Chapter XVI

Analysis of Document Viewing Patterns of Web Search Engine Users

Bernard J. Jansen
The Pennsylvania State University, USA

Amanda Spink
University of Pittsburgh, USA

ABSTRACT

This chapter reviews the concepts of Web results page and Web page viewing patterns by users of Web search engines. It presents the advantages of using traditional transaction log analysis in identifying these patterns, serving as a basis for Web usage mining. The authors also present the results of a temporal analysis of Web page viewing, illustrating that the user — information interaction is extremely short. By using real data collected from real users interacting with real Web information retrieval systems, the authors aim to highlight one aspect of the complex environment of Web information seeking.

INTRODUCTION

The Web has dramatically changed the way people locate information. One can define Web mining as:

the discovery of and analysis of useful information from the World Wide Web. This describes the automatic search of information resource available online, i.e., Web content mining, and the discovery of user access patterns from Web services, i.e., Web usage mining (Cooley, Mobasher & Srivastava, 1997).
Information viewing characteristics of users are central aspects of this view of Web mining. As the Web has become a worldwide phenomenon (Cole, Suman, Schramm, Lunn & Aquino, 2003), we need an understanding what searching trends are emerging, including both how people utilize Web search engines in the search process to locate Web documents and how searchers are visiting and viewing the documents that the search engine locates.

There is a growing body of Web research concerning how users interact with Web search engines. There are also reports on the number of results pages viewed. When a Web search engine user submits a query, the search engine returns the results in “chunks” of usually about 10 results. We refer to these “chunks” as results pages, and the search engine presents results within these pages to the user sequentially from the topmost ranked results page to the maximum number of results pages retrieved by the search engine. However, there has been little large-scale research examining the pattern of interactions between Web search engine users and the actual Web documents presented by these results pages.

In this chapter, we summarize research on the results page viewing activities of users of Web search engines. We examine general searching characteristics including the number of results pages viewed. We then examine the number of Web documents that users view, analyzing the relationship between sessions, queries, and Web pages viewed. We also explore the temporal relationships of these interactions.

We begin with a review of the literature, followed by the methodology we utilized to analyze actual Web queries submitted by users to Web search engines. We use these queries to examine trends in searching. Specifically, we examine results pages accessed, and page viewing or click-through data (i.e., the Web page/s a user visits when following a hyperlink from a search engine results page), including the temporal aspects of this viewing. Click-through data shows great promise in the area of Web mining to isolate relevant content, identify searchers’ usage patterns, and evaluate Web search engine system performance (Joachims, 2002). We then discuss the implications of these results for Web search engine users, search engine designers, and the designers of Websites. We conclude with directions of future research in this area.

BACKGROUND

There has been limited research examining the results pages and little analysis of the Web page viewing patterns of Web search engine users. There is a growing body of literature in information science that examines how people search on the Web (Hölscher & Strube, 2000; Jansen & Pooch, 2001; Jansen, Spink & Saracevic, 2000; Spink, Jansen, Wolfram & Saracevic, 2002). This research provides insight into how people search for information on the Web, and provides a framework for considering the Web document viewing and search process. Jansen and Pooch (2001) present an extensive review of the Web searching literature, reporting that Web searchers exhibit different search techniques than do searchers on other information systems.

Scalable Authoritative OWL Reasoning for the Web
www.igi-global.com/chapter/scalable-authoritative-owl-reasoning-web/37736?camid=4v1a

Demand-Driven Algorithm for Sharing and Distribution of Photovoltaic Power in a Small Local Area Grid
www.igi-global.com/article/demand-driven-algorithm-for-sharing-and-distribution-of-photovoltaic-power-in-a-small-local-area-grid/113320?camid=4v1a