

A Short Review on Different Personalization Schemes

Sandesh S. Pare, Bharati P. Vasgi

Student, Master of Engineering (IT), Sinhgad College of Engineering, Pune

Asst. Prof., Department of Information Technology, Sinhgad College of Engineering, Pune

Abstract:- Now-a-days, World Wide Web (WWW) becomes too huge, so it becomes too difficult to find out the relevant documents for user. Web Search Engine helps users to find out the relevant information; but these search engines can't provide personalized results. Different users want different kinds of information on the same query. Different people have different interest, need and intension about the same query. So, personalization is the technique which adapts the user needs and based on that, it will make the ranking decisions about the search result. Re-ranking provides more quality and effectiveness to search result; which satisfies user more. But, the main problem is how to obtain user needs or interest correctly and accurately. Based on the techniques of identifying and organizing user interest, a different type of personalization system emerges.

General Terms

Algorithms, Experimentation, Performance

Keywords:- Personalization, User Profile, Rule Based Personalization, Market Basket Analysis, Collaborative filtering.

I. INTRODUCTION

Now-a-days, the information on the Internet becomes too huge. What people lack is not the scale of the information but the way to obtain the needed information quickly and accurately? Search engine is one of the most used ways to get vast amount of Web information. But obviously, it has the problem that no matter how different the users are the search results referred to the same query keeps the same ignoring the difference of users on their personalized features such as interests, hobbies and needs. One way to solve the problem is to reorder the search results according to the user's personalization features. The key point of personalized reordering is to discover individual characters of different users. Generally speaking, the information a user truly needed typically contains the following features. First, it should be appropriate to the user's experience, interest and knowledge background. Second, it should be considered to be useful by most other users. For the first feature, a model is built considering the user's experience, interest and knowledge background to provide search results based on individual needs. For the second one, resources are recommended by collaborative filtering according to the other user's evaluation to the search results [5]. The certain pattern of web page visits by different user should be taken into consideration. By observing this pattern, different users are divided into categories according to their nature or behaviour. According to the category to which they belong, certain advertisement with the use of recommendation technique are made, which are targeted only for particular user. Different techniques to implement personalization are discussed below. These techniques can be combined with each other. Each of these technique having certain merits and demerits associated with it. According to the application or the situation, particular technique is chosen which suit them most.

II. RELATED WORK

Till date, there is lot of work done in the personalization. In this section, major strategies of personalization reviewed with a focus on how the personalization information is acquired.

a. Cookies

Cookies are small data files that are stored on user's local machine. They are created when user first interacts with website [8]. Cookies contain the set of queries and their respective results. So, it can provide the same result to the repeated query in the same session [13]. As the user provides the information such as name, address, other form of identification, the server running on the website stores this information on user's machine. In follow-up visits of the website, the server can identify needed information of the user without requiring the user to retype it. Cookies are usually small containing no more than simple user identification, e.g. name associated with the computer id. The rest of the user information is usually obtained from the web server's database. Many websites use cookies as a basic technology for personalization. Users see this in the form of a welcoming address that uses their name, e.g., "Hello Sandesh, Welcome back!"

b. Profile Based Personalization

Now-a-days, User has to log-in multiple sites and register their personal information and preferences separately. Instead, a single FOAF file (user profile) stores all information which is updated at every HTTP-GET request [9]. In order to be able to purchase or receive advanced services from most websites, users are required to register and enter personal information (user profiles) such as gender, age, interests, etc. Websites store this information in the database of the web server [5]. This information is primarily used to support the user with “type once” operations such as maintaining the user’s shipping address, etc. Websites also use user profiles for personalized services. The user’s postal code provides economical information so that the websites can recognize product access according to a user’s economic profile. For example, a featured wine sale might not be advertised to users residing in an area known to have a depressed economy. So, by maintaining the personal information of the user, the company has great benefits in the targeted marketing like advertisements, etc.

c. Personal tools

Personal informatics (PI) tools that support reflection are “personalized” insofar as the data consists of an individual’s data. Typically, this data is presented in visualizations that are generic and non-individuated. The proposed technique is deep personalization, where the visualizations are constructed by individuals. Such functionality gives individuals the power to build visualizations that are personally meaningful, allowing them to meet and address personal needs [10]. Some websites allow users to create shortcuts (links) to the information that interests them most. In most online stock trading websites such as E*TRADE and Datek, for example, users can configure their own ‘personal tools’ that trace the stock prices they are interested in. In portal sites such as Yahoo and MSN, users can create a page containing personally chosen links. Personal tools differ from profile-based personalization because it is the user, not the software that creates the personalization.

d. Opportunistic Links

According to the J. Goecks, “NuggetMine is an intelligent groupware application that collaborates with a workgroup to increase information links sharing among the group. Information links are small amounts of self-contained information, such as the URL of an interesting news article, a book title, or the time and location of a local art event. NuggetMine and the workgroup work together to build, maintain, and utilize a repository—or “mine”—of information links” [11]. Suppose, one traveller might be an outdoors person interested in visiting rainforests and taking snorkelling tours. Another traveller might be more interested in history and visits to Pearl Harbor and Diamond Head. With opportunistic links, both travellers will get the exact same ads that are based solely on destination [1]. Opportunistic links are automatic links generated by a set of rules triggered by user’s online activity. For example, if a user purchases an airline ticket to Honolulu, Hawaii from a travel agent website, he/she may get additional links added to their ticket order web page that advertise hotels, excursions and rental car deals in Hawaii, or an email with these same links may be sent to the customer. Although opportunistic links can be combined with other types of personalization technologies, simple opportunistic links do not consider a user’s personal preferences or prior purchase behaviour.

e. Recommender System

The objective of any recommender system is to recommend to a user, a new product/ article which the user has not already used or experienced but is very likely to choose from the plethora of options available to him/her. A comprehensive and state-of-the-art survey of recommender systems can be found in Chandrashekhar & Bhaskar [12] wherein the authors explain the recommendation problem both intuitively and mathematically. Recommender systems provide personalized recommendations based on user’s preferences. Collaborative filtering and content filtering are the two common technologies used to support the recommendations [7]. Collaborative filtering relies on the purchase and evaluation behaviour of past users. The preference profile of the current website user is matched to the profiles of all previous users of the website. Recommendations are made for product selection based on the products selected by the previous users that match the ‘preference profile’ of the current user. The preference profile for the user is either obtained implicitly by observation of the user’s purchase selections or explicitly through a user provided profile. Content filtering generates recommendations similar to collaborative filtering, but instead of matching a user to other users, the user’s preference profile is matched to information known about each of the products on the website. The closest matches are then recommended to the user. Several recommender systems combine collaboration filtering and content filtering to improve recommendations. Recommender systems use the most sophisticated software techniques to accomplish personalization. They are also the most likely to provide a finer grain of personalization tuned for each unique user.

III. DIFFERENT APPROACHES OF PERSONALIZATION

To understand the user needs, is the major issue in personalization. The user interest can be specified explicitly by user or automatically i.e. implicitly build by the system. Web based personalization can be implemented in a variety of ways, such as:

a. **Rule-based**

“Rule-based personalization” is one of the forms of personalization. It builds the user interest model. The user interest model is very useful to find out certain patterns; which are most probably repeated by all users. This approach makes decision based upon predefined business rules as they apply to the segmentation of a user. It classifies human interaction based on their past activities. To classify the users into different categories, it needs the information related to each user like browsing history containing all visited pages, certain downloads, etc. It uses an “if-then” process and decides what contents should provide to the user. Rules can be based on the promotional logic, dependencies between visits, demographic analysis or on simple pairing like computer and printer. E.g. if a person is rich and buys a computer, then he will buy a printer [6].

b. **Market Basket Analysis**

This is also one of the personalization approaches, which is generated mathematically. It also provides recommendations, based on pages visited together by past users. Since, page-to-page association involves less knowledge about the user than page-to-user association. Therefore, this personalization mechanism can have an almost an instantaneous response. This has ‘historically’ identified some interesting combination of web pages, on the same query. The decision of combining the web pages is totally dependent on the type of the information and the content inside the document. If the web pages representing information of same subject, they combined together. If most of users accessing both the pages in sequence, then combine them. It will beneficial to provide combined pages to the user. Because, the chances of finding the other document also relevant is more [3].

c. **Collaborative Filtering**

This is the personalization approach that is based upon mathematical analysis of user preference or activity data. This analysis leads to predictions of what a given user will like or will be likely to do next. This method involves gathering data on user preferences and behavior and then using that data algorithmically to produce recommendations for new users. These similar users are often known as “mentors” or “neighbors”. This method known as “collaborative Filtering”, solves the problem of personalizing content for an essentially unknown user, avoids the problem of grouping users in categories to which they really don’t belong, and allows for unique combination of tests in one individual. This is often referred a “Community-based recommendation”. Collaborative filtering allows the creation of “web of one”, that is, individuals to whom web pages are custom targeted. In fact, the process occurs naturally with this method: Users are automatically recommended information that suit them and them alone, because collaborative draws recommendation from a variety of the other users with similar, but not exactly corresponding tastes. Furthermore, collaborative filtering occurs in levels of recommendations: aside from simply telling users that they’d like particular information, it indicates how much they might like them. Inevitably, this provides the user with new ideas that don’t radically differ from the old ideas, increasing the likelihood that the user will find the new ideas good ones [4].

d. **User Profile Based**

This is the process of making decisions based upon predefined group membership or stored user profile information. In this technique, the browsing history of user plays an important role. For each user, a separate model called user profile model is maintained; in which all the information related to user is stored. This information contains submitted query, visited URLs, weight of a document, number of accesses, date and time of visit, etc. These contents are extracted from visited documents filtering process. It will extract out all non-necessary words and only important keywords are stored in user profile. Maintaining this information and updating it time to time is the task of personalization system. The particular document with high weight kept at a high rate in search result next time. The rating in case of explicit specification, time last visited and number of access affects the weight of a document. Weight of a keyword is decided by its position in the document. Hence, each user is represented in his user profile.

IV. PERFORMANCE RELATED ISSUES

All the approaches of the personalization discussed above have certain benefits and some limitations. According to the application or situation, one can choose any technique which suits them most [2].

a. **Benefits**

The benefit of using the particular technique is expressed in detail below:

i. **Rule Based**

- I. Because it requires little user information, rule based recommendation can be used in conjunction with other technique as “fall back” method.
- II. It works efficiently in situations where particular result providing knowledge must be leveraged, where a “blind”, algorithm based method will not work.
- III. This method is great when search engine know its perspective user base really well and it is relatively stable.
- IV. This method is useful for user if he fell accurately into broad category. In this case, users get effective result, which is targeted only at him by search engine.

ii. **Market Basket Analysis**

- I. MBA improves on rule-based recommendation, in that identifies some less obvious combinations of web pages.
- II. Since, it can be used with less data, but doesn't require rule configuration. MBA acts as an effective “fall back” method for other recommendation methods. Because personalization can be provided “first visit”, this method of personalization is very popular in sites with less loyal repeat visitors and anonymous browsers.

iii. **Collaborative Filtering**

- I. Since, it doesn't attempt to make broad judgments about groups of people; it recommends pages more likely to specifically suit a user wants and needs.
- II. It can tell users not simply that they will like a document, but how much rating they will give it.
- III. Once a CF system is put in place, the system itself selects web pages to “push”. This removes mystery from the process of deciding which web pages to provide.
- IV. Once the system is in place, it automatically generates new recommendations, without any additional input besides a continual stream of new users.

iv. **User Profile Based**

- I. It can provide the effective and qualitative output to the user; which will satisfy user more.
- II. Users can select exactly what they want to see, they control their own profile and the information in that profile.

b. **Limitations**

The problem associated with each of these techniques is expressed in detail below:

i. **Rule Based**

- I. It involves an extensive, research-intensive setup, which can be difficult to manage, particularly with a large inventory. Essentially, new rules must be added for each new web contents.
- II. It involves complex rule maintenance. Effectively, search engine must change rules to account for every situation, or risk failing to meet user expectations and decreasing site's usefulness.

ii. **Market Basket Analysis**

- I. MBA assumes that people tend to think alike- since, user x visits two pages in sequence, user y will also desire to visit these pages in sequence.
- II. Requires a base of transactional data from which to work. Since, most search engine track user behavior.
- III. In some situations, the concept of user behavior is not immediately apparent, and must be developed programmatically.

iii. **Collaborative Filtering**

- I. In this technique, the system has to track out the different documents containing similar information and bind them.
- II. And for that CF requires a body of data before it can make recommendations for new web documents. This is one half of problem with “cold-starting” CF on a particular site.
- III. Similarly, CF poorly suits an environment with rapid inventory churn; it depends on web pages being currently stored in order to rate them.
- IV. New users represent the other half of the “cold-starting” problem; since, with new users, like it doesn't know anything about the preferences of user; it can't make new recommendations for these users.
- V. This technique requires another technique which supports it to provide the efficient result to the user.

iv. User Profile Based

- I. There is no guarantee of correctly identification of user interest and provide personalized result.
- II. The process is extremely manual. Users may often miss new contents if profile gets stale, etc.
- III. It should provide the security to the user profile to eliminate the misuse of it.

V. COMPARISON

In this section, the overall comparisons of all the techniques are discussed. It is represented in the tabular form below:

Table 5.1: Comparison between Different Techniques

Techniques	Merits	Demerits
I. Rule Based	<ol style="list-style-type: none"> 1. Require less data about the user. 2. Can act as a “fall back” method. 3. Can work properly only if all the information is correct. 4. Useful if users accurately belong to the particular category. 	<ol style="list-style-type: none"> 1. It involves an extensive, research-intensive setup, which can be difficult to manage, particularly with a large inventory. 2. New Rules must be added for each new web contents. 3. It involves complex rule maintenance.
II. Market Basket Analysis	<ol style="list-style-type: none"> 1. Provide Recommendation based on combinations of web pages. 2. Doesn't require rule configuration. 3. Require less data about the user. 4. Can act as a “fall back” method. 	<ol style="list-style-type: none"> 1. It requires a base of transactional data from which to work. 2. In this technique, performance of search engine totally depends on user behaviour which is not immediately apparent, and must be developed programmatically.
III. Collaborative Filtering	<ol style="list-style-type: none"> 1. It provides the relevant document with proper weights. 2. Once CF system is put in place, the system itself automatically selects web pages to “push”. 3. It doesn't require any extra effort from the user. 	<ol style="list-style-type: none"> 1. It requires sufficient user data to make recommendation for new web document, also called “cold-starting” 2. It poorly suits an environment with rapid inventory churn. 3. Another situation of “cold starting” arises when user is new.
IV. User Profile	<ol style="list-style-type: none"> 1. It can provide the effective and qualitative output to the user; which satisfy the user more. 2. Users can update their own profile by entering their interesting document. 	<ol style="list-style-type: none"> 1. There is no guarantee of correctly identification of user interest. 2. This process is extremely manual. Users may often miss new contents if profile gets stale, etc. 3. It should provide the security to user profile to avoid its misuse.

ACKNOWLEDGMENTS

The authors thank D. Wu, M. Tremaine, K. Instone and M. Turoff for their valuable comments and edits in refining ideas for the paper. The research was supported by the information technology department of Sinhgad College of Engineering, Pune.

REFERENCES

- [1] D. Wu, I. Im, M. Tremaine, K. Instone and M. Turoff, “A Framework for classifying Personalization Scheme Used on e-commerce Websites”, 36th Hawaii International Conference on System Science, 2003.
- [2] Z. Dou, R. Song, J. Wen and X. Yuan, “Evaluating the Effectiveness of Personalized Web Search”, IEEE, Vol. 21 No. 8, 2009.
- [3] K. Sugiyama, K. Hatano and M. Yoshikawa, “Adaptive Web Search Based On User Profiles”, A Nara Inst. Of Science and Tech., 2007.
- [4] X. Wu, Y. Fu, S. Tian, Q. Zheng and F. Tian, “A hybrid Approach to Personalized Web Search”, IEEE, Vol. 12 No. 7, 2012.
- [5] A. Sieg, B. Mobasher and R. Burke, “Web Search Personalization with Ontological User Profiles”, ACM Press, 2007.

- [6] Sung Ho Ha, Seong Hyeon Joo and Hyun U. Pae, “Searching for Similar Informational Articles in the Internet Channels”, Kyungpook National University, Daegu, Korea, 2006.
- [7] M. J. Pazzani and D. Billsus, “Content-based Recommendation Systems”, Rutgers University, 2007.
- [8] N. Matthijs and F. Radlinski, “Personalizing Web Search Using Long Term Browsing History”, ACM Press, 2011.
- [9] A. Ankolekar and D. Vrandecic, “Personalizing Web Surfing with Semantically Enriched Personal Profiles”, Inst. AIFB, University of Karlsruhe, 2006.
- [10] B. A. Aseniero, S. Carpendale and A. Tang, “Deep Personalization in Tools for Reflection”, University of Calgary, 2011.
- [11] J. Goecks and D. Cosley, “NuggetMine: Intelligent Groupware for Opportunistically Sharing Information Nuggets”, ACM Press, 2002.
- [12] H. Chandrashekhar and B. Bhaskar, “Personalized Recommender System Using Entropy Based Collaborative Filtering Technique”, Journal of Electronic Commerce Research, Vol.12, No. 3, 2011.
- [13] A. Hannak, A. Kakhki and D. Lazer, “Measuring Personalization of Web Search”, International World Wide Web Conference, 2013.