Context-Based profile Personalization

From: AAAI Technical Report WS-98-08. Compilation copyright © 1998, AAAI (www.aaai.org). All rights reserved.

Chanda Dharap

Philips Multimedia Center
1040 Arastradero Road
Palo Alto, California CA 94304-1336
Chanda@pmc.philips.com

Why Personalization ?

This position statement outlines our work to increase the effectiveness of an information service by personalizing user profiles built based on user access patterns, and dynamic extraction of user context information

Rapidly expanding information archives provide access to gigabytes of electronic data, via electronic museums, newspapers, musical archives, digital libraries, software archives, mailing lists, up-to-date weather information and geographic data. Consequently current advances in information technology are driven by the need to increase the personalization of information for the user. We believe that the key to effective information manipulation lies in strategies that improve the users' ability to cluster and recluster the information space in a manner effective for highly personalized browsing.

Traditionally information providers try to overcome the inadequacies of information retrieval by providing fast and powerful search engines. Examples of searching systems are commonly available search engines, databases and library lookup systems, where the user provides a query with sufficient information and gets back a set of documents that match the query. One problem with existing search systems is that if the query is not very precise, the user is left with the task of scanning through a large of amount of result data to identify documents of interest. As a result, a large percentage of information that is retrieved is not very usable. From our investigations in this area we find that unlike a real life approach similar to "browsing at a book store", where an interesting cover may catch your eye and redirect your path, the retrieval approach is limited in that the result set contains information precisely related to the query and not much else. Consequently, the information provider is unable to retain or divert the users' focus to other yet related works of interest.

Copyright © 1998, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

User Context

In order to increase the effectiveness of browsing wide-area information, we focus primarily on the users interest as defined by access history. The key idea is to use access history to establish a context for search. By understanding and establishing the users's context we use it towards personalizing a profile for the user. Further, this profile is used as a recommendation for mapping relevant information from the provider's topic space into the users search space. In order to achieve this, the profile is partitioned into static and dynamic partitions. This approach is unique in that it allows the user to retain a constant theme in her profile (static partition) as well as to influence the profile by new interests generated while browsing the provider's space (dynamic partition). This gives a unique mechanism for providers to attract and retain the consumer's interest while browsing at their sites.

This approach combines two strategies in a unique way to provide highly personalized browsing. Context is established by information mined from user profiles and dynamic extraction of users' interactions with the system. To establish context, we extract and analyze user query logs in a log analysis module. First we created a mechanism to identify the minimal set of queries that define the context of the user's query. Context is a list of keywords associated with the topic under consideration. Since queries are logged as keyword pairs, we can easily extract keywords to define the context.

In the next step we identify contexts via the context detection phase. End of one context is the begin of another context. A user typically has several context shifts. The following example illustrates the idea of a context shift. A user looking for ideas on investing, may have started off by using the keywords investments, IRAS and banking. The query is then enhanced by adding and dropping a few keywords. During the course of the search the user may get attracted to the idea of on-line banking and may add the keyword Quicken and drop IRAs. This leads to articles about Wells Fargo's Bill Pay system and the query is

further enhanced by adding Bill Pay. After arriving at desired results, the user either quits the search, or shifts interest to another topic altogether, say, Microprocessor Architectures. We define this as a context shift. That is, a shift in the query that indicates a change of interest. By detecting such context shifts it is possible to narrow the users' search path. For example a search for ATM could imply either information regarding the ATM networking protocol or Automated Teller Machines. Based on past access history, we establish a context for this query.

Profiles

Next is the profile creation/updation phase. The results generated by a search are analyzed and a set of concept keywords are generated from them. The generation algorithm is based on topical partitioning of the provider's space and a weighted topical dictionary. Profiles are partitioned into static and dynamic portions. The static partition reflects the users' long term interests, for example this may be related to the users' research. The dynamic partition reflects the path taken by the search within the provider's information space. Updates to the static partition are made manually by the user and are expected to be occasional. On the other hand, updates to the dynamic partition are made during an ongoing search and are based on the concept words generated by the context detection phase. The user can then observe changes to the dynamic portion of the profile and select the dynamic profile as a recommendation for enhancing further queries.

Conclusion and Related Work

Traditional approaches have devised mechanisms to map a user's query to a document based on overlapping concept words between the query and the document terms. One approach [2] is to use a quantitative measure of semantic similarity between index terms for queries and documents. Another recent approach[3] introduces concept-based query expansion, where each concept is expanded to a disjunctive set of concepts on the basis of conceptual relationships pointed out by the user. Yet another idea[4] is to compute relevance feedback techniques that can cope with shifts in user interest patterns over a period of time. Earlier work we have done related to information retrieval and browsing can be found in the Proceedings of the Fifth International Conference on Information and Knowledge Management[1].

Our work demonstrates mechanisms that can be used by the user to personalize the path of search in two ways:
1) by modifying her profiles or 2) letting the system guide the way. This use of context enables us to add a highly personalized browsing capability to the system.

References

[1] Chanda Dharap and Martin Freeman. Information Agents for Automated Browsing. In Proceedings of the Fifth International Conference on Information and Knowledge Management.

[2] Alan F. Smeaton and Ian Quigley. Experiments on Using Semantic Distances Between Words in Image Caption Retrieval. In Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. August 1996, Zurich, Switzerland.

[3] Kalervo Jarvelin, Jaana Kristensen, Timo Niemi, Eero Sormunen and Heikki Keskustalo. A Deductive Data Model for Query Expansion. In Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. August 1996, Zurich, Switzerland.

[4] James Allan. Incremental Relevance Feedback for Information Filtering. In Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. August 1996, Zurich, Switzerland.