International Journal of Computing Science and Information Technology, 2015, Vol.3,Iss.3, 18-26
ISSN: 2278-9669, July 2015 (http://ijcsit.org)

Web Personalization using web usage mining: applications, Pros and Cons, Future.

Sneha Prakash
Dept of Computer Science, DePaul Institute of science and Technology,
Angamaly,Ernakulam,kerala,683573,Swans.inn@gmail.com

Abstract

Now a day's browsing web has become an important part of our life style. The data flow inside the web is very easy and speedy as a result the web has become the Centre of data transactions. The different users have different experience in web. So we can say that the web is a personal experience for each user. Hence we need to personalize the web according to the view of each users making the web a personal experience. We do this web personalization with the help of a technique called web usage mining. Web usage mining means analyzing the data generated by web surfer's sessions or behaviours.The behavior of a surfer is the clickstream data generated by the surfer.

Keywords: web usage mining, clickstream, web Personalization.

1. Introduction

Web Personalization means making the web usage a personal experience for the user. This is done by suggesting the user some links, sites ,text ,products or messages.so the user can easily access the information he needs which will provide the user a feel that he is using his personal web. According to [1]Web personalization can be described as any action that makes the Web experience of a user customized to the user's taste or preferences. Principal elements of Web personalization include modeling of Web objects (such as pages or products) and subjects (such as users or customers), categorization of objects and subjects, matching between and across objects and/or subjects, and determination of the set of actions to be recommended for personalization.so we can say Web personalization can be defined as any action that tailors the Web experience to a particular user, or set of users [2].

2. Web usage mining

Web usage mining is a method in which we use the data mining techniques to identify patterns of users web usage .we make use of the clickstream data which reveals the users behavior and interest .The clickstream data means the data generated while the user do mouse clicks on various links, pictures or products. According to [1] the process of Web personalization based on Web usage mining can consist of three stages

- Data preparation and transformation.
- Pattern discovery.
- Recommendation.

2.1. Data preparation and transformation

In this phase we transform raw Web log files into transaction data which is processed with the help of Data mining tools. This phase also includes data integration from multiple sources, such as backend databases, application servers, and site content.

The data obtained through this process can be divided into four

- Usage data.
- Content data.
- Structure data.
- User data.

2.2. Pattern discovery

In the pattern discovery phase we discover the user behavior patterns by applying variety of techniques on the data obtained from data preparation and

transformation phase. The techniques we apply here is the clustering of data, association rules for mining and sequential pattern discovery.

2.3. Recommendation

In the recommendation phase the web personalization is done on the basis of user's active content and discovered patterns. This is done by analyzing the patterns obtained in pattern discovery phase and identifying the user's interests from them.

3. Web personalization

Web Personalization is one of the major application area of web usage mining. The web is a huge repository of millions of data .the data includes web pages, links, products etc. The ease and speed of web access has entertained the e-commerce systems. Now the web has become a major market of business and center of information systems. Therefore the number of users driven to the web increases day by day. The users need a customized/personalized web experience so that they can easily access the web they need. The business systems also entertain this because by making web personal they can easily track the user's interest and they can suggest the products according to users taste. Recent studies say that this has helped e-commerce systems to accelerate their business.

So many works are performed in this field of web personalization. according to [3] Web personalization can be seen as an interdisciplinary field that includes several search domains from user modeling, social networks, web data mining, human machine interactions to web usage mining. Web personalization is the process of customizing a Web site to the needs of each specific user or set of users, taking advantage of the knowledge acquired through the analysis of the user's navigational behavior [4]. Integrating usage data with content, structure or user profile data enhances the results of the personalization process [5].ultimately we can say that web personalization is done to provide each user their personal web.

Web personalization has been recently gaining great momentum in research and in various commercial web applications. One of the interesting applications of personalization on Web is the recommender systems [6].the recommender system analyze the users usage patterns using techniques association rules, clustering and markov models. According to the analyzed pattern a result is produced and the recommender system recommends this result to the users as a result the personalization of the web is acquired**.**

4. Web personalization strategies

The different strategies that we apply for web personalization are

4.1. Memorization

In this strategy of personalization, user information such as name and browsing history is stored (e.g. using cookies), and it is used to recognize and greet the returning user. It is usually implemented on the Web server. This mode depends more on Web technology than on any kind of adaptive or intelligent learning. It can also pose a threat to user privacy

B 4.2 .Customization

This form of personalization takes as input a user's preferences from registration forms in order to customize the content and structure of a web page. This process tends to be static and manual or at best semi-automatic. It is usually implemented on the Web server. Typical examples include personalized web portals such as My Yahoo and Google.

C 4.3. Guidance or recommendation system

A guidance based system tries to automatically recommend hyperlinks that are deemed to be relevant to the user's interests, in order to facilitate access to the needed information on a large website. It is usually implemented on the Web server, and relies on date that reflects the user's interest implicitly or explicitly. This approach forms the focus of Web personalization.

D 4.4. Task performance support

In these client side personalization systems, a personal assistant executes actions on behalf of the user, in order to facilitate access to relevant information. This approach requires heavy involvement on the part of the user, including access, installation, and maintenance of the personal assistant software. It also has very limited scope in the sense that it cannot use information about other users with similar interests.

5. Approaches to web personalization

There are mainly three approaches to web personalization [3].

5.1. Filter bubble[8]

Filter bubble is a result of personalized search in which the web site identify some patterns according to users behavior and as a result the user is given a personalized web in which he is isolated from the data's that disagrees his view point .eg:google personalized search results . According to [8] users get less exposure to conflicting viewpoints and are isolated intellectually in their own informational bubble. Pariser related an example in which one user searched Google for "BP" and got investment news about British Petroleum while another searcher got information about the Deepwater Horizon oil spill and that the two search results pages were "strikingly different".

5.2. Crowd turfing[9]

There is a growing underground market on the Web for malicious crowd-sourcing. For just a few cents, you can buy Facebook likes, Twitter followers, bulk social networking accounts, and fake reviews on Yelp. These types of social spam are extremely difficult for existing security systems to stop because the damage is caused by real people, not automated programs. By using web personalization we can try to stop crowd turfing..

5.3. Social sybil's[10]

Fake accounts, otherwise known as Sybil's, are a pervasive threat on the social web. Sybil's generate a large portion of the spam on social networks, and steal personal information that is used to power targeted phishing attacks. Examples: Spam on social networking websites like Facebook and Twitter.

6. Applications of web personalization

Web personalization had brought tremendous changes in users experience in the web. Personalization, also known as customization, involves using technology to accommodate the differences between individuals. Web pages are personalized based on the characteristics (interests, social category, context ...) of an individual. Different areas were personalization is applied are [11]:

6.1. E-Commerce Web Sites

Websites can identify visitors that have purchased in the past and address them with a much more personal experience, and display relevant ads, banners and product recommendations, based on each visitor's interests and previous purchases, to increase overall customer satisfaction and customer loyalty.

E.g.:www.amazon.com Due to personalization when we visit this site (if we are a registered customer) the site shows name of the person in the home page e.g.: Edwin's amazon.in (even we are not logged in) and instead of orders it shows Hello, Edwin your orders.

6.2. Content websites

E-newspaper websites and portals can display highly-relevant ads based on the visitor's segment to increase CTR and revenue from advertising, as well as relevant news and articles, based on the visitor's favorite categories and subjects to increase visit duration, number of articles read, and percentage of returning visitors

.

6.3. Social Medias

Online social media have become respected tools for content sharing and relationship maintenance. People create digital identities to distribute videos or photos, share opinions about books and movies, or just maintain contact Information of their friends. Social Medias (e.g. Facebook) now use the personalization technique on its users based on his interest and his likes and give suggestions to users like people he may know and customizes his home page according to user's interest there by increasing popularity of the media. Now a days the social Medias have become a tool for online marketing also.

6.4. Gaming websites

Websites can identify visitors that are interested in a specific subject, product or service and target them with a much more personal experience with relevant offers, banners and recommendations, based on each visitor's interests to increase registration and sign-up rates.

7. Pros and cons of web personalization

Web site personalization can be termed as "a website per visitor".it is all about customizing the web for visitors.

7.1. Pros for web personalization

Personalization strategies and personalized services may help facing and overwhelming the problem of information overload and mass customization. The different advantages of personalization: customer satisfaction and consumer loyalty, big data, key performance indicators and problems in e-commerce [12].

7.1.1. Customer Satisfaction and Consumer Loyalty

Personalization in the area of marketing and E-Commerce tends to increase customer orientation .If these impacts on customer satisfaction and consumer loyalty are known in detail, the process of personalization can be efficiently improved and adapted to consumer needs and preferences.

Successful personalization strategies guarantee unique values and benefits for each customer. Only if customers are seen on an individual level, it is possible to propose special and individualized offers to them. The longer a customer is interacting with the firm's website, the more accurate information can be collected and subsequently analyzed. Proposals for personalized offers can then improve customer loyalty and hence increase the profits. Trust is an important element when focusing on personalization in the fields of marketing and hence on consumer loyalty. Customers should believe that a service provider "will not take advantage of the relationship to enrich himself at the customer's expense, and will deliver what is required by the customer, not just what is convenient for the firm" [13]. Personalization hence helps marketers and businesses to enhance customer satisfaction, create Additional consumer loyalty and thus to improve the binding of the customers.

7.1.2. Reducing the problem of big data

The mass of data constitutes a certain challenge for businesses. There is a need for smart algorithms to reduce the abundance of the resulting data. Only by accurate search parameters a reasonable data output for these amounts can be guaranteed leading to useful information. This is one task of personalization; it tries to over- come the problem of big data and filters only relevant information for business on the one hand and for customers on the other hand. One clearly recognizes the need for personalized systems in order to provide customers the information that is tailored to their needs [14]. The digitization process in general and big data in particular advances personalization. Rutherford and Botha [16] highlight the positive impacts that personalization may have: alleviating information overloads and describing the additional benefits in E-Commerce. The ultimate goal is to filter unnecessary information out of the research results and perfectly provide only the information the user wants and needs. Parekh [17] also states that personalization contributes positively to overcome the information overload. He concentrates on content personalization and states that a user can set his personal settings. According to his point of view, it is the user who has to react actively, not the system itself.

7.1.3. Raise Business Key Performance Indicator

The traditional parts of economic success, as for example turnover, profit, gross margin or customer-lifetime-value, are naturally important indicators for the performance of the business. Riemer and Totz [18] investigate the impacts of personalization on customer retention by means of switching costs. Switching costs or switching barriers are used in microeconomics or marketing in order to describe to costs of customers when changing the supplier or service provider. Following [18], switching costs are divided into direct switching costs, opportunity costs and sunk costs. By the construction of switching barriers for customers, personalization contributes valuable to the rise of business key performance indicators. Hence, it can be said that the willingness to migrate to other providers and vendors decreases with an increase in switching costs and with an increase in customer loyalty and relationship. Personalization is hence utilized to save costs

7.1.4. Overcome challenges in E-commerce

The personalization process aims at providing personal product recommendations depending on various criteria as for example the customer's purchase behavior. For marketers in E-Commerce sectors, it is therefore necessary to create recommenda- tion techniques. Goy et al. [19] describe the usage of so-called "dynamic taxono-mies" which goal is to gradually reduce the search space of an E-Commerce website. The idea behind this approach is, that the user knows better about his Preferences and that the system should not try to find his preferences on its own. Other Ecommerce providers use the feature "products similar to products already bought" or "customers who bought" as for example Amazon. As

the sector of E-Commerce has experienced a non-deniable evolution over the last decades, many nations have adopted special laws and regulation in order to put a certain control on it and regulate, which beneficial approaches of personalization can be legally utilized by businesses which are active in E-Commerce.

7.2. Cons of web personalization

7.2.1. Privacy Concerns

When discussing the research area of personalization, it becomes clear that users often are uncomfortable to share private information with websites. User profiles are created through special algorithms representing the user's personal interests. The information can be based on the browsing history, emails or other kind of user-related data. Items that can be found frequently are regarded as being interesting to users. Xu et al. [20] propose to create a hierarchical user profile according to the items' frequency of occurrence. The information delivered to the consumer is personalized and coordinated to his personal interests. Anton et al. [21] prove that privacy concerns of personalized web search have experienced an increase, which might be the result of an intensification of online Advertising and E-Commerce.

7.2.2. Echo Chambers

The Echo Chamber Effect was described by Weinberger [22] in 2004. This effect is based on the well-known phenomenon that people like to surround with people who share related views or even the same point of view. Since the rise of the Internet , some people tend to mainly reside in those online communities in which their own views are represented. In other terms, they listen to their own echo.

This can result in a situation where, after a while, people believe that almost all other persons or at least the nice and reasonable, share their own views. This sounds relatively harmless -nevertheless, the implications need to be assessed critically. The consequence may be that people are fixed to their own views, because they tend to believe that a majority is thinking in the same way, which approves their individual opinion and view. Weinberger [22] states that the Echo Chamber Effect on the Internet is not that frequent. He convincingly explains that the milder form has always existed: in the selection of the papers and television broadcasts that people read and see [22]. The huge challenge of avoiding echo chambers is to become proactive by looking for other information sources and opinions

7.2.3. Technological Risk

Zhang [23] describes a technological approach towards the risk of personalization, where "behavior-based personalization can hurt the profits of competing firms". Following this contribution, behavior-based personalization constitutes a peril because the result is price discrimination between firms ending in an even more intensified price competition. Behavior-based personalization influences companies' profits in a negative way.

7.2.4. Legal Concerns

Finally, the importance of privacy laws in the online sector has risen – not only because of the application of personalization by businesses. Online advertising consists nowadays more and more of personal advertising, so that certain Internet users respond annoyed because of the personal information which is known about them. Special technologies to collect the online history of a user and analyze the online behavior are utilized. Especially the European Union contributes in harmonizing legal standards throughout the member countries. The directive on data protection has been amended and intensified recently.

8. Existing research works

Lots of research have been conducted in the area of personalizing web. Some of the research works are

Xuwei Pan et al., [24] proposed context-based adaptive personalized Web Search for improving information retrieval effectiveness. Recent Web search/meta search engines are constructed to serve all users, independent of the particular requirements of any individual user in dissimilar situations. Personalization of web search is to perform retrieval for each user integrating his/her interests. In this approach, the authors proposed a novel adaptive personalized technique based on context to adapting search outputs consistent with each user's requirement in different situations for relevant information with slight user effort**.** Experimental observations prove that the adaptive personalized search system is executed by most of users and the approach to personalize Web search is effective.

jie Yu et al., [25] suggested mining user context based on interactive computing for personalized Web search. Personalized Web search is a successful way of same query.

How to achieve user's real-time information requirement is a key subject in personalized search. Existing approaches focus more on constructing user profile which depends on Web pages/documents which influences the effectiveness [26] of search engine. Additionally, dynamics of user profile is frequently ignored. To deal with this problem, the authors have introduced a technique that acquires the user context to perfectly present preferences of users for successful personalized search in this paper. Initially, small-term query context is created from Web-snippets to take part a role of semantic background of user's search behavior, recognizing associated concepts of the query. Then, user context snap is constructed depending on query context based on user's interactive search behavior. Finally, development of user context is taken into account by introducing forgetting factor to combine the independent user context snap in a user session. The experimental outputs completely reveal that this technique can effectively construct user context based on individual user information need.

Kyung-Joong Kim et al., [27] developed a personalized Web search engine using fuzzy concept network with link structure. Most of the famous search engines make use of link structure to discover precision result. Typically, a link-based search engine provides superior-quality outputs than a text-based search engine. On the other hand, they have complexity in providing the result that satisfies the specific user's preference. Personalization is necessary to maintain a more suitable result. Among many approaches, the fuzzy concept network according to a user profile can characterize a user's subjective interest appropriately. The paper proposes another search engine that utilizes the fuzzy concept network to personalize the outputs from a link-based search technique. Depending on a user profile, the fuzzy concept network rearranges five outputs of the link-based search engine, and the system presents a personalized superior quality result. Experimental observations with the three subjects show that the system proposed searches not only appropriate but also personalized Web pages on a user's preference.

F. Akhlaghian et al. [28] proposed a personalized search engine using ontology-based fuzzy concept networks. At the moment, personalization of search engines as the only web search tool plays significant role in raising the speed of access to web information. Since the users may have various backgrounds and anticipations for a specified query, personalization of search engines outputs based on user's profile can assist to better match the overall interests of an individual user. In this paper the authors personalize the search engine outputs with the help of automatic fuzzy concept networks. The main objective is to make use of the concepts of ontology to improve the common fuzzy concept networks built according to user's profile. Experimental output shows enhancement in personalized search engine outputs using enriched fuzzy concept networks contrast to common fuzzy concept networks.

C. Biancalana e t al., [29] proposed a new way for personalized Web search using social tagging in query expansion. Social networks and collaborative tagging systems are quickly attaining more recognition as most important means for categorization and sharing data: users tag their bookmarks so as to make things easier for information distribution and later visit. Social Bookmarking approaches are helpful in two essential respects: first, they can permit an individual to remember the browsed URLs, and second, tags can be made by the group of people to direct users towards important content. This paper concentrates on the latter use. The authors present a novel method for personalized web search with the use of query expansion [30, 31]. This paper additionally extends the group of recognized co-occurrence matrix approach models by using a new method of exploring social tagging services. The proposed approach illustrates its power especially in the case of disambiguation of word contexts. This paper shows steps to plan and execute such a system in practice and performed numerous experiments on a real web-dataset. This is the first study focused on the use of social bookmarking and tagging approaches for personalization of web search and its performance in a real-world application.

9. Experimental methodologies

In [15] the authors described about the experimental methodology that they used.

9.1. Experiment design

According to[32] seeks to answer two broad questions. First, what user features influence Google's search personalization algorithms? This question is fundamental: outside of Google, nobody knows the specifics of how personalization works. Second, to what extent does search personalization actually affect search results? Although it is known that Google personalizes search results, it is not clear how much these algorithms actually alter the results. If the delta

between"normal"and"personalized"results is small, then concerns over the Filter Bubble effect may be misguided. In [15] the authors focused on measuring Google Search, as it is the most popular search engine. However, their methodology is Web service agnostic, and could be repeated on other search engines like Bing or Google News Search. At a high-level, their methodology was to execute carefully controlled queries on Google Search to identify what user features trigger personalization. Each experiment follows a similar pattern: first, create x Google accounts that each vary by one specific feature. Second, execute q identical queries from each account, once per day for d days. Save the results of each query. Finally, compare the results of the queries to determine whether the same results are being served in the same order to each account. If the results vary between accounts, then the changes can be attributed to personalization linked to the given experimental feature. Note that we run some experimental treatments without Google accounts (e.g., to simulate users without Google accounts). **Sources of Noise.** Despite the simplicity of the high level experimental design, there are several sources of noise that can cause identical queries to return different results.

• **Updates to the Search Index:** Web search services constantly update their search indices. This means that the results for a query may change over time.

• **Distributed Infrastructure:** Large-scale Web search services are spread across geographically diverse datacenters. Our tests have shown that different datacenters may return different results for the same queries. It is likely that these differences arise due to inconsistencies in the search index across datacenters.

• **Geolocation:** Search engines use the user's IP address to provide localized results [33]. Thus, searches from different subnets may receive different results. • A/B Testing: Web search services sometimes conduct A/B testing [34], where certain results are altered to measure whether users click on them more often. Thus, there may be a certain level of noise independent of all other factors.

The Carry-Over Effect. One particular source of noise comes from the influence of one search on subsequent searches. In other words, if a user searches for query A, and then searches for query B, the results for B may be influenced by the previous search for A. they term this phenomenon the carry-over effect. Prior research on user intent while searching has shown that sequential queries from a user are useful for refining search results [26,35] so it is not surprising that Google Search leverages this feature. They started two different browser instances: in one we search for the first query, wait, and then for the second query, while in the other they searched only for the second query. they repeat this experiment with different wait times, and re-run the experiment 50 times with different query pairs. Finally, they compared the results returned in the two different browser instances for the second term. The carry-over effect can be clearly observed: the results share, on average, seven common results (out of 10) when the interval between the searches is less than 10 minutes (in this case, results pertaining to Turing Tests are included). After 10 minutes, the carry-over effect disappears. Thus, in all experiments in the following sections, they waited at least 11 minutes between subsequent searches in order to avoid any carry-over effects. In their testing, they observed carry-over for both logged in users and users without Google accounts.

Controlling Against Noise. In order to mitigate measurements errors due to these factors, they performed a number of steps (some borrowed from [32]): First, all of their queries are executed by the normal Google Search webpage, rather than Google's Search API. It has been shown that search engine APIs sometimes return different results than the standard webpage [25]. Second, all of their machines execute searches for the same query at the same time (i.e., in lock-step). This eliminates differences is query results due to temporal effects. This also means that each of their Google accounts has exactly the same search history at the same time. Third, they used static DNS entries to direct all of their query traffic to a specific Google IP address. This eliminates errors arising from differences between datacenters. Fourth, they wait 11 minutes in-between subsequent queries to avoid carry-over. A 11 minute wait is sufficient to avoid the majority of instances of carry-over. Fifth, unless otherwise stated, they send all of the search queries for a given experiment from the same /24 subnet. Doing so ensures that any geolocation would affect all results equally. Sixth, they included a control account in each of their experiments. The control account is configured in an identical manner to one other account in the given experiment (essentially, they run one of the experimental treatments twice). By comparing the results received by the control and its duplicate, they can determine the baseline level of noise in the experiment (e.g., noise caused by A/B testing). Intuitively, the control should receive exactly the same search results as its duplicate because they are configured identically, and perform the same actions at the same time. If there is divergence between their results, it must be due to noise.

10. Conclusion & future scope

According to latest studies in future the techniques will be developed to do deep personalization and by this the user who are logged in will get result according to his behavior which will seem to him like the whole web is personalized especially for him and if the user is not logged in then the results will be based on the location from where he is searching.in future our web will be much more like us. Imagine opening up any web page or application and being presented with an experience that's entirely personalized to you. Go to ESPN.com and see stories about the sports you love and teams you follow featured on the top. This is where things are headed. It's about shifting from you trying to find the right information to the right information finding you.

In the past, we lacked the data and the technology to make this type of personal experience a reality. But that's changing quickly. The abundant social data that's overwhelming our social streams not only presents a problem but the solution. Using natural language processing and semantic analysis to evaluate your tweets, status updates, like, shares, and check-ins, it's possible to build a holistic understanding of who you are and what you're interested in.

Once the web knows your interests, it can start to change… Any website or app can use knowledge of your interests in order to give you a personal experience.

REFERENCES

1. Mobasher, B., "Web Usage Mining andPersonalization", in Practical Handbook of InternetComputing, M.P. Singh, Editor. 2004, CRC Press. p. 15.1-37.
2. Bamshad mobasher,Robert cooley,jaideep Srivastava,"Automatic personalization based on web usage mining" August 2000/Vol. 43, No. 8 COMMUNICATIONS OF THE ACM.
3. Beatric,ryan Fernandez,leo j peo,nikhila kamat,sergius Miranda ," New Approaches to Web Personalization UsingWeb Mining Techniques" Beatric et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 2195-2201
4. Eirinaki, M.,Vazirgiannis, M.,"Web mining for web personalization", ACM Transactions on Internet Technology, Vol.31, pp. 1-27, 2003.
5. Eirinaki, M., Vazirgiannis, M.,Varlamis, I.,"SEWeP: using site semantics and a taxonomy to enhance the Web personalization process", Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, Washington DC, USA, pp. 99-108, 2003
6. Faten Khalil Jiuyong LiHua Wang," Integrating Recommendation Models for Improved Web Page Prediction Accuracy", Conferences in Research and Practice in Information Technology (CRPIT), 2008,Vol. 74.
7. EirinakiVazirgiannis M., "Web mining for web personalization.ACM Transactions on Internet Technology (TOIT)", 3(1), 1-27, 2003.
8. Eli Pariser," The Filter Bubble: What the Internet Is Hiding from You", Penguin Press (New York, May 2011) ISBN 978-1-59420-300-8
9. Gang Wang, Christo Wilson, Xiaohan Zhao, Yibo Zhu, Manish Mohanlal,Haitao Zheng and Ben Y. Zhao ," Serf and Turf: Crowdturfing for Fun and Profit" WWW 2012, April 16–20, 2012, Lyon, France. ACM 978-1-4503-1229-5/12/04 or.arXiv:1111.5654v1 [cs.SI] 24 Nov 201
10. Zhi Yang, Christo Wilson,Xiao Wang, Tingting Gao, Ben Y. Zhao ,yafei dai "Uncovering Social Network Sybils in the Wild" https://www.cs.ucsb.edu/~ravenben/publications/pdf/sybil-imca11.
11. http://personyze.com/categories/pers_examples
12. Johannes Frank, Michael J. Harnisch," Review on Benefits and Risks of Personalization and Solutions for Privacy Concerns*"* Computer Communication & Collaboration (Vol. 2, Issue 1, 2014) Submitted on 22/July/2013 DOIC: 2292-1036-2014-01-004-17.
13. Ball, A., Coelho, P.,and Vilares, M. J. (2006). "Service Personalization and Loyalty", Marketing Department Faculty Publications, 13:1-42.
14. Anand, S., and Mobasher, B. (2005). "Intelligent Techniques for Web Personalization", Lecture Notes in Computer Science, 3169:1-36.
15. Aniko Hannak ,Arash Molavi Kakhki, Alan Mislove , Christo Wilson ,David Lazer Northeastern University, Balachander Krishnamurthy AT&T Labs-research," Measuring Personalization of Web Search".
16. Rutherford, A., and Botha, R. (2003). "Towards Personalized Content", Proceedings of SAICSIT, 111-113.
17. Parekh, K. (2011). "Managing the Information Overload – a Findability Solution", Enterprise Search COE, White Paper, 1-5.
18. Riemer, K., and Totz, C. (n.d.). "The many faces of personalization – An integrative economic overview of mass customization and personalization", [Online] Available at:

http://pdf.aminer.org/000/306/617/personalization_ meets_mass_customization_support_for_th e_ configuration_and_design.pdf (August 8th,2012).
19. Goy, A., Ardissono, L., and Petrone, G. (2007). "Personalization in E-Commerce Applications", [Online] Available at: http://www.dcs.warwick.ac.uk/~acristea/courses/ CS411/2009/Book%20-%20The%20Adaptive%20 Web/PersonalizationECommerce.pdf (August 8th, 2012).
20. Xu, Y., Zhan, B., Chen, Z., and Wang, K. (2007). "Privacy-Enhancing Personalized Web Search", [Online] Available at: http://www2007.org/papers/paper247.pdf (August 13th, 2012).
21. Anton, A., Earp, J., and Young, J. (2009). "How Internet Users' Privacy Concerns Have Evolved Since 2002", North Carolina State University Computer Science Technical Report, 111.
22. Weinberger, D. (n.d.). "Echo Chambers = democracy", [Online] Available at: http:// rebooting.personaldemocracy.com/files/DavidWeinberger. pdf (August 13th, 2012).
23. Zhang, J. (2011). "The Perils of Behavior-Based Personalization", Marketing Science, 30(1): 170-186.
24. Xuwei Pan, Zhengcheng Wang and Xinjian Gu, "Context-Based Adaptive Personalized Web Search for Improving Information Retrieval Effectiveness," International Conference on Wireless Communications, Networking and Mobile Computing, Pp. 5427 – 5430, 2007.
25. Jie Yu and Fangfang Liu, "Mining user context based on interactive computing for personalized Web search," 2nd International Conference on Computer Engineering and Technology (ICCET), Vol. 2, Pp. 209-214, 2010.
26. Fang Liu, C. Yu and Weiyi Meng, "Personalized Web search for improving retrieval effectiveness," IEEETransactions on Knowledge and Data Engineering, Vol. 16, No. 1, Pp. 28 –40-2004.
27. Kyung-Joong Kim and Sung-Bae Cho, "A personalized Web search engine using fuzzy concept network with link structure," Joint 9th IFSA World Congress and 20th NAFIPS International Conference, Vol. 1, Pp. 81 – 86, 2001.
28. F. Akhlaghian, B. Arzanian and P. Moradi, "A Personalized Search Engine Using Ontology-Based Fuzzy Concept Networks," International Conference on Data Storage and Data Engineering (DSDE), Pp. 137 –141-2010
29. C. Biancalana and A. Micarelli, "Social Tagging in Query Expansion: A New Way for Personalized Web Search," International Conference on Computational Science and Engineering (CSE '09), Vol. 4, Pp. 1060 –1065-2009.
30. Zhengyu Zhu, Jingqiu Xu, Xiang Ren, Yunyan Tian and Lipei Li, "Query Expansion Based on a Personalized Web Search Model," Third International Conference on Semantics, Knowledge and Grid, Pp. 128 – 133, 2007.
31. P. Palleti, H. Karnick and P. Mitra, "Personalized Web Search Using Probabilistic Query Expansion,"International Conferences on Web Intelligence and Intelligent Agent Technology Workshops (IEEE/WIC/ACM), Pp. 83 – 86, 2007.
32. S. Guha, B. Cheng, and P. Francis. Challenges in Measuring Online Advertising Systems. IMC, 2010.
33. X. Yi, H. Raghavan, and C. Leggetter. Discovering Users' Specific Geo Intention in Web Search. WWW, 2009.
34. A. Pansari and M. Mayer. This is a test. This is only a test. Google Official Blog, 2006. http://bit.ly/Ldbb0.
35. Y. Shen, J. Yan, S. Yan, L. Ji, N. Liu, and Z. Chen. Sparse Hidden-Dynamics Conditional Random Fields for User Intent Understanding. WWW, 2011.