Exploiting Implicit User Feedback

Charlie Clarke
David R. Cheriton School of Computer Science,
University of Waterloo

Abstract

The implicit feedback users provide while interacting with a system – including clickthroughs and browsing behavior – is an important resource that often goes uncaptured and ignored. As an alternative, we present two examples of how this information may be exploited to improve the performance of systems that depend on subjective judgments of quality and relevance. Both examples are taken from search engines – one from a Web search context and the other from an enterprise search context. The long-term goal of this work is to develop search engines, and related systems, that self-tune as they are used.

Web search engines present lists of "captions", comprising title, snippet, and URL, to help users decide which search results to visit. Understanding the influence of features of these captions on Web search behavior may help validate algorithms and guidelines for their improved generation. We have developed a methodology to use clickthrough logs from a commercial search engine to study user behavior when interacting with search result captions. The findings of our study suggest that relatively simple caption features such as the presence of all terms query terms, the readability of the snippet, and the length of the URL shown in the caption, can significantly influence users' Web search behavior [1].

In an enterprise search context, retrieval accuracy may be improved by considering which document types should be filtered out and which should be ranked higher in a result list. Hence, document type may be used as a key factor for building a re-ranking model. We take a simple approach for considering document type in the retrieval process, by adapting a standard scoring function to weight term frequency based on document. The values for these weights are estimated from clickthrough data. Preliminary experiments indicate a possibility for significant improves in search performance [2,3]. Based on these examples, a general plan for developing systems of this type can be outlined. Developers are encouraged to exploit implicit user feedback, while protecting the privacy of their users.

accuracy by weighting document types with clickthrough data. SIGIR 2007: 759-760.

**Presenter Biography**

Charles CLARKE is an Associate Professor in the Department of Computer Science at the University of Waterloo. His research interests include information retrieval, software development tools, and programming language implementation. He has published on a wide range of topics, including papers related to question answering, XML, program analysis, file system search, user interfaces, statistical natural language processing, and the evaluation of information retrieval systems. He is a Program Co-Chair of the 2007 ACM SIGIR Conference on Research and Development in Information Retrieval; he was the General Co-Chair of SIGIR in 2003. He received his Ph.D. from Waterloo in 1996. From 1996 to 1999 he was an Assistant Professor in the Department of Electrical and Computer Engineering at the University of Toronto. He has previously held software development positions at a number of computer consulting and engineering firms.