Design and Implementation of Architectural Framework of Recommender System for e-Commerce

Sanjeev Kumar Sharma
Research Scholar
Devi Ahilya University Takshashila Campus
Khandwa Road Indore (M.P.) India

Dr. Ugrasen Suman
Associate Professor
Devi Ahilya University Takshashila Campus
Khandwa Road Indore (M.P.) India

***Abstract*—The rapid development of internet technologies in recent decades has imposed a heavy information burden on users. The popularity of recommender systems has evolved to provide suggestions and recommendations to the user for relevant information from web according to their preferences. Most recommender systems use collaborative-filtering or content-based methods to predict new items of interest for users. While both methods have their own advantages, individually they fail to provide good recommendations in many situations. In this paper, we propose a standard architectural framework Semantic Enhanced Personalizer (SEP) which integrates three recommendation techniques i.e., original, semantic and category-based. This framework fulfills user-based and item-based approach of recommendations. The original recommendation will be based on contextual information and the ratings provided by users explicitly, while, the semantic and category-based recommendation will be based on various data mining techniques such as, association rule mining, clustering and similarity measures. This framework overcomes the problem of cold-start and sparsity.**

Keywords- Recommender system; Data mining; similarity; ratings; SEP

I. INTRODUCTION

The explosive growth and variety of information available on the Web and the rapid introduction of new e-business services (buying products, product comparison, auction, etc.) frequently create the confusion in the mind of users. Therefore, Recommender Systems (RSs) are those software tools and techniques which gives the suggestions for items to the user according to his/her preferences [1] [2] [3]. These suggestions are related to various decision-making processes such as what items to buy or what online news to read or what music to listen. A RS normally focuses on a specific type of item (e.g., CDs, or news, movies) and accordingly its design, its graphical user interface, and the core recommendation technique is used to generate the recommendations are all customized to provide useful and effective suggestions for that specific type of item. RSs are very useful for those individuals who are lacking sufficient personal experience to evaluate the number of alternative items available on website. For example, a book recommender system that assists users to select a book to read. In the popular website such as Amazon.com, this site employs a RS to personalize the online store for each customer [4].

Recommender systems emerged as an independent research area in the mid-1990s [5][6][7][8]. In recent years, the interest in recommender systems has dramatically increased, as the following facts indicate. Recommender systems play an important role in such highly rated internet sites as Amazon.com, YouTube, Netflix, Yahoo, Tripadvisor, Last.fm, and IMDB. Moreover many media companies are now developing and deploying RSs as part of the services they provide to their subscribers. For example Netflix, the online movie rental service awarded a million dollar prize to the team that first succeeded in improving substantially the performance of its recommender system [9].

In fact, there are various reasons as to why service providers may want to use this technology. Firstly, RS increases the number of items sold and it is also capable to sell an additional set of items compared to those usually sold without any kind of recommendation. Another major function of a RS is to sell more diverse items. It means RS enable the user to select items that might be hard to find without a precise recommendation. For example, in a movie RS such as Netflix, the service provider is interested in renting all the DVDs in the catalogue, not just the most popular ones. This could be difficult without a RS since the service provider cannot afford the risk of advertising movies that are not likely to suit a particular user's taste. Therefore, a RS suggests or advertises unpopular movies to the right users. The other function of RS is to increase the user satisfaction.

A well designed RS can also improve the experience of the user with the site or the application. The RSs are also capable to increase the user-fidelity. A user should be loyal to a Web site which, when visited, recognizes the old customer and treats him as a valuable visitor. There are various types of recommendation techniques but six different classes of recommendation techniques are describing here.

The first class of recommendation is content based recommendation. In this approach, a set of documents are analyzed on the basis of ratings given by the previous users. User interests are based on the features of the objects rated by that user [10]. The second class of recommendation is collaborative-based recommendation. This approach recommends the items to active clusters with similar tastes of items for other users liked in the past [11]. The similarity in taste of two users is calculated based on the similarity in the rating history of the users. Therefore, this approach is also known as people-to-people correlation [12]. Collaborative filtering is considered to be the most popular and widely implemented technique in RS. The third class of recommendation is demographic-based recommendation. In this approach, system recommends items on the basis of demographic profile of the user. These approaches are popular in the marketing literature; there has been relatively little proper RS research into demographic systems [13]. The next approach of recommendation is knowledge-based recommendation. In this approach, knowledge-based system recommends the items based on specific domain knowledge. Notable knowledge based recommender systems are case-based [14] [15].

The other class of recommendation is **c**ommunity-based recommendation. In this approach, system recommends items, based on preferences of user's friends. It has been observed that people tend to rely more on recommendations from their friends rather than anonymous individuals [16][17][18]. This observation generates a rising interest in community-based systems or, as or as they usually referred to, social recommender systems [19]. Another class is context-based recommender system. The majority of existing approaches to recommender systems focus on recommending the most relevant items to individual users and do not take into consideration any contextual information, such as time, place and the company of other people (e.g., for watching movies or dining out). It is also important to incorporate the contextual information into the recommendation process in order to recommend items to users under certain circumstances. For example, a travel recommender system would provide a vacation recommendation in the winter that can be very different from the one in the summer. Therefore, accurate prediction of consumer preferences depends upon the degree to which the recommender system has incorporated the relevant contextual information into a recommendation method.

The other class of recommendation is hybrid recommender systems. These RSs are based on the combination of the above mentioned techniques. A hybrid system combining techniques A and B tries to use the advantages of A to fix the disadvantages of B.

The basic outline of this paper is as follows. Section 2 describes the review of related researches, recommender systems for various domains and new challenges of RSs. Section 3 presents the analysis of various data mining techniques used for RSs. Section 4 describes the components of SEP framework for RS. Section 5 describes the algorithm for recommendation. Section 6 presents the experimental evaluation. Section-7 gives conclusion of the proposed approach.

II. RELATED REVIEW

There have been a lot of research works accomplished in area of web-recommendation. In this section we will respectively address the concept of recommender system and the types of RS for various domains. Several recommender systems have been developed so far according to different recommendation techniques discussed above. These recommender systems are related to various fields of applications, such as news, music, e-commerce, movies, etc. Each domain presents different problems that require different solutions. In the area of content based Web recommenders, famous systems in literature are Letizia, WebMate [20][21].

Letizia is a web-browser that tracks the user's browsing behavior and builds a personalized model consisting of keywords related to the user's interests. It uses implicit feedback to infer the user's preferences. For example, bookmarking a page is interpreted as strong evidence for the user's interests in that page. In a similar way, another approach for representing user interests is WebMate that keeps track of user interests in different domains by learning a user profile that consists of the keyword vectors that represents positive training examples. A profile of n keyword vectors can correctly represent up to n independent user interests.

The news filtering recommender systems are INFOrmer, NewsDude [22][23]. NewsDude learns a short-term user model based on TF-IDF (cosine similarity), and a long-term model based on a naive Bayesian classifier by relying on an initial training set of interesting news articles provided by the user. INFOrmer uses a semantic network for representing both user profiles and articles.

A variety of content-based recommender systems exist in other application domains. LIBRA is a book recommender that obtains the product description from the web pages of the Amazon on-line digital store [24]. ReAgent is an intelligent email agent that can learn actions such as filtering, downloading to palmtops, forwarding email to voicemail, etc. using automatic feature extraction [25]. Re:Agent users are required only to place example messages in folders corresponding to the desired actions. ReAgent learns the concepts and decision policies from these folders. Citeseer performs a scientific literature search using word information and analyzing common citations in the papers [26].

INTIMATE learns the user preferences from movie synopses obtained from the Internet Movie Database (IMDB) and recommends movies using text categorization techniques [27]. In the same way, Movies2GO learns user preferences from the synopsis of movies rated by the user [28]. In the music domain, the commonly used technique for providing recommendations is collaborative filtering. The most noticeable system using (manual) content-based descriptions to recommend music is Pandora. The main problem of the system is scalability, because the music annotation process is entirely done manually. Conversely, FOAFing the music is able to recommend, discover and explore music content, based on user profiling via Friend of a Friend (FOAF) descriptions, context-based information extracted from music related RSS feeds, and content-based descriptions automatically extracted from the audio itself [29]. SiteIF is a multilingual news website that adopts a sense-based document representation in order to build a model of the user interests [30]. ITR (Item Recommender) provide the recommendations for items in several domains (e.g., movies, music, books) on the basis of descriptions of items e.g. plot summaries, reviews, short abstracts) [31]. SEWeP (Semantic Enhancement for Web Personalization) uses both the usage logs and the semantics of a Web site's content for web personalization [32].

Quickstep recommends on-line academic research papers, in which item-profile matching is realized by computing a correlation between the top three interesting topics in the user profile and papers classified as belonging to those topics. Foxtrot adds the features such as, profile visualization interface and an email notification to extend the Quickstep system [33].

A domain of recommendation is the set of items that the recommender will operate over, but may also include the set of aims or purposes that the recommender is intended to support. We have identified six important characteristics of the domain that an implementer should consider: heterogeneity, risk, churn, interaction style, preference stability, and scrutiny. A heterogeneous item space encompasses many items with different characteristics and most importantly, different goals they can satisfy. For example, an e-commerce recommender system as found at Amazon.com has a large number of heterogeneous items that can be recommended. Even within a single category like books, such disparate categories as home repair, romance novels, cooking, and children's fantasy all coexist in the database. A homogeneous recommendation space means that content knowledge relative to the domain will be easier to acquire and maintain. Consider a site that only recommends digital cameras versus one that has all kinds of electronics. The camera-only site would be able to invest in content knowledge specific to photography.

Recommendation domains can be distinguished by the degree of risk that a user incurs in accepting a recommendation. Risk determines the user's tolerance for false positives among the recommendations. In some domains, false negatives may also be important if there is a cost or risk associated with not considering some options. A high churn domain is one in which items come and go rapidly. In such a domain, a recommender system faces a continual stream of new items to be integrated into its knowledge sources. This greatly increases the sparsity of any kind of opinion data, as new items will necessarily have been seen by very few users.

User preferences can also have varying degrees of duration. For example, when one's favorite basketball team is in a big tournament, stories about it become highly preferred, but if they are knocked out or when the tournament is over, the user's preferences will change.

Table 1.0 illustrates 10 different domains where recommendation applications exist. High-risk domains generally considered under knowledge-based recommendation. Scrutiny is also a good predictor of knowledge-based recommendation. Heterogeneous domains are handled largely with collaborative recommendation. Web page recommendation looks a bit contradictory when we consider high churn and preference instability, which would seem to militate against collaborative methods. However, database size can compensate for preference instability and these recommenders collect large amounts of implicit preference data in each session. It has been observed that the recommender systems with social knowledge, requires high heterogeneity.

I. ANALYSIS OF DATA MINING METHODS FOR RECOMMENDER SYSTEMS

Recommender Systems (RS) apply techniques and methodologies of Data Mining (DM) for information extraction such as Similarity measures, Sampling, Dimensionality Reduction, Classification, Association-Rule-Mining (ARM) and Clustering. Similarity Measures is one of the preferred approaches to collaborative filtering (CF) recommenders is to use the kNN classifier. This classification method is highly dependent on defining an appropriate similarity or distance measures. Another technique of data mining is sampling, it is used in DM for selecting a subset of relevant data from a large data set. It is used both in the preprocessing and final data interpretation steps.

Sampling may be used because processing the entire data set is computationally too expensive. It can also be used to create training and testing datasets. Dimension Reduction is a another data mining technique which shows the notions of density and distance between points, which are critical for clustering and outlier detection, become less meaningful in highly dimensional spaces. This is known as the Curse of Dimensionality. Dimensionality reduction transforms the original high-dimensional space into a lower-dimensionality. There are two most relevant dimensionality reduction algorithms in the context of RS such as Principal Component Analysis (PCA) and Singular Value Decomposition (SVD).

Principal Component Analysis (PCA) is a classical statistical method to find patterns in high dimensionality datasets [34]. Singular Value Decomposition (SVD) is a powerful technique for dimensionality reduction [35].

TABLE I. RECOMMENDER SYSTEMS WITH DIFFERENT DOMAINS

Domain	Risk	Churn	Heterogeneous	Preferences	Interaction Style	Scrutiny	Examples	Recommendation Technology
E-commerce	Low	High	High	Stable	Implicit	Not required	Amazon.com, eBay	Collaborative-Filtering
Financial-services and Life-insurance	High	Low	Low	Stable	Explicit	Required	Koba4MS[36], FSAdvisor[37] [38]	Knowledge-Based
Job search Recruiting	High	Low	Low	Stable	Explicit	Required	CASPER [39] and [40]	Content-based
Movie	Low	Low	Low	Stable	Implicit	Not required	Netflix[41],[42], INTIMATE[27], Movies2Go[28]	Collaborativeand Content Filtering
Music	Low	Low	Low	Stable	Implicit	Not required	Pandora and [43, 44, 45]	Content-based, Hybrid
News	Low	High	Low	Stable?	Implicit	Not required	Yahoo news[46], ACR news[47], and [48] Google news[49], INFOrmer[29],NewsDude[30]	Content-based, Collaborative-Filtering
Real Estate	High	Low	Low	Stable	Explicit	Required	RentMe [50], FlatFinder[51] and [52]	Knowledge-based
Scientific Research papers	Low	Low	Low	Stable	Explicit/implicit	Not Required	QuickStep system[53], Citeseer[26]	Content-based
Software Engineering	Low	Low	Low	Stable	Explicit/Implicit	Required	[54] and [55]	Hybrid and Content-based
Tourism	High	Low	Low	Unstable	Explicit	Required	Travel Recommender [56] [57]	Content-based, Knowledge-based
TV Program	Low	Low	Low	Unstable	Implicit/Explicit	Not required	AVTAR[58]	Content-based
Web Page Recommender	Low	High	High	Unstable	Implicit	Not required	[59], [60], [61],Letzia[20]	Collaborative-Filtering, Hybrid

Another technique of data mining is classification, it is a mapping between a feature space and a label space, where the features represent characteristics of the elements to classify and the labels represent the classes. A restaurant RS, for example, can be implemented by a classifier that classifies restaurants into one of two categories (good, bad) based on a number of features that describe it. There are many types of classifiers, but in general we will talk about either supervised or unsupervised classification. In supervised classification, a set of labels or categories is known in advance and we have a set of labeled examples which constitute a training set. In unsupervised classification, the labels or categories are unknown in advance and the task is to suitably organize the elements at hand.

The next technique of data mining for recommender systems is clustering. It is a process that assigns items to groups so that the items in the same groups are more similar than items in different groups. The goal is to discover meaningful groups that exist in the data [62]. Similarity is determined using a distance measure. The goal of a clustering algorithm is to minimize intra-cluster distances while maximizing inter-cluster distances. There are two main categories of clustering algorithms such as, hierarchical and partitional. Partitional clustering algorithms divide data items into non-overlapping clusters such that each data item is in exactly one cluster while Hierarchical clustering algorithms successively cluster the items within found clusters, producing a set of nested cluster organized as a hierarchical tree. The *k*-means clustering and DBSCAN algorithm is a partitional clustering. There is another semantic document clustering technique, in which annotated pages will be created semantically and these are grouped into clusters. Such type of categorization is achieved by clustering the web document based on semantic similarity between the ontology terms that characterize them [63]. The other technique of data mining is association rule mining. It finds the rules that will predict the occurrence of an item based on the occurrences of other items in a transaction.

II. SEP ARCHITECTURE FOR RECOMMENDER SYSTEM

The proposed approach consists of an architectural framework of Semantic Enhanced Personalizer (SEP) for web personalization. This framework is consists of three modules of recommendation such as, original recommendation, semantic recommendation and category-based recommendation. The original recommendation will be based on context-based, collaborative based recommendation methods. The semantic recommendation will be performed using various data mining techniques such as clustering, association-rule-mining, similarity measures. In the category-based recommendation frequent keyword based recommendation will be performed using data mining techniques, which is shown in Fig. 1 given below.

A. Original Recommendation

The original recommendation consists of two components of recommendation such as context-based filtering and collaborative-based filtering. In the context-based filtering, recommendation of items will be based on extra contextual information related to item provided by the user (item-based recommendation) and the similarity of items will be calculated using Pearson correlation coefficient. In this process, the system first asks the user to provide context information. The system does not expect the user to provide the complete information. The system then asks the user to type any free keywords. These keywords are first fed into the query engine, which makes use of the context information to narrow down the search results. The free keywords are fed to the Synonym Finder engine. Thereafter, different senses of the entered keywords will be return to the user.

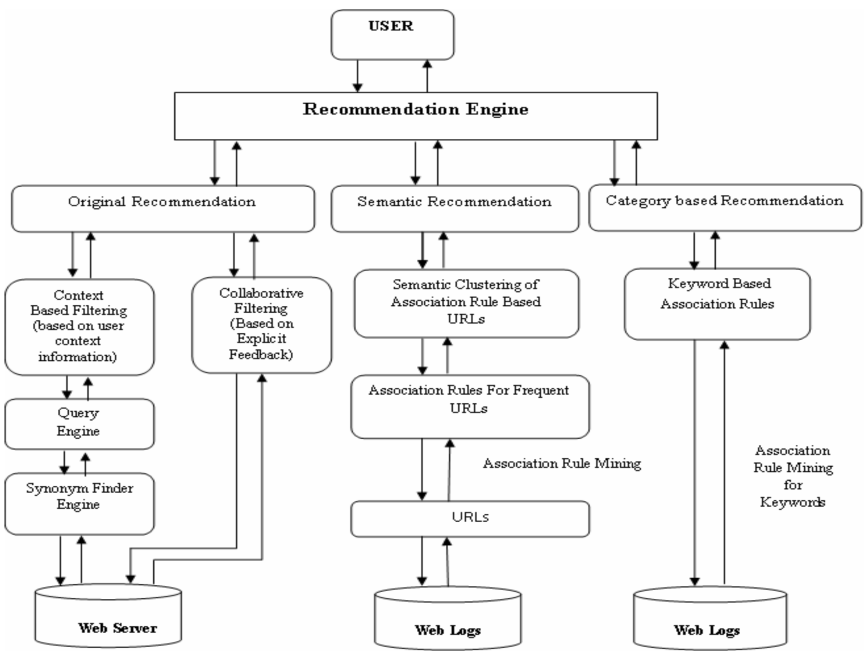

Figure 1. SEP Architecture

Here main objective is to find out the correct sense of the keyword used. All the results of the query parser and synonym finder senses are then shown to the user. Simultaneously a web service call is made to the Web Services to capture the reviews of the product shown to the user earlier. The parser searches for these senses in the web database. If a match is found then the results will be shown in that category. The advantage of using this approach is that it helps to cover the disadvantages of the User based collaborative filtering engine like lack of user ratings, false ratings etc and deliver accurate predictions to the users.

The next component of original recommendation is user based collaborative filtering. In this component, similar users will be calculated on the basis of numerical ratings of common items rated by the active users and other users of the system using Pearson's correlation coefficient. It is a way to find out correlation or similarity between two or more users.

B. Semantic Recommendation

In this approach, all URLs will be stored in database on the basis of implicit feedback. That means if user satisfies the condition for implicit feedback then URL of that page will automatically stored in database. Thereafter Performance Based Transposition Algorithm (PBTA) will be applied on to this database to generate the frequent URLs. The PBTA algorithm is improved version of Apriori algorithm [64]. Strong association rules, which satisfy the threshold value of minimum support and minimum confidence, will be generated using PBTA. Then the larger labeled clusters of similar and strong rules will be formed using Efficient Semantic Clustering (ESC). The recommendation of items will be based on to the implicit ratings. It means, if a page visited by a user included in the ratings then this URL will be matched to the Left-Hand-Side (LHS) portion of association rule in the labeled cluster, if it matches then, URLs present at the Right-Hand-Side (RHS) portion of association rule will be recommended to the user.

C. Category-based Recommendation

In this component, all keywords of respective URL will be stored in database on the basis of implicit ratings. Thereafter, Performance Based Transposition Algorithm (PBTA) will be applied on to this database to generate the frequent keywords and strong association rules based on these keywords. When user types the keywords in the query box for searching the item, these keywords will be matched with the Left-hand-Side portion of association rule. If these keywords are matching with one or more rule, then all URLs related to the keywords at the Right-hand–side portion of rules will be recommended to the user. Rather than searching for quality web pages, the users of this system would be directly taken to quality web pages matching their personal interests and preferences. The theme behind the category-based recommendation is same as semantic recommendation that incorporate the content and usage data in the recommendation process.

The proposed approach delivers quality web pages as it is not just dependent on the rating given by other users which could be deceiving at times. This approach is very useful for web page prediction.

The propose approach provides a solution of Cold-Start(new-tem and new-user) using Context-based filtering. The new-user problem arises with content-based systems. In order to make accurate recommendations, the system must first learn the user's preferences from the ratings that the user makes. The proposed approach has given the facility of context-based recommendation. In which, additional information related to item provided by the user in the query box and recommendation of items will be based on synonyms or metadata related the contextual information. The other problem is new-item problem, in which new items are added regularly to recommender systems. Collaborative systems rely solely on users' preferences to make recommendations. Therefore, until the new item is rated by a substantial number of users, the recommender system would not be able to recommend it. The proposed approach has given the facility of Context-based Recommendation in which if new item is added with website, then this item will be recommended on the basis of contextual information provided by the user.

This system is also capable to recommend the items not yet rated by any user. Therefore, it doesn't suffer from the first rater problem. This system has also the support of collaborative recommendation. Hence, the recommendation of items based on the other user with similar taste liked in the past.

This SEP is very interactive, intuitive and relatively simple to implement. The SEP is very efficient because it does not require costly training phases and it is little affected by addition of users, items, ratings, which are typically observed in large commercial applications. Besides the original recommendation, the SEP also follows the concept of Semantic Web. It is also capable of providing recommendations for items in several domains such as, Movies, Music, and Books etc.

III. ALGORITHM FOR ORIGINAL RECOMMENDATIONS

Original Recommendations (USERS, RATINGS, TYPE_OF_RECOMMENDATION)

This algorithm consists of three parameters such as, USERS, RATINGS, and TYPE_OF_RECOMMENDATION. The first parameter USERS denotes the number of users. The second parameter RATINGS denotes the ratings provided by the users on items and TYPE_OF_ RECOMMENDATION denotes the filtering scheme of recommendation select by the user.

BEGIN

1. Initialize the USERS and RATINGS with database.
2. Accept TYPE_OF_RECOMMENDATION from the user.
3. IF (TYPE_OF_RECOMMENDATION = Collaborative_based_filtering) THEN
    a. Reduce the ratings for any product between zero and one and generate the user v/s product ratings and calculate the similar users to active users.
    b. Apply fuzzy logic to generate user-user similarity matrix or item to item similarity matrix using Pearson Correlation Coefficient.
    c. Calculate the similarity of active user with other user and product liked by similar users (similarity should be more than equals to threshold value) will be recommended to the active user.

   ELSE

   IF(TYPE_OF_RECOMMENDATION= Context_based_filtering) THEN
    a. Contextual information and keywords related to particular product will be taken as input by the active user.

d. Keywords with contextual information will be input to the query engine.
e. Free keywords are loaded in the Synonym Finder Engine to find the different senses of entered keywords.
f. Synonym Finder shows all the senses of keywords on to the user screen.
g. User will select the correct sense of keyword.
h. Correct value of sense will be taken as input in the query parser.
i. Parser searches these senses in database of products.
j. If a match found then the required product will be recommended to the user.

ENDIF

END

IV. EXPERIMENTAL EVALUATION

The proposed SEP architecture consists of three modules such as, Original Recommendation, Semantic Recommendation and Category-based Recommendation. In the first module i.e., Original Recommendations, Collaborative and Context based filtering has been implemented through extensive experiments. All the experiments are performed on 2.13 Ghz Intel Pentium (P6200) Laptop machine with 3 GB main memory and Windows-XP operating system using Bookcross dataset. The dataset is described in following paragraphs.

This dataset is collected by Cai-Nicolas Ziegler in a 4-week crawl (August / September 2004) from the Book-Crossing community with kind permission from Ron Hornbaker, CTO of Humankind Systems. Contains 278,858 users (anonymized but with demographic information) providing 1,149,780 ratings (explicit / implicit) about 271,379 books [65]. The Book-Crossing dataset comprises 3 tables.

1. BX-Users: This table contains the users and all user IDs (`User-ID`) have been anonymized and map to integers. Demographic data is provided (`Location`, `Age`) if available. Otherwise, these fields contain NULL-values.

2. BX-Books: This table consists of details of books. Books are identified by their respective ISBN. Invalid ISBNs have already been removed from the dataset. Moreover, some content-based information is given (`Book-Title`, `Book-Author`, `Year-Of-Publication`, `Publisher`), obtained from Amazon Web Services. In the case of several authors, only the first Author is provided. URLs linking to cover images are also given, appearing in three different flavors (`Image-URL-S`, `Image-URL-M`, `Image-URL-L`), i.e., small, medium, large. These URLs point to the Amazon web site.

3. BX-Book-Ratings: This table contains the book rating information. Ratings (`Book-Rating`) are either explicit or implcit, expressed on a scale from 1-10 (higher values denoting higher appreciation), or implicit, expressed by 0.

The experimental work is described in subsequent paragraphs.

Firstly, User registers himself as a new user in SEP; various categories of books will be filled as choice tags along with all details are filled by the users as shown in Fig. 2.

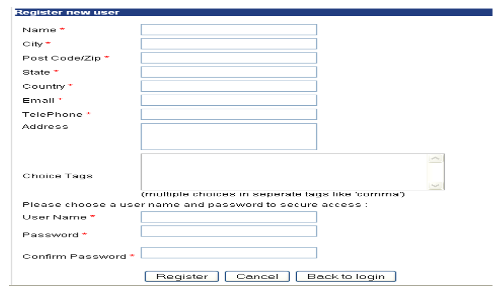

Figure 2. User registration page

There will be three categories of users are shown such as, Random Surfer, Information Seeking Visitor, Topic Oriented visitor as shown in Fig. 3 after successful login process. The user has to select Random Surfer with user name and password for original recommendation, Information Seeking Visitor for semantic recommendation and Topic Oriented Visitor for category-based recommendation.

Figure 3. User login screen with three categories of users

There are two options i.e. context based filtering and collaborative based filtering is shown in Fig. 4 with necessary details of the user with his history of previous purchase books. All the books are shown with percentage of similarity in descending order and average rating after the selection of collaborative filtering option. This similarity will be calculated on the basis of ratings given by other similar users as shown in Fig. 5. This similarity of users will be calculated using Pearson-Correlation coefficient.

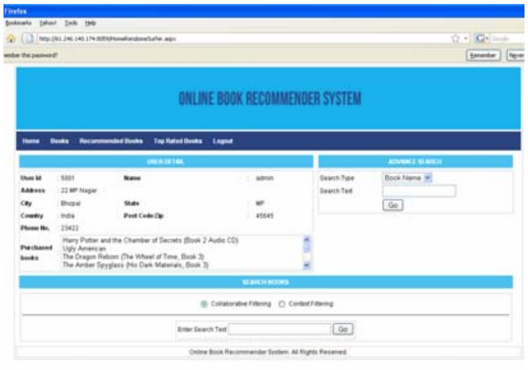

Figure 4. Categories of recommendation with user details

Figure 5. Results of collaborative filtering with similarity and average ratings according to his/her choice tags

The Top- 10 recommended books will be shown in Fig. 6 after the selection of recommended book option. These recommendations will be according to the similar user's choices to the active user on the basis of similarity value in decreasing order and average rating on each book using Pearson Correlation Coefficient.

Figure 6. Results of recommended books with collaborative filtering on the basis of similar user's profile

The Top-Rated books will be display according to the average ratings in descending order as shown in Fig. 7.

Figure 7. Results of top rated books with collaborative filtering on the basis of highest average ratings given by the user

If the user selects the context based filtering option, than the user should provide any contextual information and keyword related to the books. As the user types the contextual information and keywords in the query box, the synonyms of particular keywords will be display on the webpage using synonym web service [66]. The user will select an appropriate synonym and click on search button. All the relevant books related to synonym and contextual information will be display according to item-based similarity in descending order as shown in Fig. 8.

V. CONCLUSION AND FUTURE WORK

All existing recommender systems employ one or more of a handful of basic techniques such as content-based, collaborative, demographic, utility-based and knowledge-based etc. A survey of these techniques shows that they have complementary advantages and disadvantages. Recommender systems made a significant progress over the last decade when numerous content-based, collaborative and hybrid methods were proposed and several industrial-strength systems have been developed. In this paper, we reviewed various limitations of the current recommendation methods and discussed possible extensions that can provide better recommendation capabilities.

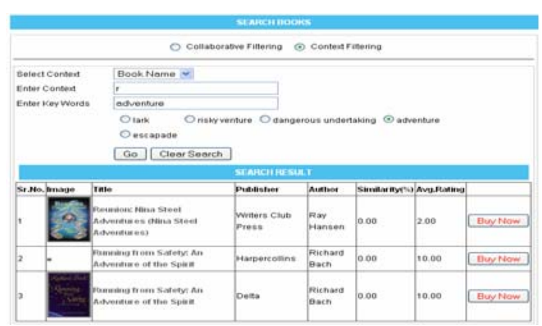

Figure 8. Results of recommended books with context-based filtering on the basis of contextual information and keywords

With these limitations, we proposed an architectural framework of SEP for recommendation of items. This framework overcomes the problem of Cold-Start problem(new-user, new-items) and sparsity because in this framework context-based, collaborative-based, semantic recommendation and category-based recommendation technique has been added. We have implemented the first component of this framework i.e. original recommendation. This recommendation consists of context-based and collaborative based recommendation. We have shown the results of these recommendations through extensive experiments. The proposed approach will be helpful to generate useful, trustful product recommendation.

In the future work, the other two modules of SEP i.e. semantic recommendation and category-based recommendation will be developed and importance of these recommendations techniques will be shown through various experiments.

REFERENCES

[1] Mahmood, T., Ricci, F. "Improving recommender systems with adaptive conversational strategies." In: C. Cattuto, G. Ruffo, F. Menczer (eds.) Hypertext, 2009, pp. 73–82.

[2] Resnick, P., Varian, H.R. Recommender systems. Communications of the ACM 40(3), 1997, pp. 56– 58.

[3] Burke, R. Hybrid web recommender systems. In: The Adaptive Web, 2007, pp. 377–408. Springer Berlin / Heidelberg .

[4] Jannach, D. "Finding preferred query relaxations in content-based recommenders." In: 3rd International IEEE Conference on Intelligent Systems, 2006, pp. 355–360.

[5] Goldberg, D., Nichols, D., Oki, B.M., Terry, D. Using collaborative filtering to weave an information tapestry. Commun. ACM 35(12), 1992, pp. 61–70.

[6] Mahmood, T., Ricci, F. "Improving recommender systems with adaptive conversational strategies." In: C. Cattuto, G. Ruffo, F. Menczer (eds.) Hypertext, 2009, pp. 73–82.

[7] McSherry, F., Mironov, I. "Differentially private recommender systems: building privacy into the net." In: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, New York, NY, USA, 2009, pp. 627–636.

[8] Anand, S.S., Mobasher, B. "Intelligent techniques for web personalization." In: Intelligent Techniques for Web Personalization, 2005, pp. 1–36., Springer.

[9] Koren, Y., Bell, R.M., Volinsky, C. "Matrix factorization techniques for recommender systems." IEEE Computer 42(8), 2005, pp. 30–37.

[10] Mladenic, D. "Text-learning and Related Intelligent Agents: A Survey." IEEE Intelligent Systems 14(4), 1999, pp. 44–54.

[11] Middleton, S.E., Shadbolt, N.R., De Roure, D.C. "Ontological User Profiling in Recommender Systems." ACM Transactions on Information Systems, 22(1), 2004, pp. 54–88.

[12] Schafer, J.B., Frankowski, D., Herlocker, J., Sen, S. "Collaborative filtering recommender systems." In: The Adaptive Web, Springer Berlin / Heidelberg, 2007, pp. 291–324.

[13] Deshpande, M., Karypis, G. Item-based top-N recommendation algorithms. ACM Transaction on Information Systems 22(1), 2004, pp.143–177.

[14] Mahmood, T., Ricci, F. "Towards learning user-adaptive state models in a conversational recommender system." In: A. Hinneburg (ed.) LWA 2007: Lernen - Wissen - Adaption, Halle, September 2007, Workshop Proceedings, 2007, pp. 373–378., Martin-Luther-University Halle-Wittenberg .

[15] Bridge, D., G¨oker, M., McGinty, L., Smyth, B. Case-based recommender systems. The Knowledge Engineering review 20(3), 2006, pp.315–320.

[16] Ricci, F., Cavada, D., Mirzadeh, N., Venturini, A. Case-based travel recommendations. In: D.R. Fesenmaier, K.Woeber, H.Werthner (eds.) Destination Recommendation Systems: Behavioural Foundations and Applications, 2006, pp. 67–93.

[17] Arazy, O., Kumar, N., Shapira, B. Improving social recommender systems. IT Professional 11(4), 2009, pp.38–44.

[18] Ben-Shimon, D., Tsikinovsky, A., Rokach, L., Meisels, A., Shani, G., Naamani, L. Recommender system from personal social networks. In: K. Wegrzyn-Wolska, P.S. Szczepaniak (eds.) AWIC, Advances in Soft Computing, vol. 43, 2007, pp. 47–55., Springer.

[19] Sinha, R.R., Swearingen, K. Comparing recommendations made by online systems and friends. In: DELOS Workshop: Personalisation and Recommender Systems in Digital Libraries, 2001.

[20] Lieberman, H. Letizia: "an Agent that Assists Web Browsing." In: Proceedings of the International Joint Conference on Artificial Intelligence, 1995, pp. 924–929.

[21] Chen, L., Sycara, K. "WebMate: A Personal Agent for Browsing and Searching." In: K.P. Sycara, M. Wooldridge (eds.) Proceedings of the 2nd International Conference on Autonomous Agents, ACM Press, New York, 1998, pp. 9–13.

[22] Sorensen, H., O'Riordan, A., O'Riordan, C. Profiling with the INFOrmer Text Filtering Agent. Journal of Universal Computer Science 3(8), 1997, pp.988–1006.

[23] Billsus, D., Pazzani, M.J. "A Hybrid User Model for News Story Classification." In: Proceedings of the Seventh International Conference on User Modeling.Banff, Canada, 2000.

[24] Collins, A.M., Loftus, E.F. A Spreading Activation Theory of Semantic Processing. Psychological Review 82(6), 1975, pp.407–428.

[25] Mooney, R.J., Roy, L. "Content-Based Book Recommending Using Learning for Text Categorization." In: Proceedings of the 5th ACM Conference on Digital Libraries, ACM Press, New York, US, San Antonio, US. 2000, pp. 195–204.

[26] Boone, G. Concept Features in "Re:Agent, an Intelligent Email Agent." In: K. Sycara, M. Wooldridge (eds.) Proceedings of the Second International Conference on Autonomous Agents, ACM Press, 1998, pp. 141–148.

[27] Bollacker, K.D., Giles, C.L. "CiteSeer: An AutonomousWeb Agent for Automatic Retrieval and Identification of Interesting Publications." In: K. Sycara, M. Wooldridge (eds.) Proceedings of the Second International Conference on Autonomous Agents, 1998, pp. 116–123., ACM Press.

[28] Mak, H., Koprinska, I., Poon, J. "INTIMATE: A Web-Based Movie Recommender Using Text Categorization." In: Proceedings of the IEEE/WIC International Conference on Web Intelligence, ISBN 0-7695-1932-6, 2003, pp. 602–605., IEEE Computer Society.

[29] Mukherjee, R., Jonsdottir, G., Sen, S., Sarathi, P. "MOVIES2GO: an Online Voting based Movie Recommender System." In: Proceedings of the Fifth International Conference on Autonomous Agents, 2001, pp. 114–115., ACM Press.

[30] Celma, O., Serra, X. FOAFing the Music: Bridging the Semantic Gap in Music Recommendation. Web Semantics 6(4), 2008, pp.250–256.

[31] Magnini, B., Strapparava, C.: "Improving User Modelling with Content-based Techniques." In: Proceedings of the 8th International Conference of User Modeling, Springer, 2001, pp. 74–83.

[32] Semeraro, G., Basile, P., de Gemmis, M., Lops, P. "User Profiles for Personalizing Digital Libraries." In: Y.L. Theng, S. Foo, D.G.H. Lian, J.C. Na (eds.) Handbook of Research on Digital Libraries: Design, Development and Impact, IGI Global,ISBN 978-159904879-6, 2009, pp. 149–158.

[33] Eirinaki, M., Vazirgiannis, M., Varlamis, I. "SEWeP: Using Site Semantics and a Taxonomy to enhance the Web Personalization Process." In: Proceedings. of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2003, pp. 99–108.

[34] Jolliffe, I.T. Principal Component Analysis. Springer, 2002.

[35] Hartigan, J.A. 1975. Clustering Algorithms (Probability & Mathematical Statistics). John Wiley & Sons Inc.

[36] Felfernig, A. "Koba4ms: Selling complex products and services using knowledge-based recommender technologies." In: CEC '05: Proceedings of the Seventh IEEE International Conference on E-Commerce Technology (CEC'05), pp. 92–100., IEEE Computer Society .

[37] Felfernig, A., Kiener, A. "Knowledge-based interactive selling of financial services with advisor." In: Proc. of the National Conference on Artificial Intelligence, 2005, pp. 1475–1482.

[38] Tartakovski, A., Schaaf, M., Bergmann, R. Retrieval and configuration of life insurance policies. In: Lecture Notes in Computer Science, Case-Based Reasoning Research and Development, 2005, pp. 552–565., Springer Berlin / Heidelberg

[39] Lee, D.H., Brusilovsky, P. "Fighting information overflow with personalized comprehensive information access: A proactive job recommender." In: ICAS '07: Proceedings of the Third International Conference on Autonomic and Autonomous Systems, IEEE Computer Society, Washington, DC, USA, 2007, p. 21.

[40] Malinowski, J., Keim, T., Wendt, O., Weitzel, T. "Matching people and jobs: A bilateral recommendation approach." In: HICSS '06: Proceedings of the 39th Annual Hawaii International Conference on System Sciences, 2006, pp. 137c–137c., IEEE Computer Society.

[41] Paterek, A. "Improving regularized singular value decomposition for collaborative filtering." In: Proceedings. of the KDD Cup and Workshop,2007.

[42] Tak´acs, G., Pil´aszy, I., N´emeth, B., Tikk, D. "On the gravity recommendation system." In: Proceedings of the KDD Cup and Workshop 2007 (KDD 2007), 2007, pp. 22–30.

[43] Hayes, C., Cunningham, P. "Smart radio: Building music radio on the fly." In: Proceedings of Expert Systems, 2000, Cambridge, UK.

[44] Hijikata, Y., Iwahama, K., Nishida, S. Content-based music filtering system with editable user profile. In: SAC '06: Proceedings of the 2006 ACM symposium on Applied computing, 2006, pp. 1050–1057.

[45] Chen, H.C., Chen, A.L.P. "A music recommendation system based on music data grouping and user interests." In: CIKM '01: Proceedings of the tenth international conference on Information and knowledge management, 2001, pp. 231–238.

[46] Billsus, D., Pazzani, M.J. User modeling for adaptive news access. User Modeling and User- Adapted Interaction 10(2-3), 2006, pp.147–180.

[47] Mobasher, B., Cooley, R., Srivastava, J. Automatic personalization based on web usage mining. Communications of the ACM 43(8), 2000, pp.142–151.

[48] Maidel, V., Shoval, P., Shapira, B., Taieb-Maimon, M. "Evaluation of an ontology-content based filtering method for a personalized newspaper." In: RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems, 2008, pp. 91–98.

[49] Das, A.S., Datar, M., Garg, A., Rajaram, S. Google news personalization: scalable online collaborative filtering. In: WWW '07: Proceedings of the 16th international conference on World Wide Web, ACM, New York, NY, USA, 2007, pp. 271–280.

[50] Burke, R. Knowledge-based recommender systems. In: Encyclopedia of Library and Information Systems. Marcel Dekker, 2000.

[51] Viappiani, P., Pu, P., Faltings, B. Conversational recommenders with adaptive suggestions. In: *RecSys '07: Proceedings of the 2007 ACM conference on Recommender systems,* 2007, pp. 89–96.

[52] Zhang, J., Pu, P. A comparative study of compound critique generation in conversational recommender systems. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 4018 NCS, 2006, pp. 234–243., Springer Verlag, Heidelberg, D-69121, Germany.

[53] S.E. Middleton, H. Alani, and D.C. Roure, Exploiting Synergy Between Ontologies and Recommender Systems," Semantic Web Workshop, vol. 55, 2002, pp. 41-50.

[54] Castro-Herrera, C., Duan, C., Cleland-Huang, J., Mobasher, B.: "Using data mining and recommender systems to facilitate large-scale, open, and inclusive requirements elicitation processes." In: Proceedings of the 2008 16th IEEE International Requirements Engineering Conference, IEEE Computer Society, Washington, DC, USA, 2008, pp. 165–168.

[55] Felfernig, A., Kiener, A. "Knowledge-based interactive selling of financial services with fsadvisor." In: Proceedings of the National Conference on Artificial Intelligence, 2005, pp. 1475–1482.

[56] Ricci, F. Travel recommender systems. In: IEEE Intelligent Systems, 2002, pp. 55–57.

[57] Mahmood, T., Ricci, F., Venturini, A., Hpken, W. Adaptive recommender systems for travel planning. In: P. O'Connor, W. Hpken, U. Gretzel (eds.) Information and Communication Technologies in Tourism, 2008, pp. 1–11., Springer .

[58] Y. Blanco-Fernández, J.J. Pazos-Arias, A. Gil-Solla, et. Al. A flexible semantic inference methodology to reason about user preferences in knowledge-based recommender systems Knowledge-Based Systems, vol. 21, 2008.

[59] Buono, P., Costabile, M.F., Guida, T., Piccinno, A. "Integrating user data and collaborative filtering in a web recommendation system." In: Lecture Notes in Computer Science: Hypermedia: Openness, Structural Awareness, and Adaptivity, SpringerLink, 2002, pp. 192–196.

[60] Li, J., Zaiane, O.R. Combining usage, content, and structure data to improve web site recommendation. Lecture Notes in Computer Science : E-Commerce and Web Technologies, 2004, pp. 305–315.

[61] Anand, S.S., Mobasher, B. Introduction to intelligent techniques for web personalization. ACM Trans. Interet Technol. 7(4), 2007, pp.18.

[62] Paterek, A. "Improving regularized singular value decomposition for collaborative filtering." In: *Proceedings. of the KDD Cup and Workshop,2007*.

[63] Sanjeev Kumar Sharma, Ugrasen Suman, "A Semantic enhanced data mining framework for web Personalization," In: Proceedings of International conference on data analysis, data quality and metadata management (DAMD-2010), Singapore, 2010, pp. 49-57.

[64] Sanjeev Kumar Sharma, Ugrasen Suman, "A Performance based Transposition algorithm for Frequent itemsets Generation," International Journal of Data Engineering (IJDE), Vol(2), 2011, pp. 53-61.

[65] http://www.informatik.uni-freiburg.de/~cziegler/BX/.

[66] http://abbreviation.com/services/v1/syno.aspx

Author Profile

Mr. Sanjeev Kumar Sharma has received Master of Computer Applications (MCA) from Indira Gandhi National Open University Delhi. He has also done M-Tech. in computer Science and Engineering from Allahabad Agriculture institute Deemed University Allahabad. He is presently a PhD Scholar in Devi-Ahilya University Indore.

Dr. Ugrasen Suman has received his Master degree from Jabalpur University and PhD degree in Computer Science from Devi Ahilya University Indore, India. He is presently a Reader at Devi Ahilya University. He is having around 10 years teaching and research experience. His areas of research are Software Engineering, Knowledge Management & Web Mining, Information System Design, Web Services, and Cloud Computing. He is currently guiding Six PhD scholars and Eight PG research scholars. He has published more than 30 research papers and also guided more than 30 PG projects. He was being a member of ACM-SIGSE. He is also working on a UGC-SAP research project of Data Mining and Software Engineer