Journal of Theoretical and Applied Information Technology
31st December 2012. Vol. 46 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ADAPTIVE HYPERMEDIA USING LINK ANNOTATION TECHNOLOGY AND RECOMMENDER MODEL (AHLARM)

N. DEEPA , S. PRIYADARSINI, VENKATESH MAHADEVAN, JOEL WILSON,
Assistant Professor, School of Information Technlogy & Engineering, VIT University, Vellore.
Assistant Professor, School of Information Technology & Engineering, VIT University, Vellore
Swinburne University of Technology Australia
Associate-projects Cognizant Technology Solutions San Francisco CA

E-mail: deepa.rajesh@vit.ac.in, priyadarsini.s@vit.ac.in, venkatsumee@hotmail.com, joelwilson.gunasekaran@cognizant.com

ABSTRACT

Adaptive hypermedia systems (AHS) combine hypermedia, user modeling techniques and a certain type of artificial intelligence that modifies the structure and contents of hypermedia documents according to each user. We have many Internet users now a days and also there are more number of webpages, it is very difficult for the surfers to find the information they need. Users must browse through a large hierarchy of concepts to find the information for which they are looking for. Users should be provided with web content/links based on their interests and preferences, thus allowing for more efficient and relevant information access. User's interests are based on his/her browsing behavior, the sites he/she visits, his/her prior history, and detailed interactions on web sites of interest to form a user interest profile. This paper focuses on providing adaptive presentation to the surfers which involves personalization, building and updating user profiles by monitoring and storing the users' browsing habits, recommending content/links by web usage logs and providing dynamic links via graphical objects. There are many Hypermedia Systems available which are related to a specific Application. In this paper we study the general purpose Adaptive Hypermedia System (AHLARM) for an application-independent architecture.

KEYWORDS: *Adaptive Hypermedia, User Modeling, Recommender Systems, Adaptation, Link Annotation*

1. INTRODUCTION

A hypermedia system is basically one that allows users to freely scan on information network in order to retrieve items of information structured by means of nodes and links. With the increasing volume of data, however, various problems are surfacing. One such problem concerns the accuracy of information retrieval : users have difficulty in obtaining their requested information, or the information they receive does not satisfy their needs. Another is that the network structure of hypermedia is static, and users can search only in a previously to acquire the user model information and realize its user adaptation, and how to support the network structure desired by the authors and users.The adaptation can be classified into two types : adaptation of contents and adaptation of links. Whichever method is taken, many problems still remain, such as how meaningful user model information, and realizes dynamic and flexible user navigation.

Adaptive Hypermedia System (AHS) is defined as hypertext and hypermedia systems that reflect some characteristic of the user in a model and adapt several visible aspects of the system for final user.
Adaptation can take 3 forms [1]:

- Adapted Systems: in which adaptation is hard-wired by the application designer; in this case, the system is customized to a particular user profile, which is created at design time.
- Adaptable system: in which adaptation is explicitly required by the user. More precisely, the user can specify her/his own preferences or adaptation can be provided by recommender systems.
- Adaptive systems: in which adaptation initiative belongs to the system itself, based on continuous observation of user preferences and needs. The user's profile is no longer static, it is dynamically updated by the system, after tracking and analyzing user behavior.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Adaptive hypermedia systems (AHS) combine hypermedia, user modeling techniques and a certain type of artificial intelligence that modifies the structure and contents of hypermedia documents according to each user.

Adaptive Hypermedia for All! (AHA!) was created at Eindhoven University of Technology. Architecture AHA! is based on reference model AHAM [2], which was designed to provide abstract structure and functionality for Adaptive Hypermedia Systems.

In this paper, section 2 describes web personalization, section 3 explains the user modeling techniques, section 4 gives details about recommender system and section 5 explains AHLARM Architecture.

2. PERSONALIZATION

Generally, personalization is a development of matching the object to specific needs. Hanson and Crayne[3] stated the specific needs are distinct to every person since different person has different identity and personality. Personalized systems have become increasingly popular since the beginning of the 1990's with the rise of the World Wide Web (WWW).

Users may spend a large amount of time at particular types of sites, e.g. movie related, science, or finance sites. Users will expect specific personalization on these sites that cannot be provided by a general-purpose behavior mining algorithm. For example, a plugin should be able to track the user's interaction with these sites, observe which movies he/she likes and dislikes, and update his interest profile to reflect these preferences.

The website content is personalized when only the relevant content displayed on the web. The relevant content should be based on individual user preferences which can be obtained implicitly or explicitly. The information is so called to be retrieved implicitly if the information obtained without acknowledging the user. As example, to track customer purchased or usage habits in order to know the types of products or services that a customer likes, dislikes, needs or wants without user's consent. The information that obtained explicitly is where the information is retrieved by getting the input from the user.

There are several browser add-ons (toolbars) that perform data collection and user behavior mining. Perhaps the most popular among them is the Alexa Toolbar, which for each user collects a complete browsing history, search engine query list, and summary of the advertisements presented to the user. This information used by Alexa to compute a number of analytic functions, some of which are returned to toolbar users as a service. Among the analytics are traffic statistics (including a comprehensive, internet-wide ranking of popular sites), related links, audience demographics, and clickstream statistics. Similarly, Bing, Google, and Yahoo all offer toolbars, although they vary in the amount of mining and automatic personalization that they perform.

A. The Problems of Personalization

Personalization needs the user information to be processed. Eirinaki and Vazirgiannis [4] stated the techniques for collecting and analyzing the user information: (1) content-based filtering to make recommendation by tracking individual user behavior, (2) collaborative filtering to make recommendation by tracking individual user behavior and returning relevant information, (3) rule-based filtering to use decision trees of questions to determine the wants and needs of the user, (4) web usage mining using certain methods such Association rule mining, clustering and classification to derive useful information. Therefore, in order provide the Personalization service, the user should have the willingness to share the information to certain degree. If the propensity of information sharing from the user is low, no data can be processed and analyzed to build personalization. If data provided is inaccurate or false information supplied, the personalization will be useless because it would not match to anyone. This issue is arising since there is significant number of Internet users has supplied false or fictitious information in personalization application. The non-sharing behavior is the major problem and yet become the barrier to website personalization.

3. USER MODELING

User modeling, often Addressed as user profile and a simple way of Capturing his/her references, plays key role in Adaptive Hypermedia Systems[12].

Modeling user behavior is an ongoing challenge in different application domains, including Knowledge Management, e-commerce, decision support, e-learning, and marketing. Such advanced systems allow having a better knowledge of the user (in terms of needs, preferences, or goals) in order to offer him/her enhanced services and better support users. Modeling human behavior, by learning from humans, may be also applied to achieve realistic agents' models in various domains [6]. There are many dimensions and elements that need to be considered when building a user model for a specific application domain [7]. Most of the studies

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of Internet user behavior are based on statistics of user traces and traffic measures [8]. In this case, user behavior characterization takes into account statistics of user session duration, data rates, application popularity, and user mobility in order to predict trends or make recommendations to the users. Researchers often make use of theories from sociology or psychology in an attempt to reveal the hidden rationality of the user browsing behavior [9], to understand individual differences and cognitive styles. From a social perspective, user behavior has been studied in online communities in relation with behavioral patterns of collaboration in order to understand different forms of motivation [10], to evaluate knowledge-sharing dilemma in order to design interventions to successfully manage organizational knowledge [11], or to understand the different categories of users and their motivation for participation in online communities [11].

A. User Modeling Overview

User modeling is the whole process of constructing user models, and creating, updating or deleting user profiles. It contains the functions which are to incrementally build up a user model, to store, update and delete entries in instantiated user profiles, and to maintain the consistency of the model. We introduce two important processes of user modeling [13]: collecting data about users and processing the data to build or update user models.

There are also various ways for collecting data from a user. The traditional way is to let the user directly provide information (for example, age, location, gender, occupation, income) by filling in a form. However, the user may withhold information because of privacy issues. Usage information is the information that can be tracked for observing users' behavior. It is perhaps the most important user data, and is extracted from a web server log, which is the primary source of data in which the activities of web users are captured. Usage information may be described in terms of simple page views, transactions (which are "significant" events, and may combine multiple page views), and sessions (which are combinations of page views or transactions that together represent users' behavior) [14]. In addition to the simple sequence of events, information about time of access and frequency of access can also be captured as usage information. However, usage information is not fully reliable. The page clicked by a user does not guarantee that the user attentively reads its content. To make user modeling simpler and more reliable, it is necessary to involve the user in the process of user modeling to acquire additional information from the user. Widely used Machine learning techniques to build a user model include linear models, TFIDF-based models, Markov models, neural networks, classification and clustering techniques, rule induction techniques, and Bayesian theory-based techniques.

B. User Profile

A user profile is defined as a collection of information about a user, including demographic information (name, age, location, to name a few), usage information (for example, pages visited, frequency of visit), and interests or goals (either explicitly stated by the user or implicitly derived by the system). It is an instance of a user model for a particular user. There are really two types of user profile data: those that describe individuals and those that describe groups of users[15].

C. GUMSAWS

We propose the GUMSAWS (A Generic User Modeling Server for Adaptive Web Systems)[15] framework, as illustrated in Fig 1.

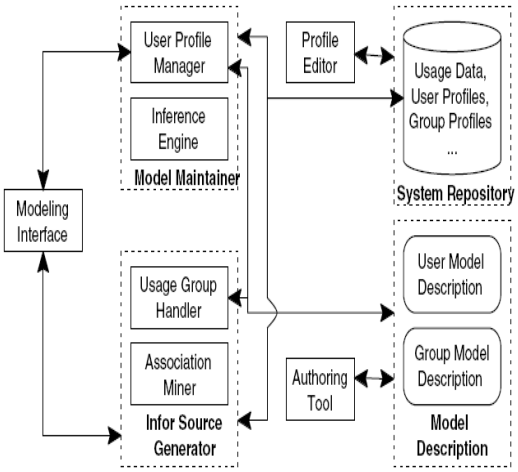

Fig.1. GUMSAWS Framework [15]

The Model Description (MD) consists of two components, the User Model Description (UMD) and the Group Model Description (GMD). These are data repositories that store domain-dependent Intermediate Format Vocabularies (IFVs) and application-dependent description for user and group models. The IFV is the schema in intermediate format for describing concepts and relationships related to an individual user or group of users existing in adaptive web systems.

The Model Maintainer (MM) offers user modeling functions, including instantiating a user or group profile; storing, retrieving, updating and deleting entries in profiles; and inferring user missing property values. Two components, the User Profile

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Manager (UPM) and the Inference Engine (IE), are grouped into this sub-system. The Information Source Generator (ISG) generates the information sources of groups information and association rules for the MM to update users' missing property values. Two components, the Usage Group Handler (UGH) and the Association Miner (AM), are included into this sub-system.

User groups are generated by the UGH according to users' visiting history. Association Miner discovers Association rules. The System Repository (SR) is used to store usage data, user profile data, and group profile data. It also stores information sources of direct information, groups information, association rules, and general facts. The SR also connects some system components together. It is constructed in the initialization stage by some system components, such as the UPM and the UGH. It will also be updated by them during the system activity. Note that some system components are not grouped into any of the four sub-systems. They provide interfaces for interactions. For example, the Modeling Interface (MI) is an interface between adaptive web systems and GUMSAWS. The main function of the MI is to forward adaptive systems' requests to components of GUMSAWS. The communication between the MI and the sub-systems is through a network. The Authoring Tool (AT) accesses model description and provides an interface for authors (i.e. system administrators) to specify the application-dependent user or group model description. Through the AT, authors may define user and group models, and default user and group profiles. The Profile Editor (PE) is implemented as an interface to allow users to see information held about them, and to modify their information. Requests from users received through the PE will be handled directly by the UPM as direct information about users. The PE makes adaptive web systems transparent in that users have full control on their information[15].

4. RECOMMENDER STRATEGIES

Increasing availability of portable web-enabled devices fuels further growth in the volume of consumer web traffic. Traditionally users can pull content from the web via mechanisms such as search engines. However, personalized delivery mechanisms, which take user behaviors into consideration when preparing web content recommendations, can be very useful. To capture web users' access habits and behaviors, one possible approach is web usage mining [16] which discovers interesting and frequent user access patterns from web usage logs. Many web usage mining techniques [17] focus on mining common access patterns, which have occurred frequently within the entire duration of all access sessions.

There are some [18] techniques which mine periodic access patterns, that occur frequently in a particular period, directly from web usage logs. Such periodic access patterns are very useful for understanding users' web access behaviors. With such knowledge, we can deduce and prepare the resources that the user is most probably interested in during a specific time period without the need of the user's current web access information.

Fuzzy logic is applied to represent real time concepts and requested resources of periodic pattern-based web access activities. The fuzzy representation is used to construct a knowledge base of the user's web access behaviors. This is then used to provide timely and personalized recommendations to the consumer.

A. Personalized Recommendation

Personalized recommendation aims at minimizing ambiguity and unwanted information that is presented to the user, thereby reducing the effect of information overload that is often encountered by web surfers. An effective personalized web content recommendation system should present the most relevant suggestions to the user in the most timely manner.

B. Web Usage Logs

The process of creating the user behavior knowledge base begins with semantically enhanced web usage log as input. As Table 1 illustrates, each requested URL in web usage logs is manually or semi-automatically annotated with semantic descriptions, i.e., one or more predefined topics, concepts or categories such as News and Sports [18].

Table 1 Semantically enhanced Web Usage Log [18]

UserID	Timestamp	URL	Semantic Annotation
User1	21/03/2011 09:20:02	URL1	#Topic1,#Topic2, #Topic5…
User2	21/03/2011 09:25:51	URL2	#Topic1,#Topic4, #Topic9…
User3	21/03/2011 09:31:08	URL3	#Topic8,#Topic2, #Topic3…

Each entry in the web usage log can be interpreted as "A certain user has accessed specific resources at a specific time". From this log entry we can predict user preferences and interest and recommend links.

Journal of Theoretical and Applied Information Technology
31st December 2012. Vol. 46 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

C. Content Recommendation

Using Web usage log content, User preference can be identified and web content can be recommended for the user/group. In order to recommend useful content for known Users/groups, SEKO. MOTEGI, YAGI, MUTO algorithm prunes useless content for the group using ratings list of individual preference between group members. Fig.2 shows system flow of Shunichi SEKO, Manabu MOTEGI, Takashi YAGI and Shinyo MUTO recommendation strategy[19]

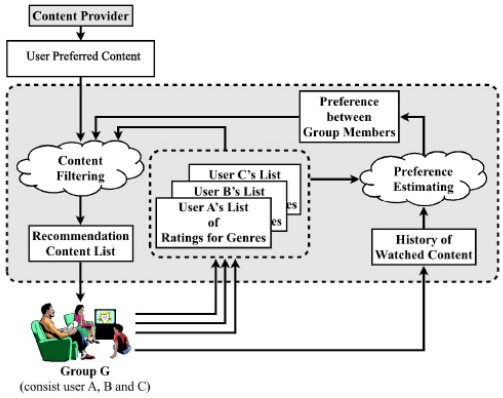

Fig.2 System Flow Of Recommendation Strategy[19]

Here along with the delivery of preferred content by the content provider, unwanted content filtering also done. History of watched content is used to generate Recommendation Content List.

5. AHLARM ARCHITECTURE

The web is a medium for accessing a great variety of information stored in various locations in the whole world. As the data on the web grows rapidly and different Internet users have different requirements for the information they want to get from the Internet. So, it has lead to several problems such as increased difficulty of finding relevant information, extracting potentially useful knowledge and learning about consumers or individual users. Web personalization improve the Web search experience by using an individual's data e.g. user's domain of interest, preferences, query history, browser history etc.

AHLARM is domain independent Adaptive Hypermedia System which provides personalized content based on users' preferences. The major components of the system are User modeling, Recommendation strategy, Content filtering, Search Engine, Link Annotation for dynamic links in the form of graphical objects.

To personalize means to make or change something so it is suitable for a particular person. Personalization is defined as the ability to deliver content and services to individuals based on their preferences and behavior. Web personalization is about personalizing aspects of web resources - the content itself, links, web page structure and navigation. User preferences are collected using web usage log records and History of user watched content. Fig.3 shows the overall design of AHLARM architecture which explains the major components.

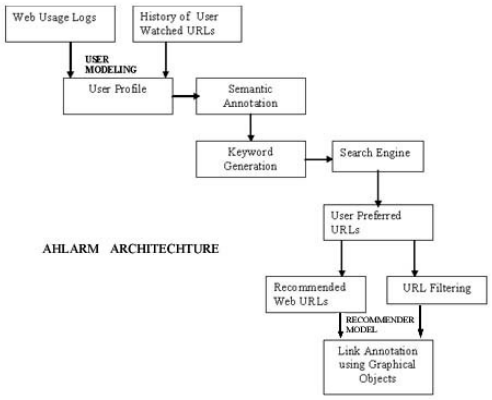

Fig. 3 Architecture Diagram Of AHLARM

User Profile is created and dynamically updated using the browsing history. From the semantic annotation collected from web usage log and history, User preferred keywords are generated. Those keywords are supplied to the search engine in order to generate User Preferred URLs. Using Filtering technique, URL filtering is done as the next step.

Personalized recommendation systems have come into the big picture now-a-days. Recommender system produces individualized recommendations based on user's individual characteristics, which can be used in many domains including books, music, web pages and news. From the Recommender strategy mentioned in this paper recommended web URL list is generated.

From the generated URL list, Link annotation technology is applied which produces dynamic links in the form of Graphical Objects. Dynamic linking allows adaptive modification of links dependent on users and provides navigation for each user.

Web pages contain links that have connections between them which represent the relations among the concepts they represent. Currently people follow links blindly without knowing what these links point to, or what kind of relationship this link represents. To make links more meaningful, instead of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

providing hypertext links, Links can be generated using Graphical Objects which gives more user friendliness.

6. CONCLUSION

Personalization requires collecting and using information about the user for showing the relevant URL while browsing. Dynamic link generation refers to the production of new links of interest to the user on the information network that have not been defined in the authorship. Several adaptive hypermedia techniques have been introduced so far. In this paper, AHLARM introduced a framework of adaptive Hypermedia system for domain independent architecture A dynamic linkage method uses two types of user model nformation. One using web usage log records and the other by gathering past history of user content. This study paper makes use of recommender model developed from previous research to generate recommended URL list from which dynamic links in the form of graphical objects can be created. The framework of the system is developed in this paper. In future, we will conduct research on the above Adaptive Hypermedia System.

REFERENCES

[1]. E.A. Edmonds (1981) Adaptive Man-Computer Interfaces, Computing skills in the user interface, Computer and People series, New York, Academic Press.

[2]. P. De Bra, et al., "AHAM: A Dexter-based Reference Model for Adaptive Hypermedia." in ACM Conference on Hypertext and Hypermedia. Darmstadt, 1999, pp.147-156.

[3]. V.L. Hanson, S.Crayne, "Personalization on web browsing: adaptations to meet the needs on an older adults," Universal Access in the Information Society, vol.4, pp.46-58, July 2005

[4]. Eirinaki.M, Vazirgiannis.M, "Web mining for web personalization," ACM Transaction on Internet Technology, Vol.3, pp.1-27, February 2003

[5]. Online Information Sharing Issues in Website Personalization, Mohd Norhisham Razali, Fariza Hanis Abdul Razak, Shah Alam, Rozita Hanapi, Puncak Alam, Selangor.

[6]. H. Hattori, Y. Nakajima, and T. Ishida, "Learning from humans: Agent modeling with individual human behaviors," IEEE Trans. Syst., Man, Cybern. A, Syst., Humans, vol. 41, no. 1, pp. 1–9, Jan. 2011.

[7]. A. Brun, A. Boyer, and L. Razmerita, "Compass to locate the user model I need: Building the bridge between researchers and practitioners in user modeling," in Proc. 18th Int. Conf. UMAP, vol. 6075, Lecture Notes inComputer Science, P. De Bra, A. Kobsa, and D. Chin, Eds., Big Island, HI, 2010, pp. 303–314, Berlin, Germany: Springer-Verlag.

[8]. L. Lancieri and N. Durand, "Internet user behavior: Compared study of the access traces and application to the discovery of communities," IEEE Trans. Syst., Man, Cybern. A, Syst., Humans, vol. 36, no. 1, pp. 208–219, Jan. 2005.

[9]. E. Brown, T. Brailsford, T. Fisher, and C. van der Eijk, "Revealing the hidden rationality of user browsing behaviour," in *Proc. 18th Conf. Hypertext Hypermedia*, Manchester, U.K., 2007, pp. 85–94.

[10]. P. Kollock, "Social dilemmas: The anatomy of cooperation," *Annu. Rev. Soc.*, vol. 24, no. 1, pp. 183–214, 1998.

[11]. A. Cabrera and E. F. Cabrera, "Knowledge-sharing dilemmas," *Org. Stud.*, vol. 23, no. 5, pp. 687–710, 2002

[12]. An Ontology-Based Framework for Modeling User Behavior—A Case Study in Knowledge Management Liana Razmerita, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS, VOL. 41, NO. 4, JULY 2011

[13]. P. Brusilovsky. Methods and techniques of adaptive hypermedia.*User Modelling and User-Adapted Interaction*, 6(2- 3):87–129, July 1996.

[14]. J. Zhang and A. Ghorbani. The reconstruction of user sessions from a server log using improved time-oriented heuristics. In *Proceedings of Communication Network and Services Research (CNSR) 2004 Conference*, pages 315–322, Fredericton, NB, Canada, May 19–21 2004.

[15]. GUMSAWS: A Generic User Modeling Server for Adaptive Web Systems Jie Zhang School of Computer Science University of Waterloo Waterloo, ON, Canada j44zhang@uwaterloo.ca Ali A. Ghorbani Faculty of Computer Science University of New Brunswick Fredericton, NB, Canada ghorbani@unb.ca

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[16]. J. Srivastava, R. Cooley, M.Deshpande and P.N.Tan, "Web usage mining: discovery and applications of usage patterns from web data". ACM SIGKDD Explorations, I(2), 2000, pp.12-23.

[17]. J.Pei, J.Han, B.Mortazavi-Asl, and H.Zhu, "Mining access patterns efficiently from web logs". Proc. 4th Pacific-Asia Conf. Knowledge Discovery and Data mining, Kyoto, Japan, 2000, pp.396-407.

[18]. Baoyao ZHOU, A.C.M. FONG, Senior Member, Siu C.HUI, Senior Member, IEEE, 2011 IEEE International Conference on Consumer Electronics (ICCE) "Effective web content recommendation based on consumer behavior modeling"

[19]. Video Content Recom mendation for Group Based on Viewing History and Viewer Preference Shunichi SEKO1, Member, IEEE, Manabu MOTEGI1, Takashi YAGI1 and Shinyo MUTO1 1NTT Cyber Solutions Laboratory, Japan. 2011 IEEE International Conference on Consumer Electronics (ICCE)