Researchers examine real-time search behavior

May 12 2010

(PhysOrg.com) -- Where do you go to find information online? Increasingly, people are looking to both traditional search engines and 'real-time' sites such as Facebook and Twitter for answers to questions and problems.

However, finding relevant content through these channels can be difficult, and real-time sites are not indexed in the same way as traditional search engines because the search results are continually updating as new content is posted, rather than remaining on one static page.

Given these challenges, Jim Jansen, associate professor of information sciences and technology at Penn State, sought to identify the characteristics of users who are looking for real-time content in a recent study involving more than 1 million queries collected over a 190-day period.

Jansen found that many users of real-time search engines are doing so from secondary applications rather than from the platform's website, that queries tend to repeat over multiple days and that the searches tend to focus on technology, entertainment and politics - rather than the pornographic or sexual queries that are popular topics for traditional search engines.

"The access of real time search results from secondary application has important implications for how these results are used and how the results

can be leveraged for marketing purposes," Jansen said.

Some 1 million searches performed over 90 days were examined in the study. About 30 percent of the queries were unique, which is very low compared to traditional search engines, where as many as 60 percent of queries are unique.

"Real time search is a compelling new area of Web interaction, with potential as a new channel for information gathering, advertising, and other uses," Jansen said. "As people become more accustomed to using real time content, real time search will become still more important. Therefore, understanding how people locate information in this context is critical."

Jansen collaborated with Gerry Campbell and Matthew Gregg of Collecta, a real-time search engine that pulls results from blogs, microblogs, news feeds, and photo-sharing services. Collecta provided the data analyzed in the study.

The paper, "Real Time Search User Behavior," was presented at the 2010 ACM Conference on Human Factors in Computing Systems, held April 10-15 in Atlanta. It can be viewed online.

Provided by Pennsylvania State University