

Comparison of Different Web Personalization Mechanisms

Enam Shah

*Department of Computer Engg
Dwarkadas J Sanghvi COE
Mumbai, India*

Hitarth Shah

*Department of Computer Engg
Dwarkadas J Sanghvi COE
Mumbai, India*

Khushali Deulkar

*Department of Computer Engg
Dwarkadas J Sanghvi COE
Mumbai, India*

Abstract- A Web is the world's widest collection of data; user may spend lot of time on the web to acknowledge relevant data or services with which they are concerned with. Web personalization is to offer clients with appropriate information and services based on the priorities of information that the client wants. Concept of personalization is to make web site sensible to the specific desires or requirements of particular user or group of users. In this paper, we have deliberated comparison of various approaches depending on different parameters.

Keywords- Personalized web search; Association Rule, content filtering, product filtering, Recommender system.

I. INTRODUCTION

The internet is flooded with a lot of useful and useless information. It is very hard to define useful information for a particular user which is varying from time to time. The useful information of one particular time may not be useful on different time or a different situation. The web itself is concerning day by day with newer technologies. Since World Wide Web(WWW) is utilizing free style medium that accepts structured, non-structured, ordered, non-ordered format to provide an information in the web, finding not only the relevant information but to plan them according to the interest of a user is also a key challenge today and is known as web personalization.

"Web personalization[1] is the approach of tailoring a website or data of website to the requirements of specific user or set of users, taking benefit of the information attained through the examination of the users browsing action [2]". Personalization means to provide services to each specific user in a tailored manner. The dissemination of information on the internet has terminated the personalization system an obligation. The personalization method must have the capability to resolve the additional data difficulties and let the customers practice at least exertion to find the data they require [3].

When the user work on the web page their activities can be classified into two wide sets: browsing and searching. If the customers want to search data on internet, they want to deliver the web scheme with a search demand. If the users don't provide the web system a clear-cut search request, the system would return huge volume of in appropriate data which can't fulfill the requirement of the user. If the

customers' request is explicit, searching data on the web might turn out to be simple [4]. To deliver exceptional recommendation to the user the web personalization system's designers must analyze what the customers' behavior when they browse on the internet.

The paper is organized as follows: the second section describes different techniques used in web personalization; Third section compares and summarizes various techniques. In the last section conclusion is stated.

II. RELATED WORK

Classification of web personalization approaches is based on the data utilized by the user, the way data is processed, interaction with users, learning model and the place where web personalization process is done.

A. Category Based Web Personalization System

In category based system [5] the two mechanisms used for web personalization are rule based personalization and collaborative filtering. Collaborative filtering allows customers to take advantage of other customer's interactive action and behavioral activities based on the degree of likeness between them. This approach requires users to disclose some personal information on their interest likes and dislikes. It should also comprehend detailed preferences, input by the user, which can then be matched against other sets of preferences input by other users. In rule based personalization instead of matching users reply to the web content or summaries of other users, this system match user query to some immobile rules or convention, about customer performance.

The following components were designed to implement all the above logic:

- System Logger
- Category Generator
- Customizer

The function of system logger is to gather users web usage information. Web Server Log collects user's visiting counts on every hyperlink on the web pages. Some Data mining techniques are applied to the log to extract data from the log. To identify which user belongs to which group category generator is used. Category Generator can

group the users into different categories on the basis of log data. To customize the throughput data to the user customizer is used.

In this schema, author proposed a new system called as novel system logger. This system is designed to stock all the data accessed by a particular user, and along with category generator, which splits the content into various categories and maps the most relevant result to the user. However, the user interest is not precisely reflected when the user possess large amount of visits on the web site.

B. Multilevel Web Personalization

To understand the two components of Web namely users and services, the author proposed a unified model to incorporate all techniques and mechanism of personalization. By means of unified model, any ideology can be inferred from multilevel views in pattern of hierarchies, and each view consists of various levels. Diverse views and levels offer improved web recommendations and personalization systems to the target user. To achieve this author made use of Web personalization function to measure the fulfillment of a service to a user [6].

The multilevel structure is emphasized, it is more extensive and appropriate for theory formation in a multilevel model can strengthen the understanding of the Web personalization concept itself, and provides more precise suggestions. If order of attributes which define the users is changed then different hierarchy can be constructed for the same attributes so, the complexity of model is increased.

C. Fuzzy Logic Based Product Filtering for Web Personalization

In this accession, the author recommended a system based on content-based product filtering contrary to rule based product filtering and collaborative filtering. In this approach two information segments and two processing segments are deliberated. To collect the customer and service information, information model is used and processing models are used to measure the client affinity and product filtering which are the vital mechanisms in this method [7].

The fuzzy logic method also reinforces preference learning which examine the imprecise data or facts from user's actions. The recommended system can handle ambiguities and uncertainties with the help of fuzzy sets and linguistics fuzzy rules and simultaneously offer user intelligible fuzzy rules to allow experts to augment the rules. Fuzzy sets are labeled as a numerical method and the representation of fuzzy sets is hinged on membership function. The membership function addresses how each distinctive fact from input space is mapped to a membership value in the scale of [0, 1]. Based on user's action and their access time, this method produces the most

relevant and meaningful results. The act of developing admissible membership function for fuzzy logic is a big challenging argument in fuzzy systems as it influences the correctness of fuzzy logic.

D. Web Page personalization Based on Weighted Association Rule

In this technique, by allotting a tangible weight to each item within the transaction, the author elongated the conventional association rule method to inspect the importance of each element within the transaction and extend a new procedure based on projected weighted association rule mining mechanism. In this weighted association rule [8] mining method, a measurable weight to each element was prescribed by the author based on the time spent on each page and visit count of each page rather than the customary binary weights. In the weighted schema, time spent on a page and visit count of a page is used to examine its significance in each transaction.

In the association rule based recommendation, the analysts purported all the access web pages are analogously important to user browsing pattern profile and neglected the distinctness in the importance of the pages in user duration. Duration is the time consumed on each page which apprises the significance of each page to the particular user, as user invest extra time on more fascinating item or else user might turn to new page. However, a rapid skip might also happen due to the least size of a page so the dimension of a web page may influence the real time spent on each item. Frequency is the computation that a web page is fetched by various clients. Higher frequency of web pages reveals the higher interest to client about those pages. Frequency count and time spent on a page are the vital statistics in estimating the user's interest on the web page.

The more interesting items as reported by the individual user are allotted higher weights. This particular weight is used to approximate the importance of a web page to a user, and try to provide improved recommendation which is highly fruitful to the user. But by considering the time spent on a page and frequency count the complexity of the system is increased.

E. Web personalization Based on User Access Patterns

In this technique, author suggested a system that uses sequential access pattern mining. A compact data model known as Pattern-tree, are constructed to store the sequential web access pattern, which is then used for user pattern matching and producing web links for recommendation. Prior to implementing sequential pattern mining preprocessing of data must be applied on the system logger which stores key information such as date-timestamp, client IP address, user ID, requested URL and HTTP status code. Data cleaning and data transforming are the main techniques used in data preprocessing. To determine the all sequence pattern with minimum support,

each order is a record of transaction arranged by transaction access time with each transaction in a sequence record [9].

On the basis of user’s current access sequence the most relevant access track in the pattern tree is found with the help of Recommendation Rule Generation module. The relevancy of the suggested model is approximated by considering some estimation measures such as Satisfaction and Precision. The recommended system has accomplished a frequent sequence however it is problematic to determine which child node should suggested by the system to the user and the support can’t identify the actual significance of the pages.

III. SUMMARY OF VARIOUS PERSONALIZATION APPROACH

In this part, we have delineated the various techniques used with its advantages and disadvantages and based on different parameters such as output correctness and perplexity we have summarized the above techniques. Basically output correctness means the appropriate data delivered to the user which is decided as low, medium and high and perplexity means the number of iterations required for computation which is also decided as low, medium and high. Various techniques, parameters and pros and cons of each the above mentioned techniques are shown in Table.

Table 1comp

Sr. No	System	Focus	Mechanisms Used	Output Correctness	Perplexity	Pros	Cons
1	Category Based web personalization [5]	Web Content, Rules	Collaborative Filtering, Observational personalization	Low	High	User can gain benefits from other user’s if similar interest	Rule based personalization depend upon the users astuteness
2	Multilevel web personalization [6]	Hierarchal Structure	Personalization function	Medium	High	Any ideology can be understood by multiple views	Individual attribute may have the many structures
3	Fuzzy Logic Based Product Filtering for web Personalization [7]	Product Filtering	Fuzzy logic	High	Medium	Deal with ambiguities for improved personalization	Correctness of fuzzy system is difficult to maintain
4	Web Page Personalization Based on Weighted Associations Rules [8]	Hyperlink structure with time constraint	Weighting Schema	High	High	Recommend the highly weighted pages to the customers based on their priorities	No pattern is followed
5	Web Mining Based On User Access Patterns For Web Personalization[9]	Extract Useful Pattern	Data Mining Techniques	Medium	High	Suggest web pages based on the customers access sequence	Support of the sequence cannot identify the true significance of sequence

IV. CONCLUSION

In this paper, we have reviewed techniques for web personalization and in basis of parameters a brief summary is provided which consists of comparison of all the above discussed approaches with its loss and benefits. However, the motive of all the papers is to make more précised web page recommendations and offer relevant and satisfactory information and facilities to specific user or group of users at different point of time by these personalization systems. Some techniques rely on the content of the web pages and user’s priorities and some of the methods comprises of clustering and data mining techniques.

REFERENCES

[1] G. Alkhatib, Z. Maamar and S.K. Mostefaoui, “Context-based Personalization of Web Services Composition and Provisioning,” Proceedings of the 30th EUROMICRO Conference (EUROMICRO’04) IEEE 2004.

[2] M. Vazirgiannis and M. Eirinaki, “Web Mining for Web Personalization,”ACM Transactions on Internet Technology, Vol 3 February 2003.

[3] M. H.Sadreddini and S. Khonsha, “New Hybrid Web Personalization Framework,” IEEE 2011.

[4] S.Y. Ho and K.Y. Tam, “Web Personalization: Is it Effective?,” International Conference on Computer & Communication Technology (ICCCT)- IEEE 2003.

[5] Ching-Cheng Lee and W. Xu, “Category-Based Web Personalization System,” International Conference on Web Information Systems and Technologies, IEEE, pp.1372-1377, 2001.

[6] Y. Yao, Y. Zhao and N. Zhong, “MultilevelWeb Personalization,” International Conference on Web Intelligence (WI’05) IEEE 2005.

[7] C.C.Fung, K. Wai Wong and B.Hua, “Fuzzy Logic Based Product Filtering for Web Personalization In E-Commerce,” IEEE 2007.

[8] A. Ghari Neiat, R. Forsati and M.R. Meybodi, “Web Page Personalization based on Weighted Association Rules,” International Conference on Electronics Computer Technology, IEEE, pp.130-135, 2009.

[9] L. Yue and W. Xiao-Gang, “Web Mining Based on User Access Patterns for Web Personalization,” IEEE, pp.194-197,2009.