UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Distributed, Community Information Filtering

Permalink

Journal

Author

Recker, Mimi

Publication Date

1997

Peer reviewed

Distributed, Community Information Filtering

Mimi Recker (mimi.recker@vuw.ac.nz)
Victoria University of Wellington
P.O. Box 600, Wellington, New Zealand

You've arrived in a brand new city, and hunger pangs have erupted. How do you decide that all-important question: Where to dine? You might consult restaurant guides, newspapers, or ask friends. But, in the end, you want trusted sources to provide you with information about the quality of restaurants in order to help you with your selection.

Finding and choosing information on the Web has much in common with the "restaurant" problem. The explosive increase in the number of World-Wide Web resources has seen a corresponding growth in the problem of finding relevant, quality, and validated information. The Web lacks the structure and strong typing found in more closed database system. Moreover, its distributed nature precludes the implementation of filtering and reviewing conventions typically provided by libraries and publishers.

Within information systems, there have been several promising approaches to the problem of labelling, categorising, and filtering information. Malone et. al. (1987) describe three types of information filtering activities: cognitive, economic, and social. **Cognitive** filtering is based on indexing content (and is what most Web search engines do). Content-based filtering depends on a machine-readable and parseable format. Unfortunately, this can be difficult to implement in a multimedia environment. **Economic** filtering is based on a cost-benefit analysis of searching activities. While a powerful approach in large information repositories, it generally prevents serendipitous discovery of information. **Social** information filtering is based on word-of-mouth and recommendations from trusted sources (Maltz & Ehrlich, 1995; Shardanand & Maes, 1995)

In this research, we are developing an infrastructure for supporting collaborative, distributed information filtering of Web resources. Such a system has several important requirements, which we briefly describe in turn.

First, the system must easily integrate into the existing Web infrastructure. A system imposing additional cognitive and technical overhead is much less likely to be used. To this end, we build on widespread Web technology by using a simple relational database that communicates with the user's Web client via a Common Gateway Interface (CGI) compliant script.

Second, it must employ a useful yet simple rating scheme. As a first pass, we are focussing on resources within a particular community of users, specifically, Asia-Pacific Law and Business (an area well-known to a member of the research group). Within this area, we have developed a partially domain-specific rating vocabulary, complemented by dimensions applicable to Web resources generally. Domain-specific rating dimensions include geographical focus, subject category, the reputation and authoritativeness of the information source, and language(s) used. More general rating dimensions include the nature of the organisation running the particular site, the richness of the information, the ease of navigation, graphical reliance, and overall site layout.

Third, the system must make it easy for users to add ratings of Web documents. Again, a simple Web-Forms interface embedded into a Web-Frame allows users to add their ratings as they view source documents.

Fourth, a critical mass of users must participate to ensure rating reliability. Naturally, Web availability ensures a large potential user pool.

Fifth, it must be easy for information seekers to see and understand the ratings of other users. We are currently experimenting with several client-side displays of "community-relevant representations." For example, if the database has several relevant ratings, it presents a composite picture of the document, thereby capturing community knowledge. Ratings are also augmented with contextual information, such as title of the document, author of the rating, and usage history (Hill et al.; 1995; Maltz & Ehrlich, 1995). Together, this meta and contextual information should help users evaluate the value and quality of particular Web resources.

We are currently conducting pilot evaluations of our system involving several users. These studies examine the usability and usefulness of the approach, from both HCI and social information filtering perspectives.

Acknowledgments

References

Hill, W., Stead, L., Rosenstein, M., and Furnas, G. (1995). Recommending and evaluating choices in a virtual community of use. *ACM Conference on Human Factors in Computing Systems* (pp 194-201). New York, NY.

Malone, T., Grant, K., Turbak, F., Brobst, S., and Cohen, M. (1987). Intelligent information sharing systems. *Communications of the ACM, 30*(5).

Maltz, D. & Ehrlich, K. (1995). Pointing the way: Active collaborative filtering. *ACM Conference on Human Factors in Computing Systems* (pp 202-209). New York, NY: ACM.