Trust-enhanced peer-to-peer collaborative web search

Smyth, Barry; Briggs, Peter

2008-12

Beyond Search: Computational Intelligence for the Web Workshop at the twenty-Second Annual Conference on Neural Information Processing Systems (NIPS 2008)

http://hdl.handle.net/10197/1300
Abstract

We spend a lot of our time online using web search services, but even the leading search engines frequently fail to deliver relevant results for the vague queries that are commonplace among today’s web searchers. Interestingly, when we look at the search patterns of link-minded searchers (perhaps friends or colleagues) we do find considerable overlap between their queries and result-selections. This motivates a more collaborative approach to web search, one in which the past search experiences of friends and colleagues can be used to usefully influence our new searches. In this talk we will describe how a novel combination of case-based reasoning, web search, and peer-to-peer networking can be used to develop a platform for personalized web search, which benefits from better quality results, improved robustness against search spam, while offering an increased level of privacy to the individual user.

1 Motivations

Web search is far from perfect and many recent studies have highlighted the extent to which leading search engines struggle to provide users with relevant results. For example, Smyth et al. [1] describe how as many as 56% of Google Web searches fail to attract any result selections. Over the past few years, as “the business of search” has matured in to a major market sector, researchers have continued to look for new ways to enhance existing search engine technology. In this regard the idea of “social search” — that result-lists might usefully be influenced by interests, preferences, or activities of other searchers — has gained some considerable attention as a way to improve search quality by personalizing result-lists.

In our research we have been developing social search technologies that combine ideas from social networking and conventional web search in order to improve existing web search services. Specifically we have looked at different ways to harness the search activities of communities of like-minded users to improve result quality. For example, collaborative web search (CWS) (Balfe & Smyth [2]) demonstrates how the search experiences (queries and result selections) of communities of like-minded users can be reused as a source of result recommendations (promotions) during future searches; in short, for a new target query, results that have been frequently selected by community members for similar queries in the past are promoted during the new search.

There are limitations to this standard approach to collaborative web search, however. For example, it relies on some explicit representation of a search community even though users may not find it convenient to select a community context beforehand. Another limitation is that individual community members cannot be identified. In fact this is often cited as a privacy benefit, but in truth it has a downside when it comes to auditing the source of a promoted result. As the seminal work Leake & Whitehead highlights, the origin or provenance of an experience can be an important quality indicator [3]. This is especially true in CWS because it is possible for malicious users to influence
result promotions [4]. By recording the source of a promotion (the searcher who originally selected the result), it is possible to present provenance information alongside the promotion as a form of explanation. But, this is only possible if individual users can be distinguished within a community.

2 Core Contributions

In this paper we present an alternative model of collaborative web search. One that avoids the need for explicit communities, and which facilitates the identification of individual searchers to determine the provenance of promotions. We demonstrate how this can be used to enhance the conventional CWS interface and show how provenance information can help to improve the quality of results in two important ways. First, we will show how it can be used as the basis for a computational model of user-trust, which can be used as a promotion filter. Second, we argue that exposing the provenance of promotions through the search interface helps to build social relationships between searchers, helping them to avoid making spurious result selections. Furthermore, we explain how the advantages of this trust-based approach can be achieved while preserving the privacy of individual searchers by implementing a distributed peer-to-peer search network (see Figure 1). In this network, the search histories of individuals are maintained by their own local search-agent and only shared on the basis of trusted relationships between search peers. As an added benefit, we explain how this peer-to-peer architecture facilitates a more flexible approach to CWS by avoiding the need for explicit communities; essentially, an individual’s search community evolves as they develop implicit relationships with other online searchers via the sharing and promotion of search experiences.

References


