International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Survey of Privacy Protection in Personalized Web Search Using Metric Prediction

Sneha M. Durugkar[1], S. S. Barde[2]

[1]Savitribai Phule Pune University, Smt.Kashibai Navale College of Engineering, Vadgaon (BK), Pune-41

[2]Savitribai Phule Pune University, Smt.Kashibai Navale College of Engineering, Vadgaon (BK), Pune-41

Abstract: ***The utilization of World Wide Web is expanding quickly throughout the years. These days, web has gotten to be immeasurable wellspring of Information. In methodology of securing that information, search engines assume a paramount part. Number of search results are acquired or demonstrated via search engines, yet low quality and less exactness of search results make troublesome for client to increase information that is required. As of late search engines has thought of cutting edge systems. Users are not intrigued to invest time for characterizing queries for search. It has discovered that if search results are late and mistaken then users are not intrigued. Need of flow circumstance is to give prompt and exact search results. This need can be fathomed actualizing personalized web situations. As it is getting to be essential angle to give personalized web situations numerous methodologies and systems have advanced. At the same time in the meantime security of personalized web searches has additionally picked up significance in which personal/private information cannot be uncovered by web searches. Personalized web search is demonstrating its adequacy additionally raising issue of privacy and securing personal information. Numerous personalization methods have been found and executed. Yet it is not certain that those procedures will guarantee their viability in distinctive queries for diverse users. In this paper we will examine on need of personalized web search and securing personalized information.***

Keywords: Personalized web search, Personalization Techniques, Privacy, Information Retrieval, Immediate and Accurate Results.

1. Introduction

In widespread development of Internet use, obtaining obliged or planned information is likewise a vital point. Results of search ought to be reliant on user behavior. For that execution of personalized web environment is needed. Web search results ought to adjust to user necessities. Primary issue in search engines is same results are produced for distinctive users. Now and again these results are short and confounding [1]. Imperativeness ought to be given to regions like inquiry answer system, better perusing, virtual results, confinement and adjusted web search. Search engines ought to show results relying on the user. Numerous advances have been actualized for distinctive needs of users. At present search engines are serving all users same search results paying little heed to uncommon needs of user.

In spite of the fact that indexing algorithms are utilized as a part of customary search engines experiences issues in attaining productivity expected by user. To acquire information identified with user's advantage personalization is utilized. Generic web search varies from personalized web search in light of the fact that it gives same search results to diverse users for comparable queries. Information expected from diverse users may be distinctive while utilizing same query. Blended sort of search results will be recoverd. In the light of this time will be devoured to pick information that user needs. By knowing a few things about user, a search motor may refine user results to make them more important is the idea driving personalized search.

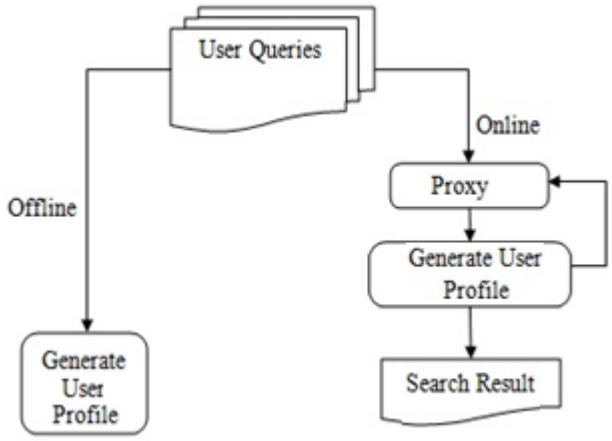


Figure 1: Personalized Web Search – General Architecture

Personalizing web search is the methodology of acquiring web search results relying on history or past behavior of user. Contingent on conditions, personalization could be possible at server or client side [12]. User may be meaning to accomplish numerous objectives in single query. Subsequently effectiveness of personalization of web search is relied on upon user behavior and query. In web personalization, perusing personalities of users is testing assignment to perform. Likewise there is impediment of words used to search. Two three words are utilized to search over web. That is expanding the test in personalizing web search. There is need of security in the personalized web search. Users are not quick to disclose their data amid web search [13]. This has gotten to be real issue in profiling the user in personalized web search. There ought to be a system which sum up profiles as indicated by data gave by user [7].Really more the search motor thinks about user, more exact search results will be acquired via search supplier. Anyhow users can't trust on search motor that data gave by user is not abused. Search motors can give more exact and

Licensed Under Creative Commons Attribution CC BY

particular data if users trust search motor and give more data. Consequently, search motors ought to give security system such that user will be guaranteed of its security and its data ought to be kept safe [8].There are mainly two areas where research can be focused. 1) By using personalization of the user profile, improve search result quality. 2) For keeping privacy risk under control, they need to hide private information of user present in user profile. Typically search is performed by providing queries to retrieval system in form of set of words. If different users enter same query, the system will produce same results without considering the user. But search results should be produced by taking the user in the equation, so that different users can get different search results for same query by keeping track of user's personal information and interests. There are fundamentally two zones where research can be centered. 1) By utilizing personalization of the user profile, enhance search result quality. 2) For holding protection chance under control, they have to conceal private information of user present in user profile. Ordinarily search is performed by giving queries to recovery framework in manifestation of set of words. In the event that diverse users enter same query, the framework will deliver same results without considering the user. In any case search results ought to be created by taking the user in the comparison, so that diverse users can get distinctive search results for same query by staying informed concerning user's close to personal information and interests.

2. Literature Review

This paper concentrates on a personalized web search and systems to attain personalization of client's web search.

1) Personalized Web Search:

There are primarily two sorts of personalized web search:
A. Click-Log-Based
B. Profile-Based

A. Click-Log-Based:
In this kind of personalized web search, personalization is completed on the premise of clicks. The data get recorded through clicks in query logs to mimic client experience. In this approach, the web pages as often as possible clicked by client in past history for a specific query is recorded and personalized score is registered for specific web page and web search results are as indicated by that score[9].

Click log based personalized web search methodologies will perform steady and significantly well when it is chipping away at rehashed questions. Inconvenience of click log based methodology is the point at which a never asked query is entered by client; it won't deliver exact and related search results. It has observed that one third of questions are rehashed and it will function admirably on that [1].

B. Profile-Based:
In this methodology, search results are sorted on the premise of personal enthusiasm of user profile. Fundamentally there are two techniques are utilized for making user profile. One is that utilizing words which are every now and again utilized as a part of documents. Extensive profiles are made by this methodology. It gives less precision and less proficiency in search result. Second approach is that utilizing predefined ontology, for example, DMOZ [1][2]. This methodology facilitates creation and upkeep of issue. This has observed that profile based personalized web search has ended up precarious when users history increments.

2) Chasing Personalization

At search engine, results must be downloaded where user profile is put away or a user profile must be communicated to server where web corpus is put away, to incorporate information about user. There are few explanations behind re-ranking of results by regional standards, 1) as we are taking a shot at generally little data; incorporation of computationally-intensive methodology is achievable, 2) for privacy reasons as users are uncomfortable with sending their data on the internet or to an obscure end of the line, 3) direct assessment can be given by re-ranking techniques. As opposed to gathering rating for all documents on the internet which is infeasible in nature, we have to consider just top results [3].

Three critical segments of our model are [3]:
1. Corpus Representation
2. User Representation
3. Document/Query Representation

1. Corpus Representation
As we are managing web search, our corpus will be web. There are sure parameters like number of documents present on web, number of documents that contains searched keyword. As users don't have immediate access to the points of interest of web it is turning to be an impediment of personalization. From all documents on web or documents identified with search keyword insights of web can be created. The corpus representation can be concentrated by the query entered from user.

2. User Representation
A rich index of personal substance is utilized to speak to a user. An index comprises of user's hobbies and computational activities. Email messages which are seen or sent, logbook sections, web pages went by user and documents exhibit in customer machine this can be incorporated in index content. By utilizing this information, rich however unstructured profile of user can be made. By treating each report as a wellspring of confirmation related the user's advantage, autonomous of the query, this index can be utilized.

3. Document and Query Representation
The representation of document is fundamental for figuring out what terms are incorporated and how regularly they are utilized. To get to full content of documents it will require significant investment so by getting to just title and snippets of documents is utilized to make representation of documents. Snippets of documents got via search engines are focused around query.

3) Information Retrieval:
There are numerous Information Retrieval frameworks. In the majority of these algorithms, information retrieval issue treated as one single query and documents sets. A large

Licensed Under Creative Commons Attribution CC BY

portion of existing retrieval models settles on choice built just with respect to query and document gathering. Information of real user and search setting is generally disregarded in existing information retrieval frameworks. Web search engines give search results relying on the query put together by user. To enhance retrieval exactness, extra connection information ought to be abused by an ideal retrieval framework at whatever point it is accessible. Significant test in information retrieval is setting delicate retrieval of information.

4)Privacy in Personalized Web Search:
Personalized web search is picking up more prevalence. However keeping up privacy is not kidding issue in personalized web search. As personalizing search obliges assembling and preparing of user information, which prompts privacy issue. This is turning into the principle obstruction in conveying personalized web search applications. Personalized obscurity is a security procedure which is executed to give privacy in personalized web search in which individual can determine level of privacy [5]. Anonymizing user profile is additionally method by which privacy of user can be kept up [11].

3. Conclusion

We have arranged a survey report for distinctive themes of Personalized Web Search. This report spreads issues like need of personalized web search, how personalized web search can be actualized, what are difficulties in it, privacy and security issue of it and existing arrangement of personalized web search. Based on such report, we surveyed Click-log-based PWS, Profile-based PWS, Pursuit of personalization, Information Retrieval, Privacy in Personalized web search.

References

[1] Z. Dou, R. Song, and J.-R. Wen, “A Large-Scale Evaluation and Analysis of Personalized Search Strategies,” Proc. Int’l Conf. World Wide Web (WWW), pp. 581-590, 2007.
[2] K. Ramanathan, J. Giraudi, and A. Gupta, “Creating Hierarchical User Profiles Using Wikipedia,” HP Labs, 2008.
[3] J. Teevan, S.T. Dumais, and E. Horvitz, “Personalizing Search via Automated Analysis of Interests and Activities,” Proc. 28th Ann.Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), pp. 449-456, 2005.
[4] X. Shen, B. Tan, and C. Zhai, “Context-Sensitive Information Retrieval Using Implicit Feedback,” Proc. 28th Ann. Int’l ACM SIGIR Conf. Research and Development Information Retrieval (SIGIR), 2005.
[5] X. Xiao and Y. Tao, “Personalized Privacy Preservation,” Proc.ACM SIGMOD Int’l Conf. Management of Data (SIGMOD), 2006.
[6] K. Ja¨rvelin and J. Keka¨la¨inen, “IR Evaluation Methods for Retrieving Highly Relevant Documents,” Proc. 23rd Ann. Int’l ACM SIGIR Conf. Research and Development Information Retrieval (SIGIR), pp. 41-48, 2000.
[7] Y. Xu, K. Wang, B. Zhang, and Z. Chen, “Privacy-Enhancing Personalized Web Search,” Proc. 16th Int’l Conf. World Wide Web (WWW), pp. 591-600, 2007.
[8] X. Shen, B. Tan, and C. Zhai, “Privacy Protection in Personalized Search,” SIGIR Forum, vol. 41, no. 1, pp. 4-17, 2007.
[9] M. Spertta and S. Gach, “Personalizing Search Based on User Search Histories,” Proc. IEEE/WIC/ACM Int’l Conf. Web Intelligence (WI), 2005.
[10] J. Pitkow, H. Schu¨ tze, T. Cass, R. Cooley, D. Turnbull, A. Edmonds, E. Adar, and T. Breuel, “Personalized Search,” Comm. ACM, vol. 45, no. 9, pp. 50-55, 2002.
[11] Y. Zhu, L. Xiong, and C. Verdery, “Anonymizing User Profiles for Personalized Web Search,” Proc. 19th Int’l Conf. World Wide Web (WWW), pp. 1225-1226, 2010.
[12] J. Teevan, S.T. Dumais, and D.J. Liebling, “To Personalize or Not to Personalize: Modeling Queries with Variation in User Intent,” Proc. 31st Ann. Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), pp. 163-170, 2008.
[13] X. Xiao and Y. Tao, “Personalized Privacy Preservation,” Proc.ACM SIGMOD Int’l Conf. Management of Data (SIGMOD), 2006.

Author Profile



Sneha M. Durugkar received the B.E. degree in Computer Science and Engineering from Walachand Institute of Technology, Solapur University in 2013. Currently doing M.E. from Smt.KashibaiNavale college of Engineering, Vadgon, Savitribai Phule Pune University.

Licensed Under Creative Commons Attribution CC BY