

Personalized E-commerce Recommendations

Penelope Markellou^{1,2}, Ioanna Mousourouli², Spiros Sirmakessis^{1,2,3}, Athanasios Tsakalidis^{1,2}

¹ *Research Academic Computer Technology Institute, Patras, Greece*

² *University of Patras, Department of Computer Engineering & Informatics, Patras, Greece*

³ *Technological Education Institute of Messolongi, Department of Applied Informatics in Administration & Economics, Messolongi, Greece*

markel@cti.gr, mousouro@ceid.upatras.gr, syrma@cti.gr, tsak@cti.gr

Abstract

Recommendation systems are special personalization tools that help users to find interesting information and services in complex online shops. Even though today's e-commerce environments have drastically evolved and now incorporate techniques from other domains and application areas such as web mining, semantics, artificial intelligence, user modeling and profiling, etc. setting up a successful recommendation system is not a trivial or straightforward task. This paper argues that by monitoring, analyzing and understanding the behavior of customers, their demographics, opinions, preferences and history, as well as taking into consideration the specific e-shop ontology and by applying web mining techniques, the effectiveness of produced recommendations can be significantly improved. In this way, the e-shop may upgrade users' interaction, increase its usability, convert users to buyers, retain current customers and establish long-term and loyal one-to-one relationships.

1. Introduction

Nowadays, e-commerce features as an extremely dynamic economic sector and at the same time, as one of the most appealing ways of beginning or expanding a business activity. Statistics for its future are quite optimistic indicating that the business-to-consumer (B2C) e-commerce will grow rapidly in the years to come [5]. In fact, despite the recent talking about businesses cases that went online but did not manage to profit, looking at the big picture reveals that such cases comprise just a small portion of the market and most inhabitants of the e-commerce world are not only surviving, but also doing very good business. This is a strong indication for assuming that there are certain distinctive features that make the difference and, which

result in either success or failure. Moreover, surveys and studies have shown that consumers are increasingly giving online commerce a chance. Taking also into account the rapidly growing number of those that have access to the Internet and those that gain familiarity with the medium, we have a quite promising picture.

In this paper we argue that one of the factors that may block or urge online sales concerns the *individuality* of Internet shoppers. There is no dispute about whether or not e-commerce has a future; it is a new philosophy in conducting business and it is here to stay. The question is how easily Internet users become e-consumers and which are the internal “mechanisms” and external factors that participate in an e-purchase. The problem arises from the fact that shoppers with varying needs, preferences and background navigate through large and complex web structures and are confronted with too many options, missing in many cases the goal of their inquiry. Generally, search engines are used for filtering pages according to explicit users' queries. However, their results are often poor since the produced lists are long, unmanageable and contain irrelevant pages [17].

Currently, *web personalization* is one of the most promising approaches to alleviate this overload problem and to provide users with tailored experiences. According to [9] “*personalization is defined as any action that adapts the information or services provided by a web site to the knowledge gained from the users' navigational behavior and individual interests, in combination with the content and the structure of the site*”. In this direction, the recent web technological advances help online companies to acquire individual customer's information in real-time and with low cost. Based on this information, they construct detailed profiles and provide personalized services. Thus, e-shops have now

the opportunity to improve their performance by addressing the individual's needs and preferences, increasing satisfaction, promoting loyalty, and establishing one-to-one relationships.

Recommendations systems (RSs) that comprise the most popular forms of personalization are becoming significant business tools. They have emerged in the middle of 1990's and from novelties used by a few web sites have changed to important tools incorporated to many e-commerce applications (e.g. Amazon.com, eBay.com, CDNow.com). Specifically, these systems take advantage of users' and/or communities' opinions in order to support individuals to identify the information or products most likely to be interesting to them or relevant to their needs and preferences. The recommendations may have various forms e.g. personalized offers/prices/products/services, inserting or removing paragraphs/sections/units, sorting/hiding/adding/removing/highlighting links, explanations or detailed information, etc.

The initial efforts were limited to *check-box personalization*, where portals allowed the users to select the links they would like on their "personal" pages. This has proved deficient since it depends on the users knowing in advance the content of their interest. Moving towards more intelligent approaches, *collaborative filtering (CF)* was deployed for implementing personalization based on knowledge about likes/dislikes of past users that are considered "similar" to the current one (using a certain similarity measure). These techniques required users to input personal information about their interests, needs and/or preferences but this posed in many cases a big obstacle, since web users are not usually cooperative in revealing this type of data. Due to such problems, researchers resorted to *observational personalization*, which is based on the assumption that we can find "clues" about how to personalize information, services or products in records of users' previous navigational behavior [20].

This is the point where *web mining* comes into play. Web mining is defined as the use of data mining techniques for discovering and extracting information from web documents and services and is distinguished as *web content*, *structure* or *usage mining* depending on which part of the web is mined [13]. In the majority of cases, web applications base personalization on web usage mining, which undertakes the task of gathering and extracting all data required for constructing and maintaining user profiles based on the behavior of each user as recorded in server logs [15].

A relatively recent development that is foreseen to greatly affect personalization (and more specifically the web mining subtasks) is the creation of the

semantic web [4]. *Semantic web mining* [3] combines the two fast-developing research areas of semantic web and web mining with the purpose of improving web mining by exploiting the new semantic structures. The web will reach its full potential when it becomes an environment where data can be shared and processed by automated tools, as well as by people. The notion of being able to semantically link various resources (e.g. documents, images, people, concepts, etc.) is essential for the personalization domain. With this we can begin to move from the current web of simple hyperlinks to a more expressive, semantically rich web, where we can incrementally add meaning and express a new set of relationships (e.g. *isAuthorOf*, *dependsOn*, *worksFor*, *hasSubjectOf*, *hasLocation*, etc.) among resources, and making explicit the particular contextual relationships that are implicit in the current web. The semantic web will allow the application of sophisticated mining techniques, which require more structured input [16]. This will open new doors for effective information integration, management and automated e-services. Especially e-shops will become smarter and more comprehensive, enhance searching and information retrieval and provide intelligent recommendations [8].

This paper introduces an approach for generating personalized e-commerce recommendations. The knowledge about customers and products is extracted from usage mining and ontological data in conjunction with customer-product ratings and matching techniques between similar customers. This integration provides additional knowledge about customers' preferences and allows the production of successful recommendations. Even in the case of "*cold-start problem*" where no initial behavioural information is available, it can offer logical and relevant recommendations to the customers. The provided recommendations are expected to have higher accuracy in matching customers' requirements and thus higher acceptance by them.

The remainder of this paper is structured as follows. Section 2 reviews related work. Section 3 describes the e-shop ontological schema. Section 4 and 5 present data acquisition and analysis tasks correspondingly. Section 6 introduces the approach for finding recommendations based on web usage mining and ontological data. Section 7 summarizes the most interesting conclusions of the paper and presents open issues and thoughts for future research.

2. Related work

The need to provide personalized services over the web has been a strong driving force behind numerous

research efforts that combine algorithms and techniques from web usage mining, semantic web and ontologies. The variety of available applications is a solid indication of the area maturity.

Berendt in [2] proposes concept hierarchies for analyzing complex web usage data of visitors. These hierarchies are used as a basic method for aggregating web pages. Moreover, interval-based coarsening and its inverse zooming are described as a technique to mine web usage at different levels of abstraction. Basic and coarsened stratograms have been exploited to visualize web usage at different degrees of detail. Using a case study of online shopping with an anthropomorphic agent has been demonstrated that this kind of abstraction offers new possibilities of understanding complex paths through a semi-structured, interaction-rich environment.

Eirinaki et al., in [10] propose a recommendation method that integrates content semantics and navigational data. Their technique uses the terms of a domain-ontology in order to uniformly characterize both the content and the users' navigational patterns, and thus produce a set of recommendations that are semantically related to the user's current visit. The work deals with the important problem of multilingualism, which arises when the content of a web site appears in more than one language and is integrated in the personalization system SEWeP [11].

Middleton et al., in [18] describe an ontological approach to user profiling within recommender systems, working on the problem of recommending on-line academic research papers. The papers are classified using ontological classes. CF and content-based recommendation algorithms used to recommend papers seen by similar people on their current topics of interest. They use the ontological relationships between topics of interest to infer other topics not yet browsed and recommend them to the users.

Mobasher et al., in [19] introduce an approach for semantically enhanced CF. In this work, structured semantic knowledge about items, extracted automatically from the web based on domain-specific reference ontologies, is used in conjunction with user-item ratings (or weights) to create a similarity measure for item comparisons. Provided experimental results on two different data sets demonstrate that their approach yields significant advantages both in terms of improving accuracy, as well as in dealing with very sparse data sets or new items.

Choa and Kimb in [7] propose a recommendation methodology based on web usage mining, and product taxonomy to enhance the recommendation quality and the system performance of current CF-based systems. Web usage mining populates the rating database by

tracking customers' shopping behaviors on the web, thereby leading to better quality recommendations. The product taxonomy is used to improve the performance of searching for nearest neighbors through dimensionality reduction of the rating database.

The basic contribution of our approach is that combines usage and ontological data in the off-line phase in order to generate the user navigational model. It distinguishes the cases of new and old user and further incorporates ratings and similarities with other users in the on-line phase to sort the recommendations. By combining different available information channels we succeed to better meet users' needs and preferences.

3. E-shop ontological schema

In many online shops a basic classification schema for the provided items is available. This schema can be depicted as a tree, where the root represents the most general class, the internal nodes all intermediate classes and the leaves the specific items. Its role is important especially in the knowledge discovery process, since it represents e-shop's dependent knowledge and may affect the results. In this framework, the proposed recommendation approach is based on an ontological schema and has been incorporated in a pilot e-shop that rents movies to customers. Specifically, the web site consists of a collection of pages containing information about movies attributes such as title, category, studio, actor, director, producer, awards, year, duration, audience, story, etc. The initial data has been retrieved from the Internet Movie Database (<http://www.imdb.com>), organised in the ontological schema and enhanced in order to comprise a rich superset covering all needed concepts and relations.

The underneath ontological schema formulates a representation of the e-shop domain by specifying all of its concepts, the relations between them and other properties, conditions and "regulations". It allows semantic annotation and has the ability to perform semantic querying and ontology-based browsing. The ontology "building" was a complex and time-consuming task and was based on our intuition in order to depict all e-shop notions, organize their taxonomic hierarchies and represent their relationships. The development was akin to the definition of a set of data and their structure. In this way, the ontology can be considered as a knowledge base that is used further for extracting useful knowledge. Specifically, its role is to be used as an input for the mining phase in order to extract, combine and transform the existing implicit

and explicit knowledge (user class, history profile, e-shop content and structure) into new forms. The output of this task is a list of possible recommendations.

A part of the e-shop ontology is depicted in Figure 1. In particular, the ontology creates connections between movies according to different attributes that characterize them. Using now the specific customer's history file, his preferences can be figured out. For example, it can be founded if the customer likes or dislikes watching movies:

- from a certain category e.g. "action", "drama", "comedy", "westerns", "musicals", etc.

- by a certain director e.g. "Steven Spielberg", "Francis Ford Copolla", "Sydney Pollack", etc.
- with a certain actor e.g. "Sean Penn", "Nicole Kidman", "Brad Pitt", "Julia Roberts", etc.
- or other attributes' combinations.

When one or more matching criteria are met, then other movies can be discovered according to the ontological schema that have similar attributes with those that the customer has already rented. In the case of a new customer (history file= \emptyset), the information from his registration form is analyzed and the recommendations are based on the ontology.

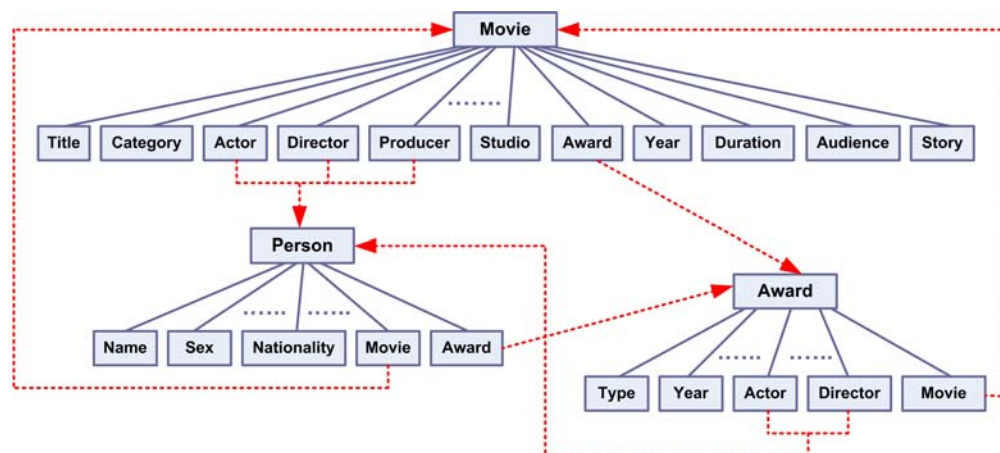


Figure 1. Part of the e-shop ontology (extended version of [19]).

4. Data acquisition

The objective of data acquisition is the identification of available customer's information. This task is in continuous execution and leads (after cleaning and appropriate transformation) to the construction of customers' representation models that will allow for further processing and updating. Specifically, it encompasses data about:

- the *user* (demographics, user knowledge, skills, capabilities, interests, preferences, goals, etc.),
- the *usage* (selective actions, ratings, usage frequency, action correlations or sequences, etc.),
- and the *usage environment*, hardware, software and physical (browser version, available plug-ins, bandwidth, display/input devices).

There are two general approaches for acquiring user data: either the user is asked *explicitly* to provide them (using questionnaires, fill-in preference forms, etc.), or the system *implicitly* derives such information without initiating any interaction with him (using acquisition rules, plan recognition, and stereotype reasoning).

For our approach we use the model introduced in [14], which helps us to better understand the interactions between consumers and e-businesses. This model represents the *overall consumer purchase decision cycle* and investigates the issues that affect web users, from selecting a specific e-shop to the delivery of the product and the overall assessment of the shopping experience. This process has been divided into three successive stages: outside the e-shop, inside the e-shop and after sales, with each stage analyzed on the basis of customer states and transition conditions. Special focus is set on identifying the data that the web site gathers from customers' interactions and will be used for mining purposes.

5. Data analysis

The step following after the data acquisition and the building of user profiles is data analysis. Initially, some preparation activities take place in order to clean the data and facilitate their manipulation. For instance, entries that do not reveal actual usage information are removed and missing data are completed. Then follows

the application of statistical and data mining techniques in order to detect interesting patterns in the pre-processed data. The most well-known techniques used for data analysis include clustering [12], classification [6] and association rules mining [1].

Our recommendation approach is depicted in Figure 2. As we have mentioned before, the pilot e-shop rents movies to the customers and based on web usage mining techniques and semantic metadata, as well as customer-product ratings and matching techniques between customers produces recommendations. The operation of the e-shop is straightforward. In the case of a new customer the e-shop motivates him to become a member by filling in the registration form. In this

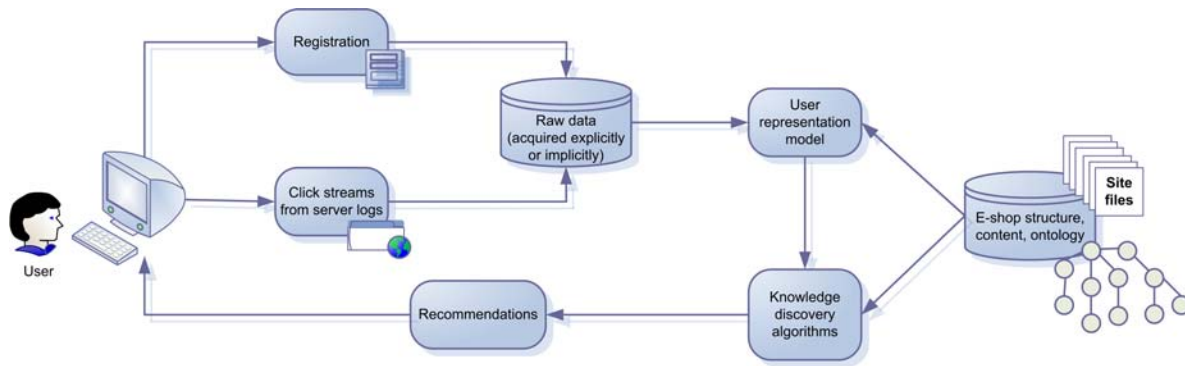


Figure 2. Proposed recommendation approach.

5.1. Customers classification

The main objective of classification step is to assign customers to a set of predefined classes. These classes represent different user profiles and classification is performed using selected features with high discriminative ability as refers to the set of classes describing each profile. The features that determine a specific class can be tuned and typically include userID, age, sex, nationality, occupation, as well as education and preferences. For example, from the class “rent comedies”, customers that like to rent comedies and customers that don’t can be extracted. Table 1 depicts the profile of an active buyer. Consider the attributes and the values presented in Table 2 and suppose we want to decide if the next unknown user X rents or not comedies: $X = \{userID=101, age=22, sex=w, occupation=student, education=medium\}$.

Classification requires a training data set with pre-assigned class labels, since it is categorized as a supervised machine learning technique. Then the classifier (by observing the class assignment in the training set) learns to assign new data items in one of the classes. It is often the case that clustering is applied before classification to determine the set of classes.

Table 1. Part of a user profile.

userID	“67”
age	“30”
sex	“male”
occupation	“seller”
education	“higher”

We have based our categorization model on a naïve Bayes classifier, where each snapshot X consists of a set of attributes x_1, x_2, \dots, x_n . We have defined m classes C_1, C_2, \dots, C_m . Given an unknown snapshot X for which we don’t know its class, the classifier predicts that X belongs to the class with the higher probability. The classifier assign X in class C_i if $P(C_i|X) > P(C_j|X)$ for $1 \leq j \leq m, j \neq i$. According to Bayes theorem $P(C_i|X) = P(X|C_i) * P(C_i) / P(X)$. $P(X)$ is constant for all snapshots, so we need to maximize the expression $P(X|C_i) * P(C_i)$. For categorizing X, we compute the probability $P(X|C_i) * P(C_i)$ for each class C_i . X will be assigned in class C_i if $P(X|C_i) * P(C_i) > P(X|C_j) * P(C_j)$ for $1 \leq j \leq m, j \neq i$. This means that X is assigned in the class with the maximum probability.

This classification model may be revised as time passes and updated based on collected transactions data. User assignment to classes might also be used for

provoking interactions among users and enhancing this way collaboration and communication, as well as for allowing e-shop to perceive a useful insight of the “virtual communities”. For each class a set of recommendations is attached. These recommendations are lists with e.g. the best and worst movies that other users “close” to the current have contributed.

Table 2. A snapshot of the database.

user ID	age	sex	occupation	education	class: rent comedies
1	30	w	engineer	high	yes
2	25	m	student	medium	no
3	34	m	salesman	low	yes
4	18	m	student	low	yes
5	45	w	director	high	no
...
100	28	w	artist	high	yes

5.2. Association rules mining on click-streams

Association’ rules connect one or more events. Their aim is to discover associations and correlations between different types of information without obvious semantic dependence. Moreover, correlations between pages not directly connected and previously unknown associations between groups of users with specific interests may be revealed.

The model tracks all user actions as successive page requests recorded in server logs. Log files are then cleaned from all redundant information (such as secondary, automatically generated requests for page images, etc.). Combining the remaining requests with information about the way web site content is structured, the system distills user accesses to movies pages. The set of movies that have been accessed by a certain user during all past visits to the e-shop are stored in the user profile, and this is where the generator seeks for discovering association rules.

An association rule example is $\{movie_i, movie_j\} \rightarrow \{movie_x\}$, with support=0.02 and confidence=0.68. This rule conveys the relationship that users who accessed $movie_i$ and $movie_j$ also tend (with a confidence of 68%) to be interested in $movie_x$. Support represents the fact that the set $\{movie_i, movie_j, movie_x\}$ is observed in 2% the sets of movies accessed. Other examples of association rules are the following:

- 60% of users that rent the movie “Lord of the Rings – The Fellowship” in sequence select “Lord of the Rings – The Two Towers” and “Lord of the Rings – The Return of the King”.

- 75% of users who visited the “Shrek” movie belong to the 15-20 age group.
- 25% of users who rent the movie “Harry Potter”, were in the 20-30 age group and lived in Athens.

The discovered association rules may use as input, either the sets of movies accessed by all users, or just the ones accessed by users that belong to the same class as the current one. Another option is to use both approaches and suggest the union of discovered topics. This scenario is very useful when association rules inside classes fail to produce reliable recommendations due to lack of adequate input.

6. Production of recommendations

This section describes in detail the algorithm used for the production of movies recommendations. These recommendations are suggested to the customer in the form of “customers who liked/rented this movie also rented...”. The filtering procedure is based on different kind of data. More precisely it depends on the cluster, in which the user was classified, his click-streams, his transaction history and ratings, and the web site ontology metadata. The main idea is to generate an initial set of recommendations combining all the aforementioned data and then prune (using a pre-defined threshold $-\alpha$) this set and sort the remaining items in an ascending order (first the movie with the higher predicted rate).

We separate between two cases according to the existence of user’s history data:

Case 1: when the user is an old one and system has kept track of his history. We define as $M_{History}$ the set of movies that the user has seen and rated. In this case, the system according to $M_{History}$ finds the users “close” to him with similar history and ratings. We define this set of users as U_L . Every user U_i that belongs to U_L is associated with a number s_i which depicts the similarity of U_i with the current user. Also we define as U_M the set of users that belong to U_L and have seen movie M . Moreover, according to $M_{History}$ the model uses the ontology to discover the associations between the items that the user has rated. Therefore, we define as $R_{ontology}$ the list of recommendations that derive from the ontology based on user’s history file.

Case 2: when the user is a new one. In this case, the system uses information from user’s registration and consequently the initial class in which he has been assigned to. We define as U_K the set of users in the class that have similar attributes with the current one. These attributes are the information that the user gave to the system after filling the registration form. For example age, sex, occupation, education, preferences,

etc. Every user U_i that belongs to U_K is associated with a number s_i which depicts the similarity of U_i with the current user according to the aforementioned attributes. We define as U_M the set of users that belong to U_K and have seen movie M . Every class is attached with a set of movies. We define as $R_{cluster}$ the recommended movies according to user's class.

Continuing, we define as $W_{current}$ the current session window which contains the movies that the user has accessed in the current transaction. We can produce recommendation in accordance to user's current click-stream by using association rules and the site ontology. So $R_{rules-current}$ are the recommendations that derive from the current session window and association rules mining and $R_{ontology-current}$ are those that derive from the current session window and the ontology.

We use two methods for processing the information of current session because by their combination we achieve better predictions. Each method uses different ways to predict user's preferences, i.e. association rules encapsulate the navigational behavior of users, while the ontology reflects the connections between items' attributes and discovers patterns that explain user's preferences. In sequence, we name $R_{initial}$ the initial recommendation set that will be filtered by the algorithm and ordered in ascending form. It is computed by the union of $R_{ontology-current}$, $R_{rules-current}$, and in the first case (when the user has historical data) we use the $R_{ontology}$, while in the second case (when the user is a new one) we use the $R_{cluster}$. Consequently, we compute $R_{initial}$ for each case as follows:

Case 1: $R_{initial} = R_{ontology-current} \cup R_{rules-current} \cup R_{ontology}$

Case 2: $R_{initial} = R_{ontology-current} \cup R_{rules-current} \cup R_{cluster}$

The algorithm used for producing recommendations to the e-shop customers is presented in Table 3. As far as similarity among users is concerned, we consider that every user has a vector with the movies he has rated (i.e. a history vector). For example, let X and Y be two users. We define as X_R and Y_R their history vectors respectively and R_X and R_Y the sets with the rated movies. Then C is the set with the rated movies that they have in common, so $C = R_X \cap R_Y$. If the users do not have in common any movie then $C = \emptyset$. Additionally, we define X_C and Y_C the vectors with the movies that both users have rated (movies that belong to C). A limitation that arises in this step concerns the percentage of common rated movies in relation to all rated movies. For example, if the user X has rated 60 movies, $R_X = 60$, and the user Y has rated 40 movies, $R_Y = 40$, and their common movies are only 5, $C = 5$, then these users are not similar. For considering two users as similar they should have in common more than 50% of the one with the fewer rated movies.

Table 3. The algorithm used for producing recommendations.

Input

U : all users, U_i : user with ID i

M : all movies, M_j : movie j

$M_{History}$: movies that the customer has rated

$R_{initial}$: initial recommendations

Algorithm

For each M_i in $R_{initial}$

 If $M_{History} = \emptyset$ then

 Find U_m in U_K

 Else Find U_m in U_L

 For each U_j in U_m

$r_{Mi} = R_{Mi,Uj} * s_j$

 if $r_{Mi} > \alpha$ then

$R = R \cup \{M_i\}$

 end

end

Output

A list R of movies recommendations in ascending order (first the movie with the biggest r) that user U_i has not seen and probably would like to see

Then the similarity between users X and Y is the distance between vectors X_C and Y_C and can be computed from the following type:

$$sim(X_C, Y_C) = \sqrt{\sum_{i=1}^N (R(X_C, M_i) - R(Y_C, M_i))^2}$$

When the user is a new one, we compute the similarity among users in the same class. In this case, we use the attributes that derived from the registration form. These attributes have a certain position in the vector which describes the user's profile. For example, if the attributes are sex, age, occupation and education, an instance of this vector for user X could be $\{0, 22, 3, 2\}$. This means: $\{\text{sex}=\text{male}, \text{age}=22, \text{occupation}=\text{student}, \text{education}=\text{medium}\}$. In other words, every attribute takes values that correspond to certain information e.g. $\text{occupation}=3$ denotes a student.

To sum up, we have used the explicit data that the user gives during his registration in order to categorize him into a class of predefined groups. These groups (classes) are static and generated according to system's characteristics. A user can belong in many different classes. This classification can not be altered in accordance on user's navigational activity. It is static and it can only be modified if the user changes his preferences, i.e. his registration data. On the other hand, we have used the association rules and the ontology of the site. This information derives dynamically from user's navigational activity. Rules' extraction is done during the off-line process while the discovery of the proper ones (according to current session window) is done on-line.

7. Conclusions and future work

In this paper we introduce a recommendation approach for e-shops that extends web usage-based techniques by incorporating semantic annotations, in order to better meet the customers' preferences. To evaluate the proposed approach we measure the *recommendation accuracy* as a ratio of successful recommendations R_s among all recommendations R . As R_s we consider those that the user actually clicks in his navigation. The preliminary evaluation experiments show that the whole process is been enhanced by the combination of site's ontology, user's history files, user's registration data and association rules, which encompasses all necessary knowledge about users' navigational attitudes. Currently we are working on collecting further data in order to extensively evaluate the approach using more metrics such as *coverage* that measure the ability of the system to produce all pages that are likely to be visited by users.

An important point of future consideration relates to the close dependence of the algorithm from the e-shop ontological schema, which can be overbalanced by the association rules mining. Rules catch all users' navigational activity and their relative behavior in relation on products. Another interesting area of work concerns the improvement of algorithm performance by decreasing its complexity, which is augmented analogically with the number of products, of users and their history files. Finally, it will be interesting to investigate different approaches to combine various information channels in order to generate more accurate recommendations.

8. References

- [1] Agrawal, R., and Srikant, R., "Fast Algorithms for Mining Association Rules", *Proc. of the 20th VLDB*, J. Bocca, M. Jarke, & C., Zaniolo, (Eds.), Morgan Kaufmann, 1994, pp. 487-499.
- [2] Berendt, B., "Understanding Web Usage at Different Levels of Abstraction: Coarsening and Visualizing Sequences", *Proc. of the Mining Log Data Across All Customer TouchPoints*, San Francisco, USA, 2001.
- [3] Berendt, B., Hotho, A., and Stumme, G., "Towards Semantic Web Mining", *Proc. of the ISWC 2002*, I. Horrocks and J. Hendler (Eds.), Springer-Verlag, 2002, pp. 264-278.
- [4] Berners-Lee, T., Hendler, J., and Lassila O., "Semantic Web", *Scientific American*, 2001, 284(5), pp. 34-43.
- [5] Browne, J., Higgins, P., and Hunt I., "E-business Principles, Trends, Visions", *E-business Applications, Technologies for Tomorrow's Solutions*, J. Gasos & K.D. Thoben (Eds.), Springer-Verlag, 2003, pp. 3-16.
- [6] Chakrabarti, S., "Mining Web: Discovering Knowledge from Hypertext Data", Morgan Kaufmann, 2003.
- [7] Choa, Y.H., and Kimb, J.K., "Application of Web Usage Mining and Product Taxonomy to Collaborative Recommendations in E-commerce", *Expert Systems with Applications*, 2004, 26, pp. 233-246.
- [8] Ding, Y., Fensel, D., Klein, M., Omelayenko, B., and Schulten, E., "The Role of Ontologies in eCommerce", *Handbook on Ontologies*, S. Staab and R. Studer (Eds.), Springer-Verlag, 2004, pp. 593-615.
- [9] Eirinaki, M., and Vazirgiannis, M., "Web Mining for Web Personalization", *ACM Transactions on Internet Technology*, 2003, 3(1), pp. 1-27.
- [10] Eirinaki, M., Lampos, C., Paulakis, S., and Vazirgiannis, M., "Web Personalization Integrating Content Semantics and Navigational Patterns", *Proc. of 6th ACM International Workshop on Web Information and Data Management*, Washington, USA, 12-13 November 2004, pp. 72-79.
- [11] Eirinaki, M., Vazirgiannis, M., and Varlamis, I., "SEWeP: Using Site Semantics and a Taxonomy to Enhance the Web Personalization Process", *Proc. of 9th ACM Conference on Knowledge Discovery and Data Mining*, Washington, USA, 2003, pp. 99-108.
- [12] Hand, D., Mannila, H., and Smyth, P., "Principles of Data Mining", MIT Press, 2001.
- [13] Kosala, R., and Blockeel, H., "Web Mining Research: Survey", *SIGKDD Explorations*, 2000, 2(1), pp. 1-15.
- [14] Markellou, P., Rigou, M., and Sirmakessis, S., "An Online Consumer Purchase Decision Cycle", *Advances in Electronic Marketing*, I. Clarke III and T. Flaherty (Eds.), Idea Group, 2005, pp. 287-299.
- [15] Markellou, P., Rigou, M., and Sirmakessis, S., "Mining for Web Personalization", *Web Mining: Applications and Techniques*, A. Scime (Ed.), Idea Group, 2005, pp. 27-48.
- [16] Markellou, P., Rigou, M., Sirmakessis, S., and Tsakalidis, A., "Personalization in the Semantic Web Era: a Glance Ahead", *Proc. of 5th International Conference on Data Mining, Text Mining and their Business Applications, Data Mining*, A. Zanasi, N.F. Ebecken and C.A. Brebbia (Eds.), Malaga, Spain, PWIT Press, 2004, pp. 3-11.
- [17] Middleton, S.E., De Roure, D.C., and Shadbolt, N.R., "Ontology-based Recommender Systems", *Handbook on Ontologies*, S. Staab and R Studer (Eds.), Springer-Verlag, 2004, pp. 477-498.
- [18] Middleton, S.E., Shadbolt, N.R., and De Roure, D.C., "Ontological User Profiling in Recommender Systems", *ACM Transactions on Information Systems*, 2004, 22(1), pp. 54-88.
- [19] Mobasher, B., Jin, X., and Zhou, Y., "Semantically Enhanced Collaborative Filtering on the Web", *Proc. of the European Web Mining Forum*, B. Berendt et al. (Eds.), Lecture Notes in Artificial Intelligence, Springer-Verlag, 2004.
- [20] Mulvenna, M., Anand, S., and Behner, A., "Personalization on the Net Using Web Mining", *Communications of ACM*, 2000, 43(8), pp. 122-125.