Dynamically Selecting an Appropriate Context Type for Personalisation

Tomáš Kramár
Faculty of Informatics and Information
Technologies
Slovak University of Technology
Bratislava, Slovakia
kramar@fiit.stuba.sk

Mária Bieliková
Faculty of Informatics and Information
Technologies
Slovak University of Technology
Bratislava, Slovakia
bielik@fiit.stuba.sk

ABSTRACT

Narrowing down the context in the ranking phase of information retrieval has been shown to produce results that are more relevant to searcher's need. We have identified two types of contexts that could be used in the process of personalisation. We research these contexts in the domain of personalised search, but show that our approach can be used for any kind of personalisation or recommendation. We focus on two aspects of the context: temporal context and activity-based context and describe a more general personalisation framework based on lightweight semantics, that can leverage any type of context.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Information filtering

1. INTRODUCTION

Recommendation systems and search engines play a crucial role in accessing the amount of content that can be found of the Web nowadays. Users are usually able to filter the information by issuing short keyword queries, but this model has several known disadvantages. The number of keywords is usually low, typically 1-3 keywords [8] and many of the words are ambiguous. The queries are almost never accurate [5], they are either too generic or too specific, but almost never exactly aligned with the specific intent the user has in mind. When we combine the impact of each of the described problem, we come to a conclusion, that finding the relevant document is indeed a difficult task, both for the user and the search engine.

A number of approaches have been researched, each with the specific goal of helping people find relevant content, preferably without changing the established paradigm of searching by keywords. The approaches range from modifying the query in-place to better capture the specific intent the user has in mind, i.e., query reformulation [3], learning to rank [9], or relevance feedback [12].

Each of the existing approaches leverages some sort of context to infer user's search intent and personalize the search results accordingly. Most of the time, the search context is long-term, built from all available data, e.g. learning to rank approaches use clickthrough data to learn user's preference based on her past search result clicks. The other extreme, short-term context is used in relevance feedback, where search results are modified according to user's last action, with a basic assumption, that a click represents a form of implicit feedback on the relevance of the clicked search result.

In all cases, the time by which the context is bound is determined statically and does not adapt to what is appropriate in the given situation. Long-term search context has the advantage that much data about the user is available, so the personalisation system can make reliable model of user's long-term interests. However, when a new, previously unseen goal appears, the information in the long-term model can shadow it, leading to bad personalization. E.g., an aquarium guide looking to buy a new *Barracuda* hard drive - the term *Barracuda* is associated with a sea fish in the model of his interests and the search results are personalized towards results that deal with sea and fish, while the guide is in fact looking for information on the specific model of the hard drive. On the other hand, the short-context can contain so little information, that no personalisation is possible, because no search goal can be inferred. We believe that the solution to this problem is to select the size of context dynamically at the time of search, and that the search context should be as specific as possible.

We have identified two types of search context that may be helpful in personalizing the search:

- *Activity-based context*, which is based on user's actions; what is she doing, how is she interacting with the search engine in terms of clicking on the search results. The activity-based search context is basically bounded by the changes in information-seeking goals. This context starts, and ends with each change in information-seeking goal.

- *Recurrent temporal context*, which is based on user's preferences in time. We assume that users form some kinds of recurrent habits that correlate with time, e.g., someone may be searching for work-related goals most of the time during work-week, while on weekends, the search goals are very different. The temporal habits

may happen on a larger scale, e.g. searching for asparagus recipes in spring, during the asparagus season.

2. CONTEXT MODELING AND PERSONALISATION FRAMEWORK

We model user's context using lightweight metadata extracted from accessed documents. Normally, a personalisation system would have access to this type of information in a direct way, but for the purposes of our research, we are collecting this information using a logging proxy server [1]. Usage of the proxy server is completely transparent for the users and does not affect system performance. Users have to explicitly opt-in for the proxy and they are aware that their activity is logged and we provide tools to selectively remove logged data or opt-out.

For each accessed document we extract document metadata, along with other user interaction indicators (mouse and keyboard activity and time spent on page), which are used to calculate implicit feedback on the page.

The context model is a hypergraph $H :=< V, E >$, with a set of vertices $V = A \cup P \cup T$, where A represents a set of users accessing the pages: $A = (a_1, a_2, \cdots, a_k)$, P represents a set of pages $P = (p_1, p_2, \cdots p_l)$ and T represents a set of terms $T = (t_1, t_2, \cdots, t_m)$; and a set of edges $E = (a, p, t)|a \in A, p \in P, t \in T$ and $P \cap T = \emptyset, A \cap P = \emptyset, A \cap T = \emptyset$. Using this representation is advantageous, as it allows for good denormalization and allows us to track each of the vertices independently. It may seem intuitive to merge accesses and pages, but this models allows us to abstract page from access, and if the document represented by the URL (page) changes, we can create new vertex in the graph and connect it respectively.

We use following series of steps to produce an interest model that can be used for recommendation or personalisation.

- First, search context can be selected by any arbitrary method. Context selection is based on selection of actions, e.g., we can use a long-term context and select all documents ever viewed by the user, or we can leverage activity based search context and select only the documents that she viewed for the purposes of her current information goal.

- Next, we extract a subgraph H_c from the original hypergraph H, by selecting only vertices $A_c \in A$, that match the context restriction established in previous step. The subgraph also contains vertices adjacent to A_c, i.e., $P_c \in P$ and $T_c \in C$ and their connecting edges $E_c \in E$.

- The subgraph H_c now represents the interest model for current context. We further enrich it by finding its nearest neighbors in terms of similarity in access patterns and similarity of the acquired document metadata T. This interest model and models of similar interests can now be used for personalisation or recommendation.

To validate this context acquisition approach we ran synthetic experiments [10] and experiments in a live system [11]. In both cases we used a simple form of long-term context, by using $H_c = H$. In both cases we evaluated our approach

in the domain of personalised search and used two simple query reformulation approaches:

- search the query history in the interest models (extracted from URLs of accesses to search engine, find reformulation patterns of the current query and select the one with highest satisfaction score, calculated by considering implicit feedback on clicked results, assuming, that unsatisfactory search results lead to clicks with low implicit feedback scores and to a further query reformulation until the user is satisfied.

- search the metadata in the interest models and find terms that are often co-occurring with the terms from the current query. Select the most frequent co-occurrences and extend the original query.

Using this experimental setup, we were able to substantially increase the relevance of search results and thus validate our approach. The proposed framework has two important properties: both context, and the personalisation method can be selected independently and we believe that using a different form of context than the long-term context, one could improve the relevance of recommendations even more.

3. ACTIVITY-BASED CONTEXT

Basic idea of an activity-based context is that information from previously accessed documents can help to understand the current information goal. E.g., imagine a series of queries: *watermelon; used cars; bentley; jaguar* – it is easy to see that the current query *jaguar* is connected with two previous queries, *used cars* and *bentley* and that the search goal in this case is to buy a car. With this knowledge, the personalisation can be moved towards search results that deal with used jaguar cars.

The problem of finding activity-based search context is stated as follows. The user is issuing queries and clicking on search results. The goal of activity-based search context detection is, for a current query, select the previous queries (and their respective clicked search results), that are part of the same search goal as the current query. This problem is equivalent to the problem of detecting the search session, when we define the search session as a set of queries following the same goal. Note that the queries that are part of the same search session may not be issued sequentially, but can be interleaved.

The problem of detecting search sessions is well researched, although the definition of a search session usually varies. The approaches could be categorized into three classes. The methods that are based on time, e.g. [6], are not appropriate for this definition of search session. They assume that a search session is bounded by a higher inactivity time, i.e. if the query is issued after a specified time from the last query elapsed, the query is considered to start a new session. The pause in the interaction with search engine can be caused by external factors (e.g., phone call, lunch) and user can return to the original task afterwards. Similarly, the search goal can change rapidly, without a pause in the interaction. The other class of approaches is based on lexical similarity between queries [7]. Two queries are considered part of the same session if they share common terms. However, this class of approaches does not deal with semantic similarity, e.g., queries *information retrieval* and *IR*. The third class of

Table 1: Results of search session segmentation using various approaches.

approach	precision	recall	f-score
lexical	0.985 ± 0.078	0.499 ± 0.410	0.662
temporal30m	0.932 ± 0.153	0.489 ± 0.380	0.641
temporal5m	0.974 ± 0.101	0.346 ± 0.373	0.510
metadata	0.992 ± 0.055	0.276 ± 0.312	0.432
lexical+meta	0.969 ± 0.108	0.540 ± 0.411	0.694

approaches uses signals from the search results page, such as snippets from the search results page [4].

The existing methods have two downsides: firstly, most of them do not consider multitasking in search, and restrict the search session to sequential queries. Once the search session is terminated, in cannot be reopened to add new queries. Secondly, existing methods that are using search results consider all search results equal.

We designed a method that considers multitasking in web search and incorporates implicit feedback on the search results. We build a query model, where we leverage the same lightweight semantic document metadata as in our personalisation framework. Each query is associated with the respective clicked search results. We leverage a negative implicit feedback and assume that a search result that was viewed for less than a specified number of seconds was not relevant to the search goal and therefore closed. We remove these documents from the query model.

To put the query into a session, we look at the clicked search results and their respective metadata. For each query, a metadata vector is built and compared to metadata vectors of currently active search session. When the vectors are similar (in terms of shared metadata), the query is considered part of the session and its metadata is merged into session model. At any given time, a stack of previous search sessions is being maintained and if the query model does not match the current session model, the stack of sessions is searched, until a match is found or the stack is exhausted. When a match is found, the matching search session is moved to the top of the stack and becomes active search session. Old search sessions are removed from the stack automatically, so that they cannot become active.

We have experimentally evaluated this approach against a set of 245 manually segmented queries of 3 different users. We have compared the results with other approaches (temporal with a window of 5 and 30 minutes and lexical). The results are described in Table 1. *Precision* indicates the internal coherence of the session. First, queries from automatically detected session are linked to the manually detected sessions and precision is calculated as the ratio of the cardinality of the best match to the cardinality of the whole session. *Recall* indicates the completeness of the session. It is calculated as the ratio of cardinality of the best match against the manually created sessions to the cardinality of the best-matched session. Best results were obtained by combining the lexical approach with our session matching method (referenced in the table as *metadata*) as a fallback, i.e., the first heuristics of a session match is the lexical similarity of queries and if it is negative, the document metadata are compared.

The preliminary experiments seem promising and we be-

lieve that applying activity based context to a personalisation process will yield better results that using a long-term context.

4. RECURRENT TEMPORAL CONTEXT

We postulate that there are patterns in user behavior changes that are related to time. For example searching for summer resorts before summer holidays, or searching for which movies to watch on Sunday night. We believe that these patterns are recurrent and appear more or less regularly during the same periods of day, week, month or year.

To the best of our knowledge, there are no studies that would support or disprove our hypotheses. There are many works that study the temporal dynamics of search, but all of them analyze it only globally, and do not focus on individual patterns. For example, in [13] authors analyze transactional logs from the Excite search engine using Markov matrices and Poisson sampling and explore variations in aggregated users' searches related to changes of time in day. Or similarly, Beitzel et al. [2] analyzed a partially manually topically labeled large-scale transactional log of a search engine and analyzed topical patterns in the queries and results, but again, only in an aggregated form.

In our work, we analyze changes in user interest for the top-100 users from the publicly available AOL search engine. We focused on analyzing the patterns of interest changes during work hours and free time. For each user in the top-100 log, we have separated the clicked search results into two distinct clusters using two setups: work-week vs. weekend and work-hours (9:00 - 17:00) vs. free-time (the rest). Next, we compared how well were the two clusters separated from each other in each of the two setups.

For this analysis, we used HTML code of the search results and extracted keywords, description and title of each page (provided by the author) and it's respective ODP categories. For each user, these vectors of textual metadata were converted to binary vectors, 1 indicating presence of the metadata in the vector, 0 otherwise. These vectors were then treated as vectors in Euclidian space and we calculated standard Davies-Bouldin score for each user's clusters in both setups. Davies-Bouldin score represents how well are the clusters separated from each other and how coherent are they internally. The results for both setups are displayed in Figure 1.

The Davies-Bouldin scores suggest, that our hypotheses are not valid in general, but there exist some users, for which they are true. In both setups, there are some users for whom the score is low, which means that they indeed search for semantically different information during weekend and during work-week or during work-time and free-time hours. What is interesting, is that the calculated scores cover the whole range of scores nearly linearly, without any peaks. This suggests that the probability of a user changing interests on the two temporal conditions exhibits uniform distribution.

5. CONCLUSIONS AND FUTURE WORK

In our work, we try to design a framework for personalisation based on context constraints, with the aim of selecting the proper type of context dynamically. So far, we have designed and evaluated a framework based on lightweight semantics and nearest neighbors. We have identified two types of context that might be helpful in constraining the
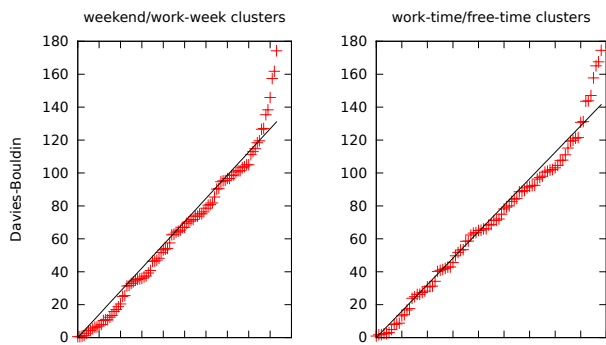
Figure 1: Davies-Bouldin cluster scores calculated for the two setups: work-week/weekend clusters and work-time/free-time hours assigned to each of the top 100 active users, scored ascendingly by the value of the score. Lower values of the score denote more tight and separated clusters. Each plot contains a best linear fit.

personalisation process and have done some initial work in order to understand them better. We have designed and evaluated a novel approach for search session detection that is based on implicit feedback and designed with multitasking in mind. We have also done an initial analysis of temporal recurrent patterns in user searching behavior and are now working on extending our work to automatically detect different types of recurrent patterns in clickthrough data automatically. Eventually, we plan to design a method for selecting an appropriate type of context (e.g., long-term, activity-based, or temporal) based on the analysis of the contexts found by each of the methods.

Our approach is not limited to search personalisation only, it can be used for any kind of personalisation or recommendation. We believe that the two identified context types can be found in other domains, e.g. a user researching a certain topic in a news portal is apt for an activity-based context recommendation [14], or a user searching for information on Olympic games is apt for a temporal context based recommendation.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] M. Barla. Towards social-based user modeling and personalization. *Information Sciences and Technologies Bulletin of the ACM Slovakia*, 3:52–60, 2011.

[2] S. M. Beitzel, E. C. Jensen, A. Chowdhury, D. Grossman, and O. Frieder. Hourly analysis of a very large topically categorized web query log. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '04, pages 321–328, New York, NY, USA, 2004. ACM.

[3] C. Biancalana, A. Micarelli, and C. Squarcella. Nereau: a social approach to query expansion. In *Proceeding of the 10th ACM Workshop on Web Information and Data Management*, WIDM '08, pages 95–102, New York, NY, USA, 2008. ACM.

[4] M. Daoud, M. Boughanem, and L. Tamine-Lechani. Detecting session boundaries to personalize search using a conceptual user context. In S.-I. Ao and L. Gelman, editors, *Advances in Electrical Engineering and Computational Science*, volume 39 of *Lecture Notes in Electrical Engineering*, pages 471–482. Springer Netherlands, 2009.

[5] D. Downey, S. Dumais, D. Liebling, and E. Horvitz. Understanding the relationship between searchers' queries and information goals. In *Proc. of the 17th ACM conference on Information and knowledge management*, CIKM '08, pages 449–458, New York, NY, USA, 2008. ACM.

[6] E. Herder. *Forward, back and home again. Analyzing user behavior on the Web*. (phd) dissertation, University of Twente, 2006.

[7] B. J. Jansen, A. Spink, C. Blakely, and S. Koshman. Defining a session on web search engines: Research articles. *J. Am. Soc. Inf. Sci. Technol.*, 58:862–871, April 2007.

[8] B. J. Jansen, A. Spink, and T. Saracevic. Real life, real users, and real needs: a study and analysis of user queries on the web. *Inf. Process. Manage.*, 36:207–227, January 2000.

[9] T. Joachims. Optimizing search engines using clickthrough data. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '02, pages 133–142, New York, NY, USA, 2002. ACM.

[10] T. Kramár, M. Barla, and M. Bieliková. Disambiguating search by leveraging the social network context based on the stream of user's activity. In *UMAP '10: Proc. of the 18th Int. Conf. on User Modeling, Adaptation, and Personalization*, pages 387–392, Hawaii, USA, 2010. Springer.

[11] T. Kramár, M. Barla, and M. Bieliková. Personalizing search using metadata based, socially enhanced interest model built from the stream of user's activity. *Journal of Web Engineering*, 2012. (submitted).

[12] Y. Lv and C. Zhai. Adaptive relevance feedback in information retrieval. In *Proceedings of the 18th ACM conference on Information and knowledge management*, CIKM '09, pages 255–264, New York, NY, USA, 2009. ACM.

[13] S. Ozmutlu, A. Spink, and H. C. Ozmutlu. A day in the life of web searching: an exploratory study. *Inf. Process. Manage.*, 40:319–345, March 2004.

[14] D. Zeleník and M. Bieliková. News recommending based on text similarity and user behaviour. In *Proceedings of the 7th International Conference on Web Information Systems and Technologies*, WEBIST, pages 302–307, Noordwijkerhout, The Netherlands, 2011. SciTePress.