A recommender system for online shopping based on past customer behaviour

Katerina Pramatari
Athens University of Economics and Business

Olga Papaemmanouil
Brandeis University

e-Everything: e-Commerce, e-Government, e-Household, e-Democracy

14th Bled Electronic Commerce Conference

Bled, Slovenia, June 25 - 26, 2001

A Recommender System for Online Shopping Based on Past Customer Behaviour*

George Prassas, Katherine C. Pramataris, Olga Papaemmanouil, Georgios J. Doukidis

eLTRUN (Electronic Trading Research Unit)
Athens University of Economics and Business
47A Evelpidon & 33 Lefkados Str., 113 62, Athens, Greece
Prassas@aueb.gr, K.Pramatari@aueb.gr, Papaemma@aueb.gr, GJD@aueb.gr

Abstract

With current projections regarding the growth of Internet sales, online retailing raises many questions about how to market on the Net. While convenience impels consumers to purchase items on the web, quality remains a significant factor in deciding where to shop online. The competition is increasing and personalization is considered to be the competitive advantage that will determine the winners in the market of online shopping in the following years. Recommender systems are a means of personalizing a site and a solution to the customer's information overload problem. As such, many e-commerce sites already use them to facilitate the buying process. In this paper we present a recommender system for online shopping focusing on the specific characteristics and requirements of electronic retailing. We use a hybrid model supporting dynamic recommendations, which eliminates the problems the underlying techniques have when applied solely. At the end, we conclude with some ideas for further development and research in this area.

1. Introduction

As the World Wide Web becomes increasingly important as an information source and a place to conduct commerce, Web surfers face the daunting challenge on how to sift through a morass of information to get to the needed one. One solution to this

* This study was partly conducted within the context of the ACTIVE project (EP 27046), ESPRIT Programme (Framework IV), Commission of the European Union.

information overload problem is the use of recommender systems [20]. Recommender systems are changing from novelties used by a few e-commerce sites to serious business tools that are re-shaping the world of e-commerce. Many of the largest commerce Web sites are already using recommender systems to help their customers find products to purchase. The products can be recommended based on the top overall sellers on a site, based on the demographics of the customer, or based on an analysis of the past buying behaviour of the customer as a prediction for future buying behaviour [22]. Broadly, these techniques are part of personalization on a site, because they help the site adapt itself to each customer. Thus, recommender systems automate personalization on the Web, enabling individual personalization for each customer [20][23].

Personalization is one of the elements that may well improve the interaction between people and computers and offer possibilities for establishing long-term customer relations. Now more than ever, the promise of electronic commerce and online shopping will depend to a great extent upon the interface and how people interact with the computer [12]: the online shopping experience. This feature is especially important for sites that offer a large spectrum of products and services to their customers, as is the case for grocery retail sites.

In this paper, we study the application of recommendation systems for electronic retail sites, focusing on the peculiar characteristics and requirements of this environment. More specifically, we look into the various forms that online recommendations may take in the eyes of the online retail shopper, the services associated with them, and the technical ground upon which they are built. Based on this discussion, we present a hybrid model for dynamic recommendations in electronic retail sites and conclude by illustrating its application in a prototype development. The work presented in this paper has been the direct or indirect result of the ACTIVE (Advertising and Commerce Through Internet in the context of Virtual Enterprise) project (EP 27046), funded by the ESPRIT Programme (Framework IV) of the Commission of the European Union.

In the remainder of this document and more specifically in section two we describe the application of the proposed solutions in the case of the ACTIVE research project. In section three we review recommender systems and their underlying technologies. Section four discusses in more detail the implementation of the recommendation sub-modules. . Finally, section five concludes with some general thoughts and suggestions regarding further research and development in this area.

2. Online recommendations in the ACTIVE e-store

The ideas and solutions presented in this paper have been tested through the development of a prototype application as part of the ACTIVE (Advertising and Commerce Through Internet in the context of Virtual Enterprise) project (EP 27046), funded by the ESPRIT Programme (Framework IV) of the Commission of the European Union. The ACTIVE project introduces a global electronic commerce platform providing an intelligent interface upon which the involved players establish a tied and trusted relationship. In addition to the common shopping functionality, as offered by most online retail sites, the ACTIVE platform has certain advanced features, including online recommendations. More specifically, the ACTIVE platform consists of the following main components:

- *Home Shopping Tool*: The Home Shopping Tool implements the basic shopping functionality including: dynamic construction of product catalogues, uploading of product descriptions, shopping basket, search facility, payment services, etc. In addition to that, the tool supports more advanced personalised services, such as shopping lists and the quick shopping facility. The shopping lists functionality offers consumers a generic mechanism for customizing the structure of the online retail store, while the quick shopping facility restructures the product catalogues based on what the consumer has bought in the past.
- *Consumer Behaviour Tool*: The Consumer Behaviour Tool is the main component supporting personalisation. It captures consumers' behaviour information while they navigate through the ACTIVE store, monitors sales data, response to promotions, response to questionnaires, after-sales support requests etc. The information captured includes: consumer demographics, navigation data and sales data. The collected information is then processed by an analysis module and the results are stored in the Consumer Information Model, used further to create/update consumer profiles.
- *Advertising Tool*: The Advertising Tool offers to the e-tailer and the product suppliers the ability to advertise their products in the ACTIVE virtual shop. Advertisement takes place in the form of banners. The Advertising Tool supports the targeting of an advertisement campaign to a group of consumers based on their demographic characteristics and shopping behaviour. In addition, it supports the online booking of advertisement space, the uploading of advertising content and the measurement of advertisement effectiveness.
- *POS (Point Of Sales) Analyser*: The POS Analyser allows both retailer and product suppliers to access past sales data and analyse the effectiveness of specific brands, product segments, product categories, etc. in terms of sales, turnover or profit generated. In addition, the POS analyser combines sales data with consumer information, in order to analyse performance by target group, show the product preferences of specific consumer profiles and facilitate the definition of the target consumer groups.
- *On-line Sales Negotiator:* The consumer is able to make automated on-line negotiations with the sellers for the purchase of various goods and commodities. This functionality is implemented through the use of agent technology.
- *Shopping Recommender:* Last but not least, the Shopping Recommender tool assists the user during the shopping process, while at the same time promotes the products on behalf of the retailer and the suppliers. This component implements online recommendations based on ideas and solutions presented in

[19], tailoring them to the online customer's individual profile. The Shopping Recommender consists of the following sub-modules:

- *Shopping List Recommender:* Suggests products of regular use that may be of interest to the customer to add to his/her shopping list
- *Check-out Recommender:* Reminds the customer to buy products that he/she may have forgotten to add to his/her shopping cart at the "check-out"
- *E-store Navigation Recommender:* Guides the customer through the "electronic aisles" by recommending the next relative product category to visit
- *Product Assortment Recommender:* Sorts the lists of products according to each customer's preference

The ACTIVE's Shopping Recommender offers two alternatives on generating the recommendations, manual and automatic generation of recommendation rules.

The first alternative refers to the manual definition of targeted recommendations by the system administrator / virtual retailer. The retailer can provide targeted recommendations to his customers by specifying target groups based on demographics and navigational information and applying a set of rules (i.e. constraints) to each target group. These recommendations refer to discounts, special offers, contests, lotteries etc. An example of a recommendation rule would be: "if the customer is a woman, has children less than 3 years old, and buys milk of brand XYZ, provide her with a discount offer of 20% on the family size".

The second alternative refers to the automatic generation of recommendation rules. The consumer behaviour information is analysed so that two types of rules are generated: association rules, linking related product categories with each other, and collaborative filtering rules, linking related brands within different product categories. The rules are then stored and executed in the same way as for the manual system.

The architecture of the ACTIVE Shopping Recommender system is graphically depicted in Figure 1 below.

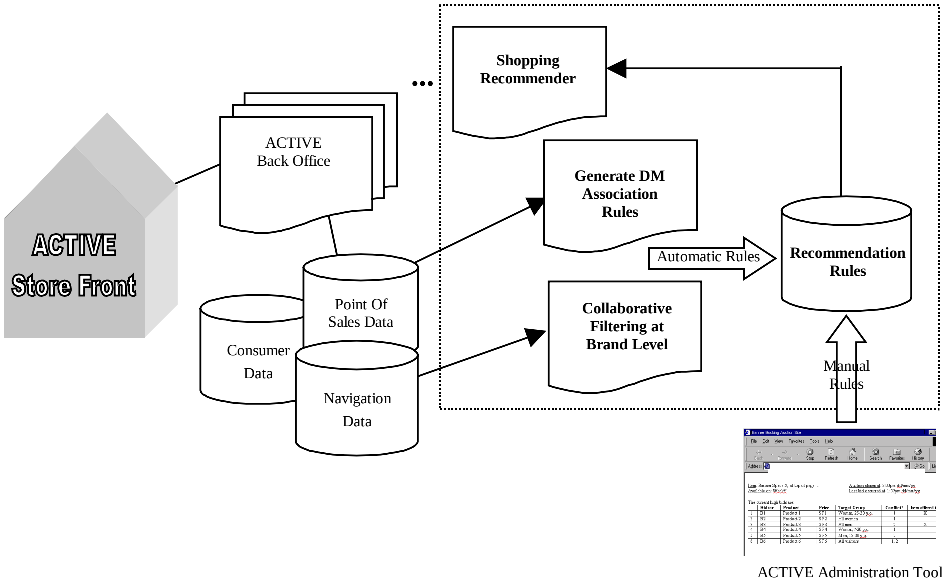

Figure 1: *The ACTIVE Shopping Recommender*

As Figure 1 depicts, the Shopping Recommender, which is a component of the ACTIVE platform's back office, is fed with recommendation rules that are the basis for the recommendation process. These recommendation rules are either generated automatically through association rules and collaborative filtering or manually through the ACTIVE administration tool by the administrator. In either way the administrator must set priorities on the recommendation rules, so that the Shopping Recommender can distinguish the most important rules and apply them first. These priorities have to be set manually regarding the manually generated rules. For the automated ones, the administrator can set 'meta-rules' that generate priorities between rules. Such meta-rules could be 'sort by ascending support', 'sort by ascending confidence' and so on. The administrator can further program more complex rules through the administration tool. In the following sections we will describe in more detail the underlying technologies of the recommendation system and relate it to a broader framework and taxonomy of recommender systems in e-commerce.

3. Recommender Systems in e-Commerce

Recommender systems are already used by many e-commerce sites under various formats, interfaces and underlying technologies. Many sites use them as a means to stimulate buying and perform a cross-sell, others as a way to offer personalized customer service and build loyalty. The techniques used to support them range from the manual definition of recommendations by a system administrator to the automatic generation of recommendation rules based on the user behaviour. The interface used for presenting the recommendations may be that of a simple e-mail message, a pop-up window, a special site section etc. Obviously, each of these alternatives has a different usability for the user, effectiveness for the site owner and complexity to implement.

Shafer et al. [22] introduce a taxonomy of recommender systems met currently in e-commerce sites based on two key dimensions: a) the degree of automation, which depends on the effort the customer has to put in order to get the recommendation, and b) the degree of persistence in recommendations, which depends on whether the recommendations are based on data regarding previous customer sessions with the system or not. The automation axis ranges from completely Automatic recommendations, where the customer has to spend no effort at all in order to get the recommendation, to completely Manual recommendations. The persistence axis ranges from completely Ephemeral recommendations, which are based on a single customer session, to Persistent recommendations, which are based on the site recognizing the customer and exploiting information previously gathered on him/her. This taxonomy is depicted in Figure 2 below.

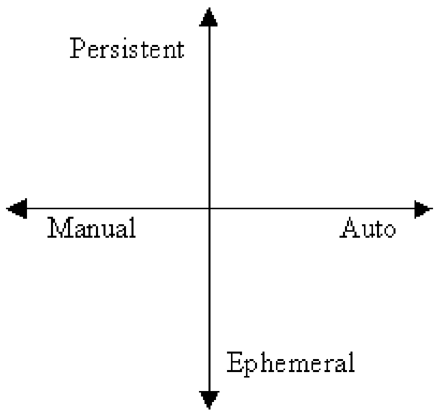

Figure 2: *A taxonomy of recommender systems in e-commerce [22]*

Taking the above taxonomy a step further and moving from a user-focused analysis to a more generic one, we introduce an additional axis that takes into account the implementation technology behind online recommendations. We believe that this axis plays an important role in automating the functionality of a site towards an increased degree of personalization offered to the user. On the other hand, such automation may require the application of intelligent techniques that increase the implementation complexity. We define this axis by identifying the following two sides of the spectrum:

a) on one side there is hardly any automation behind the creation of recommendations; these are created manually by a system administrator as entries in a database. For example, these can have the form of 'if-then' rules such as "If the customer is male, teenager and adds product A to the shopping cart, then recommend product B".

b) on the other side of the spectrum the recommendations are generated automatically by the system through the application of intelligent techniques, such as data mining, that extract information from the customer behaviour data. The recommendations may again have the form of 'if-then' rules that are stored in a database.

Thus, this third axis also ranges from Automatic to Manual, depending on the degree of automation in the generation of recommendations. Adding this third axis to the taxonomy of recommender systems in e-commerce presented above renders the graphical presentation of Figure 3.

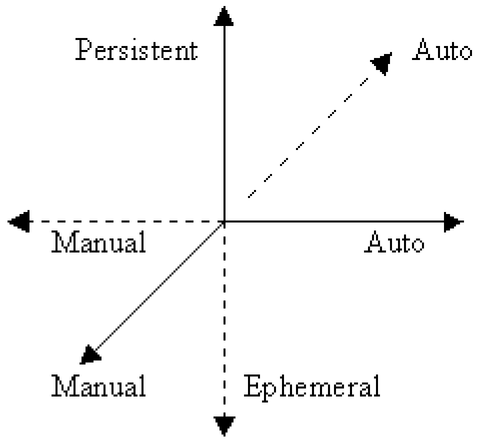

Figure 3: *A taxonomy of recommender systems in e-commerce*

In this paper we focus on the last aspect of recommender systems introduced above, i.e. the automatic generation of recommendations, as we believe that this is more closely linked to an e-commerce site's effectiveness, efficiency and degree of personalisation that can be achieved. Regarding the two other axes, we consider persistent and automatic recommendations, as being closer to this concept. In the following paragraphs we describe in more detail the techniques mainly used for the automatic generation of recommendations, that is data-mining and collaborative filtering. In section three we then look into how these are specifically applied in the case of recommender systems in electronic retail sites, given the specific characteristics and shopping behaviour met in these sites.

Relating the functionality of the ACTIVE platform to the taxonomy of recommender systems presented above, we can say that ACTIVE's Shopping Recommender is an automatic persistent system where the only manual effort from the customer is his/her shopping process, and the recommendations are always based on previous customer behaviour. Regarding the third dimension of the taxonomy, the automatic generation of recommendation rules, ACTIVE offers both alternatives sitting on each side of the spectrum, since it provides both manual and automatic generation of recommendation rules.

3.1 Data Mining

Data mining, which is also referred to as knowledge discovery in databases, has recently become an important area of research. It focuses on the extraction of implicit knowledge from large amounts of data. The most common data mining problems include finding clusters, associations and classifications [7]. Resnick and Varian [20] argue that there is a need for personalized recommendation systems and the data mining approach offers an opportunity for implementing such systems that are free of the need for explicit user input [23]. For example, the SurfLen system uses association rules to recommend web pages that possibly interest the Internet users based on their former navigation history [9].

The problem of discovering association rules, introduced in [5], is this area of data mining research that is more closely linked to our case. The original problem applies to market basket data and is the basis of market basket analysis. The goal is to discover buying patterns like "30% of transactions that contain item A also contain item B; 2% of all transactions contain both of these items". Association rules describe the association between the items in a specific transaction database.

More specifically, given a set of transactions, where each transaction is a set of items, an association rule is an expression of the form $X \Rightarrow Y$, where X and Y are sets of items. An example of association rule is: "c% of transactions that contain X also contain Y; s% of all transactions contain both of these items". Here c% is called the confidence of the rule and s% the support of the rule. The extraction of association rules with minimum support *minsup* and minimum confidence *minconf* is also referred to as the Boolean Association Rules Problem. Many efficient algorithms have been presented for the fast discovery of association rules [5][2]. These algorithms focus on the efficient retrieval of all the large item sets, which are all the item sets that have support larger than *minsup*.

Since the first formation of the problem, many other similar problems have appeared. Agrawal and Srikant introduced quantitative [4] and generalized [3] association rules, whilst Aggarwal et al. [1] introduced a specific case of quantitative rules, the profile association rules.

3.2 Collaborative Filtering

Techniques like collaborative filtering were used in the first recommendation systems [10][21], which filtered information based on the preferences of each user. The concept originated with the Information Tapestry project at Xerox [10]. Among its other features, Tapestry was the first system to support collaborative filtering in that it allowed its users to annotate the documents they read. All users could retrieve documents to read based not only on the content of the documents, but also on what other users had said about them. Collaborative filtering was defined as the situation of people collaborating to help one another perform filtering, by recording their reactions to documents they read. It is based on the premise that people looking for information should be able to make use of what others have already found and evaluated. Since then collaborative filtering has been used in many areas, such as newsgroups, user profiling and others. Similar systems have been implemented, such as Usenet Net News Collaborating Filtering System [14][13] and GroupLens [20][11].

The functionality of collaborative filtering systems is based on the collection of user ratings regarding the products of a particular market. As we have already seen, a 'product' can even be a document in an information database, used by a search engine. The ratings that are collected are used to identify the users that have similar preferences. Such users become members of the same group. Collaborating filtering recommends to people in the same group the products that their co-members prefer and they have not used/bought yet, based on the hypothesis that their common ratings imply similar preferences. The users of a collaborative filtering system collaborate in order to gain from each other's ratings.

The success of collaborative filtering systems depends on the collection of ratings from a large number of users. There are two ways to collect these ratings [15]: explicit rating and implicit rating.

Explicit Rating: Explicit rating requires that the evaluator examines the item and assigns it a value on the rating scale. Collaborative filtering systems that use explicit ratings require a large number of ratings to remain viable.

Implicit Rating: Implicit rating uses implicit sources like purchase data, navigating behaviour (bookmarks, time spent on a page etc) etc. Collaborative filtering systems that use implicit ratings remove the cost to the evaluator/user of examining and rating an item.

Collaborative filtering systems' main weakness was the need for explicit data [15], i.e. the requirement for explicit rating from a large number of users. At present the collaborative filtering technique is used by recommendation systems of well-known e-commerce sites [22] that use both implicit and explicit data.

4. Automatic Generation of Recommendations in Electronic Retailing

Issues of technology usage become critical as businesses and retailers attempt to exploit the boom in electronic commerce and marketing. There are large differences between a physical store and its electronic counter-part. A help button on the home page of the Web shopping site replaces the sales clerk’s friendly advice and service. The familiar layout of the physical store becomes a maze of pull-down menus, product indices, and search features. Now more than ever, the promise of electronic commerce and online shopping will depend to a great extent upon the interface and how people interact with the computer [12]: the online shopping experience.

Personalisation is one of the elements that may well improve the interaction between people and computers and offer possibilities for establishing long-term customer relations [18]. Recommender systems are a basic means of personalising a site, since recommendations are based on a customer's previous online behaviour and profile characteristics. In the following paragraphs we look more specifically into the problem of offering dynamic recommendations in the online retail environment, what form these can take and how we can exploit data mining and collaborative filtering techniques in order to support them.

Fu et al. [9] have made the first effort of using data mining techniques and collaborative filtering to generate recommendations of web pages. Their system mines the browsing history of the users from two perspectives. Firstly, each user u is represented as a set of URLs $\{p_1, p_2, \ldots, p_m\}$, which can lead to association rules between pages. Secondly, each page is represented as a set of users $\{u_1, u_2, \ldots, u_m\}$, which can lead to association rules between users. A rule like $\{p_1, p_2, \ldots, p_m\} \Rightarrow p$, implies the recommendation of page p to the users that have read $\{p_1, p_2, \ldots, p_m\}$ but not p. Additionally, large 2-item user sets like $\{u_1, u_2\}$, can be used for recommending pages that one user has read and the other has not, and vice versa. The last idea is a form of collaborative filtering, since the user item sets imply users with the same preferences. Then, this knowledge is used to recommend to each other pages that the one has read but the other hasn't. In this particular case, the large space of the web poses a significant problem, the weakness of generating association rules due to sparse datasets.

In this paper we tackle a similar problem to the one of web pages recommendation, the recommendation of products in an electronic retail environment. The similarity is obvious; in both cases there are users (customers in retail) and items (web pages/products). In addition, the space of items is also large in a retail store, especially in grocery retailing, so the problem of generating rules remains, while the phenomenon of revisiting a web page is associated with the need in retailing to purchase the same product on a regular basis. The main difference between the two problems, however, is the existence of a multi-level hierarchy of product categories, that is used as a product taxonomy. This fact is paired with the different user behaviour and functionality found in electronic retail sites, which asks for our contribution, as we will describe in more detail below.

We start by describing the various forms that online recommendations can take in an electronic retail store, viewed through the perspective of the ACTIVE platform.

4.1 Forms of Online Recommendations in ACTIVE's Shopping Recommender

Shopping List Recommender

Many electronic retail stores use shopping lists to facilitate the shopping process for their customers [17]. The success of shopping lists is based on the fact that the consumers' shopping basket doesn’t differ a lot from one purchasing trip to the next, especially as far as grocery shopping is concerned. There are some products that the consumers buy on a regular basis and there is no need to search them in the store each time. The online shopping list metaphor "saves" these products, simulating the handwritten shopping list everybody makes before visiting a traditional store. In the majority of electronic retail stores customers can create and save their own shopping lists. Then, each time they visit the store they can directly move the products from the shopping list(s) to the shopping cart metaphor and then browse the store and add any additional product to the cart. The shopping lists thus become a very useful source of knowledge for the customer's purchasing habits.

An attractive service to the customers, that can increase sales, would be the recommendation of products that are associated with those that are included in the shopping list. The recommended products are likely to interest the customer, due to the strong correlation they have to the products the customer buys on a regular basis. This service usually appears as a suggestion for specific product items, categories or segments that might interest the customer ("Are you interested in…?").

Check-out Recommender

The “check-out” metaphor is this step in the online buying process where the customer reviews the contents of the basket and fills-in the details for the payment and delivery of the products. In most of the e-stores the customer has to confirm the contents of the cart at this stage before proceeding.

Just like the shopping list, described in the previous paragraph, the cart is a set of products. In proportion to the previous service, another service would be to recommend products at the checkout, when the customer is ready to pay. The system must check the products that are already chosen and recommend associated products –perhaps complementary products-, that are likely to interest the customer. In this case, as above, the recommender system can use a combination of knowledge from the analysis of both the transactions and the consumer profiles.

The basic difference between this and the previous service is that the processing must take place online, since the cart is created dynamically, whilst the shopping list is stored in the database and can be processed off-line. This service usually appears as a reminder for products that the customer might have forgotten to buy ("Don't forget to buy…").

E-store Navigation Recommender

Despite the frequent use of shopping lists, especially in online grocery retailing, there is always a number of customers that prefer to do their shopping by navigating the “aisles” of the virtual store. This is also the case for all the new customers visiting the store for the first time and existing customers that want additional products besides the ones stored in their shopping lists. During this process, the customer navigates through the hierarchy of product categories and subcategories in the store until he/she reaches the lower level in this hierarchy, i.e. the list of specific product items. The recommender system could monitor the browsing through the categories and the products that the customer adds to the cart and make recommendations regarding related categories or products for the customer to visit. This service facilitates browsing and could be characterized as "aisle guidance".

Product Assortment Recommender

Another type of recommendation service refers to the order in which the products appear in a list. The recommendation system can change the order of the products in the list of categories / subcategories / product items etc. so that it better matches the profile of each individual customer. For example, if a customer usually buys "low fat products", then it would be positive, when he/she enters the section "milks" to show at the top of the list the "low fat milks". This service will probably reduce the search time for the customer, and consequently increase his/her satisfaction. In addition, the possibility for the customer to add products to the cart each time he/she visits a new category is increased, as the most attractive products for him/her are shown at the top of the list. This service might not be a "visible" value-adding service to the customer, but it can have significant indirect benefits for the e-store.

From the description of the recommendation services above a crucial question raises: whether to recommend specific product items or product categories/subcategories. From a marketing perspective, recommendation of products at the brand-name level could be considered a type of advertisement or effort to promote specific brands by the customers. As such, it is possible to create a negative impression regarding the usefulness of the service. On the contrary, recommendations of product categories/subcategories are more likely to be considered as a value-adding service and, thus, be better accepted. However, the objective of this paper is not to justify which of the two cases would be more effective for the e-store and value-adding for the customer, but to analyse the implementation problems of each alternative, as discussed below.

4.2. Generating Online Recommendations Based on Data Mining

The solution for the automatic generation of the recommendations for the retail store is similar to the one given in [9] for the recommendation of web pages. We use as input the online purchasing data and extract association rules of the form $\{i_1, i_2, \ldots, i_m\} \Rightarrow i_k$, where the item i_n may stand for specific product items, brand-names, categories or other segmentation level in the products hierarchy. An example of a ‘route’ in the products hierarchy is 'Dairy'⇒'Milks'⇒'Low Fat Milks'⇒'Low Fat Milk A'. The lower level in the hierarchy lists all the product items, characterised by a specific brand name, size, colour etc.

Adapting the solution of Fu to the retail case means that whenever the itemset $\{i_1, i_2, \ldots, i_m\}$ appears, the system can recommend the item i_k. Regarding the fact that the shopping list, the cart and every product list are itemsets of the form $\{\{i_1, i_2, \ldots\}$, the itemset $\{i_1, i_2, \ldots, i_m\}$ may be either a sub-itemset of a shopping list or of the current cart content or of the final product list at the checkout. Proportionally, the recommendation made to the customer may be to:

- ✓ add product i_k to the shopping list, if the customer browses the shopping list and i_k is a specific product
- ✓ navigate through products of category i_k and add one to the shopping list, if the customer browses the shopping list and i_k is a product category
- ✓ add product i_k to the cart, if the customer browses the store (and maybe has just added a product to the cart) and i_k is a specific product
- ✓ visit the relative category i_k, if the customer browses the store (and maybe has just added a product to the cart) and i_k is a product category
- ✓ add product i_k to the final list before continuing with payment, if the customer is at checkout and i_k is a specific product
- ✓ visit category i_k in order to add a product of that category to the cart before continuing with payment, if the customer is at checkout and i_k is a product category

The basic problem of this approach is identical to the problem of Fu's approach [9]. Sparse datasets produce a small number of association rules, in regard to the total number of discrete products. Regardless of the absolute number of the association rules, the quality of the majority of the rules will be poor. This means that most of the rules will have small support and confidence and are not to be trusted. In addition, many products will be absent from the rules, because they don't appear enough times to extract associations for them (non-frequent products). Despite the fact that purchasing data are usually vast, the combinations are so many that it is possible the purchasing histories of customers to intersect infrequently. By extracting associations at a higher level of the products hierarchy (e.g. using product categories instead of brand names), the mining process is more successful [6].

4.3 The hybrid approach of ACTIVE's Shopping Recommender to Dynamic Recommendations

In order to overcome the above inefficiency, in case we want to recommend specific product items to a customer, we propose a hybrid recommendation model that combines both data mining and collaborative filtering techniques. The proposed model may generate recommendations at any level of the products hierarchy, even at the lower level of product items, as required by the service of product assortment recommendations described above. Thus, contrary to the weakness of the data mining solution, our model's main feature is that it can always propose a recommendation.

The model works in two phases. During the first phase, the product category is selected using the data mining approach discussed above. More specifically, from all the available rules we isolate those rules that can be applied to our recommendation case. These rules are the ones that their left itemset is a sub-itemset of the customer's shopping list/cart/final list according to the recommendation service. From these rules, one outclasses depending on the constraints set by the system administrator (for example "the stronger rule is the one with the highest confidence, support", "Don't recommend products that already exist in the shopping list/cart/final list"). The rule that 'survives' provides the product category to be recommended. Let this category be C_{rec}. Then, we find all the categories that are associated with C_{rec}. These categories can be sought among all the large 2-itemsets extracted from the database or from the rules with one item on the left side of the rule and C_{rec} on the right one. According to the measure that the administrator sets (support, confidence), we find the category with the strongest association. We note that this category must be a category that the customer has already purchased from. Let this category be C_{assoc}.

Until now we have used only data mining to extract the associations between the categories. During the second phase we use collaborative filtering in order to identify the specific product(s) that we should propose. Collaborative filtering is based on the common behaviour between the specific customer and other customers, in order to propose the products that the customer hasn't bought yet but are preferred by the customers belonging to the same group. In this case we use implicit rating for extracting customer behaviour and preferences based on purchasing data. More specifically, we first seek for the product(s) p_{assoc} that the customer has already chosen from category C_{assoc} and isolate all the customers that have bought p_{assoc} in the past. Constraints such as buying frequency and total number of times the product has been bought in the past can be applied to this selection of peer customers. Then, we find which products of the category C_{rec} these customers prefer to buy. The most frequently purchased products are those that the system will finally recommend.

The main assumption in the above process is that the products the customer would be interested in are the ones that customers with similar purchasing habits prefer. For example, if a customer usually buys expensive products, then the customers that are part of the same group also buy expensive products. So, if the target customer has bought an expensive product from category C_{assoc}, our system will search for the products that customers with the same profile usually buy in category C_{rec}, which are probably the more expensive ones.

The following table summarizes the recommendation services described above, the product hierarchy level at which these are applied and the underlying implementation technologies.

WHERE	WHAT	HOW	SERVICE
Shopping list	Category	*Association rules*	"Are you interested in…?"
	Product	Association rules/Hybrid	*"Are you interested in…?"*
At checkout	Category	Association	*"Don't forget to buy…"*
	Product	Association rules/Hybrid	*"Don't forget to buy…"*
Navigation	Category	Association	Aisle guidance
	Product	Association rules/Hybrid	*"Are you interested in…?"*
Assortment	Category	Association	Not visible
	Product	Association rules/Hybrid	Not visible

Table 1: *Recommendation services and underlying technologies in the ACTIVE platform*

5. Conclusions and Further Research

Shafer et al. [22] argue that recommendation systems are one solution to the need for customisation of companies to serve multiple needs, as Pine states in his book Mass Customization [16]. In this paper we focused on the area of Internet retailing with our main perspective being the automatic generation of recommendations. Our system takes advantage of the product category hierarchies and combines both data mining techniques and collaborative filtering in order to solve problems that these techniques have when applied separately [15][9].

A key issue that emerges from this study, asking for future research, refers to the effectiveness of the recommendations generated by such a system and how this can be improved using the customer reactions to the recommendations as feedback. In this way the system will not only use purchasing data as input, but also the customer reactions to the recommendations, which is the most basic measure of its effectiveness. Another pending research issue is the efficient implementation of recommendations, which in most cases must be made online. The application of a rule on an itemset is a high order complexity task and the existence of a vast amount of rules may cause problem to their online application, where the speed of response to the user is critical.

Recommendation systems have become serious business tools and are re-shaping the world of e-commerce. Effective recommendations are a valuable service to the customers and a profitable service to the retailer. We hope that the work presented in this paper can be used to extend the knowledge in this area and open up new perspectives for future research.

References

[1] Aggarwal, C.C.; Sun, Z.; Yu, P.S. Online Algorithms for Finding Profile Association Rules. Proceedings of the International Conference on

Knowledge Discovery and Data Mining, KDD-98, pp122-133, AAAI Press, 1998

[2] Agrawal, R. and Srikant, R. Fast Algorithms for Mining Association Rules. Proceedings of the 20 the VLDB Conference, Santiago, Chile, 1994

[3] Agrawal, R. and Srikant, R. Mining Generalized Association Rules. Proceedings of the 21st VLDB Conference. 1995

[4] Agrawal, R. and Srikant, R. Mining Quantitative Association Rules in Large Relational Tables. Proceedings of the ACM SIGMOD Conference. 1996

[5] Agrawal, R.; Imielinski, T.; Swami, A. Mining association rules between sets of items in large databases. In SIGMOD'93, pages 207-216, Washington D.C., USA, May 1993

[6] Cabena, P.; Choi, H.H.; Kim, I.S.; Otsuka, S.; Reinschmidt, J.; Saarenvirta, G. Intelligent Miner for Data Application's Guide. IBM Corporation. International Technical Support Organization. http://www.redbooks.ibm.com, 1999

[7] Chen, M.S.; Han, J.; Yu, P.S. Data Mining: An Overview from a Database Perspective. IEEE Trans. on Knowledge and Data Engineering, Vol. 8, No.6, 1996

[8] Ernst & Young. Global Online Retailing. An Ernst & Young Special Report. Ernst & Young LLP, 2000

[9] Fu, X.; Budzik, J.; Hammond, K.J. Mining Navigation History for Recommendation. Proceedings of the 2000 international conference on Intelligent user interfaces , 2000, Pages 106 – 112.

[10] Goldberg, D.; Nichols, D.; Oki, B.M.; Terry, D. Using collaborative filtering to weave an information tapestry. Communications of the ACM, 35(12): 61-70, December 1992

[11] Konstan, J.A.; Selman, B.; Miller, N.; Maltz, D.; Herlocker, H.L.; Gordon, L.R.; Riedl, J.. GroupLens: Applying collaborative filtering to Usenet news. Communications of the ACM, 40(3): 77-87, March 1997

[12] Lohse, P.L and Spiller, P. Electronic Shopping. Communications of the ACM. Vol. 41(7). 1998

[13] Maltz, D. and Ehrlich, K. Pointing The Way: Active Collaborative Filtering. In Proceedings CHI'95, 202-209. 1995

[14] Maltz, D.A. Distributing Information for Collaborative Filtering on Usenet Net News. MIT Department of EECS MS Thesis. May 1994

[15] Nichols, D.M. Implicit Rating and Filtering. In Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering, 10-12.Budapaest, Hungary, ERCIM, 1997

[16] Pine, B.J. (1993). Mass Customisation. Harvard Business School Press. Boston, Massachusetts

[17] Pramataris, K.C.; Vrechopoulos, A.P.; Doukidis, G.I. The transformation of the promotion mix in the virtual retail environment: an initial framework and comparative study. Journal of New Product Development and Innovation Management, Vol. 2, No. 1, June 2000.

[18] Pramataris, K.C.; Vrechopoulos, A.P.; Mylonopoulos, N.; Papamichail, G.; Poylimenakou, A. Personalized Services and Promotions in Internet Retailing. In Proceedings of eBusiness and eWork 2000 Conference, Madrid, Spain, Oct 2000.

[19] Prassas, G.; Pramataris, K.; Papaemmanouil, O.; Doukidis, G.I. Dynamic Recommendations in Intrnet Retailing, ECIS 2001, Bled, Slovenia, 27-28 June 2001

[20] Resnick, P. and Varian, H.R. Recommender Systems. Communications of the ACM, Vol. 40, No., Pages 56 – 58. 1997

[21] Resnick, P.; Iacovou, N.; Suchak, M.; Bergstrom, P.; Riedl, J. Grouplens: An open architecture for collaborative filtering of netnews. In Proceedings of ACM CSCW'94 Conference on Computer Supported Cooperative Work, pages 175-186, 1994

[22] Shafer, J.B.; Konstan, J.A.; Riedl, J. Recommender Systems in E-Commerce. In ACM Conference on Electronic Commerce (EC-99), pages 158-166. 1999

[23] Yu, P.S. (1999). Data Mining and Personalization Technologies. DASFAA 1999, 6-13.