International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 3, Issue 2, February 2014

HYBRID RECOMMENDATION USING ASSOCIATION RULE MINING BY PARTIAL EVALUATION OF WEB PERSONALIZATION FOR RETRIEVAL EFFECTIVENESS

Richa Soni[1] (Student) Department of Computer Science & Engineering. Chandigarh University (Gharuan, Mohali) India,

Gurpreet Kaur[2](Asstt. Prof.), Department of Computer Science & Engineering. Chandigarh University (Gharuan, Mohali) India,

Abstract- **World Wide Web is the biggest source of information. Though the World Wide Web contains a tremendous amount of data, most of the data is irrelevant and inaccurate from users' point of view. Consequently it has become increasingly necessary for users to utilize automated tools such as recommender systems in order to discover, extract, filter, and evaluate the desired information and resources. Most recommendation algorithms attempt to alleviate information overload by identifying which items a user will find worthwhile. Web page recommender systems predict the information needs of users and provide them with recommendations to facilitate their navigation. Web content and Web usage mining techniques are employed as conventional methods for recommendation. In this paper, we proposed hybrid recommendation systems in m-commerce, which could integrate multiple association rules together to improve recommendation performance. The effects of the hybrid recommenders are examined by comparing the results of hybrid system against the results of single recommendation method. Result shows that the hybrid recommender provides successful recommendation when the recommended page is generated by all the systems of the hybrid. Our proposed approach based on both straight and meandering rules are attached into one set of global association rules, which might be used for the recommendation of web pages and for personalization. These experiments have shown that the use of reduced datasets saves computational time, and neighbor information improves performance.**

Keywords: **Association rules, Data mining, hybrid recommendation, Hybridization Methods, Personalization, Recommendation systems. Web usage mining.**

I. INTRODUCTION

Recommender systems apply data mining techniques and prediction algorithms to predict users interest on information, products and services among the tremendous amount of available items. The vast growth of information on the Internet as well as number of visitors to websites add some key challenges to recommender systems. These are: producing accurate recommendation, handling many recommendations efficiently and coping with the vast growth of number of participants in the system. Therefore, new recommender system technologies are needed that can quickly produce high quality recommendations even for huge data sets. AS INTERNET users surf the Web to find information or products of interest, they are faced with terminologies, such as personalized search, retrieval, filtering, intelligent agent, products that meet customers' needs. Wireless Web faces various advances in hardware technologies; access to application is often very difficult. Most of the application produced for web is not compatible to wireless web due to some constraints. Such recommendations are an essential part of attracting customers. Different Web usage mining techniques have been used to develop efficient and effective recommendation systems. User satisfaction is the most important part of the recommender system. Today the quality of recommendations and the user satisfaction with such systems are still not most favorable. Recommender systems are not favorable for quality of recommendations and user satisfaction. Methods used for the recommender system focuses on the different characteristics of the user. Online companies have the capability to acquire customers' preferences, and then, use them to recommend products on a one-to-one basis in real time and, more importantly, at a much lower cost to company. The software that makes such customized responses possible is commonly called recommendation systems. Recommender system can be running either remotely in a server, or locally in a fixed or mobile consumer device. In both scenarios, the personalization tool selects automatically items that match the customers preferences and needs, which are previously modeled in their personal profiles.

ISSN: 2278 – 1323

In current approaches, the profiles store items which are unappealing to the customers, along with their main attributes (named content descriptions) and their ratings (i.e., the customer's levels of interest). These ratings can be explicit or implicit. In the first case, customers are required to explicitly specify their preferences for any particular item, usually by indicating a value in a continuous range. Negative values commonly mean disliking, while positive values express liking. As explicit ratings impose additional efforts on customers, recommender systems can also infer information about their interests from their behaviour in a much less obtrusive way. Owing to the difficulty of acquiring explicit ratings, some providers of recommendation services adopt hybrid approaches: they compute recommendations based on explicit ratings whenever possible; in case of unavailability, inferred implicit ratings are used instead. The most common Web usage mining techniques used for recommender system are Markov models, Association rules and Clustering. These techniques have strengths and weaknesses. Once the customer's preferences have been modeled, the recommender system elaborates suggestions by resorting to different recommendation strategies. After a recommendation is received, the customer can provide information about its accuracy in an Association rules mining is one of the most important and widespread data mining techniques. They reflect regularities in the co occurrence of the same items within a set of transactions. Association rules that reveal similarities between web pages derived from customer behavior can be simply utilized in recommender systems. The main goal of such a recommendation is to suggest to the current customer some web pages that appear to be useful. Classical association rules, here called straight||, replicate associations alive between items that comparatively often co-occur in ordinary transactions. Association rule mining is a major pattern discovery technique. The main limitation of association rule mining is that many rules are generated, which result in contradictory predictions for a user session. Second limitation is that association rule mining is a nonsequential mining technique that does not preserve the ordering information among pageviews in user sessions.

Recommendation system based clustering can capture a broader range of recommendations, though this is sometimes at the cost of lower prediction accuracy. combining different systems to overcome disadvantages and limitations of a single system may improve the performance of recommenders. Hybrid recommender systems can be used to avoid the drawbacks or limitations of previous recommendation method. They combine two or more systems to improve recommender performance. This research proposes a hybrid algorithm of the hybrid recommenders. This is achieved by comparing the results of hybrid system against the results of single recommendation method and its performance is evaluated based on the correct prediction of the next request of a user, namely Hit-Ratio. The study has focused on developing a hybrid approach that is to suggest a high quality recommendation method for a tremendous volume of data.

The remainder of this paper is organized as follows: Section 2 introduces related work. Section 3 presents partial evaluation of association rule mining for retrieval effectiveness of web personalization. Section 4 shows the experiments and section 5 conclude the paper.

II. RELATED WORK

There is a large body of work on Web usage mining, recommender system and hybridization of recommendation system. there are many researches on how to gain recommendation. In accordance with different data sources, we have divided related work into three categories straight association rule , meandering association rules, Global association rules the missing rating based on existing rating, computing implicit rating based on market basket data, and computing implicit rating based on browsing behavior. The comprehensive review of some main work is done as follows.

Data mining associated with the Web, called Web mining, is divided into three domains: Web usage mining, Web content mining, and Web structure mining. Web mining refers to the automatic discovery of interesting and useful patterns from the data associated with the usage, content, and the linkage structure of Web resources. It has quickly become one of the most popular areas in computing and information systems because of its direct applications in e-commerce, e-CRM, Web analytics, information retrieval/filtering, Web personalization, and recommender systems. Employees knowledgeable about Web mining techniques and their applications are highly sought by major Web companies such as Google, Amazon, Yahoo, MSN and others who need to understand user behavior and utilize discovered patterns from terabytes of user profile data to design more intelligent applications. The primary focus of this course is on Web usage mining and its applications to e-commerce and business intelligence. Specifically, we will consider techniques from machine learning, data mining, text mining, and databases to extract useful knowledge from Web data which could be used for site management, automatic personalization, recommendation, and user profiling. Various Web usage mining techniques have been used to develop efficient and effective recommendation systems. Resnick and Varian proposed the term recommender system to represent a system that takes user recommendations of items as inputs and uses these recommendations as a basis for making recommendations to other users. The six hybridization techniques are surveyed in this work: weighted, mixed, switching, feature combination, feature

ISSN: 2278 – 1323

augmentation and meta-level. In order to generate the association rules, we have used WEKA software [13]. WEKA software provides machine learning algorithms to implement several data mining tasks. It is open source software.

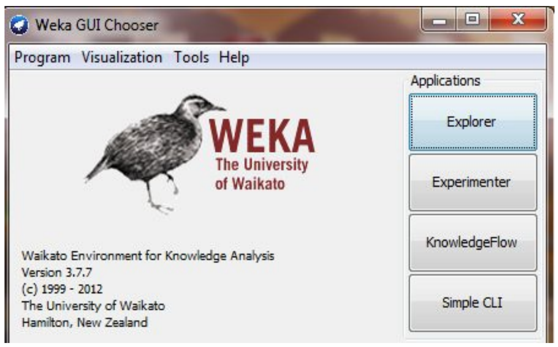

Fig. 1. WEKA Tool

The WEKA crashed and failed to produce any association rules due to a lack of memory issue. The Apriori algorithm scans the database each time that the algorithm mines over the dataset, and it produces a large number of candidate itemsets [9]. The Apriori algorithm is not efficient to work on two dimensional space (User _ Item) with a huge number of items in the space. The algorithm takes an insufficient amount of time to generate the association rules. Additionally, in a machine with a limited memory size and a huge dimensional space, software like WEKA will not be able to generate the association rules due to a memory issue. Our proposed framework consists of two parts. The first part is to generate a set of association rules using the Apriori algorithm. The second part is to apply the generated association rules to recommend items for a user. Association rule mining algorithm We have implemented an association rule mining algorithm oriented to education which is based on the following algorithms: 1) Predictive Apriori for association rule discovery without parameters; and 2) IAS for subjective analysis and classification of unexpected rules by comparing them to a previously defined knowledge database on the field. The algorithm also includes a new weight-based interestingness measurements presented in the section 3.2, to recommend to the teacher any rules according to:

Algorithm

The Proposed Recommendation Framework Part I: Generate the association rules using Apriori Algorithm Part II: for each target user m do find the items that the user m has ranked before group the items that the user m has ranked into two classe : Favorite Items Class (rating of the items >= 3) Non-Favorite Items Class (rating of the items < 3) for each item n in the Favorite Items Class do if the item n is in the associated items then if the user m has not ranked the item u that is derived from item n then recommend the associated item u to the user m end if end if end for end for for each item k in the Non-Favorite Class do use the Item-Based approach to find similar items for the target user m end for Association rule mining algorithms normally discover a huge quantity of rules and do not guarantee that all the rules found are relevant. Therefore, they must be evaluated in order to find the best rules for a specific problem. Traditionally, the use of objective measures has been suggested (Tan and Kumar, 2000), such as support and confidence, mentioned previously, as well as other measures such as Laplace, chi square statistics, correlation coefficients, entropy gain, interest, conviction, etc. These measures can be used to rank the rules obtained so that the user can select those with the highest values for the most appropriate measures. On the other hand, subjective measures are becoming increasingly important (Silberschatz and Tuzhilin, 1996). These measures are based on subjective factors controlled by the user. Most subjective approaches involve user participation in order to express which rules are of the most interest for clarifying and updating previous knowledge. An Interestingness Analysis System (IAS) was proposed by (Liu et al., 2000). IAS compares the newly discovered rules to the user's current knowledge about the area of interest. Using their own specification language, they indicate their level of knowledge about the matter in question through relationships between the fields or items in the database. Let U be the set of user's specifications representing his knowledge space, and A be the set of newly found association rules. This algorithm implements a pruning technique to remove redundant or insignificant rules by ranking and classifying them into four categories:

Conforming rules

A discovered rule $Ai \in A$ conforms to a piece of user's knowledge $Uj \in U$ if both the conditional and consequent parts of Ai match those of $Uj \in U$ well. They use conformij to denote the degree of the conforming match.

Unexpected consequent rules

A discovered rule $Ai \in A$ has unexpected consequents with respect to a $Uj \in U$ if the conditional part of Ai matches that of Uj well although the consequent part does not. They use unexpConseqij to denote the degree of unexpected consequent match.

Unexpected condition rules

A newly found rule $Ai \in A$ has unexpected conditions with respect to a $Uj \in U$ if the consequent part of Ai does matches that of Uj well while the conditional part does not. They use unexpCondij to denote the degree of unexpected condition match.

ISSN: 2278 – 1323

Both-side unexpected rules

A discovered rule Ai ∈ A is unexpected on both-side with respect to a Uj∈ U if neither the conditional nor the consequent parts of rule.

III. PARTIAL EVALUATION OF ASSOCIATION RULE MINING FOR RETRIEVAL EFFECTIVENESS OF WEB PERSONALIZATION

Personalization doesn't just have to be product recommendations: it can also include inserting any content like images or text (e.g. displaying a golf-orientated banner for a returning golf supplies buyer), or customizing content that is already there (e.g. "Hi Joe, we've got some great movie suggestions for you!").

Our strategy includes two steps. The first step is to map a user query to a set of categories which represent the user's search intention and serve as a context for the query. The second step is to utilize both the query and its context to retrieve Web pages. Our goal is to improve retrieval effectiveness. To accomplish it, we propose the following modes of retrieval:

A. The user query is submitted to a search engine without specifying any category. In fact, this is not a mode of personalized search and will be considered as the baseline mode in our experiment.

B. As discussed before, our system determines the three categories which are most likely to match the interests of the user with the given user query. From these three categories, the user can either pick the ones which are most suitable or he/she can decide to see the next three categories. The process continues until the desired categories are chosen by the user. The user usually finds the desired categories within the first three categories presented by the system. Let us call this the semiautomatic mode.

C. In the automatic mode, the system automatically picks the top category or the top two categories or the top three categories without consulting the user. Thus, the two-step personalization of Web search can be accomplished automatically, without the involvement of users. In the last two modes, the user query is initially submitted without specifying any category. Then, the query is submitted by specifying each of the chosen categories as a context.

Personalization technology enables the dynamic insertion, customization or suggestion of content in any format that is relevant to the individual user, based on the user's implicit behavior and preferences, and explicitly given details or also ' "…in any format" – it isn't restricted to the web. It can be implemented for any medium or touch point, such as emails, apps, in store kiosks, etc."…that is relevant to the individual user, based on the user's implicit behavior and preferences, and explicitly given details" – finally, the most important part. Personalization uses both implicit and explicit information, derived in two ways. Firstly, a visitor might explicitly declare some information, such as their gender or date of birth.

There have been several prior attempts on personalizing Web search. A comprehensive survey on personalized search can be found in [21]. In the following sections, we will summarize previous personalized search strategies, including personalized search based on content analysis, personalized search based on the hyperlink structure of the Web, and personalized search based on user groups. Personalized Search Based on Content Analysis One approach of personalized search is to filter or re rank search results by checking content similarity between returned web pages and user profiles. User profiles store approximations of user interests. User profiles are either specified by users themselves [9], [16] or are automatically learnt from a user's historical activities. As the vast majority of users are reluctant to provide any explicit feedback on search results and their interests [22], many works on personalized Web search focus on how to automatically learn user preferences without the user being required to directly participate [5], [9], [15], [23]. In terms of how user profiles are built, there are two groups of works: topical categories [9], [15], [24] or keyword lists (bags of words) [5], [10], [13], [23], [25]. Several approaches represent user interests by using topical categories. In [9], [16], [26], [27], [28], and [29], a user profile is usually structured as a concept/topic hierarchy. User-issued queries and user-selected snippets/documents are categorized into concept hierarchies that are accumulated to generate a user profile. . The documents are re ranked based upon how well the document categories match user interest profiles.

Mass personalization and Predictive personalization

Mass personalization is defined as custom tailoring by a company in accordance with its end users tastes and preferences. The main difference between mass customization and mass personalization is that customization is the ability for a company to give its customers an opportunity to create and choose product to certain specifications, but does have limits. Clothing industry has also adopted the mass customization paradigm and some footwear retailers are producing mass customized shoes.

A website knowing a user's location, and buying habits, will present offers and suggestions tailored to the user's demographics; this is an example of

ISSN: 2278 – 1323

mass personalization. The personalization is not individual but rather the user is first classified and then the personalization is based on the group they belong to. Behavioral targeting represents a concept that is similar to mass personalization. Predictive personalization is defined as the ability to predict customer behavior, needs or wants - and tailor offers and communications very precisely. Social data is one source of providing this predictive analysis, particularly social data that is structured. Predictive personalization is a much more recent means of personalization and can be used well to augment current personalization offerings.

Analysis of Personalized Web Search

In this paper, we reveal that personalization should not be used for all queries in the same manner. Some researchers have also noticed that personalization varies in effectiveness for different queries. For instance, Teevan et al. [7] suggested that not all queries would be handled in the same manner. For less ambiguous queries, current Web search ranking might be sufficient, and thus, personalization is unnecessary. Chirita et al. [16], [25], [35] divided test queries into three types: clear queries, semiambiguous queries, and ambiguous queries. They concluded that personalization significantly increased output quality for ambiguous and semiambiguous queries, but for clear queries, one would prefer a common Web search. Tan et al. [10] divided queries into fresh queries and recurring queries. They found that the recent history tended to be much more useful than the remote history, especially for fresh queries, whereas the entire history was helpful for improving the search accuracy of recurring queries. These conclusions inspired our work of detailed analysis on these kinds of problems. Teevan et al.'s recent work [36] is quite relevant to the work in this paper. They also revealed that personalization does not work equally well on all queries. They examined the variability in user intent using both implicit and explicit measures and further proposed several features to predict variation in user intent.

Discovery Of Community Web Directories From Web Usage Data

The construction of community Web directories is a fully automated process, resulting in operational personalization knowledge, in the form of user models.

User communities are formed using data collected from Web proxies as users browse the Web. The goal is to identify interesting behavioral patterns in the collected usage data and construct community Web directories based on those patterns. The process of getting from the data to the community Web directories is summarized below: Usage Data Preparation comprises the collection and cleaning of the usage data, as well as the identification of user sessions. Web Directory Initialization provides the characterization of the Web pages included in the usage data, according to the categories of a Web directory. We compare two different approaches for the characterization of the Web pages. The first approach organizes Web pages into an artificial Web directory using hierarchical document clustering. The second approach classifies them onto an existing Web directory, like ODP. Community Web Directory Discovery is the main process of discovering the user models from data, using machine learning techniques and exploiting these models to build the community Web directories.

IV. EXPERIMENTAL RESULT

The system is based on client-server architecture with N clients, which applies an association rule mining algorithm locally on students' data using an online course. In the server application are included two modules. The first is a web application server so the experts can manage a knowledge base (KB) and can add, delete or edit tuples, as well as being able to vote on the contributions made by other experts in the team. The second module is a web service, which allows the server to share the updated KB with the client in PMML format (Data Mining Group, 2006). PMML (Predictive Model Markup Language) is an XML-based language that enables the definition and sharing of predictive models between applications, establishing a vendor-independent means of defining these models, so that problems with proprietary applications and compatibility issues can be circumvented The main phases used in the CIECoF (Continuous improvement of e-learning course framework) architecture are (Figure 2)

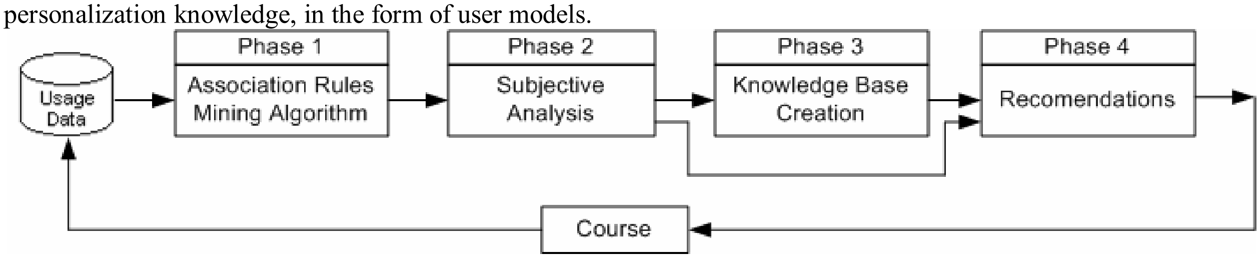

Fig. 2 Main phases of CIEC oF architecture.

ISSN: 2278 – 1323

Association rules mining

This phase aims to find association rules on the data set generated as the students complete the course. Once the data has been pre-processed, it is used as input of the Predictive Apriori algorithm, the nucleus of this phase. Also, the teacher could select specific data and attributes in order to restrict the search domain. The output of this module (rules found) is then analyzed by the subjective analysis module.

A. Subjective analysis

This phase uses a subjective rule evaluation measure to determine the interestingness of the rules found by association rule mining. It also applies the IAS algorithm to classify the rules in expected or unexpected comparing them with the rules stored in the knowledge base.

B. Knowledge base creation

This phase combines collaborative filtering techniques with knowledge based techniques to create and to manage the rules repository. The information in the knowledge base is stored in form of tuples (rule-problem-recommendation-relevance) which are classified according to a specific course profile. In order to avoid the cold start issue of collaborative filtering systems, the experts propose the first tuples of the repository and also vote for those tuples proposed by other experts. On the other hand, the teachers could discover new tuples that must be validated by the experts before being inserted in the repository and also votes for the others tuples.

C. Recommendations

The expected rules found by the phase 2 joined to the more intuitive tuples format mentioned in phase 3, are then used in this last phase to show the teacher, in most of the cases non expert in data mining, possible solutions to some problems detected in the course The system is based on client-server architecture with N clients, which applies an association rule mining algorithm locally on students' data using an online course. In the server application are included two modules. The first is a web application server so the experts can manage a knowledge base (KB) and can add, delete or edit tuples, as well as being able to vote on the contributions made by other experts in the team. The second module is a web service, which allows the server to share the updated KB with the client in PMML format (Data Mining Group, 2006). PMML (Predictive Model Markup Language) is an XML-based language that enables the definition and sharing of predictive models between applications, establishing a vendor-independent means of defining these models, so that problems with proprietary applications and compatibility issues can be circumvented. So, once the updated version of the KB has been downloaded from the server, the client can apply the mining algorithm offline. Client application is part of the iterative methodology (García et al., 2006) that teachers use to develop courses. It is capable of detecting possible problems in the design and content of an e-learning course by adding a feedback or maintenance stage to the course.

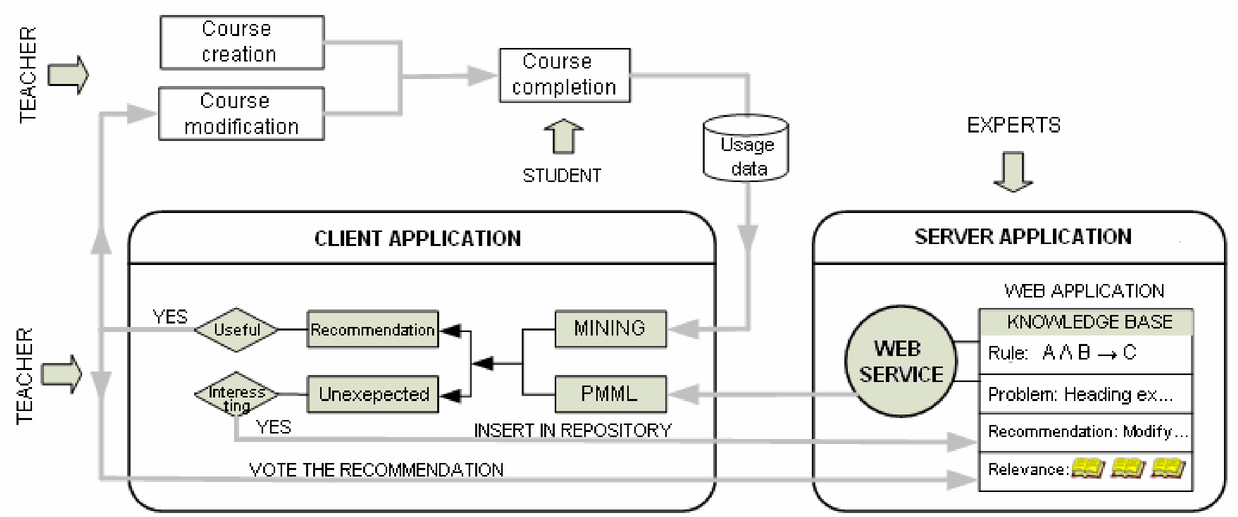

Fig. 3 Client and Server Application

Results of Mapping User Queries to Categories
First, we investigate the effectiveness of the four batch learning algorithms based on only the user profiles. Fig. 3 show their accuracy results. As can be seen from Fig. 3, pLLSF, kNN, and bRocchio have similar effectiveness

ISSN: 2278 – 1323

and all of them perform well; their accuracy ranges from 0.768 to 0.975 with the exception of user 1. These three algorithms outperform LLSF .This indicates that dimension reduction with SVD is worthwhile. In the evaluation framework, we use query logs of Windows Live Search to simulate and evaluate personalized reranking strategies. We organize a log entry for a query as the format. In Windows Live Search query logs, each user is identified by "Cookie GUID"

Time, Cookie GUID, Query String, Browser GUID
{
  $(\text{Position}_1, \text{URL}_1)$, $(\text{Position}_2, \text{URL}_2)$, ... $(\text{Position}_n, \text{URL}_n)$
}

that remains the same in a machine as long as a cookie is not cleaned. For each query, the Windows Live search engine logs the query string and all click-through information, including clicked web pages and their corresponding ranks. A "Browser GUID" is assigned when a browser is opened and expired when the browser is closed. "Browser GUID" is used as a simple identifier of a session that contains a series of related queries made by a user within a small range of time. A session is usually meaningful in capturing a user's attempts to fulfill certain information needs [1], [2].

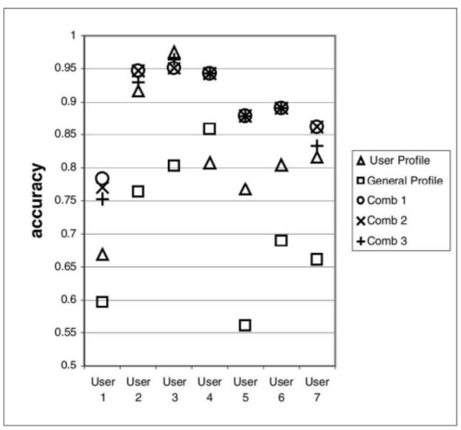

Fig. 4 Comparison of different mapping methods on seven users.

We use static classification to compare teachers, so similar profile refers to an exact coincidence between one profile and other.b) A team of validating experts has voted for in terms of interest or validity.

The algorithm implemented is especially useful in collaborative recommender systems, which can take advantage of the synergies offered by the network, in order to produce recommendations that are increasingly useful and precise. The main algorithm is interactive and iterative (see Table II). In each iteration, the teacher runs the mining algorithm in order to find the rules that will act as a basis for recommendations; this can be done as often as necessary.

Table II. Main algorithm.

Input: Topic, Level, Difficulty: teacher profile;
N: number of rules to discover;
1) Iters = 0;
2) KB = Get_Rules_fromServer(Topic, Level, Difficulty);
3) While (teacher doesn't stop) do
4) Re, Rne = Rules_Mining_Algorithm(N, KB, Iters);
where Reiters ≠ Reiters+1, Rneiters ≠ Rneiters+1
5) For each i-rule in Re do
6) Teacher_Vote_Recommendation(Rei)
7) End
8) For each i-rule in Rne do
9) If (Interesting(Rnei)) then
10) Add_to_KnowledgeBase(Rnei);
11) End if
12) End
13) Iters ++;
14) End while
15) End all

In step 1) the variable Iters, which counts the number of iterations, is initialised at zero; in step 2) the teacher downloads the knowledge base (KB) from the server corresponding to his/her course profile; in step 3) the main loop starts and all its instructions will be executed until the teacher decides to stop it. Step 4) calls up the rule mining algorithm described , which returns the sets of recommendations (Re) and unexpected rules (Rne) discovered where Re and Rne are different from one iteration to another. From steps 5) to 7), the teacher votes on whether the recommendation has been useful or not, and in steps 8) to 12), he/she evaluates unexpected rules to determine whether or not they are useful; unexpected rules might be added to the knowledge base (KB), subject to prior validation by the experts. Finally, in step 13), the Iters variable is incremented. The rule mining algorithm implement is described as follows Let accRi (i=1,2,…n) be the predictive accuracy of Ri; R the set of rules discovered by the current teacher, Re the set of expected rules, and Rne the set of unexpected rules, then R = Re ∪ Rne; KB is the set of rules that makes up the knowledge database concerning this field. In step 1), the GenRules function reveals the association rules; this function is provided with the desired number of rules and calls on the PA

ISSN: 2278 – 1323

algorithm. In step 2), the rule found is classified as being expected if it syntactically matches rule in the current knowledge database, that is, if it has both the same antecedent and consequent. The rule is classified as unexpected if it does not. From steps 3) to 5), for each rule Ri ∈ Re, the new weight-based interestingness measurement WAcc is calculated. Rule Mining Algorithm Input**:** N: number of rules to discover; Iters: number of iterations

KB: knowledge base; Output: Re: recommendations set; Rne: unexpected rules;

1) R, accR = GenRules (N, Iters); // Call to Predictive Apriori
2) Re, Rne = Classify(R);
3) For each i-rule in Re do
4) R i WAcc = CalculateWeightedAccuracy (Ri);
5) End
6) For each i-rule in Rne do
7) For each j-rule in KB do
8) conformij, unexpConseqij, unexCondij, bsUnexpij = IAS( );
9) End
10) End
11) Order all the rules in Re from largest to smaller Wacc
12) Output the set Re as the set of recommendations
13) Ouput the unexpected rules Rne according to IAS
14) End all

From steps 6) to 10) the IAS algorithm is used to calculate the degree to which each unexpected rule Rne coincides with the rules stored in the knowledge base (KB). In our system, all the unexpected rules are ordered as follows: a) the conformed rules that are the basis of recommendations to the professor; b) unexpected both-sided rules whose antecedent and consequence have never been mentioned in our knowledge base; c) the unexpected consequent rules that show us those rules found to be contrary to our existing knowledge; and d) the unexpected condition rules show us that there are other conditions outside of our specified knowledge range that could be pertinent and conducive to learning. In step 11), the set Re is ordered from highest to lowest based on the previously calculated WAcc. Step 12) displays all the recommendations corresponding to each of the previously ordered rules. Finally, in step 13), the teacher is given the chance to view the set of unexpected rules in order to assess which candidates are feasible and desirable for our knowledge database.

Analysis of the recommendation effectiveness

In order to verify the effectiveness of the changes made by the teachers in the course, based on the recommendations suggested by the system, it is important to bear two points of view in mind: 1) the teacher's perspective, in terms of the percentage of apparently corrected problems, based on initial recommendations, that reappear in successive courses with different groups of students; and 2) the perspective of the students with respect to how the removal of those problems based on the recommendations influences their final score. Two hypotheses can initially be drawn from these aspects. Firstly, if the changes made by the teacher are 100% effective, then these problems should not be detected again in subsequent groups of students doing a course that has already been updated by applying the corrections. And secondly, if these problems do not happen again, then students' scores should improve. We have implemented an iterative methodology to improve the course gradually with use. Using the recommendations obtained from the usage data of different groups of students, successive corrections to the course improve it step by step. In order to calculate the effectiveness of these recommendations (EfecRec1,i), we use equation where TotalNew1 represents the total number of recommendations found when the usage data of the first group of students were analysed, which led to changes in the structure or content of the course. TotalRep1,i is the total number of recommendations that are repeated in consecutive runs of the same course, always applying the corrections with each different group of students. Thus, the effectiveness of the changes made can be calculated, based on the recommendations proposed in the initial stage (the first course run) with respect to stage i (i=2,3...N),

V. CONCLUSION AND FUTURE SCOPE

We described a strategy for personalization of Web search:

A. A user's search history can be collected without direct user involvement.
B. The user's profile can be constructed automatically from the user's search history and is augmented by a general profile which is extracted automatically froma common category hierarchy.
C. The categories that are likely to be of interest to the user are deduced based on his/her query and the two profiles.
D. These categories are used as a context of the query to improve retrieval effectiveness of Web search. It should be noted that the experimental results reported here include seven users, a few hundred queries, and a limited number of relevant documents. There is also room for obtaining higher levels of improvement than reported here as we choose reasonable (but not exhaustive) values for a number of parameters (e.g., the weight associated with each list of retrieved documents). Future research in the area consists of a much larger scale of experiments as well as optimization of parameters In most previous work on personalized Web search, all queries were usually personalized in the same manner.Another important conclusion we revealed in this paper is that personalization does not work equally well under various

ISSN: 2278 – 1323

situations. We defined the click entropy to measure variation in information needs of users under a query.

Experimental results showed that personalized Web search yields significant improvements over generic Web search for queries with a high click entropy. For the queries with a low click entropy, personalization methods performed similarly or even worse than generic search. As personalized search had different effectiveness for different kinds of queries, we argued that queries should not be handled in the same manner with regard to personalization. Our proposed click entropy can be used as a simple measurement on whether a query should be personalized.

REFERENCES

[1] J. Allan, "Incremental Relevance Feedback for Information Filtering," Proc. 19th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 270-278, 1996.

[2] M. Balabanovic and Y. Shoham, "Learning Information Retrieval Agents: Experiments with Automated Web Browsing," Proc. OnLine Working Notes AAAI Spring Symp. Series on Information Gathering from Distributed, Heterogeneous Environments, pp. 13-18,1995.

[3] K. Bollacker, S. Lawrence, and C.L. Giles, "A System for Automatic Personalized Tracking of Scientific Literature on theWeb," Proc. Fourth ACM Conf. Digital Libraries, pp. 105-113, 1999.

[4] J. Budzik and K. Hammond, "Watson: Anticipating and ContextualizingInformation Needs," Proc. 62nd Ann. Meeting Am. Soc.Information Science, 1999.

[5] U. Cetintemel, M.J. Franklin, and C.L. Giles, "Self-Adaptive User Profiles for Large-Scale Data Delivery," Proc. 16th Int'l Conf. Data Eng. (ICDE), pp. 622-633, 2000.

[6] L. Chen and K. Sycara, "WebMate: A Personal Agent for Browsing and Searching," Proc. Second Int'l Conf. Autonomous Agents and Multi Agent Systems, pp. 132-139, 1998.

[7] S. Deerwester, S.T. Dumais, G.W. Furnas, T.K. Landauer, and R. Harshman, "Indexing by Latent Semantic Analysis," J. Am.Soc. Information Science (JASIS), vol. 41, no. 6, pp. 391-407, 1990.

[8] R. Dolin, D. Agrawal, A. El Abbadi, and J. Pearlman, "UsingAutomated Classification for Summarizating and Selecting HeterogeneousInformation Sources," D-Lib Magazine, 1998.

[9] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar, "Rank Aggregation Methods for the Web," Proc. 10th Int'l World Wide Web Conf., pp. 613-622, 2001.

[10] P.W. Foltz and S.T. Dumais, "Personalized Information Delivery:An Analysis of Information Filtering Methods," Comm. ACM, vol. 35, no. 12, pp. 51-60, 1992.

[26] J. Chaffee and S. Gauch, "Personal Ontologies for Web Navigation," Proc. ACM Int'l Conf. Information and Knowledge Management (CIKM '00), pp. 227-234, 2000.

[27] S. Gauch, J. Chaffee, and A. Pretschner, "Ontology-Based Personalized Search and Browsing," Web Intelligence and Agent Systems, vol. 1, no. 3/4, pp. 219-234, 2003. J. Trajkova and S. Gauch, "Improving Ontology-Based User Profiles," Proc. Recherche d'Information Assiste´e par Ordinateur (RIAO '04), pp. 380-389, 2004.

[29] M. Speretta and S. Gauch, "Personalized Search Based on User Search Histories," Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence (WI '05), pp. 622-628, 2005.

[30] L. Page, S. Brin, R. Motwani, and T. Winograd, "The Page Rank Citation Ranking: Bringing Order to the Web," technical report, Computer Science Dept., Stanford Univ., 1998.

[31] T.H. Haveliwala, "Topic-Sensitive Pagerank," Proc. 11th Int'l World Wide Web Conf. (WWW), 2002.

[32] T. Sarlo´ s, A.A. Benczu´ r, K. Csaloga´ny, D. Fogaras, and B. Ra´cz,"To Randomize or Not to Randomize: Space Optimal Summariesfor Hyperlink Analysis," Proc. 15th Int'l World Wide Web Conf. (WWW '06), pp. 297-306, 2006

38] E. Yom-Tov, S. Fine, D. Carmel, and A. Darlow, "Learning to Estimate Query Difficulty: Including Applications to Missing Content Detection and Distributed Information Retrieval," Proc. 28th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '05), pp. 512-519, 2005.

[39] D. Carmel, E. Yom-Tov, A. Darlow, and D. Pelleg, "What Makes a Query Difficult?" Proc. 29th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '06), pp. 390-397,2006.

[40] Y. Zhou and W.B. Croft, "Query Performance Prediction in Web Search Environments," Proc. 30th Ann. Int'l ACM SIGIR Conf.Research and Development in Information Retrieval (SIGIR '07),pp. 543-550, 2007.

[41] T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. Gay, "Accurately Interpreting Clickthrough Data as Implicit Feedback," Proc. 28th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '05), pp. 154-161, 2005.

ISSN: 2278 – 1323

[42] Z. Guan and E. Cutrell, "An Eye Tracking Study of the Effect of Target Rank on Web Search," Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '07), pp. 417-420, 2007.

[43] M. Klusch and P. Kapahnke, "iSeM: Approximated Reasoning for Adaptive Hybrid Selection of Semantic Services," Proc. IEEE Fourth Int'l Conf. Semantic Computing (ICSC '10), pp. 184-191, 2010.

[44] G. Linden, B. Smith, and J. York, "Amazon.com Recommendations: Itemto-Item Collaborative Filtering," IEEE Internet Computing, vol. 7, no. 1,pp. 76-80, Jan. 2003.

[45] L. Liu, F. Lecue, N. Mehandjiev, and L. Xu, "Using Context Similarity for Service Recommendation," Proc. IEEE Fourth Int'l Conf. Semantic Computing(ICSC '10), pp. 277-284, 2010.

[46] Hongwei Liu, Zhouyang Liang,|| Implicit Rating Model in Mcommerce Recommendation System|| 978-1-4244-4507-3/09/IEEE.- 2009.

[47] Ching-Huang Yun and Ming-Syan Chen, Fellow, IEEE," Mining Mobile Sequential Patterns in a Mobile Commerce Environment"- ieee transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 37, no. 2, march 2007.

[48] Chengzhi Liu, Caihong Sun, Jia Yu,|| The Design of an Open Hybrid Recommendation System for Mobile Commerce|| ISE 2008.

[49] Shu Liao, Tengyue Zou, Huiyou Chang,|| An Association Rules and Sequential Rules Based Recommendation System|| 978-1-4244-2108-4/08-2008 IEEE.

[50] Daniel Gallego and Gabriel Huecas,|| An Empirical Case of a Context-aware Mobile Recommender System in a BankingEnvironment|| Third FTRA International Conference on Mobile,Ubiquitous, and Intelligent Computing- 2012.

[51] M. Hall and E. Frank and G. Holmes and B. Pfahringer and P. Reute-mann and I. H. Witten, The weka data mining software: An update, SIGKDD Explorations, 11:10-18, 2009.

[52] M. Welch and J. Cho,Automatically identifying localizable queries, In 31st annual International ACM SIGIR conference on Research and development in information retrieval, pp. 1185-1186, ACM, 2008.

[53] FENG ZHANG, HUI-YOU CHANG,|| On A Hybrid Rule Based Recommender System|| The Fifth International Conference on Computer and Information Technology (CIT'05).

[54] Guoling Liu, He Jiang, Runian Geng, Huijie Li,|| Application of Multidimensional Association Rules in Personal Financial Services|| International Conference On Computer Design And Appliations (ICCDA 2010).

[55] Chellatamilan T. ,Suresh R. M.|| An e-Learning Recommendation System using Association Rule Mining Technique|| ISSN 1450-216X Vol.64 No.2 (2011), pp. 330-339.

[56] R. Forsati, M. R. Meybodi, A. Ghari Neiat,|| Web Page Personalization based on Weighted Association Rules|| International Conference on Electronic Computer Technology-2009.

[57] M. Kamvar and S. Baluja : Deciphering trends in mobile search. Com- puter, 40(8), 58-62, 2007

[58] Ourdia Bouidghaghen, Lynda Tamine, Mohand Boughanem, Personalizing Mobile Web Search for Location Sensitive Queries 12th IEEE International Conference on Mobile Data Management-2011.

[59] Shuk Ying Ho, Sarah Bull, Customers' —Trust and Intention to Adopt Personalized Mobile Services|| Journal of Next Generation Information Technology, volume 2, Number 1, February, 2011.

ISSN: 2278 – 1323