An Innovative Approach for Implementing Web Personalization Using Web Mining Techniques

Sheetal*

Abstract: Web mining extract useful information from the server logs and its techniques web usage mining and web content mining specifies the personalization process. Contents on the web are increasing dramatically which create confusion for the users and the customer of the e-business website. Most of the Search engine does not provide users personalized search but offers them a lot of irrelevant information. In this paper we present personalized web search to the user on the basis of their navigation pattern and recommend content after filtering the navigation links and contents available on the system by using appropriate web mining techniques. This paper mainly focus on the filtering techniques which firstly apply on the user’s navigation link namely collaborative filtering which discover interesting user’s navigation pattern. After this system apply content filtering to both on the result come from collaborative filtering and content available on the site. As a result, user get personalized web page containing Collaborative and Content Recommendation.

Keywords: Personalization, Usage mining, content mining, content filtering, collaborative filtering.

1. INTRODUCTION

In the world of internet, there is fast and explosive growth of information which flood users with large amount of information and various sources of information which pose great challenge of information Overloading. Now days to become popular many web applications accept this great challenge and fascinate its repeat users with the content they want to. Web Personalization cope this information overloading problem and become powerful tool to store, present, analyze and filter this information so to deliver the required content to the user according to its specific need. A Successful e business website provide special treatment to its repeat visitors by analyzing its navigation pattern and filtering them to deliver appropriate content to the user of its Interest. Personalization is a process of gathering and storing information about site visitors, analyzing the information, and, based on the analysis, delivering the right information to each visitor at the right time [5]. Although each application area may need tailoring, especially in the areas of user interface and data collection, the core techniques for personalization used are web usage mining and web Content mining which are discussed in the paper.

Personalization has gone through different phases. Initially, personalization was used to identify visitor’s profile and retrieve information from visitor profile. The next phase is to select content that match visitor preference by applying filtering techniques. In the last phase, personalization is increasingly used as a means to expedite the delivery of information to a visitor, making the site useful and attractive to return to. To improve the quality of web site it is necessary for web designer to know what user actually want by analyzing its behavior, filter its available content on the basis of this analysis to predict user’s interested page and then deliver this personalized web page to user.

Creating a personalized web application is challenging task because it involves complete picture of the client as every different user have different interest, need different attributes that need to retrieved, analyzed and processed to deliver complete user personalized profile, which involves many security and usability issues. In this paper through the implementation of personalized web application we tackle all these issues and make use of two important web mining techniques which deliver peer profile to user containing content and collaborative Recommendations.
2. WEB MINING TECHNIQUES

Web mining is the technique to discover and analyze patterns available on the web [6]. It is the application of data mining and based on the analysis web mining is divided into three broad categories which are:

(a) **Web Content Mining**: Web Content Mining is the process of extracting useful content in the form links, categories, data from web documents. These contents contain useful facts and data which are most relevant to the user of the specific domain in all respects. The technologies which are mainly used for the web content mining are Information retrieval and Natural language Processing.

(b) **Web Usage Mining**: Web Usage Mining is the process of extracting useful information from server logs, history of web clients and proxy servers. Web usage mining is the process of finding out what users are looking for on the Internet by analyzing their navigation link or browsing history so to make proper future recommendation to the user with respect to their specific interests.

(c) **Web Structure Mining**: Web structure mining it the process of extracting structured information from the web documents which is in the form of graphs, tress, linked list. This type of web mining helps in deriving useful links which are in the form of nodes and useful links which are present on the edges of structured web documents [5].

3. WEB PERSONALISATION

Web Personalization is key technology used in many e-business web sites for delivery right information to the right user at right time. There are mainly two common aspects to personalize are structure and Content Personalization. (a) Structure Personalization change the position of available links and make them available at most prominent place on the basis web history of each user. (b) Content Personalization changes or update the content on the user personal page after analyzing its specific need or preferences and delete the relevant topics which are timely outdated.

How to Personalize Web: Several techniques are involved in the process of personalization from collecting information to analyzing and delivery at user end. The whole process complete through specific steps like Collection of user information, analyzing an, filtering information and delivery of content [9].

(a) **Collection of user information**: The main objective of collecting user information is to develop profile on the basis user interest and preference where user can get its relevant information that can be automatically updated with changes of user interest. The most common methods of information collection are explicit profiling, implicit profiling.

- *Explicit profiling*: This information can be collect during the time of registration by asking user interests or preferences for information or separate questionnaires are also provided to clients that’s can be filled by users and accordingly the accounts information personalized for particular user.

- *Implicit profiling*: Implicit information is indirectly obtained from the user itself, through tracking user’s browsing pattern and other habits like buying and selling habit or navigation pattern. For example, amazon.com saves the buying history of specific customer and based on that history, recommended specific product.

(b) **Analyzing the client information**: When the user profile is available, the next step is to analyze the profile information in order to present or recommended document, product, and action to specific visitor. Providing this sort of information and recommendation is a challenging task. Many techniques are used to represent and recommend content and filtering is the one of common and best.

(c) **Filter the analyzed content**: Filtering techniques compares user’s tastes with those of others in order to build up a picture of like-minded people. Three most common filtering techniques:

- *Simple filtering*: It relies on predefined groups, or a class, of visitors to determine what content is displayed or what service is provided.

- *Content-based filtering*: It works by analyzing the content of the objects to form a representation of the visitor’s interests. Generally, the analysis needs to identify a set of key attributes for each object and then fill in the attribute values and recommend contents to user based on the user interests and content available on the system.

- *Collaborative Filtering*: Collects visitors’ opinions on a set of objects, using either explicit or implicit ratings, to form like-minded peer groups and
then learns from the peer groups to predict a particular visitor’s interest in an item and deliver favourite or most visited link on the user peer profile.

(d) **Delivery of content:** Providing personalization for real-time applications, such as serving dynamically constructing Web pages based on the visitor’s profile, affects system performance. The basic approach to handling personalized pages is to serve his in basic HTML page from the server, which doesn’t require new technology and high-level programming.

There are three broad methods of personalization:
With implicit personalization the personalization is performed by the web page (or information system) based on the different categories mentioned above. With explicit personalization, the web page (or information system) is changed by the user using the features provided by the system. Hybrid personalization combines the above two approaches to leverage the best of both worlds.

4. **PROPOSED ARCHITECTURE FOR WEB PERSONALIZATION**

Web Personalization is divided into four categories
(a) Gathering of useful web data (b) Pre-processing of the collected data (c) Interpretation of the web data (d) Decision making or Final Recommendation.

When user sends request for a particular link or web page in a form of a query to the Search Engine, the Search Engine return the response by giving URL’s of all the matching links depending on the algorithm used by the Search Engine. When user enter query for its particular search request, the Query is first processed to filter out irrelevant data and then analyzed to discover interesting navigation patterns and relation between web page and user groups. The Personalized data is then given as an input to the ranking algorithm or temporal algorithm for Personalizing the result according to most visited links or favorite links of the particular user by evaluating web mining results.

This figure depicts the overall architecture of the web personalization.

After building Knowledge Base, System can give recommendation based on the navigation and most visited links by particular user in featured domain. The System help the user to retrieve his/her interested domain by clicking on the list of his/her favorite links with more convenience the current Recommendation is based on the last selections.

5. **PROCESS WORKFLOW**

*Implementation of the personalization consists of the following workflow:* Firstly user enters to the system and collects raw information i.e. Information which is available to the all visitors of the website. After analyzing this information, user login to the system and get its peer profile which contain user personal details and most visited link recommended by the system based on the URL ranking algorithm (collectively most visited favorite links of all users).

This figure depicts overall process workflow for personalizing web.

![Figure 1: System Architecture for Personalizing Web](image1)

![Figure 2: Process Workflow for Web Personalisation](image2)
simultaneously update already exiting links on the basis of periodical temporal update.

In the last step system again perform analysis on the result of collaborative filtering and deliver recommend those content in which user is most interested in and the contents available on the system by performing content filtering.

The result of the overall process workflow is that system deliver web page to user which contain its personal detail, most visited link of the site, user’s favorite link and the contents recommended by the system of its interested domain in single customized format.

6. CONCLUSION

The main research goal of this study is to show the implementation of personalization process by using web usage mining and web Content mining techniques. Personalize web search is delivered by the system by creating separate Homepage for each user, which contain its Personal details, most visited link and favorite links of specific user, which are timely updated on the basis of URL ranking algorithm and this personalize page also recommend content to the user after filtering and analyzing user’s favorite link and content available in the system. Personalization process is accomplished through analysis of user’s goals and then filters these goals to give collaborative recommendation. And the result of above analysis is also used to filter out the contents available on system to give Content recommendation.

References