Research Paper　　　　　　　　　　　　　　　　　　　　　　**Engineering**

Evaluation of Web Personalization

*Nitin Nara ** Komal Arora

* M.Tech Final year Student, Satpriya College of Engg. & Management, Jind Road Rohtak, Haryana, India

** Lecturer In Satpriya College Of Engg. & Management, Jind Road Rohtak, Haryana, India.

ABSTRACT

With millions of pages available on web, it has become difficult to access relevant information. One possible approach to solve this problem is web personalization. Web personalization is the process of customizing a Web site to the needs of specific users, taking advantage of the knowledge acquired from the analysis of the user's navigational behavior (usage data) in correlation with other information collected in the Web context, namely, structure, content and user profile data. Personalization of a web site may be performed by the provision of recommendations to the users, highlighting/ adding links, creation of index pages, etc.

The web personalization systems are mainly based on the exploitation of the navigational patterns of the Web site's visitors. In this research paper, we explore the integrated process of website personalization, various available techniques and recommender systems and their applicability in particular business context. We also explore the novel techniques that use the content semantics and the structural properties of a web site in order to improve the effectiveness of web Personalization.

Keywords : Web Personalization, Web Mining, Web Usage Mining, User Profiling.

1. INTRODUCTION

The continuous growth in the size and use of the World Wide Web imposes new methods of design and development of on-line information services.. On the other hand, the e-business sector is rapidly evolving and the need for Web marketplaces that anticipate the needs of the customers is more evident than ever [2]. Web personalization is a technique to personal-ize web pages according to the needs of customer, without asking them explicitly. It provides an organization, a variety of Services to increase their sales, customer retention and to compete in the market. Web Personalization is a strategy rather than it is a technique. It is a bundle of various tech-niques and tools. Web Personalization is simply defined as the task of making Web-based information systems adaptive to the needs and interests of individual users [3]. Web person-alization is a process and a strategy to compete in the online market. Web site personalization can be defined as the pro-cess of customizing the content and structure of a Web site to the specific and individual needs of each user by knowing their preferences implicitly or explicitly.

Principal elements of Web personalization include (a) the cat-egorization and pre-processing of Web data, (b) the extrac-tion of correlations between and across different kinds of such data and (c) the determination of the actions that should be recommended by such a personalization system [2]. The per-sonalization of services offered by a Web site is an important step in the direction of alleviating information overload, mak-ing the Web a friendlier environment for its individual user and hence creating trustworthy relationships between the Web site and the visitor-customer [3].

Personalization is not limited to online shopping, and it is ap-plied to search engines. In 2004, two personalized search engines, A9.com by an Amazon subsidiary and MyJeeves by Ask Jeeves, were launched to let users store individual search results and then provide personalized web searches. Recently, the My Yahoo service has been enhanced with per-sonalized searching..

1. WEB PERSONALIZATION TECHNIQUES

Basically the Web Personalization is implemented using the following four basic technique viz handicraft decision tech-nique, Hyperlink-based technique, Content-based filtering, Collaborative filtering.

1.1 Rule Based

Rule-based techniques provide a visual editing environment for the business administrator to specify business rules to drive personalization [4]. In rule-based filtering the users are asked to answer a set of questions. These questions are de-rived from a decision tree, so as the user proceeds to answer them, what he finally receives as a result (e.g. a list of prod-ucts) is tailored to his needs [5].

1.2 Simple Filter

Simple filtering relies on predefined groups, or classes, of visitors to determine what content is displayed or what ser-vice is provided [4].

1.3 Content-based filtering

Content-based filtering works by analyzing the content of the objects to form a representation of the visitor's interests [4]. Content-based filtering examines the relationship between resources and users. It takes advantage of the similarity of information and users' interests to filter information. When a user visits the website and seeks resources that match the established profile, the related information is retrieved. Con-tent-based filtering systems are solely based on individual us-ers' preferences. The system tracks each user's behavior and recommends items to them that are similar to items the user liked in the past [5].

1.4 Collaborative filtering

Collaborative filtering collects visitors' opinions on a set of objects, using either explicit or implicit ratings, to predict a

particular visitor's interest in an item [4]. Collaborative filtering compares the relationship between one user and another. It makes use of similarities among users to filter information and provides the similar user with commonly-sought information. Its advantage is that it can discover new information a user may be interested in but is not asking for. Collaborative filtering systems invite users to rate objects or divulge their preferences and interests and then return information that is predicted to be of interest to them. [5].
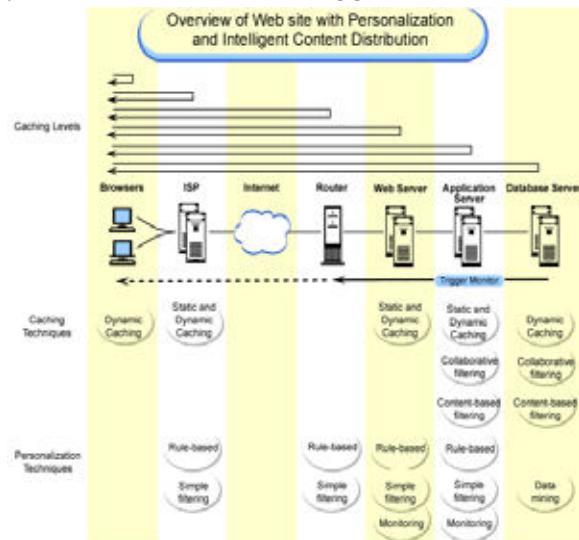
Fig 1 Overview of Web site with personalization and intelligent content distribution

2. PROCESS OF WEB PERSONALIZATION

2.1 Collection of Web data –
Data collection can be performed in two ways, i.e. implicitly and explicitly. Explicit data collection has typically been modelled as ratings of items, personal demographics and preference (including utility) data. Implicit data collection refers to any data that can be collected on the user inconspicuously by "watching" their interaction with the system. *Implicit* data includes past activities/click streams as recorded in Web server logs and/or via cookies or session tracking modules. Explicit data input has a cost associated with it as it requires users to detract from their principle reason for interacting with the system and provide data, the benefits of which are intangible to the user [6].

2.2 Preprocessing of Web data –
 The first step in data preprocessing is to clean the raw Web data. During this step the available data are examined and irrelevant or redundant items are removed from the dataset. Data is frequently pre-processed to put it into a format that is compatible with the analysis technique to be used in the next step [6]. The first issue in the preprocessing phase is data preparation. Depending on the application, Web log data may need to be cleaned from entries involving pages that returned an error or graphics file accesses. In some cases such information might be useful, but in others such data should be eliminated from a log file. Furthermore, crawler activity can be filtered out, because such entries do not provide useful information about the site's usability[5].

2.3Analysis of Web data –
Also known as *Web Usage Mining* , this step applies machine learning or *Data Mining* techniques to discover interesting usage patterns and statistical correlations between web pages and user groups. This step frequently results in *automatic user profiling*, and is typically applied offline, so that it does not add a burden on the web server [6].

2.4 Decision making/Final Recommendation Phase –
The last phase in personalization make sure of the results of the previous analysis step to deliver recommendations to the user. The recommendation process typically involves generating dynamic Web content on the fly, such as adding hyperlinks to the last web page requested by the user. This can be accomplished using a variety of Web technology options such as CGI programming [6].
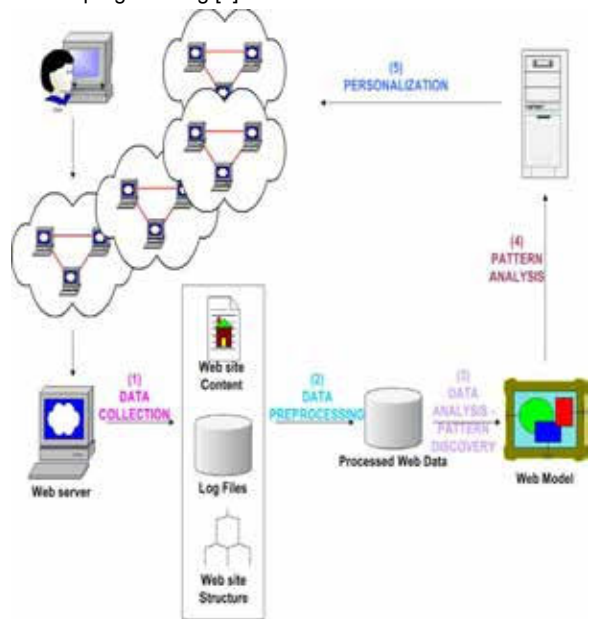
3.USER PROFILING
User profiling is the process of collecting information about the characteristics, preferences, and activities of a Web site's visitors[5. A user profile can be either static, when the information it contains is never or rarely altered or dynamic when the user profile's data change frequently[2]. In the Web domain and especially in e-commerce, user profiling has been developed significantly because Internet technologies provide easier means of collecting information about the users of a Web site, which in the case of e-business sites are potential customers[5].

4.Web Usage Mining
The purpose of Web usage mining is to apply statistical and data mining techniques to the preprocessed Web log data, in order to discover useful patterns. Originally, the aim of Web usage mining has been to support the human decision making process and, thus , the outcome of the process is typically a set of data models that reveal implicit knowledge about data items, like Web pages, or products available at a particular Web site[7]. The purpose of Web usage mining is to apply statistical and data mining techniques to the pre-processed Web log data, in order to discover useful patterns. In the context of Web mining, we can distinguish two cases, user clusters and page clusters. [2].

5.Web Log
Each access to a Web page is recorded in the access log of the Web server that hosts it. The entries of a Web log file consist of fields that follow a predefined format. The fields of the common log format are:

remotehost rfc931 authuser date "request" status bytes

where is not available, *rfc931* is the remote log name of the user, *authuser* the username as which the user has authenticated himself, available when using password protected WWW pages, *date* the date and time of the request, *"request"* the request line exactly as it came from the client (the file, the name and the method used to retrieve it), *status* the HTTP status code returned to the client, indicating whether or not the file was successfully retrieved and if not, what error message was returned, and *bytes* the content-length of the docu-

ments transferred. If any of the fields cannot be determined a minus sign (-) is placed in this field[2,5].

6. CONCLUSION
The overall process of Web personalization consists of five modules, namely: user profiling, log analysis and Web usage mining, information acquisition, content management and Web site publishing. Web personalization is a domain that has been recently gaining great momentum not only in the research area, where many research teams have addressed this problem from different perspectives, but also in the industrial area, where there exists a variety of tools and applications addressing one or more modules of the personalization process.

Web personalization is a domain that has been recently gaining great momentum not only in the research area, where many research teams have addressed this problem from different perspectives, but also in the industrial area, where there exists a variety of tools and applications addressing one or more modules of the personalization process. I started from the meaning of web personalization, its characteristics, issues related to the web personalization, viz viability of the web personalization, user's privacy issues, process of web personalization, and classification of web personalization techniques according to their applicability.

REFERENCES

[1] Ayesha Ameen, Khaleel Ur Rahman Khan , B.Padmaja Rani, Vol 2, No.6, 2012, Semantic Web Personalization: A Survey, Information and Knowledge Management ,ISSN 2224-5758 (Paper) ISSN 2224-896X (Online), www.iiste.org | [2] Magdalini Eirinaki, Michalis Vazirgiannis, ACM Transactions on | Internet Technology, {Vol. 3, No. 1, February 2003}, Web Mining for Web Personalization, http://doi.acm.org/10.1145/643477.643478 | [3] | [4] R.Khanchana, Dr.M.Punithavalli, Volume 11 Issue 7 Version 1.0 May 2011, Web Page prediction for Web Personalization: A Review, Global Journal of Computer Science and Technology, Global Journals Inc. (USA) | [5] MAGDALINI EIRINAKI and MICHALIS VAZIRGIANNIS, February 2003, Pages 1–27, Web Mining for Web Personalization, Vol. 3, No. 1, ACM Transactions on Internet Technology | [6] A.JEBARAJ RATNAKUMAR, AN IMPLEMENTATION OF WEB PERSONALIZATION USING WEB MINING TECHNIQUES, Journal of Theoretical and Applied Information Technology © 2005 - 2010 JATIT. www.jatit.org | [7] DIMITRIOS PIERRAKOS1,GE ORGIOS PALIOURAS1,