



Personalized Web Search with Improved User Profile Using Decision Tree

Rashmi Chaudhari

ME Student, Department of Computer Science Engineering,
G. H. Raisoni Institute of Engineering and Management, Jalgaon, Maharashtra, India

Abstract— *Now-a-days everyone uses web engines to find certain amount of information in very less amount of time. Also the customers use some online shopping sites to shop goods such as electronic gadgets, cloths, books, cosmetics, household things, etc. It's just because of the user friendliness of the web search engines and less time it takes to work. Users just have to fire the query to the search engine and within few seconds he gets various links related result. This paper presents the personalized web search with improved user profile that works on the classification of the search. That means here a user profile is created and whatever user searches is classified into different categories. For that purpose classification method: Decision tree is used here. And on the basis of that when next time that particular user login with his profile and searches for information, the results are provided. At the same time, some solutions are proposed in this paper to overcome the problem of ambiguity in the queries.*

Key Words: *Personalized web search, User Profile, web search engine, web results, search queries, ambiguity in queries.*

I. INTRODUCTION

Generally people search information on the web search engine. But sometimes users experience failure as they get inappropriate and irrelevant results that they do not meet their intentions. Typical search engines do not know who give the query. It just provides the result according to the query given. So, the requirement arises to Personalized Web Search i.e. PWS. In Personalized Web Search i.e. PWS gives results appropriate to the user. Even though there are many research papers published on this PWS previously, but still there are some chances to improve the quality of search.

Until now there is no solution provided on the problem of inappropriate results provided by the web search engine. In this paper, solution to meet the user's real aim behind the query is given. This is done by using the Decision Tree: an effective classification method. In spite of this problem here the solution for the problem of ambiguous queries is proposed.

II. PROBLEM STATEMENT

A PWS i.e. Personalized Web Search is proposed to make the search easy and secured. But along with that the appropriateness of the result should be maintained. That means the search engine should understand the exact intension of the user behind entering the query. The appropriateness of the result is directly depends on the query entered by the user. That means if the user entered an ambiguous word as a term in query then the search engine should understand what result exactly the user wants. If this will get achieved there will be two advantages: one will be the user get the appropriate result and secondly this will decrease the response time. That means the user get the exact result what he want in minimum time.

III. RELATED WORK

There have been several prior attempts to personalize Web search. One approach to personalization is to describe the user's general interests [7].

A study by Teevan *et al.* [8] suggests that rather than fully specifying their search goals up front, people often browse to their targets via pages identified by less precise but more easily specified queries.

Another approach to make search user friendly is to generalize the online user profile using two greedy algorithms namely GreedyDP and GreedyIL [1].

D. Benjamin and M. J. Atallah [9] provides a protocols for the safe and private outsourcing of linear algebra computations, that enable a client to securely outsource expensive algebraic operations like the multiplication of huge matrices, to two remote servers, such that the servers get nothing about the customer's private input or the result of the computation.

Z. Nematzadeh Balagatabi [10] has classified the researchers based on Decision Tree and Naïve Bayes techniques and finally selects the best method of classification based on the highest accuracy to help the researchers to have the best feedback based on their demands in the digital libraries.

IV. PROPOSED TECHNIQUE

Now-a-days everyone uses web engines to find certain amount of information in very less amount of time. It's just because of the user friendliness of the web search engines and less time it takes to work. Users just have to fire the query to the search engine and within few seconds he gets various links related result. The users use online shopping sites to shop goods such as electronic gadgets, cloths, books, cosmetics, household things, etc. So, to minimize the response time for a query and to make ambiguity in queries less the following solutions are proposed:

A. Improved User Profile Using Decision Tree

Here the paper presents the personalized web search with improved user profile that works on the classification of the search. That means here a user profile is created that maintains the user privacy along with the search history. This profile maintains the search history and classify it according to its categories. For example, suppose a user searches for the power bank then for ethnic wear and then for sunscreen lotion then the user profile classifies these into categories Electronic, Clothing, Cosmetics respectively.

For the classification purpose a classification method named Decision tree is used here. Decision Tree is the faster than other classification methods such as k-Nearest Neighbor. The classification time of Decision Tree is fast because there is no calculation in its classification [5]. The decision tree model is created outside the application using Weka data mining tool. And the model is converted into rules before being incorporated into the application. Classification by way of following the tree rules is faster than the ones that need calculation as in the case of Naïve Bayes and k-NN. So here it is advised to use the Decision Tree as a classification method.

On the basis of such classification, when next time that particular user login with his profile and searches for information, the results are provided.

B. Solution for Ambiguous Queries

Ambiguous words means the words that are open to more than one interpretation. Or we can also define ambiguous words as the words that are not having one obvious meaning [11]. For example, the word "bear". Bear is 'a large and heavy animal that have thick hair and sharp claws that can stand on two legs like a person'. Another meaning of bear is 'something that is difficult to do or deal with' [12].

Sometimes when user enters the queries that are ambiguous then the search engine has to show the results for all meanings of that word. Then the user again has to check which meaning is required to him. This take much time. So to overcome this problem following solutions are proposed:

- a. Finding keywords related to query: This can be done by using the terms in the query or can be based on the commonly clicked URLs.
- b. Query URL reference: Here the URLs having similar context are referred.
- c. Query grouping: Here the groups of similar queries are created to recommend the URLs to frequently asked queries of a search engine.

V. CONCLUSION AND FUTURE WORK

In this paper the personalized web search with improved user profile is presented that works on the classification of the search. Here a user profile is created and whatever user searches is classified into different categories. For that purpose classification method: Decision tree is used here. And on the basis of that when next time that particular user login with his profile and searches for information, the results are provided. This will help the user to get the proper result in minimum time. At the same time, some solutions are proposed in this paper to overcome the problem of ambiguity in the queries.

In future, new concepts can be applied to make the ambiguity in query less and to improve the response time. Also, new classification method that can perform than decision tree can be applied.

ACKNOWLEDGEMENT

This work was guided by Miss. Sonal Patil, and H.O.D of Computer Science Department, Prof P.P. Rewagad, from GHRIEM, Jalgaon. I would like to thank both of them for their valuable guidance, help, and appreciate to work and constant source of inspiration.

REFERENCES

- [1] Lidan Shou, He Bai, Ke Chen, and Gang Chen "Supporting Privacy Protection in Personalized Web Search", *Knowledge and Data Engineering, IEEE Transactions on Year: 2014, Volume: 26, Issue: 2*, 2014.
- [2] S. Radha, P. Veera Prakash "A Customized Framework for Improving the Quality of Web Search", *International Journal of Advanced Research in Computer Science and Software Engineering*, Volume 5, Issue 12, December 2015. ISSN: 2277 128X
- [3] Archana Ukande, Nitin Shivale "Supporting Privacy Protection in Personalized Web Search with Secured User Profile", *International Journal of Science and Research (IJSR)*, Volume 3 Issue 12, December 2014
- [4] Ms. A. S. Patil, Prof. M.M.Ghonge, Dr. M. V. Sarode," User customizable Privacy-preserving Search Framework-UPS for Personalized Web Search", *International Journal of Research in Advent Technology*, Vol.2, No.4, April 2014 E-ISSN: 2420-14141

- [5] Ahmad Ashari, Iman Paryudi , A Min Tjoa “Performance Comparison between Naïve Bayes, Decision Tree and k-Nearest Neighbor in Searching Alternative Design in an Energy Simulation Tool” (*IJACSA International Journal of Advanced Computer Science and Applications, Vol. 4, No. 11, 2013*)
- [6] Jaime 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).
- [7] Google Personal. <http://labs.google.com/personalized>
- [8] Teevan, J., Alvarado, C., Ackerman, M. S. and Karger, D. R. “The perfect search engine is not enough: A study of orienteering behavior in directed search”, In *Proceedings of CHI ‘04*, 415-422, (2004).
- [9] D. Benjamin and M. Atallah, “Private and cheating-free outsourcing of algebraic computations”, in Proc. Conf. PST, 2008, pp. 240_245.
- [10] Z. Nematzadeh Balagatabi, “Comparison of Decision Tree and Naïve Bayes Methods in Classification of Researcher’s Cognitive Styles in Academic Environment,” *Journal of Advances in Computer Research. Vol. 3(2), 2012, 23 – 34.*
- [11] <http://www.merriam-webster.com/dictionary/ambiguous>
- [12] <http://www.merriam-webster.com/dictionary/bear>