

Chapter 4

Ambiguity and Clarification: A Visual and Textual Search

ABSTRACT

By the time web engines were developed, the number of queries prompted by users had grown exponentially. This fast growth shows the high demand of users from web search engines. This high demand made search engines responsible for the users' satisfaction during a search session. One way to improve a user's satisfaction is to visualize search engine result page (SERP). Recent studies for meeting this aim focused on a whole page thumbnail for assisting users to remember recently visited web pages. This chapter explores how a specific visual content of a page can allow users to distinguish between a useful and worthless page within results in SERP especially in an ambiguous search task.

INTRODUCTION

Search behavior of users over Web has been extended during time. For instance, when Google was developed by September 1998, it had to deal with ten thousands of searches each day (Battelle, J. (2005)), however, by year 2012 this amount has grown to 1.2 trillion¹. This dramatic change means that search engines should do their best to improve a user's experience during a search session especially for an ambiguous query. Ambiguity in a search query is the case that a user's search query belongs to more than one search category. Despite the fact that there are vast researches for dealing with this

DOI: 10.4018/978-1-7998-0961-6.ch004

issue, we believe that these methods usually need a user's data engagement tightly to personalize a web search and/or it neglects visual cues for assisting a user in disambiguation task. In addition, there is a possibility that the user intended meaning of the search query may not appear in the Search Engine Result Page (SERP) immediately.

To tackle this, there are numerous approaches. For instance, personalization and keyword extension are two approaches to increase users' satisfaction toward an ambiguous search task. Regardless of effectiveness of such techniques, this can be argued that these techniques still need a user's data engagement which brings privacy issues and neglected visual features for a disambiguation task. To overcome these drawbacks, we introduced a novel technique which gives a user more insight of each result's page content by extracting the most relevant visual content of the page. Moreover, it improves the textual part of the search snippet by engaging naïve Bayesian classifier to retrieve most relevant portion of a Web page regarding a user's ambiguous query.

There are two methods to engage visuality in search snippets to provide users with a better search experience; By 2010 Google has introduced a visual and textual page preview (Aula, A., et al. 2010) to improve a user's chance to make a successful decision to distinguish a previously visited Web apge. This task has been achieved by making a thumbnail of the *whole page* for each URL in search result page underneath each page description/search snippet. This has been concluded from this study that if SERP includes both textual and visual features of a page, it would give the most accurate relevancy prediction for users. However, we believe this can be distractive to face a whole-page thumbnail for every result in SERP.

The concept of Aggregated Search (Paris, C et al. 2010) is another effort to make SERP richer in terms of information representation. The aim of an aggregated search is to collect results from different verticals (such as Image Search Results, News, Videos, etc), while a user searches a query using Web search engines. This way of representation could be confusing and result in a failure if the query has multiple underlying meaning. This is caused by the fact that an aggregated search will not consider the different taxonomies when an ambiguous query is prompted and hence, would return and merge results from different verticals based on the default meaning of the ambiguous keyword. For instance, as we can see in Figure 1, if we search the ambiguous keyword, *Kingfsher* by Bing, the aggregated results only included the default

13 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/ambiguity-and-clarification/268295

Related Content

A Model-Based Approach for Diagnosing Fault in Web Service Processes

Yuhong Yan, Philippe Dague, Yannick Pencolé and Marie-Odile Cordier (2010). *Web Technologies: Concepts, Methodologies, Tools, and Applications* (pp. 1970-1993). www.irma-international.org/chapter/model-based-approach-diagnosing-fault/37725

A New Approach in Cloud Computing User Authentication

Alaa Hussein Al-Hamami and Rafal A. Al-Khashab (2016). *Web-Based Services: Concepts, Methodologies, Tools, and Applications* (pp. 2402-2418). www.irma-international.org/chapter/a-new-approach-in-cloud-computing-user-authentication/140906

Identifying Users Stereotypes for Dynamic Web Pages Customization

Sandro José Rigo (2009). *Handbook of Research on Web Log Analysis* (pp. 284-306). www.irma-international.org/chapter/identifying-users-stereotypes-dynamic-web/22006

An Efficient and Accurate Discovery of Frequent Patterns Using Improved WARM to Handle Large Web Log Data

Sahaaya Arul Mary and M. Malarvizhi (2014). *International Journal of Information Technology and Web Engineering* (pp. 47-64). www.irma-international.org/article/an-efficient-and-accurate-discovery-of-frequent-patterns-using-improved-warm-to-handle-large-web-log-data/115934

Computer Forensic Investigation in Cloud of Things

A. Surendar (2019). *Computational Intelligence in the Internet of Things* (pp. 256-270). www.irma-international.org/chapter/computer-forensic-investigation-in-cloud-of-things/224452