Sharing Bookmarks among Same Interest Persons

Mikihiko Mori and Seiji Yamada
CISS, IGSSE, Tokyo Institute of Technology
4259 Nagatsuta, Midori, Yokohama 226-8502, Japan
mori@ymd.dis.titech.ac.jp

Abstract: We propose Bookmark-agent system that enables users to share information in each user's bookmarks with others. This system is more effective in finding specific information that is a common interest of all the members in a group than existing search engines, such as Yahoo and AltaVista. When a user wants to find certain URLs, the Bookmark-agent will find them on behalf of the user. In addition to that, the Bookmark agent requests other Bookmark agents to search their own bookmarks. As a result, the user can obtain information that best matches his/her needs from the Bookmark-agent. Since users filter the selected information beforehand, it is possible for the Bookmark-agent to satisfy a users' requests more precisely than existing search engines. After conducting experiments on six users, we have found that the Bookmark-agent system is a promising approach to share URLs in a small community.

1 Introduction

In this paper, we propose Bookmark-agent system. The agent, which represents each user, uses user's bookmarks to get he/her interest. The bookmark is URL as filtered information from all Web pages in the WWW. In a small user group having the same interest, the agents share bookmarks of each user as their interest through communication in between them. The group has been gathered beforehand. A Bookmark-agent assists a user to search for the sites that resemble a Web page visited. The agent refers user's bookmarks and others through communication in between them if users allowed the bookmarks to be open to public. There are no need to give any keywords to the agent, because the agent can extract some keywords from the Web page a user has just visited. Note that a Bookmark-agent gives its user a smaller number of more significant URLs than the result given by ordinary search engines, because the agents utilize bookmarks, which are filtered information of the WWW by users, and are only used in a small group sharing the same interest. Keywords that represent users' interest make the number of presenting URLs less than the result given by ordinary search engines.

Siteseer (Rucker & Polanco 1997) utilizes each user's bookmarks as an implicit declaration of the interest in the underlying content. The difference from Bookmark-agent is that Siteseer measures the degree of overlap in URLs between each user's folder and other people's folders. WebMate (Chen & Sycara 1998) learns user's interest incrementally and automatically sends mails that match user's interests. Moreover, it adds keywords from a user profile to search for URLs of interesting sites. WebMate monitors user's browser to get user's profile, while Bookmark-agent makes profile from bookmarks. Beehive (Huberman & Kaminsky 1996) designs a distributed system for social sharing and filtering of information on email system. Information of interest by a user is sent to members of listed groups through Beehive system. The member who does not communicate is deleted from the member list, if the frequency of his writing messages is lower than a threshold. Content of message is no concern of this system.

2 Bookmark-agent

An Bookmark-agent is invoked for each user and assists searching the WWW. Functions of Bookmark-agent are the followings:

- Extracting keywords from a new URL.
- Indicating URLs of similar pages[1].
  Searching for similar pages requires the following processes:
  - Searching the user's keyword-database.
  - communicating and requesting to other agents in order to search other users' keyword-databases.
- Accepting a search request.

We define a set of keywords for a Web page. Each page consists of words surrounded by tags, which weighs the words. We prepare a set of Keyword-weight pairs. The set has no duplication of the same keyword, and weights for the same word are added together. The top 5 keyword-weight pairs are selected to be used as the keyword set of the Web page. The following tags and weight assigned to each tags (in parenthesis) are used to calculate the weight for each keywords: `CONTENT` attribute in `<META>` tag (10), `<TITLE>` tag (10), `<H1>` tag (6), `<H2>` tag (5), `<H3>` tag (4), `<EM>` tag (1), `<U>` tag (1).

---

[1]defined as a page whose similarity is over a threshold.

The similarity between two Web pages is defined as the number of elements in the product set in between the keyword sets of these two Web pages. The similarity varies from 1 to 5.

3   Experiments

We calculate precisions when one to three keywords in the databases are given to an agent. We chose the keywords that are surely contained in the keyword-databases of all users.

For comparison with Bookmark-agent, the search engine Goo[2] is used, since Goo answers a lot of URLs and uses a Web robot for retrieval. Our assumption is that Goo's capability of answering a lot of URLs must have covered all area of ordinary interest. It is not easy for users to find necessary URLs out of many URLs search engines recommend. To solve this problem, some search engines present URLs in the order sorted by priority to a user. To evaluate precision from the proportion of the retrieved documents that is actually relevant, if the search engine answered 20 URLs or more, we decided to use the URLs from the top to the 20th.

For evaluating the performance of Bookmark-agent, we prepared six users as testees who had 782 bookmarks as a whole. These users used Bookmark-agent and had already their own keyword-database.

In these experiments, a user inputs keywords to Bookmark-agent or the search engine, Goo. The relevance of each URL presented was judged from his point of view.

4   Discussion on The Experimental Results

A keyword used in a group sharing concept is presented higher precision than one is not used in the group. For instance, 'agent' usually means that the person who takes user's place or give service for users who cannot do. However, it is clear that the 'agent' in the group of testees means agent system in AI. The precision of keyword 'agent' in the search engine Goo was half or less than average. However, the precision of 'agent' in Bookmark-agent did not change, regardless of the number of keywords. Therefore, it is reasonable to suppose that users have common sense regarding the keyword, and it is considered as that the users share preference.

The rank sum test between precision of each keyword from Bookmark-agent system and that from Goo was conducted for each case of Table 1 to 3. As a result, we found a difference between the result of Bookmark-agent and that of Goo on significance level of 5% in each test. Thus, the method of the search engine Goo did not have ability of that of Bookmark-agent.

First, we pay attention to the preference that is shared in the group. Although the user in the experiments does not share preference in general sciences, they have common sense in AI, robot and agent system. Which is the reason that the precision of 'ai', 'robot' and 'agent' are higher than that of 'science'. The keyword in which they share preference presents more URLs than the one in which they do not share.

Next, we pay attention to the preference that is not shared. The keywords 'robot' and 'linux', the query of which returned 10 and 22 URLs respectively, are the words which concepts are shared among the users. On the other hand, they do not share the concept of the word 'perl', because only the user C have preference on 'perl.' As a result, the keyword 'perl' returned only 3 URLs. Normally, a user must have been indicated more URLs to 'unix' than to 'linux'. The keyword 'linux', however, returned more URLs than 'unix', because the number of the users using Linux is more than that of using Unix.

5   Conclusion

We proposed Bookmark-agent, which uses bookmarks to support users' searching for information, and conducted experiments on six users. Sharing URLs of each user's bookmarks, which helps indicating necessary Web pages, Bookmark-agent provides a few but significant URL as a search result. Bookmark-agent is clearly different from general search engines on the point that search engines often give a large number of URLs that often contains irrelevant information to users. By using bookmarks, Bookmark-agent is able to provide a small number of URLs to a user who shares preference with other users.

References

Chen, L. and Sycara, K. (1998). *WebMate: A Personal Agent for Browsing and Searching*, Proc. 2nd Int. Conf. Autonomous Agent, 132–139.

Huberman, B.A. and Kaminsky, M. (1996). *Beehive: A System for Cooperative Filtering and Sharing of Information*, http://www.parc.xerox.com/spl/groups/ dynamics/www/beehive.html.

Rucker, J. and Polanco, M.J. (1997). *Siteseer: Personalized Navigation for the Web*, Commun. ACM, 40 (3), 73–75.

---

[2]http://www.goo.ne.jp/