An Exploration in Personalized and Context-Sensitive Search

Jing Su & Mark Lee
School of Computer Science
University of Birmingham
B15 255
http://www.cs.bham.ac.uk/~jxs
{jxs|mgl}@cs.bham.ac.uk

Abstract

It is often claimed that the use of personalized context sensitive search will greatly improve the majority of Information Retrieval (IR) tasks. However, most search engines still follow the model of "one size fits all" - all users receive the same result for a given query leading to low retrieval precision. This paper will report on ongoing research concerning the incorporation of web personalization techniques and user active context into the IR process to provide a personalized and context-sensitive search service.

1. Introduction

As previous research has shown [e.g., Leake & Scherle, 2001, McGowan et al., 2002], one of the key reasons for low user satisfaction with the results of search engines is that the searching process is not personalized or context-sensitive. This causes two kinds of difficulty. First, different users have distinct goals and backgrounds and so will have different views of what is relevant. Secondly, even a single user will have different needs at different times according to his or her currently active task.

Yang et al. [2000] outline three aspects to be improved in the third generation of search engines. First, search engines must have the ability to detect user conceptual intention; secondly, they must have the ability to allow the user to provide more information as required; and thirdly, they must be able to allow the user to have more sophisticated interactions with the system. All of these criteria suggest the importance of research towards personalized and context-sensitive search.

Much research has been done toward providing a personalized retrieval result by constraining the search results by a user's profile that can represent user's long-term information needs. However, the majority of previous work in this area has not considered the use of the users' active tasks which could be used to provide a rich context of the users' short-term information needs. Consider the following two possible scenarios when a user forms a query such as "Intel Pentium 4 Chip." In scenario one, he or she is looking for documents as references for a paper about a hardware design; in scenario two, he or she is looking for a cheap price for the component while writing a recommendation to his or her friend in an email message. Clearly, the user has different information needs in these two scenarios although the queries are expressed in the same terms. In such a situation, even so-called personalized services could not provide satisfying performance, because the same result will be returned according to this user's specific query and specific profile. However, such problems could be solved by using knowledge of the user's active context document to assist in discovering his or her explicit short-term information needs. A search process that is aware of a user's active tasks can be considered a form of context-sensitive search. For example, an engine could use the contents of the active document that user is viewing or editing when he or she initiates a search query as additional contextual information to aid the search process. We believe this contextual information can be used to sharply cut down irrelevant results.

The structure of the rest of the paper is as follows: in section 2, we will propose some possible solutions targeted to the limitations of existing approaches of processing the user's active task document. In section 3, we will describe an experiment we have performed to confirm our basic assumptions and in section 4, we will provide some further discussions and detail future work.

2. Previous work

In recent work in order to make the search service personalized and context sensitive, researchers have used data from the user local machine to assist the searching process, including the analysis of both user behaviour to infer his profile for detecting his long term information goal, and users active document to infer his short term information goal.

These systems vary from each other in the methods used to gain the user context, the methods to process the context document data, and the methods they use to involve the context data into the search process. However the common aspects which need to be improved in all these approaches include: semantic understanding of the context document; cooperation between the context with other user data such as profile and actions; and allowing the possibility to expand the usage of user personal data out of the user's local machine to server side processes.

The methods used to process user local active context are similar in several projects (for example, JITIR [Bradley, 2000]). They process the user active context document into a list of weighted words and then select the top ranked N words to represent and expand the user's query. This simple text processing method ignores the relationships between words and therefore can not represent the semantic & contextual meaning of the document precisely. An alternative text processing technique such as term matching rather than word matching, could be used here to improve performance.

Such systems are based on the assumption that more similar documents to the user's local context will be more useful documents, and thus follow the underlying assumption of the conventional information retrieval model. However, we believe that the document can be used to directly expand the user query and represent his or her needs more accurately if combined with a combination of different types of personal data.

All existing usages of context in search can be placed into two categories: Pre-processing before search and Post-processing after search. However, context has not been used in the actual search process. The reason for this lies in the gap between the search engine server and user client machine. We believe that this gap could be bridged by the use of novel user interface designs allowing the search engine server to get the user's active context, and then the context could be involved in the searching process. This is one of the key tasks of this research: to explore the possible approaches to make the searching process adaptive to a user's personal data. In order to explore our assumption we performed the experiment described in the next section.

3. Experiment to evaluate the effects of context

The conventional information retrieval model assumes that similar documents to the query are equal to the useful documents to the user so the goal of searching is to maximize similarity.

In our experiment, the user's active context is introduced into the search process so that the query is expanded by terms from the active document with the hope that the retrieval results will be more related to the user's active task. The text processing method used in a typical approach treats the context document as

a list of words and selects the top ranked N words to expand the query. Essentially this approach leads to retrieval results with high similarity to the active documents besides of the query itself. The similarity assumption has been re-examined by Budzik [2000] and his experimental results show that the relevancy of retrieved documents is not always equal to the usefulness. This problem is also mentioned in [Bradley, 2000] where Bradley argues that the low quality and repeated content retrieved by such methods are the possible reasons for low usefulness.

This raises the following questions: what is the relationship between the similarity degree and usefulness degree of the retrieved documents, and what is the relationship between them and the user active context document, and how we can retrieve the most useful documents for the user with the assistance of the context document. If similar documents are equal to the useful documents to the user, then in a context-sensitive search process, once the context is obtained, the search goal could be simplified to a problem of looking for the similar documents for that *context*. However, as we shall see, the experiment below suggests that this is not the case.

Another purpose of this experiment is to evaluate the different modes of query formulation. Three different typical modes are used in this experiment, and combined with the retrieved documents so that we can have an idea of which mode will lead to more useful results. We also asked users to complete a survey so we could analyze common user habits on the query formulation.

3.1. Experimental Method

The experiment consists of comparing machine calculation of similarity of returned results and user evaluation of these results. A program was developed to rank similar documents in a database in relation to a specific document provided by the user. We asked 5 research students from the School of Computer science to use this programme in our experiment. Because our subjects were research students we assumed that they were all competent users of search engines.

We asked our subjects to provide the following data: 1) the research progress reports of the research students to be used as pseudo local context documents. These vary in length from 1000 to 3000 words; 2) 5 keywords that represent each student's current research area and topic, and 5 keywords that represent the counterpart (or opposed) methods or algorithms in the same topic. 3) A set of papers retrieved using citeseer[1] using the keywords provided by the user.

Step 1: Query Formulation.
We used three modes of query formulation with the user provided keywords user selected to retrieve relevant documents from NEC citeseer digital library. The three modes are: mode 1, to use the combination of the first 5 keywords as queries, which represents the user's research field and topic; mode 2, to use the combination of the second 5 keywords as queries, which represents the counterpart methods and algorithms of the user's topic; mode 3, to use a random combination of first 5 keywords and second 5 keywords as queries. From the results of each mode we downloaded the top 10 ranked documents, and used this to form a thirty document database for each user. This process was aimed to mimic a user's query formulation and information search action.

Step 2: Document Ranking.
Once we built up the small database, we used each user's research progress report as the source document and then ranked the top 10 documents from the database according to the similarity measure generated by our programme. The intention was to treat the user's report as a pseudo context document, and treat the top

1. http://citeseer.nj.nec.com/cs

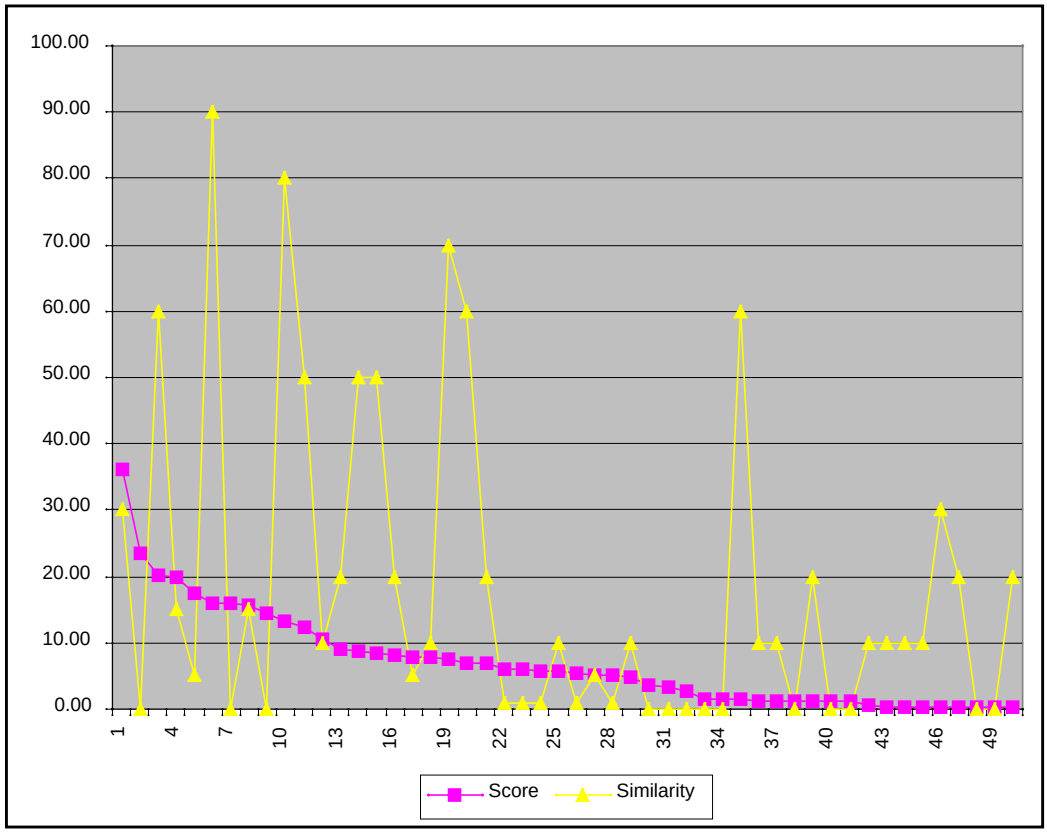

Figure 1: The comparison between Score and Similarity rates

ranked documents as the retrieved documents. In doing so, basic similarity could be guaranteed. In addition, since the keywords were selected by the user, and the report lengths relatively short, the keywords can be viewed as extractions of the context documents.

Step 3: User Evaluation.
The ten retrieved documents were returned back to the user for evaluation. We asked the users to complete a questionnaire which asked them to rate the documents in terms of similarity, relevancy, and usefulness. Active and negative factors to support the rating of usefulness were also questioned so that the user could give the reasons behind their judgements. Finally, there is a survey attached which investigated their habits of query formulation and needs in their daily information seeking activities.

3.2. Results and Analysis

The results list the rates of Similarity, Relevancy and Usefulness for 50 documents in total with the users questionnaires. Below is an analysis of the relationships among these rates.

Score and Similarity
We found the user judged similarities very differently to the program scores for the similarity for the same documents (see Figure 1). There are two possible reasons for this. The first reason is that users judgements are not objective in some aspects. This is supported by another phenomenon: the users' ratings on similarity, relevancy and usefulness followed the same extreme trend in most cases. However the program calculated scores consistently for all documents.

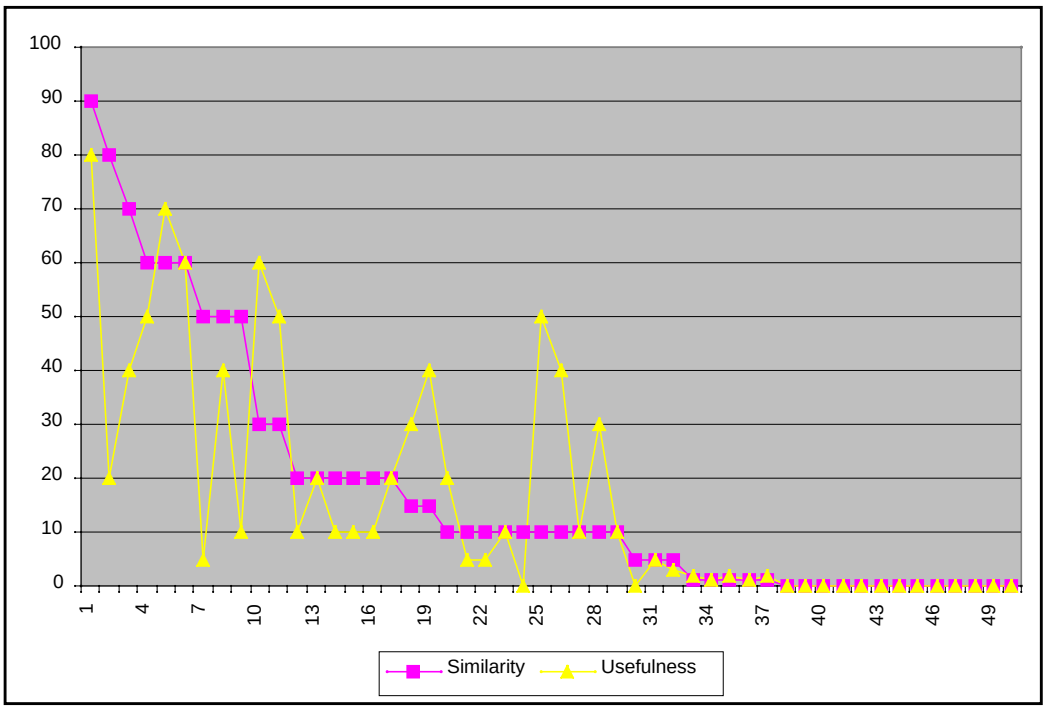

Figure 2: The comparison between Similarity and Usefulness rates

The second reason could be because humans use different criteria for making similarity judgements. Similarity scored by the program is based on the co-occurrence of words calculated through a conventional algorithm which ignores the relationship among words. However, human judgement is based upon their understanding of the entire document. A user can understand a document by terms and discourse structure, and compare the similarity using his or her understanding of the document meaning. Therefore similar documents in terms of word occurrence are not necessarily the similar documents by topic. This is also confirmed by the user's feedback.

Similarity and Usefulness

The user judgements on the similarity and usefulness of a document in general are very close to each other (see Figure 2), the line of similarity and the line of usefulness follow along a similar trend, although they do not exactly own the same value. In most cases the users judged the similarity, relevance and usefulness of a document to a similar level.

Score and Usefulness

The user's judgement on document usefulness and the similarity score of documents calculated by the program do not matched to each other (see Figure 3). We distinguish the differences into two groups: the first group are cases where the user rates the document with low usefulness while its similarity is high; the second group is where the user rates a document with high usefulness but with its similarity rated low.

The reasons for this distinction can be partially explained by the active and negative factors users reported to support their judgements. In the situation of useful documents with low similarity ratings, the common active factors that users selected are: "it talks about your topic, but from another perspective"; and, "it talks about your topic, but from another field"; and "it talks in depth about the counterpart methods of your topic." This suggests that the documents contain key knowledge for the user's information need but from another field or perspective, or that the documents concern the counterparts algorithms and theories and are

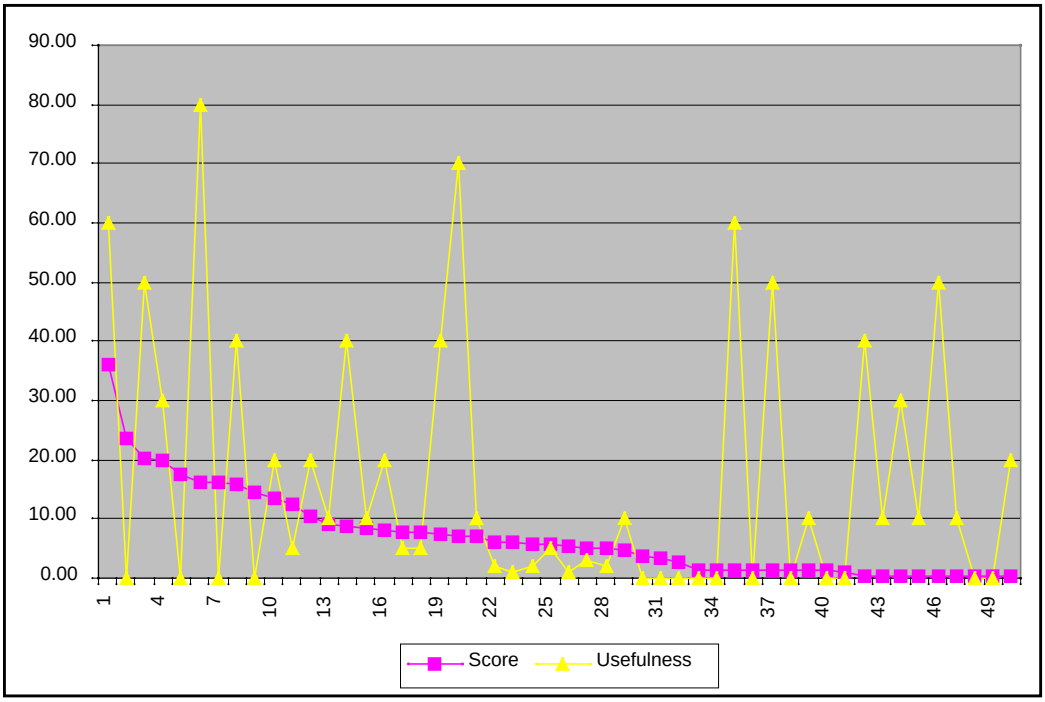

Figure 3: The comparison between Score and Usefulness rates

judged to be useful, but that they do not have a high similarity with the user's own report. The users' feedback suggested that these types of documents could provide them with additional knowledge which would help their understanding in various aspects.

In the contrary situation, the negative factors selected by the users for those documents with high similarity but low usefulness are mainly: "irrelevant document"; and, "the document content is too basic"; and "you have read it before/too old." In most cases, for these documents, users rank their similarity low too. There are four possible reasons for this: first, the users ratings of similarity are affected by their judgements of usefulness; and secondly, according to the user's feedback, we find that the words leading to high similarity for these documents did not carry the key information that the user is looking for. An interesting phenomenon is the third case, some users claim that the reason they rated a document with high similarity/low usefulness was because it talks about key information but from another perspective or field. So even if most terms have a high repeat rate, the content of the entire document is about its applications in another area, and therefore it is useless to the user. This is exactly the same reason for which some users rate low similarity documents with a high usefulness. Another factor indicated here is that we need to be aware of the user's knowledge level and browsing history to avoid too basic or too old information.

Another purpose of this experiment was to compare the effects of different modes of query formulation. We summarize the different modes of queries that leads to top the 10 ranked documents for each user briefly: the combination of the first 5 keywords which represent the user's own work retrieved 44% of total highest similarity results; while a combination of counterparts retrieved 34%, and the third mode takes a 22% share. Although the use of keywords of the counterpart methods seems quite promising, we found from the user survey that users did not use this method in forming queries.

4. Discussion & Future work

Following the findings described above, future work will focus on the user local machine. Our goals are to: 1) extract the semantic meaning from active documents which carry key information related to the user's need to improve the performance of the current context document processing method implemented in the existing related systems; 2) build up a model based on the user local machine to combine the user's personal information data in an appropriate way. The first step of this is to combine the context data and user profile data. By this approach we could access knowledge about the user's background and knowledge level that we could not get completely from the context; 3) Involve user interaction into the whole structure, to make the process more sensitive to a user's real time information intention in order to adapt the user's need adaptively.

5. Conclusions

This paper has described an experiment investigating how user profiles and context sensitive information can be used to improve the quality of information retrieval. In particular, our work argues for four findings. First, usefulness ratings and the similarity ratings of retrieved documents are not equal to each other. To use the entire context document directly as the query and then look for similar documents does not satisfy the user's information need precisely. A better approach should involve it into the search process cooperatively with other data that user provided, such as query or user action, to define the user's information need more accurately. Secondly, the context itself should be processed to amplify its effect, such as by selecting the key terms rather than isolated words to maintain its semantic meaning, and by weighting terms with additional user profile data to constrain the information need with the user's knowledge domain. Thirdly, other types of user data should be involved to assist the understanding of the user's information goal, such as user profile, which can provide a user's background and profession; or the user's browsing history, which can illustrate the user's preferences and reading record. Finally, user interaction should play an important role. Static data can not predict a user's intention completely and so a flexible approach should be adaptive to the user's real time actions.

References

[Yang et al., 2000] Q. Yang, H. F. Wang, J. R. Wen, G. Zhang, Y. Lu, K. F. Lee and H. J. Zhang. "Toward A Next-Generation Search Engine". In: *Proceedings of the Sixth Pacific International Conference on Artificial Intelligence. Pages 5-15. (The Sixth Pacific International Conference on Artificial Intelligence (PRICAI'00). Melborne, Australia, August 28-September 01, 2000) Springer 2000.*

[Bradley, 2000] J. R. Bradley. "Just-in-Time Information Retrieval". Ph.D. thesis. Massachusetts Institute of Technology. 2000.

[Budzik et al., 2000] J. Budzik, K.J. Hammond, Larry Birnbaum, and Marko Krema. "Beyond similarity". In: *Proceedings of the 17th National Conference on Artificial Intelligence: Workshop of Artificial Intelligence for Web Search*. (The 17th National Conference on Artificial Intelligence: Workshop of Artificial Intelligence for Web Search, in conjunction with the AAAI conference (AAAI 2000). Austin, Texas, USA. July 30-31, 2000) AAAI Press, 2000.

[Budzik, 2003] J. Budzik. "Information Access in Context: Experience with the Watson System". Ph.D. thesis, Northwest University. 2003.

[Leake & Scherle, 2001] D. B. Leake and R. Scherle. "Towards Context-based Search Engine Selection". In: *Proceedings of the 6th international conference on intelligent user interfaces, pages 109-112. (Interna-*

tional Conference on Intelligent User Interfaces (IUI'01). Santa Fe, New Mexico, USA. January 14-17, 2001) ACM Press, New York, NY, USA. 2001.

[McGowan et al., 2002] J. P. McGowan, N. Kushmerick, B. Smyth. "Who do you want to be today? Web Pesonae for personalised information access." In*: Proceedings of Adaptive Hypermedia and Adaptive Web-Based Systems, Second International Conference,* pages 514-517. (The Second International Conference on Adaptive Hypermedia and Adaptive Web Based Systems (AH 2002). Malaga, Spain. May 21-23, 2002) Springer. 2002.