INFORMATION RETRIEVAL MODEL IN SELF DIRECTED E-LEARNING USING DYNAMIC SEMANTIC NETWORK

***Ms.LathaKurian**
Assistant Professor
Sacred Hearts Girls First Grade college,
Bangalore.

ABSTRACT

Self- directed e-learning focuses on the independent learner one who engages in education at his own space free from curricular obligation. E-learning systems accessible through the Internet have great potential to improve education through extending educational opportunities for those who cannot use the time and place bound traditional courses and offering new services and functions that enhance the traditional classroom. Traditional information retrieval techniques have become inadequate, since significant differences exist between students, such as their learning rate, personal interest and domain knowledge. To alleviate this problem, personalization becomes a popular remedy to customize the web environment for the learners. Therefore the goal in e-learning is to be established as "turning learners into better learners".

The learner's interest model is constructed based on the dynamic semantic network, which represents the learner's level of interest in the material currently being examined by the e-learner. Thus dynamic user profiles are maintained based on which, the information retrieval model is constructed. Considering the long term interest of the learner this retrieval model takes the initiative to push any newly entered information to the learner if the learning material is judged as being relevant to the learner.

Keywords: E-learning, web usage mining, Information Retrieval.

I. INTRODUCTION

This study seeks to build new knowledge about the user's access interests in self-directed e- learning. The study utilizes web mining methods to extract patterns of user access interests. The purpose of the study is to determine if patterns of user access interests can be extracted and to associate these patterns extracted, with other attributes to recommend information items that are likely to be of interest to the user. In accordance to the characteristic and requirement of different users, personalization services classify the information resources and further provide and recommend users, the information according to the demands and preferences of the user. This enables the user to get better services. It can shorten the searching time, provides better services, and let learners study more useful and suitable resources.

II. Background of Study

The introduction of World Wide Web has had a profound impact on education, reducing the necessity of a learner and a teacher to share the same physical space and creating an entirely new form of knowledge delivery. With an ever increasing number of Internet users and

websites, on line learning, training and on-line educational multimedia – all generally referred to as 'e-learning' are becoming increasingly prevalent. E-learning systems accessible through the Internet have great potential to improve education through extending educational opportunities for those who cannot use the time and place bound traditional courses and through offering new services and functions that enhance the traditional classroom. E-learning systems provide multiple ways of learning (self-placed, collaborative, and tutorial) within a common application as well as providing the potential for rich media and complex interactions. In reality a substantial portion of available information is stored in text databases or document databases. Text databases are rapidly growing due to the increasing amount of information available in electronic form, such as electronic publications, various kinds of electronic documents, e-mail and the World Wide Web. Nowadays most of the information in government, industry, business is stored electronically in the form of text databases. The data stored in most text databases are semi structured data; it contains structured and unstructured text components. The information retrieval techniques such as text indexing methods have been developed to handle unstructured documents. Traditional information retrieval techniques become inadequate for the increasingly vast amounts of text data. Typically only a small fraction of the many available documents will be relevant to a given individual user. Therefore users need some tools to compare different documents, rank the importance and relevance of the documents or find patterns and trends across multiple documents. Therefore text mining has become an increasingly popular and essential theme in data mining.

Information retrieval is concerned with the organization and retrieval of information from a large number of text based documents. Text retrieval methods fall into two broad categories, document selection problem or a document ranking problem. In document ranking methods, the query is used to rank all the documents in the order of the relevance. The vector space method is the most popular retrieval method used. In Web-Based learning it is very difficult and time consuming for educators to thoroughly track and access all activities performed by all learners. On the learners side it would be very useful if the system could automatically guide the learners' activities and intelligently recommend on-line activities or resources that would favor and improve the learning. In the field of electronic commerce a significant research effort has been made to devise methods to take advantage of customers' access and purchase behaviors in order to enhance the purchasing experience and customer satisfaction by user profiling and smart recommendations. Amazon.com suggests books to purchase related to a current purchase based on preference information and similar user purchases. Moviefinder.com uses collaborative filtering which predicts a person's preferences as a linear weighted combination of other people's preferences. Web based course delivery systems rely on web servers to provide access to resources and applications. Every single request that a web server receives is recorded in an access log mainly registering the origin of request, a time stamp and the course requested. The web log provides a raw trace of the learners' navigation and activities on the site. These log entries can be processed and valuable patterns can be extracted and can be used to enhance the learning system. A log entry is automatically

added each time a request for a resource reaches the web server. The entries of all users are mixed in the log, but in e-learning the users need to login to the system as registered learners. In e-learning a learning session can span many access sessions whereas in e-commerce a purchase transaction is defined starting from an initial access to the web site to a purchasing or order usually in a very short time frame. In e-commerce sites the goal is to encourage the customers to buy new products but in e-learning the goal is difficult to quantify and subjective. The goals of e-commerce sites i.e., 'turning visitors into purchasers' are different from goals of e-learning i.e., 'turning learners into effective better learners'.

Since the Web is a large collection of semi-structured and structured information sources, Web users often suffer from information overload. To alleviate this problem, personalization becomes a popular remedy to customize the Web environment for users. Among all personalization tools, recommendation systems are the most employed tools in e-commerce businesses. Recommendation systems are usually used to help the customers to locate the products they would like to purchase. Recommender systems, recommendation systems, recommendation engines, recommendationframeworks, recommendation platforms or simply recommender form or work from a specific type of information filtering systems technique that attempts to recommend information items (movies, TV program, music, books, Web pages, scientific literature such as research papers etc.) that are likely to be of interest to the user. Recommender systems are a useful alternative to search algorithms as they help users discover items they might not have found by themselves. Interestingly enough, recommender systems are often implemented using search engines indexing non-traditional data. A recommender system compares a user profile to some reference characteristics, and seeks to predict the 'rating' that a user would give to an item they had not yet considered. These characteristics may be from the information item (the content-based approach) or the user's social environment (the collaborative filtering approach). Collaborative filtering is based on the idea that the active user is more likely to prefer items that like-minded people selected. Collaborative filtering based techniques compare those records with the historical records of other users in order to find out the top k users who have similar tastes or interests. The mapping of a visitor record to its neighborhood could be based on similarity in rating of items, access to similar contents or pages or purchase of similar items. Other techniques include association rules discovery, Bayesian networks and Horting. All these approaches for recommender systems must rely on the usage history of users and focus on the current demand of the user. The important drawback of all these systems is that the items or pages which are added to site a recently cannot be found. For these new items or pages there are no accessed records, no rating from users and still no relationship are found by algorithms, all of which make it difficult for people to discover them. This is referred as 'new item problem'. Cold start is a potential problem in computer-based information system which involves a degree of automated data modeling. Specifically, it concerns the issue that the system cannot draw any inferences for users or items about which it has not yet gathered sufficient information. The cold start problem is most prevalent in recommender systems. In the

content-based approach, the system must be capable of matching the characteristics of an item against relevant features in the user's profile. In order to do this, it must first construct a sufficiently-detailed model of the user's tastes and preferences through preference elicitation. This may be done either explicitly (by querying the user) or implicitly (by observing the user's behavior). In both cases, the cold start problem would imply that the user has to dedicate an amount of effort using the system in its 'dumb' state – contributing to the construction of their user profile – before the system can start providing any intelligent recommendations.

In the collaborative filtering approach, the recommender system would identify users who share the same preferences (e.g. rating patterns) with the active user, and propose items which the like-minded users favored (and the active user has not yet seen). Due to the cold start problem, this approach would fail to consider items which no-one in the community has rated previously. In recommender systems, the cold start problem is often reduced by adopting a hybrid approach between content-based matching and collaborative filtering. New items (which have not yet received any ratings from the community) would be assigned a rating automatically, based on the ratings assigned by the community to other similar items. Item similarity would be determined according to the items' content-based characteristics. The construction of the user's profile may be automated by integrating information from other user activities, such as browsing histories. If, for example, a user has been reading information about a particular music artist from a media portal, then the associated recommender system would automatically propose that artist's releases when the user visits the music store. This can be overcome by extracting semantic features from items or pages content. The item or page can be recommended to the user that seems to have the similar tastes or interests matching with the semantic feature of this item. In general the recommendations are based on an information item (the content based approach) or the user's social environment (the collaborative filtering approach). A personalization approach to recommender system is also referred. Four main approaches to recommendations are given as personalized recommendation, social recommendation, item recommendation and combination of three approaches. The personalized recommendation, recommend things based on the individual's past behavior. The social recommendation recommends things based on the past behavior of similar users. The item recommendation recommends things based on the item itself. Recommender systems are characterized with 'individualized' that separate them from search engines. Most of the e-learning environments used in the educational field today are still delivering the same educational resources and services in the same way to different learners. There is increasing need for personalization in e-learning systems to make these systems deliver the most appropriate content to learners according to their interests and needs.

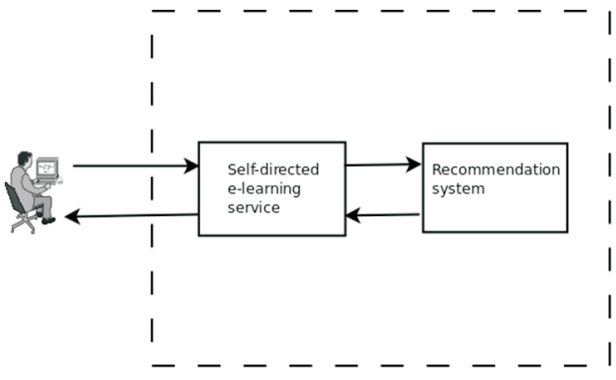

Figure 1: Relationship between services and users

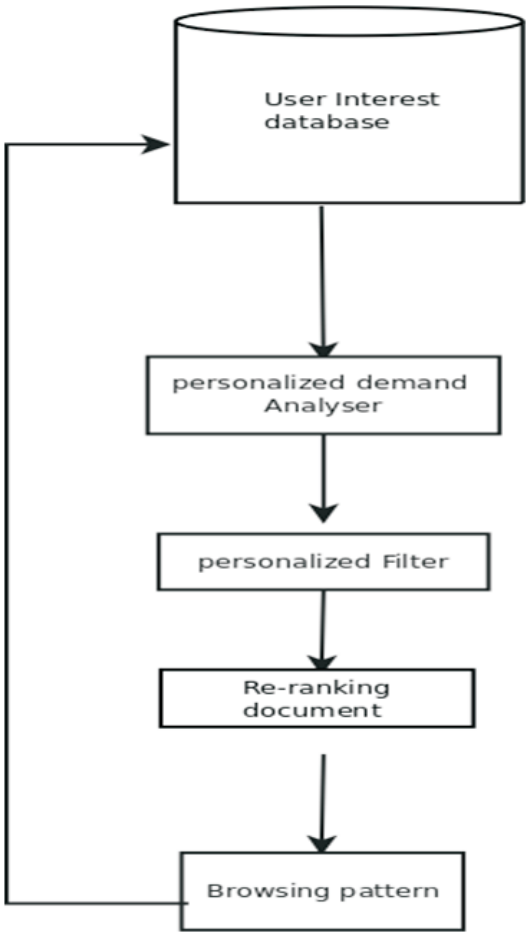

Figure 2: User Interest Model Architecture

III. Methodology

Two kinds of data sources are considered, the documents itself and the web log registers, which are easily available. The web log registers contain information about the browsing behavior of the user. The second data is the document itself and each document is defined by its content. Here the only interest is in the content of the document, particularly its keywords. The main idea is that the different documents (represented by the keywords) available can be seen as nodes in the semantic net, connected by a number of links, along which variable numbers of users travel in search for information. The resulting structure can be enhanced with weights based on the frequency of usage of each user which can be collected from the web log registers.

Web usage mining is the type of Web mining activity that involves the automatic discovery of user access patterns from one or more Web servers. Most of the recommendation approaches do not distinguish the importance of different pages and all the visiting pages are treated equally. It is quite probable that not all the pages visited by the user are of interest to him/her. Therefore it is imperfect to consider all the visited pages equally, to capture user interest and to predict user behavior. Therefore we need a weight measure for approximating the interest degree of a web page to a user. The usage data are the data that represents using documents. The data items include user-id. Document-id, duration of visit, number of times visited etc. The access log file is the major source of usage data.

The document weight calculation calculates term frequency in both local (that is frequency of the term in the document itself, tf value) and global documents (that is frequency of the term in the whole document stored in the database, idf value), and the product between the local and the global term frequency. Here term frequency (tf) is static value and inverse document frequency is a dynamic value based on the web browsing behavior of the user. For a user, when he navigates through the semantic network, the links with highest strength are visible to the user. This knowledge can be used to improve the quality of the results that users get in return for their queries, point them

towards relevant documents, reorganize the relevant documents to make it more efficient, and even discover wholly new concepts and relationships. Self-organization of semantic network and restructuring the network also possible based on the above procedure. New nodes can be added and these nodes can be placed in the network based on the relevance.

IV. Conclusion

One of the new forms of personalization in e-learning environment is to give recommendations to learners in order to support and help them through the e-learning process. This is a scalable technology which occurs in the 'online' phase to compute the recommendations against a possibly massive repository of educational resources in 'real time'.

References:

J.Han, M.Kamber, Data Mining : Concepts And Techniques, 2006

J.Chen, J.Yin, J.Huang, Recommendation Of New Items Based On Indexing Techniques, Proceedings Of 2004 International Conference On Machine Learning And Cybernetics

O.R.Zaine, Web Usage Mining For A Better Web Based Learning Environment, University Of Alberta, 2001

P. Desikan, C.Delong, K.Beemanapalli, A.Bose, J.Srivastava, Web Mining For Self-Directed E- learning, University Of Minnesota.

R.Forsati, M.R.Meybodi, A.Rahbar, An Efficient Algorithm For Web Recommendation Systems, IEEE,2009

http://en.wikipedia.org/wiki/coldstart

http://en.wikipedia.org/wiki/Recommender_system

http://www.readwriteweb.com/arcives/recommende_systems.php