Information Retrieval Using Short-term Context

Matúš VACULA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova, 842 16 Bratislava, Slovakia
vacula.matus@gmail.com

The content of the Web is continuously expanding. Several years ago it might have been problem for user to find desired information on the Web because the information might not have been published online. Nowadays user experiences slightly different kind of problem. There is too much information and it is problematic to find the most relevant piece of information for user's interest. All of the modern search engines have to accommodate to this.

Search engine developers are trying to make searching more efficient by involving the context. First to define the term "context-aware" were Schilit and Theimer [3] who defined it as location and identity of nearby persons or objects and the their changes. Dey [1] defines context as any information that can be used to characterize the situation of an entity. Entity is a person, place or object which is relevant to user's interaction with application. Utilising the context is usually done by taking into consideration the geo-location of the user, his language or previous web-searches. But there are also other forms of context that are not utilized as much as those mentioned. It is possible to take into consideration also information such as the time of the day, time of the week, user's personal data, social aspects of context and his activity. Although the resources are available and numerous methods how to implement the context into the information retrieval are known search engines are still not using these possibilities at their full potential.

In our work we focus on utilising the context from the aspect of the user's momentary activity. Information about user's current activity could lead to achieve more accurate search queries or at least more unambiguous meaning of his demand. These information are relatively easy to obtain from the web browser. We developed our method on the hypothesis that user's web-search relates to his current activity and this activity is reflected in the web-browser in form of open websites.

We propose a method that uses the context of user's activity to enhance the web-search. This method consists of extracting and selecting the proper keywords from

* Supervisor: Dušan Zeleník, Institute of Informatics and Software Engineering

Spring 2013 PeWe Workshop, April 5, 2013, pp. 67–68.

browser's open websites and modifying the search queries with use of them. The process of keyword selection is based on scoring each keyword based on these criteria:

- proximity of keyword to search-query if found in the content of open websites
- number of appearances
- significance - which HTML is the keyword from (more significant are keywords from headlines and highlighted texts)

We expect this will result in providing more accurate and relevant search results to user. The aim is to raise the user's satisfaction and shorten the time while he obtains the desired information. We are aware that proposed method will only help to improve those web-searches which would be performed with sufficient contextual data. This will be mostly applied to users who are accustomed to browse more than one website at the time. That means using tabs in web-browser.

We decided to implement the method in form of browser extension (Chrome) which gives us wide access to data describing user's activity. The architecture of this solution is based on the concept of meta-search engine [2]. Information retrieval itself is provided by Google search engine and the browser extension is responsible for optimizing the queries.

We also propose the quantitative experiment which will evaluate the success rate of the proposed method. The principle is to provide the search results from enhanced and unenhanced web-search mixed together in random order in the way that user would be unable to identify which search result come from which web-search (enhanced or unenhanced). This way the chances of bias are reduced. During the experiment we will monitor the user's choices of search results and the time spent browsing the websites navigated through the selected search result. More clicks on search results from enhanced web search and more time spent on websites recommended by enhanced web search should indicate that proposed method helped to make the search result more attractive to user and helped him to find more relevant information.

In order to evaluate the effect of proposed method more precisely we suggest to consider only those searches which were performed with use of sufficient amount of contextual data. This will provide the evaluation of effectiveness on the selected behavioural group of test subjects who tend to use more than one browser tabs at the time.

References

[1] Dey, A.K., Understanding and Using Context, Personal and Ubiquitous, Computing, Volume 5, Issue 1, pp: 4 – 7, 2001.

[2] Glover, E. J., Lawrence, S., Gordon, M. D., Birmingham, W. P., and Giles, C. L. (2001) Web Search Your Way. Communications of the ACM 44, 97-102.

[3] Schilit B. N., Theimer M. M., “Disseminating active map information to mobile hosts”, IEEE Network, 8(5):22–32, September/October 1994