An Empirical Performance Measurement of Microsoft’s Search Engine and its Comparison with Other Major Search Engines

Xiannong Meng, Bucknell University, USA
Song Xing, California State University - Los Angeles, USA
Ty Clark, Bucknell University, USA

ABSTRACT

This article reports the results of a project attempting to assess the performance of the new Microsoft search engine (MSE) from various perspectives. Specifically, the study collects statistics, such as the average user response time, average process time for a query reported by MSE itself, the number of pages relevant to a query, and comparisons with its competitors. The project also studies the quality of search results generated by MSE and other search engines using RankPower as the metric. We found MSE performs well in speed and diversity of the query results, while weaker in other statistics, compared to some other leading search engines. The contribution of this article is to use different measures to assess and compare the quality of different search engines, especially MSE.

Keywords: information search and retrieval; interact IS; relevance information; user needs

INTRODUCTION

Search engines, since their inception in the early to mid-1990s, have gone through many stages of development. Early search engines were derived from the work of two different but related fronts. One is to retrieve, organize, and make searchable the widely available, loosely formatted HTML documents over the Web. The other is the then-existing information access tools such as Archie (Emtage, 1992), Gopher (Anklesaria, McCahill, Lindner, Johnson, Torrey, & Alberti, 1993), and WAIS.
(Kahle, 1991) (wide area information servers). Archie collects information about numerous FTP sites and provides a searchable interface so users can easily retrieve files through different FTP sites. Gopher provides search tools to large number of Gopher servers on the Internet. WAIS has similar functionality to that of Archie, except that it concentrated on wide variety of information on the Internet, not just FTP sites. With the fast development of the Web, search engines designed just for the Web started to emerge. Some of the examples include WWW (World Wide Web Worm), the then-most-powerful search engine AltaVista, NorthernLight, WebCrawler, Excite, InfoSeek, HotBot, AskJeeves, AlltheWeb, MSNSearch, and of course, Google. Some of these search engines disappeared in history; others were retooled, re-designed, or simply merged; yet others have been able to stay at the front against all the competition.

Google, since its inception in 1998, has been the most popular search engine mostly because of its early success in its core algorithm for search, the PageRank algorithm (Brin & Page, 1998). In a relatively short history many aspects of search engines including software, hardware, management, investment and others have been researched and advanced. Microsoft, though a later comer in the Web search business, tried very hard to compete with Google and other leading search engines. As a result, Microsoft unveiled its own search engine on November 11th, 2004 with its Web site at http://beta.search.msn.com (Sherman, 2005). Since the new search engine does not have a specific name, we will call it the MSN search engine or MSE in this study. This article reports the results of a project attempting to assess the performance of the Microsoft search engine from various perspectives. Specifically the study collects statistics such as the average user response time, average process time for a query reported by MSE itself, the number of pages relevant to a query, the quality of the search in terms of RankPower, and comparisons with its competitors. The rest of the article is organized as follows, the second section provides an overview of search engine performance metrics. The goals and the metrics of this study are described in the third section. The fourth section discusses the method of study and the experimental settings, followed by the results and their analysis in the fifth section. Our thoughts and conclusions about the study are presented in the sixth section.

An Overview of Performance Metrics for Information Retrieval Systems and Search Engines

Everyone has their opinions about which search engine is the best. The answer most likely depends on the scenario of search and the information needs of the user. While user perceptions are important in measuring the retrieval performance of search engines, quantitative analyses provide more “scientific evidence” that a particular search engine is “better” than the other. While there
Related Content

Collusion-Free Privacy Preserving Data Mining
[www.igi-global.com/chapter/collusion-free-privacy-preserving-data/64381?camid=4v1a](www.igi-global.com/chapter/collusion-free-privacy-preserving-data/64381?camid=4v1a)

Clustering Analysis and Algorithms
[www.igi-global.com/chapter/clustering-analysis-algorithms/24291?camid=4v1a](www.igi-global.com/chapter/clustering-analysis-algorithms/24291?camid=4v1a)

Interval Cut-Set of Generalized Interval-Valued Intuitionistic Fuzzy Sets
[www.igi-global.com/article/interval-cut-set-generalized-interval/68991?camid=4v1a](www.igi-global.com/article/interval-cut-set-generalized-interval/68991?camid=4v1a)
Human Health Risk Assessment via Amalgamation of Probability and Fuzzy Numbers
