

Chapter 3

Horizontal Web Searching and Navigational Resource Identification

ABSTRACT

Generally speaking, horizontal search engines are meant to deal with general web queries. In the context of this chapter, the authors investigated the act of navigational resource identification in the light of horizontal web searching. State-of-the-art navigational resource identification is reluctant to the distinct characteristics of the navigational queries and specific users' treatments toward different searching tasks. Consequently, in this chapter, the authors discussed a new mechanism for navigational resource identification according to previous findings.

INTRODUCTION

In previous chapters, we have shown that there are different searching tasks with significant characteristics. We also have shown that these characteristics may be different for users based on their language preferences for navigational searching.

In this chapter, we introduced a novel approach for answering users navigational searching. In general, navigational searching targets only one particular Web page that is called *navigational resource* and the act of

DOI: 10.4018/978-1-7998-0961-6.ch003

assigning a user's query to its navigational resource is called *navigational resource identification*.

Google patent (Upstill, T. et al. (2012)) is an example of a framework which is specifically developed for such task. Corresponding patent is highly dependent to a query log which brings three issues: firstly, collecting a query log for a query means that a user issued multiple queries to reach to relevant page(s). This fact is stated that query reformulation decreases users' satisfaction (Agrawal, R. et al. (2015)). Secondly, collecting users' data brings privacy issue. Last but not least, using query log makes it impossible to answer newly issued queries.

Taking into account facts and motivations which we discussed here and in chapter 1 and 2, we are motivated to propose a framework which considers prior research findings regarding act of navigational resource identification. Earlier studies have shown that URL of Web pages receive more attention from users in a navigational searching compared to other types of Web searching. Therefore, fundamental characteristic of proposed method takes into account URL information of a Web page for the process of navigational resource identification.

BACKGROUND

In this section, we went through related works in respect to navigational resource identification.

Kang and Kim (2003) proposed to combine URL and content information for the purpose of navigational resource identification. This study assumed that navigational resources are merely home pages which are against prior Web usage studies (B. J. Jansen et al. (2008), B. J. Jansen and A. Spink (2005)).

Westerveld et al. (2001) also tried to distinguish between home pages and other pages by applying some probability models. An interesting finding of this study is that adding URL information to the result of mean reciprocal rank (MRR) improved it from 0.26 to 0.82 which showed the importance of URL information in navigational Web searching. In addition, a part of Google patent (Upstill, T. et al. (2012)) uses URL characteristics incorporate with click-through rate (CTR) data and other evidences in order to successfully apply navigational resource identification. Nevertheless, assumptions made

10 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/horizontal-web-searching-and-navigational-resource-identification/268294

Related Content

A Cloud-Assisted Proxy Re-Encryption Scheme for Efficient Data Sharing Across IoT Systems

Muthukumaran V. and Ezhilmaran D. (2020). *International Journal of Information Technology and Web Engineering* (pp. 18-36).

www.irma-international.org/article/a-cloud-assisted-proxy-re-encryption-scheme-for-efficient-data-sharing-across-iot-systems/264473

Sizing Web Applications

Emilia Mendes (2008). *Cost Estimation Techniques for Web Projects* (pp. 78-119).

www.irma-international.org/chapter/sizing-web-applications/7163

Research Trends on Business Process Management in Higher Education and Recommendations for Vietnam

Anh Hoang, Kha Van Phan and Vijender Kumar Solanki (2022). *International Journal of Information Technology and Web Engineering* (pp. 1-21).

www.irma-international.org/article/research-trends-on-business-process-management-in-higher-education-and-recommendations-for-vietnam/315609

Hierarchical Scheduling in Heterogeneous Grid Systems

Khaldoon Al-Zoubi (2007). *International Journal of Information Technology and Web Engineering* (pp. 1-16).

www.irma-international.org/article/hierarchical-scheduling-heterogeneous-grid-systems/2620

Heuristic Task Consolidation Techniques for Energy Efficient Cloud Computing

Dilip Kumar, Bibhudatta Sahoo and Tarni Mandal (2016). *Web-Based Services: Concepts, Methodologies, Tools, and Applications* (pp. 760-782).

www.irma-international.org/chapter/heuristic-task-consolidation-techniques-for-energy-efficient-cloud-computing/140828