AN EFFECTIVE FRAMEWORK FOR GENERATING RECOMMENDATIONS

Vandana Srivastava¹, Dr S Q Abbas², Dr Mohammad Husain³
¹Research Scholar, IT College, Lucknow
²AIMT, Lucknow
³JIT, Barabanki

Abstract- The Internet, which brought the most innovative improvement on information society, web recommendation systems based on web usage mining try to mine user’s behavior patterns from web access logs, and recommend pages or suggestions to the user by matching the user’s browsing behavior with the mined historical behavior patterns. In this paper we propose a recommendation framework that considers different application status and various contexts of each user. We successfully implemented the proposed framework and show how this system can improve the overall quality of web recommendations.

I. INTRODUCTION

The volume of information available on the internet is increasing rapidly with the explosive growth of the World Wide Web and the advent of e-Commerce. While users are provided with more information and service options, it has become more difficult for them to find the “right” or “interesting” information, the problem commonly known as information overload. Recommender systems are alternative, user-centric, promising approaches to tackle the problem of information overload by adapting the content and structure of websites to the needs of the users by taking advantage of the knowledge acquired from the analysis of the users’ access behaviors. They can be generally defined as systems that guide users toward interesting or useful objects in a large space of possible options. In recent years there has been an increasing interest in applying web usage mining techniques to build web recommender systems. Web usage recommender systems take web server access logs as input, and make use of data mining techniques such as association rule and clustering to extract implicit, and potentially useful navigational patterns, which are then used to provide recommendations. Web server access logs record user browsing history, which contains plenty of hidden information regarding users and their navigation. They could, therefore, be a good alternative to the explicit user rating or feedback in deriving user models. Unlike traditional techniques, which mainly recommend a set of items deemed to be of interest to the user base their decisions on user ratings on different items or other explicit feedbacks provided by the user. These techniques discover user preferences from their implicit feedbacks, namely the web pages they have visited. Clustering and collaborative filtering approaches are ready to incorporate both binary and non-binary weights of pages, although binary weights are usually used for computing efficiency. Association Rule (AR) mining can lead to higher recommendation precision, and are easy to scale to large datasets, but how to incorporate page weight into the AR models has not been explored in previous studies.

To help users to choose a proper service among the available services on the Internet, many brokering web sites for consultancy services such as consultant web portals and search engines have been developed. The users can use the brokering web sites as starting points and find appropriate services using them. This improvement allows the users to access information about the services much easier than before, and the consultant providers to save more lives. The brokering web sites, however, showed their limitation that more sophisticated mechanism is required in the domain of services. The most of the users who does not have any knowledge about servicers or any idea what is wrong with their choice cannot find out proper services. What they need is not organized information about services, but a professional guideline to the most appropriate services for a specific user. Therefore, recommendation systems for better services are proposed. Recommendation system for the consultancy is a web site that recommends services or provides useful information to the users considering. Service Provider Recommendation System is an example of well proposed service recommendation system.

II. Recommendation System

Recommendation system can be measured as a precious expansion of usual information systems used in industries such as travel and generosity. However, recommendation systems have mathematical roots and are more similar to artificial intelligence than any other IT regulation. A recommendation system learns from a customer’s behavior and recommends a product in which users may be concerned. At the heart of recommendation systems are machine-learning constructs. Leading e-commerce players use recommendation engines that separate users’ past purchase histories to recommend products such as magazine articles, books, goods, etc. Online companies that influence recommendation systems can increase sales by 8 percent to 12 percent. Companies that succeed with recommendation engines are those that can quickly and proficiently turn huge amounts of data into actionable information.

III. Analysis of a Recommendation Engine

The key component of a recommendation system is data. This data may be garnered by a array of means such as customer ratings of products, feedback from purchasers, etc. This data will provide as the source for recommendations to users. After data collection, recommendation systems use machine-learning algorithms to find similarities and affinities between products and users. Recommender logic programs are then used to build suggestions for specific user profiles. This procedure of filtering the input data and giving recommendations to users is also known as “collaborative filtering.”

Along with collaborative filtering, recommendation systems also use other machine-learning techniques such as clustering and classification of data. Clustering is a technique which is used to bunch large amounts of data together into
similar categories. It is also used to see data patterns and provide huge amounts of data simpler to manage. Many search engines use clustering to group results for similar search terms. Classification is a technique used to choose whether new input or a search term matches a previously observed pattern. It is also used to identify doubtful network activity. Yahoo! Mail8 uses classification to decide if an incoming message is spam. Image sharing sites like Picasa9 use classification techniques to determine whether photos contain human faces. They then offer recommendations of people that are identified in the user contacts list.

IV. Approaches to Generate Recommendations

Personalization systems fall into three basic categories: Rule-based systems, content-filtering systems, and collaborative filtering systems.

A. Rule-Based Personalization Systems

Rely on manually or automatically generated decision rules that are used to recommend items to users. Many existing e-commerce Web sites that employ personalization or recommendation technologies use manual rule-based systems. Such systems allow Web site administrators to specify rules, often based on demographic, psychographic, or other personal characteristics of users. In some cases, the rules may be highly domain dependent and reflect particular business objectives of the Web site. The rules are used to affect the content served to a user whose profile satisfies one or more rule conditions. Like most rule-based systems, this type of personalization relies heavily on knowledge engineering by system designers to construct a rule base in accordance to the specific characteristics of the domain or market research. The user profiles are generally obtained through precise communications with users. Some research has focused on machine learning techniques for classifying users into one of several categories based on their demographic attributes, and therefore, automatically derive decision rules that can be used for personalization.

The primary drawbacks of rule-based filtering techniques, in addition to the usual knowledge engineering bottleneck problem, originate from the methods used for the generation of user profiles. The input is usually the subjective description of users or their interests by the users themselves, and thus is level to unfairness. Furthermore, the profiles are often static, and thus the system performance degrades over time as the profiles age.

B. Content-Based Filtering Systems

Also called item-to-item correlation. This kind of recommendation mainly relies on association rules pattern between the goods. Association may be based on co-purchase data, co-visited data, content similarities, preference by common customers, or other measures. In Content-based filtering systems, a user profile represent the content descriptions of items in which that user has previously expressed interest. The content descriptions of items are represented by a set of features or attributes that characterise that item. The recommendation generation task in such systems usually involves the comparison of extracted features from unseen or unrated items with content descriptions in the user profile. Items that are considered sufficiently similar to the user profile are recommended to the user.

In most content-based filtering systems, particularly those used on the Web and in e-commerce applications, the content descriptions are textual features extracted from Web pages or product descriptions. As such, these systems often rely on well-known document modeling techniques with roots in information retrieval and information filtering research. Both user profiles, as well as, items themselves, as represented as weighted term vectors. Predictions of user interest in a particular item can be derived based on the computation of vector similarities or using probabilistic approaches such as Bayesian classification. Furthermore, in contrast with approaches based on collaborative filtering, the profiles are individual in nature, built only from features associated with items previously seen or rated by the active user.

The primary drawback of content-based filtering systems is their tendency to over specialize the item selection since profiles are solely based on the user’s previous rating of items. User studies have shown that users find online recommenders most useful when they recommend unexpected items, suggesting that using content similarity alone may result in missing important “pragmatic” relationships among Web objects such as their common or complementary utility in the context of a particular task. Furthermore, content-based filtering requires that items can be represented successfully using extracted textual features which are not always practical given the heterogeneous nature of Web data.

C. Collaborative Filtering Systems

Collaborative filtering approaches, also called customer-to-customer correlation. This is based on other customers’ opinions. It can recommend the products to the special customers by using their similarity. It can discover the new interested content for the customer. Collaborative filtering distinguishes the customers’ neighbors from history-information, and the information, which the customers are possibly liked through analysing these neighbors. It is different from content-based filtering. What is recommended by collaborative filtering is based on other customers’ preferences, not similar products that the customers liked in the past. Only the similarities are calculated between the customers, but not calculated between the products.

There are three steps to provide the suggestion for the new customers by Collaborative filtering:

- A new customer file is established by selecting items involved in the website.
- The new file is compared with other customers’ files to find similar files.
- For the products that new customer has not listed, those may be forecasted by using the similar customers’ files.

The second one is based on the customers’ preference in the past and content analysis, which is difficult to apply to content that, is hard to analyse and classify. The third one is to find other customers that have shown similar tendencies to the given customers and recommend what the past customers have liked, which is difficult to apply to content which have few or no preference data from the customers.

Collaborative filtering has tried to address some of the shortcomings of other approaches mentioned above. Particularly, in the context of e-commerce, recommender systems based on collaborative filtering have achieved notable successes. These techniques generally involve matching the ratings of a current user for objects (e.g., movies or products).
with those of similar users (nearest neighbors) in order to produce recommendations for objects not yet rated or seen by an active user.

Collaborative filtering usually performs best when explicit non-binary user ratings for similar objects are available. In many Web sites, however, it may be desirable to integrate the personalization actions throughout the site involving different types of objects, including navigational and content pages, as well as implicit product-oriented user events such as shopping cart changes, or product information requests. A number of optimization strategies have been proposed and employed to remedy these shortcomings. These strategies include similarity indexing and dimensionality reduction to reduce real-time search costs and remedy the sparsely problems, as well as offline clustering of user records, allowing the online component of the system to search only within a matching cluster. A model-based variant of collaborative filtering is known as item-based collaborative filtering in which, starting from the same user-rating profile databases, an item-item similarity matrix is built offline, and used in the prediction phase to generate recommendations. Rather than basing item similarity on content descriptions of the items, similarity between items is based on user ratings of these items. Each item is represented by a vector, and the similarities are computed using metrics such as cosine similarity and correlation-based similarity. The recommendation process predicts the rating for items not previously seen or rated by an active user using a weighted sum of the ratings, by that user, of items in the item neighborhood of the target item. Evaluation of the item-based collaborative filtering approach has shown that item-based collaborative filtering can provide recommendations that are, in general, of similar quality when compared to memory-based collaborative approach. Most data mining approaches to personalization can be viewed as extensions of collaborative filtering. In these approaches the pattern discovery algorithms take as input the historical rating or navigational profiles of past users and generate aggregate user models. The user models, in turn, can be used, in conjunction with the profile of an active user, to predict future user behavior or generate recommendations.

V. SERVICE ORIENTED ARCHITECTURE (SOA) TO IMPLEMENT FRAMEWORK FOR PERSONALIZED RECOMMENDATIONS

Software development has undergone various stages of paradigms - the Service-oriented Architecture (SOA), as the next generation software architecture and through the utilization of the web service, XML and other related technologies provides viable working solution to implement dynamic e-business. In this paper we develop a framework which implementation based on SOA.

A service-oriented architecture is essentially a collection of services, among which the communication can involve either simple data passing or it could involve two or more services coordinating some activity, requiring means of connecting services to each other. The first service-oriented architecture in the past was with the use DCOM or Object Request Brokers (ORBs) based on the CORBA specification.

To understand service-oriented architecture must begin with a clear understanding of the term service. A service is a function that is well defined, self-contained, and does not depend on the context or state of other services. The technology of Web services is the most likely connection


technology of service-oriented architectures. Web services essentially use XML to create a robust connection. A service consumer sending a service request message to a service provider. The service provider returns a response message to the service consumer. The request and subsequent response connections are defined in some way that is understandable to both the service consumer and service provider. How those connections are defined is explained in Web Services explained. A service provider can also be a service consumer.

As a distributed software model, an SOA is usually comprised of three primary parties: Producer (of services), Consumer (of services), Directory (of services). Web Services are considered an example of Service Oriented Architecture. Service Networks take on the properties of an SOA.

VI. RELATED WORK

Bachus et al. have been holding a U.S. patent that is Healthcare Provider Recommendation System (HPRS). A consumer who experienced good of a healthcare service can register the service’s provider on this system. Information about registered and accumulated services is provided to the other users; a user who is willing to be treated can query to the system using provider’s contexts such as location, specialty, and reputation. This system can encourage information sharing between consumers, and rating healthcare services may be honest and actually helpful for consumers’ point of view. However, information about the services may not be possible to show professionalism, because services are registered by laymen. Besides, the system cannot consider user’s contexts, especially the health status, so novice users may not receive successful recommendation results. ABC Homeopathy is a healthcare web site that proposes remedies or medicines according to the users’ symptoms. Using this web site, a user can choose multiple symptoms of each body part, and retrieve information about remedies for the selected symptoms. We can regard that this system considers health status to recommend remedies. However, results are too simple and limited to remedies only. Moreover, it does not handle enough health status that is required for fine recommendation, and still, it is not useful when users could not be aware of their specific symptoms.

Bahram Amini, Roliana Ibrahim, Mohd Shahizan Othman, faculty of Computer Science and Information Systems, University of Teknologi Malaysia, “A Framework for Personalized Information Integration in Higher Education Institutes”, International Journal of Computer Applications(0975-8887) Vol 23-No.4, June 2011, a service-oriented framework which augments recommendation approach with components of semantic-based information integration and provides interactive and contextual based information integration for decision makers in Higher Education Institutes. The underlying semantic web technology facilitates on-demand integration of information from internal sources as well as the Web and provides web service discovery and invocation for effective information analysis. In addition, the framework enables the users to analyze instances of student’s information and to receive recommendation of new information sources as well as appropriate analytical services based on the students’ status. Service orientation paradigm provides dynamic and flexible means of communication for service interoperability among the framework components.
System Service Recommendation Framework (SSRF) is a computerized system that recommends suitable services to service consumers based on their various interest. In other words, the framework acts as a mediator for business or non-business interactions between system service providers and consumers. A service provider is a Network node that provides a service interface for a software asset that manages a specific set of tasks. A service provider node can represent the services of a business entity or it can simply represent the service interface for a reusable subsystem. Therefore, it is an essential functionality for career web software such as e-career portals or search engines for system services.

For more personalized recommendation of the services, SSRF applies user’s interest to its recommendation process. Interest is the information about user’s current states or conditions, and it is the most important key to determine what specific services are suitable for the user. However, to use interest without any technical obstacle, the interest must be measurable and standardized. As a mediator, SSRF manages complex interactions between system service providers, consumers, and system administrators. The system service providers such as Professor, Director, Engineers, Technical advisor etc. can describe and register their own services on SSRF. Even if service not registered in SSRF, but SSRF provided service on-demand to incorporate new sources of data is required. Then, multiple recommendation mechanisms that are developed by the system administrators eventually search and recommend those registered or unregistered services for the users. Users can retrieve information about recommended services and evaluate them. The web portal system provides various services and information about system and also acts as an interface of SSRF for service consumers. The web portal actually triggers a recommendation process automatically delivering users’ interest to SSRF; SSRF performs recommendation process with given users’ contexts, analysis of data in runtime done by runtime analysis tools and sends recommended results back to the portal.

VIII. Framework Architecture

We designed a flexible architecture of SSRF considering extensibility and scalability of the framework. Because, a brand-new type of service and interest can emerge at any time after the system is published, SSRF should require less effort to adapt those changes. Also, SSRF must be able to handle large amount of services and consumers. A consumer should be able to receive recommended results with high quality and low delay even if there are many services or requests from other consumers. To meet requirements above, we adopted SOA (Service Oriented Architecture) design paradigm to SSRF. System services can be implemented using the Web Services technology and registered easily at runtime. Also, core logics for the recommendation can be realized to web services. For instance, we can imagine that there are number of web services available and a recommendation web service that gathers and arranges the services is deployed on the system. Likewise, there are recommendation web services that are in charge of their own categories and all the results from them are reorganized for users.

The entire architecture of SSRF is consists of three types of modules: façade, core logic, and data access object. Façade modules are outer interfaces of SSRF. Each of them is connected to its core logic that performs the main functionality of the system. Core logic modules may need to use service repository to access information about consultancy services; data access object modules are wholly responsible for all transactions to the System Service Repository and provide handy interfaces to it. There are three façade modules in SSRF: interfaces for service providers, interfaces for users, and interfaces for administrator. Each of them is in charge of SSRF operations that aid each system user properly. Moreover, they can protect the framework from incorrect operations at the same time. Interface module for consultancy providers interacts with System Service Manager, interface for administrator is on Web Service Manager, and users’ interface deals with Web Services Pool that is a set of various web services for system service recommendation. The data access object is a well-known module that is effective for handling database. SSRF puts all information about system services into the database called System Service Repository, if data is not available in database than it provided by searching data and store in database called Data provided on-demand and all transactions to the repository must be performed through the data access objects only. It is such a common design pattern for handle the database to prevent damages of data from any unwilling database operations. SSRF has two core logic modules for managing system: the Web Service Manager and the System Service Manager, and numerous web services for the recommendation logics in the Web Services Pool. The Web Service Manager manages the Web Services Pool of SSRF enabling recommendation web services to be deployed and managed at runtime by the system administrators. The only parameter that is required for deploying a new recommendation web service is URL (Uniform Resource Location) of WSDL (Web Service Description Language) document of a new web service, and the system administrator can manage deployed web services using their URL as a primary key.

The System Service Manager provides various functionalities to manage the System Service Repository; registering a new system service, modifying properties of the services, inquiring recommendation statistics, activating or deactivating the service, deregistering the services, and so on. This framework can also analysis the result through runtime analysis tools. This tool analyze result by various software like Ms-Excel, SPSS etc. before providing result to user. So that user can choose the best service.

IX. IMPLEMENTATION

We have implemented the developed framework in case study of Consultancy Services (CS).

To evaluate functionality and feasibility of SSRF, we present an example of Consultancy Service recommendation process for service consumers that may occupy most of...
transactions of the framework. A recommendation process starts, as mentioned before, from the e-Career Web Portal system. The Web Portal system requests consultancy service recommendation for a consumer sending consumer’s interest to Consultancy Service Recommendation Framework (CSRF). A process for the service consumers automatically starts during when they are using the e-Career Web Portal System, and they need some consultancy from consultancy services. To start recommendation process, the e-Career Web Portal triggers a recommendation request to CSRF using the CSRF Invoker module. As an input of CSRF, a request message contains user’s contexts including interest. Passing through the façade module, the request is propagated to multiple recommendation web services in the Web Services Pool simultaneously. CSRF receives the request, and passes it to multiple recommendation web services in the Web Services Pool. After simultaneous tasks of those web services, the framework gathers and rearranges results from them and analyze the result by runtime analysis tools. Finally CSRF returns a composed recommendation result back to the Web Portal system and the Web Portal system displays information about recommended services with analysis of result to a consumer. From the implementation result above, we have confirmed that a recommendation process of CSRF works pretty well, and that a result for consultancy recommendation is valuable and reasonable to the consumers. However, more certain evaluation of the framework would be a statistical investigation on service consumers’ satisfaction about recommendation results. Further evaluation of CSRF, which was difficult without large experimental group and long observation, will be possible when e-Career Web Portal is used by service consumers widely and frequently.

X. CONCLUSION

In this paper, we suggested a personalized consultancy service recommendation framework that considers consumers’ interest to find and analyze satisfactory services for them. Our framework gathers information about service consumer’s interest and calculates service similarities between consumer and consultancy services automatically and runtime analysis of that service. Based on these similarities of each consumer, the framework arranges and recommends proper consultancy services. Also, we implemented CSRF and evaluated its functionality and feasibility. Although the assessment was not fully assured to demonstrate all approaches of this paper, we concluded that our framework is relatively adequate to provide enhanced consultancy service recommendation to apprentice users.

REFERENCES

[37] www.service-architecture.com
[38] http://www.servicearchitecture.com/webservices/articles/web_services_explained.html