figure 8), he clicked to add Canadian lemon ice cream to his regular cart so that next time he could buy it directly form the regular cart. Also, he decided to buy another package of cookies. All he needed to do was to increase the amount for the cookie and "recalculate total price". He then clicked on submit to purchase the items with his regular credit card.

The HOPE system can also help customers with their book-keeping, as it maintains their purchasing history. Figure 9 shows Allen's purchase history, which is listed in reverse chronological order. Each entry is an order ID followed by the purchase date. The window in Figure 9 is obtained by clicking on the "Order History" menu button in Figure 8. Figure 10 shows the detailed purchase he just made, it is shown when he clicked on the first entry (the latest purchase) of his purchase history in Figure 9.

3   Implementation

Our approach uses a "self-organization tag" field to represent the structure of the hierarchical data. Table 1
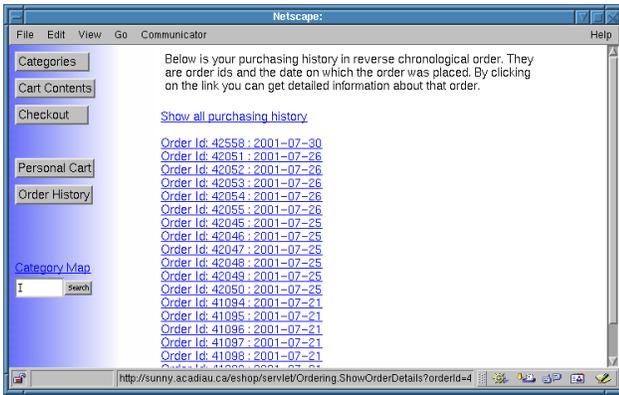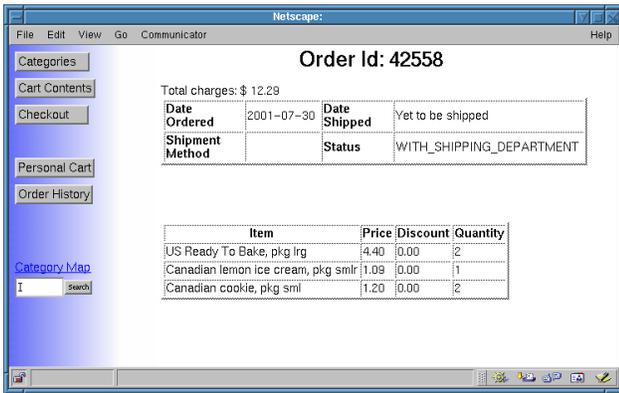
Figure 9. Customer purchase history

Figure 10. Customer purchase order log

shows the self-organization tags for a sales office and staff example. Each item (e.g., sales office or staff) is assigned a unique SOA (self organization approach) tag. For example, "Canada" has a tag of "ABA" and "US" has a tag of "ABB". To find an item's parent, simply trim the last character off the SOA tag. The tag for Canada's and US' parent is "AB" (i.e., America). Parent's parents are found the same way. Similarly, the children of a node can be easily found. For example, all of Canada's descendants SOA tags begin with "ABA", which is Canada's tag.

Recommendations are generated from three sources: "Personal Recommendation", "Target Advertisement" and "Word of Mouth Propagation". Personal Recommendation gives recommendations from the customers' own purchasing history. Customers would receive the advertisement as a convenient shortcut. Target Advertisement is for the vendors. The recommendation system can suggest a given number of potential customers for a given product. It is accurate for new products. The recommendation system can find similar products (based on its flavor, packages size, price range, etc) and make suggestions accordingly. Word of Mouth Propagation propagates suggestions among similar interest people. Different customers will see different

Table 1. Sample self-organization tag for staff and location

```
   "A","Total"
   "AA","Europe"
   "AAA","France"
   "AAAA","Paris"
5  "AB","America"
   "ABA","Canada"
   "ABAA","Ontario"
   "ABAAA","Ottawa"
   "ABAAAA","Nepean"
10 "ABAAAAA","CityView branch"
   "ABAAAAAA","Joe"
   "ABAAAAB","CentralPoint branch"
   "ABAAAABA","Sam"
   "ABAAAB","Kanata"
15 "ABAAB","Toronto"
   "ABB","US"
   "ABBA","Washington DC"
   "ABBAA","New York"
```

advertisements.

We use a kNN algorithm for generating recommendations. The kNN suggestion system is implemented in Java, in three main classes: Knn, DbKnn and DbHandler. The class Knn implements the normal kNN algorithm. It is built on top of the weka package [1]. The class DbKnn is an extended kNN class that uses the database table as the data source and destination. It is built on top of the class Knn. The DbHandler is the control class that does all the management, kNN training, self-testing, and suggesting. It is the console based command line interface. It controls the kNN training, and recommendation table generation (Figure 11). It uses the command line instead of a web interface for both security and efficiency reasons. The simplicity of the command line approach makes it ideal to put kNN learning and recommendation tasks into daily cron jobs.

Products are stored in the products table. The categories table holds products' category tree. Consumer information (user profile) is stored in tables profiles and addresses. Consumer's purchase history is recorded in table orders. It holds summary information of each consumer's purchase. Each purchase may contain many items. Each individual purchased item is stored in table one_order. These tables provide information for kNN training and suggestion. Table knn_raw holds a snapshot of the above information: the customers' purchase, customers' age, gender, products' producer, quantity, price, discount, etc. Table knn_raw is used for kNN training, i.e., generating the kNN knowledge bases for product and customer. The training results are stored in these two knowledge bases. They are used for the recommendation table generation. The generated rec-

ommendations are stored in database table `knn_suggest`. It holds suggestion information such as customer's profile ID and product ID. The recommendation system can query suggestions from it during "user recommendation look up" phase, i.e., when the customer is browsing the web pages.

Data tables are mainly inherited from eshop [2]. Most tables are changed to suit HOPE's shopping and suggestion environment. The database system is DB2 personal edition 7.1 for linux, which is the latest available version for Linux. The drawback of using it is that it only supports JDBC 1.0, not JDBC 2.0. This has hindered the possibility of developing a more powerful web interface.

4   Conclusion

HOPE's goal is to provide a user friendly customized online shopping environment. HOPE is fully implemented and operational. Details can be found in [3]. Future work involves testing the system with a large product and user base.

References

[1] I. Witten and E. Frank, Data Mining (Morgan Kaufman, 2000).

[2] V. Sharma and R. Sharma, Developing e-Commerce Sites (Addison Wesley, 2000).

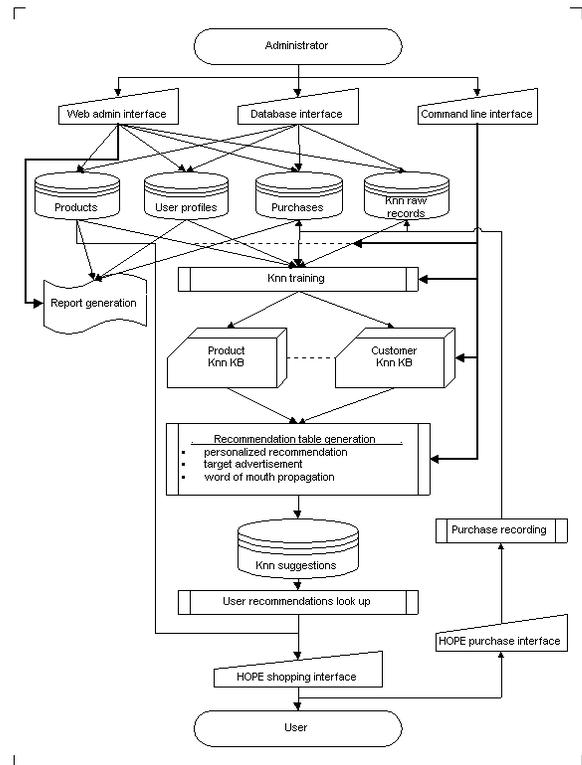[3] T. Sun, Hope: Helpful online purchase environment, Master's thesis, Acadia University, Canada, 2002.

Figure 11. Data and control flow