

Activity Based Key Search Information Re-finder

Yeswanth.L¹, Nimala.K²

¹*M.Tech, Database Systems, Dept of Information Technology,
SRM University, Chennai, India*

²*Assistant Professor, Dept of Information Technology,
SRM University, Chennai, India*

Abstract— As most of the users depend on the digital information, the extraction of documents, files, messages, emails, and various page links etc., has become a vital important in the world. People usually forget where the files existing in the personal computer but they might have some low exposed image. Considering memory of users, we propose an application called “Activity-based Information Re-finder.” It is the technique which helps to use Cluster and Association process. We provide an activity based instances which in the sense organized in an associated manner. Also helps to re-find the files and web pages based on the criteria of previous accessed files. We also propose a feedback session to this application. Based on the given feedback, we can give a ranking to the searched pages often.

Keywords— Activity based, Feedback, Page Rank, Re-finder access log.

I. INTRODUCTION

As we know that users are fascinated to read, write and more focused to gather information from desktop system and the web. Users have low exposed image in their memory. So re-finding makes easy to find what they have accessed before. A normal way is to maintain re-finder logs. Due to low exposed image in human memory, it is uneasy and time consuming for users to regain the lost information. So we propose a technique called “Activity-based re-finder”.

Scientists discover that there is a kind of memory in a human brain called episodic memory [1]. Association and context are generally viewed as two of the central ideas in the history of episodic memory research [3]. For example, it may be hard to recall a recipe’s detail encountered one month ago, but the place, time and activity associated with the happening of that access event may leave a deeper impression, which could serve as useful cues to re-find the target recipe [2].

We develop an activity based information re-finder. We create a link between present accessed information and the previous accessed information. A memory contains huge amount of associated links ordered in clusters. Develops a new interaction method called pivoting [4], which allows users to look for linked activities and come across an end piece of information.

Associations between data items are used to improve keyword search matched results. Soules and Ganger [5]

developed a file search tool combining content-based search with temporal relationships between files gathered from user’s file operations. Chau et al. [6] developed a system which supports multilevel associative retrieval of desktop information.

Page Rank is used to assign a numerical weightings to hyperlinked documents (or web pages) indexed by a search engine. Page Rank system is used by the popular search engine Google to help determine a page’s relevance or importance. Based on the user’s feedback, the Page Rank can be calculated. Page Rank relies on the unique nature of the web by using its vast link structure as an indicator of an individual page’s value. The Page Rank depend on the number and Page Rank metric of all pages that link which passes through it. A page that is linked by many pages can make high rank. If it does not have any links to a web page there is no support of this specific page.

II. RELATED WORK

It develops a new interpreter technique called pivoting, which allow user to look for accessed activities and also we represent a scripting method. Time mark permits user to access all activities that was continuous at particular minute than keeping track individual history files and bookmarks. We develop an association method to describe valid intentional table of relationship between cases in a data space [7]. It balances well when the number of association trace is enlarged. We explore the variation among questions [8] that had changes between the earlier questions and the re-finding question and those that had smallest alteration with traverse sessions. It develops an associative memory based personal desktop system, imecho [9] which increases the expected text key search with semantic relationship from user activity and also system offers connection graph to help clients improve and associate results produced by the key search. This develops a personal data space administration [10] which provides consistent personal information choice. This also develops iMemex which can be used by a Graphical User Interface (GUI) client to provide client with in a worldwide view. The implementation develops a personal information management systems requires a powerful data model which is able to stand for semi-structured and grouped data surrounded on its model. Production of iMemex data model is for private information administration [11]. It is used to represent semi- planned and planned information inside a particular model. Also represent a consequence of an estimation of previously performed in iMemex. Here

Ambient Intelligence also elevates new challenges related to preparedness and customer’s interface [12] involving us to assume existing database format. We bring out the planned strategies through an enclosed transport, consisting of data administrator and a personal data administrator.

III. EXPERIMENT

A. Re-finder access log

We propose activity based re-finder technique to find the files or web pages. When people read or write from the desktop system or pages from the World Wide Web, Re-finder makes user to find interesting documents or pages with the previous access content by providing activities like date and time.

B. Page Rank (PR) Calculation

The Page Rank of each page depends on the Page Rank of the pages pointing to it. A link to a page counts as a vote of support. If there’s no link there’s no support. We assume page A has pages B, C, D which point to it (citations). The parameter d is a damping factor which can be set between 0 and 1. We usually set d to 0.85. Also C (A) is defined as the number of links going out of page A. The Page Rank of a page A can be calculated as,

$$PR(A) = (1-d) + d (PR(B)/C(B) + \dots + PR(D)/C(D))$$

C. Page Rank Result

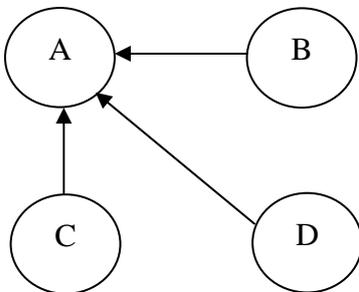


Fig.1. Page Rank example

Users can provide their feedback based on their views and the Page Rank changes as how many times the user visited the web page links. A page that is linked by many pages then it is shown as highest page rank. Otherwise the page rank will be less. By default Page Rank is measured between 0 and 10.

IV. SYSTEM ARCHITECTURE

The user’s can revisit the web page links based on the activity search links provided. Activity search is based on time, area and date. After providing activity instances the webpage links are located in the log file. The user can also provide a feedback to the application which gives page rank to the more web page links visited.

A. Re-finding Design

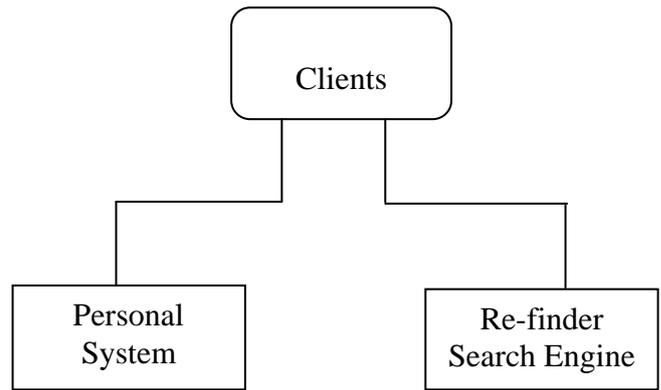


Fig.2. Re-finder Design

During the normal search in the personal system, the documents represent in the desktop and while in the Re-finder search engine the file can be shared with others. We provide automatic activity based on the system with the accessed file information. We can also make the webpage link as a document. If the page link requested is vital then it can also be shared using the social network.

V. CONCLUSION

Behind this mechanism customers should make the system recognize the accessed contents which will be recalled afterwards and which part of a web page or file is of interest. Study of user’s access performance, access history, accessed information and user’s activity should hold up through decision making. This makes the re-finder another focal point that deserves the future enhancement.

ACKNOWLEDGEMENT

The authors are grateful to express sincere thanks and gratitude to our Department of Information Technology, SRM University for encouragement and the facilities that were offered to us for carrying out this project.

REFERENCES

- [1] E. Tulving, “What is Episodic Memory?” *Current Directions in Psychological Science*, vol. 2, no. 3, pp. 67-70, 1993.
- [2] S. Dumais, E. Cutrell, J. Cadiz, G. Jancke, R. Sarin, and D.C. Robbins, “Stuff I’ve Seen: A System for Personal Information Retrieval and Re-Use.” *Proc. 26th Ann. Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR)*, 2003.
- [3] M.J. Kahana, M.W. Howard, and S.M. Polyn, “Associative Retrieval Processes in Episodic Memory”, *Learning and Memory: A Comprehensive Reference*, pp. 1-24, Academic Press, 2008.
- [4] J. Hailpern, N. Jitkoff, A. Warr, R. Karahalios, K. Sesek, and N. Shkrob, “YouPivot: Improving Recall with Contextual Search,” *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI)*, 2011.
- [5] C.A.N. Soules and G.R. Ganger, “Connections: Using Context to Enhance File Search,” *Proc. 20th ACM Symp. Operating Systems Principles (SOSP)*, 2005.
- [6] D.H. Chau, B. Myers, and A. Faulring, “What to Do When Search Fails: Finding Information by Association,” *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI)*, 2008.
- [7] M.A. Salles, J. Dittrich, and L. Blunschi, “Intensional Associations in Dataspace,” *Proc. Int’l Conf. Data Eng. (ICDE)*, 2010.

- [8] S.K.Tyler and J.Teevan, "Large Scale Query Log Analysis of Re-Finding," Proc. Third ACM Int'l Conf. Web Search and Data Mining (WSDM), 2010.
- [9] J.Chen, H.Guo, W.Wu, and W. Wang, "iMecho: An Associative Memory Based Desktop Search System," Proc. 18th ACM Conf. Information and Knowledge Management (CIKM), 2009.
- [10] E. Adar, J. Teevan, and S.T. Dumais, "Large Scale Analysis of Web Revisitation Patterns," Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI), 2008.
- [11] J.P. Dittrich and M.A. Salles, "iDM: A Unified and Versatile Data Model for Personal Dataspace Management," Proc. 32nd Int'l Conf. Very Large Data Bases (VLDB), 2006.
- [12] L. Feng, P.M. Apers, and W. Jonker, "Towards Context-Aware Data Management for Ambient Intelligence," Proc. 15th Int'l Conf. Database and Expert Systems Applications (DEXA), 2004.