

Search Engine Ranking Efficiency Evaluation Tool

Wadee S. Alhalabi

Dept of Electrical and Computer
Engineering
University of Miami
7820 Camino Real Apt J320
Miami, FL 33143-6877,
w.alhalabi@umiami.edu

Miroslav Kubat

Department of Electrical and
Computer Engineering
University of Miami,
P.O. Box 248294
Coral Gables, FL 33124-0620
mkubat@miami.edu

Moiez Tapia

Department of Electrical and
Computer Engineering
University of Miami,
P.O. Box 248294
Coral Gables, FL 33124-0640
mtapia@miami.edu

Abstract

In view of the recent progress in the field of internet search engines, there is a growing need for mechanisms to evaluate the performance of these useful and popular tools. So far, the vast majority of researchers have relied on the information-retrieval metrics of “precision” and “recall” that quantify the occurrence of “hits” and “misses” in the returned list of documents. What they fail to do is to measure the quality of the ranking that the search engine has provided. This paper wants to rectify the situation. We discuss the issue in some detail, and then propose a new mechanism that we believe is better suited for our needs.

Keywords: Internet, search engines, ranking, performance evaluation

1. Introduction

The Internet revolution has given rise to the search engine, a tool whose task is to identify among the billions of existing websites those that are relevant to a user's query. Starting from mid 1990s, hundreds of companies offering this service have appeared on the market. By now, many of them have gone out of business while others have merged, and yet others have joined this thriving market only recently, seeking either to outperform their predecessors, or to offer functions that fit previously neglected niches.

The principle is simple. When a user submits his or her query (usually expressed as a Boolean function of carefully chosen keywords), the search engine analyzes its repository of stored web sites and returns the list of hyperlinks to those that contain information requested by the query. Importantly, this list is ordered so that the most relevant web sites come up first. Many mechanisms to assess the degree of a web site's relevance have been proposed, among them keyword frequency in the document, the average time spent on a given web page by other users, interlinking with other important web sites, and various combinations of all these three approaches.

The multitude of existing solutions leads to the logical question, “Which of them provides the best ordering of the relevant documents in the sense that the more relevant the links, the higher up the relevant documents are in the list?” Each user may have his or her subjective assessment; however, a professional needs performance criteria that are objective. To develop these criteria is the task for this paper. Whereas previous research has predominantly relied on the classical information-retrieval metrics of precision

and recall, their utility for our needs is limited. The problem is that precision and recall only establish to what extent the returned list contains the predominantly relevant links and the number of relevant links missing. What they ignore is whether links that are more relevant find themselves high up in the list.

Section 2 provides a survey of the related work. Then, Section 3 presents our own formula and illustrates its behavior by simple examples that show its advantages as compared to other methods. Finally, Section 4 offers a brief discussion of the proposed method.

2. Related Studies

Many research groups have studied alternative ranking algorithms, indicating the broad interest in finding methods to improve these rankings. As for the performance evaluation of these algorithms, the vast majority of the published approaches rely on the precision and recall known from information retrieval.

Thus Zhang and Dong [1] review many ranking algorithms, discuss their deficiencies, and propose their own algorithm that they show provides an improvement in the ranking quality as evaluated by precision and recall. Shafi and Rather [2] use the same criteria to evaluate the performance of five different search engines. Chu and Rosenthal [3] who use precision and response time instead of plain precision and recall reported similar work also. Li and Danzig [4] introduce a new ranking algorithm that they argue is much better in terms of space requirements and computational costs. Again, they evaluate the algorithm's