

**REFINING USER SEARCH GOALS WITH FEEDBACK SESSION AND  
CLASSIFIED AVERAGE PRECISION**Rajesh M.Kamble<sup>1</sup>, Hridaynath P. Khandagale<sup>2</sup>,<sup>1</sup>*Department of Technology, Shivaji University, krajkamble@rediffmail.com*<sup>2</sup>*Department of Technolog, Shivaji University, k\_hriday@yahoo.com*

**Abstract**— Now a day the users are generally using the internet for several information needs. Different users may have different information on dissimilar facets upon the submission of same query to search engine. So discovering different user search goals becomes complicated. The estimation and representation of user search goals can be very useful in improving search engine applicability and user knowledge. This paper proposes an approach for concluding user search goals by analyzing user query logs from several search engines. In proposed method is to determine dissimilar user search goals for a query by clustering the user feedback sessions. The Feedback sessions are built from click- through logs of various search engines. This method generates virtual-documents to better represent feedback sessions for clustering. Finally, clustering virtual documents to discover different user search goals and describe them with some keywords. Then these user search goals are used to restructure the web search results.

**Keywords**- User Search, Feedback Session, Pseudo Documents, Classified Average Precision (CAP),

**I. INTRODUCTION**

Now a day in Web application, user has to submit their query to the search engine. Search engine list out the results related to that particular query. In this system Different user's wanted to get different aspects of information for a same query. For example, user submits query as "apple", some user want to view about the apple product and some user want to know about the nutrients of the fruit apple. So it is important to know the user search goal for a query. [2]

The Search goals of different users may be different depending upon their need. In web search engines queries are submitted to obtain the needed web pages. The meaning of a query may be wider or different user may give the same query for different information. The analysis of user search goals for a query can be very useful in improving search engine & user experience. If the necessary page is not available in the pages, the user has to search in other pages or by giving some other query words. The user has to spend a great amount of valuable time by giving different form of queries. The analysis can also be used for re-ranking of the result obtained in a user search in such that the most wanted pages can be displayed initially & then the remaining.[5].

In this paper, we introduce a method to predicting the user search goals of the query that a user given by clustering the existing feedback details. Feedback is obtained from the customer. The submitted feedback is clustered as feedback session. Feedback session only includes the URL it consists of the clicked URL & unclicked URL links & end with the last URL that was clicked in a session. From user click through logs, feedback session provides idea to cluster user care about & does not care about. Based on the feedback session the pseudo document is generated. This pseudo document consists of phrases for each URL's present in the process of grouping the data into classes or clusters. So that objects within a cluster have a high similarity in comparison to one another but are very dissimilar to object in other clusters. After construction the pseudo document the web search results are reconstructed & Classified Average Precision (CAP) is used to evaluate the performance of restructured search results.

In short, we provide a framework that allows a search engine to understand the intentions of the user's search by identifying the search goals associated with the search keywords provided by the

user through clustering of feedback sessions. A pseudo document is created by combining the URLs that are clicked by the user in the feedback sessions. This allows a good inference on what the user is looking for. Therefore the goals of the user's search can be easily identified. Finally, we make use of a criterion called Classified Average Precision to determine the user search goals effectiveness as they are inferred by the restructuring of the net search results

The Remainder of this paper is organized as follows. Section II presents the related work regarding user search goals with feedback session. The detail of Feedback Session is presented in section III. In section IV idea of Classified Average Precision is explained. At last section V present the concluding remark.

## **II. RELATED WORK**

A prior utilization of user click-through logs is to obtain user implicit feedback to enlarge training data when learning ranking functions in information retrieval. Thorsten Joachims did many works on how to use implicit feedback to improve the retrieval quality [8], [9], [10]. In our work, we consider feedback sessions as user implicit feedback and propose a novel optimization method to combine both clicked and unclicked URLs in feedback sessions to find out what users really require and what they do not care.

One application of user search goals is restructuring web search results. There are also some related works focusing on organizing the search results [6], [18], [20]. In this paper, we infer user search goals from user click-through logs and restructure the search results according to the inferred user search goals.

Efficient algorithm is used for finding user friendly environment in web search application using data mining concepts. In order to finding the automatic goal identification based on human subject study by using user-click behavior and anchor link distribution, Zheng Lu et al. [1]. In [3] Barbara Pobleto et al. to achieve better results using non-supervised tasks such as clustering and labeling.

In [4] clustering search results is an efficient way to organize search results using commercial search engine log data. In [5] Rosie Jones et al. generating a new query to replace a user's original search query using query pair algorithm. In [6] Xiao Li et al. increasing the amount of training data based on click graph using semi-supervised click graphs.

In [5] Steven M. Beitzel et al. web query classification is used to improving retrieval effectiveness and efficiency. In [8] C.K Huang et al. query suggestion by mining query patterns from search logs using query suggestion step and concept sequence suffix. In [14] Dou Shen et al. web query classification aims to classify web users queries, which are often short and ambiguous, into a set of target categories. In [12] Rosie Jones et al. analysis of typical timeouts used to divide query streams into sessions, and demonstration that they are less than optimal for this task

## **III. FEEDBACK SESSION**

Feedback session is a session for web search is a sequence of consecutive queries to satisfy a single information require and some clicked investigate results focal point on inferring user search goals for a exacting query. Consequently the single session contain simply one query is introduce, which distinguish from the conservative session. For the moment, the feedback session is based on a solitary session, though it can be comprehensive to the entire session. It consists of both clicked and unclicked URLs and ends with the last URL that be clicked in a single session. It is forced that previous to the last click, all the URLs have been scanned and evaluate by users. Each feedback session can tell what a user requires and what he/she does not care about. Moreover, there are plenty of diverse feedback sessions in user click-through logs. Consequently, for inferring user search goals, it is additional efficient to examine the feedback sessions than to examine the investigate consequences or clicked URLs in a straight line. To represent the feedback session efficiently some demonstration methods needed, because each and every user

based search goal feedback sessions are differs and their corresponding log files also changed.

Represent a feedback session to Pseudo-Documents with Binary vector technique to characterize a feedback session search consequences are the URLs return by the search engine when the question “the sun” is submits, and “0” represent “unclicked” in the click sequence. The binary vector [0110001] can be second-hand to symbolize the feedback session, where “1” represent “clicked” and “0” represents “unclicked.”

#### IV. CLASSIFIED AVERAGE PRECISION (CAP)

In the proposed system “Classical Average Precision (CAP)” is used to understand user search goals efficiently & evaluate the performance of inferring user search goals. In this user needs is highlighted & provides a user friendly search engine. Performance of inferring user search goals is evaluated using a new CAP.

The following algorithm for CAP is shows how to evaluate the performance of user search goals using feedback session

##### *Algorithm for CAP*

1. **Input:** User’s Search Query (Q)
2. **Output:** Re-Structured Results (RR)
3. **BEGIN**
4. Get query “Q” from user
5. Populate **QHistory** as a query history dataset with the retrieved results from the database - **DBQueries**
6. If **QHistory** is empty
7. **Return** Actual Web Search Results “**RW**”
8. Else
9. For each Query Instance **Qi** in **QHistory**
10. Generate Feedback Session **FSi**
11. Convert Feedback sessions to Pseudo Documents **PDi**
12. End For
13. End If
14. Set Web Search Results “**RW**” as results from web search engine
15. Set Output “**RR**” using the cluster\_function(**PDi**, **Qi**, **Q**)
16. **Return** Re-Structured Results – “**RR**”
17. **END**

The above algorithm for CAP is shows how to evaluate the performance of user search goals using feedback session.

#### V. CONCLUSION

We introduce feedback sessions to be analyzed to infer user search goals rather than search results or clicked URLs. Both the clicked URLs and the unclicked ones before the last click are considered as user implicit feedbacks and taken into account to construct feedback sessions. Therefore, feedback sessions can reflect user information needs more efficiently. Second, we map feedback sessions to pseudo documents to approximate goal texts in user minds. The pseudo-documents can enrich the URLs with additional textual contents including the titles and snippets. Based on these pseudo-documents, user search goals can then be discovered and depicted with some keywords. Finally, a new criterion CAP is formulated to evaluate the performance of user search

goal inference when users submit one of the queries, the search engine can return the results that are categorized into different groups according to user search goals online. Thus, users can find what they want conveniently.

## REFERENCES

- [1] Zheng Lu, Student Member, IEEE, Hongyuan Zha, Xiaokang Yang, IEEE, Weiyao Lin, Member, IEEE, and Zhaohui Zheng –“A New Algorithm for Inferring User Search Goals with Feedback Sessions”- IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 3, MARCH 2013.(Main Paper)
- [2]Thea Van Der Geest, Judith Ramey, Stephanie Rosenbaum, Senior Member, IEEE, and Lex Van Velsen.”Introduction to the Special Section: Designing a Better User Experience for Self-Service Systems” IEEE TRANSACTIONS ON PROFESSIONAL COMMUNICATION, VOL. 56, NO. 2, JUNE 2013.
- [3]B.Poblete and B.-Y Ricardo, “Query-Sets: Using Implicit Feedback and Query Patterns to Organize Web Documents,” Proc. 17th Int’l Conf. World Wide Web (WWW ’08), pp. 41-50, 2008
- [4] R. Baeza-Yates,C. Hurtado, and M. Mendoza, “Query Recommendation Using Query Logs in Search Engines,” Proc. Int’l Conf. Current Trends in Database Technology (EDBT ’04), pp. 588-596,2004.
- [5] R. Jones and K.L. Klinkner, “Beyond the Session Timeout: Automatic Hierarchical Segmentation of Search Topics in Query Logs,” *Proc.17th ACM Conf. Information and Knowledge Management (CIKM’08)*, pp. 699-708, 2008.
- [6] U. Lee, Z. Liu, and J. Cho, “Automatic Identification of User Goals in Web Search,” *Proc. 14th Int’l Conf. World Wide Web (WWW ’05)*, pp.391-400, 2005
- [7] H. Cao,D. Jiang, J. Pei, Q. He,Z. LiaoE. Chen, and H. Li,“Context-Aware Query Suggestion by Mining Click-Through,” Proc. 14th ACM SIGKDD Int’l Conf. Knowledge Discovery and Data Mining (SIGKDD ’08), pp. 875-883,2008.
- [8] C.-K Huang, L.-F Chien, and Y.-J Oyang, “Relevant Term Suggestion in Interactive Web Search Based on Contextual Information in Query Session Logs,” *J. Am. Soc. for Information Science and Technology*, vol. 54, no. 7, pp. 638-649, 2003.
- [9] H. Chen and S. Dumais, “Bringing Order to the Web: Automatically Categorizing Search Results,” Proc. SIGCHI Conf. Human Factors in Computing Systems (SIGCHI ’00), pp. 145-152, 2000.
- [10] T. Joachims, “Evaluating Retrieval Performance Using Click-through Data,” Text Mining, J. Franke, G. Nakhaeizadeh, and I. Renz, eds., pp. 79-96, Physica/Springer Verlag,2003.
- [11] T. Joachims, “Optimizing Search Engines Using Clickthrough Data,” Proc. Eighth ACM SIGKDD Int’l Conf. Knowledge Discovery and Data Mining (SIGKDD ’02), pp. 133-142, 2002.
- [12] T. Joachims, L. Granka, B. Pang, H. Hembrooke and G. Gay, “Accurately Interpreting Clickthrough Data as Implicit Feed-back,” Proc. 28th Ann. Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR ’05), pp. 154-161, 2005.
- [13] B. Poblete and B.-Y Ricardo, “Query-Sets: Using Implicit Feedback and Query Patterns to Organize Web Documents,” Proc. 17th Int’l Conf. World Wide Web (WWW ’08), pp. 41-50,2008.
- [14] D. Shen J. Sun, Q. Yang, and Z. Chen, “Building Bridges for Web Query Classification,” Proc. 29th Ann. Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR ’06),pp. 131-138 2006.
- [15] X. Wang and C.-X Zhai, “Learn from Web Search Logs to Organize Search Results,” Proc. 30th Ann. Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR ’07), pp. 87-94, 2007.[19] J.-R Wen, J.-Y Nie, and H.-J Zhang, “Clustering User Queries of aSearch Engine,” Proc. Tenth Int’l Conf. World Wide Web (WWW ’01),pp. 162-168, 2001.
- [16] H.-J Zeng, Q.-C He, Z. Chen, W.-Y Ma, and J. Ma, “Learning to Cluster Web Search Results,” Proc. 27th Ann. Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR ’04), pp. 210-217, 2004

