CCIT Report #779 November 2010

Context-Sensitive Query Auto-Completion*

Naama Kraus
Department of Computer Science
Technion, Haifa, Israel
nkraus@cs.technion.ac.il

Ziv Bar-Yossef
Department of Electrical Engineering
Technion, Haifa, Israel
and Google Haifa Engineering Center
Haifa, Israel
zivby@ee.technion.ac.il

November 22, 2010

Abstract

Query auto completion is known to provide poor predictions of the user's query when her input prefix is very short (e.g., one or two characters). In this paper we show that context, such as the user's recent queries, can be used to improve the prediction quality considerably even for such short prefixes. We propose a context-sensitive query auto completion algorithm, NearestCompletion, which outputs the completions of the user's input that are most similar to the context queries. To measure similarity, we represent queries and contexts as high-dimensional term-weighted vectors and resort to cosine similarity. The mapping from queries to vectors is done through a new query expansion technique that we introduce, which expands a query by traversing the query recommendation tree rooted at the query.

In order to evaluate our approach, we performed extensive experimentation over the public AOL query log. We demonstrate that when the recent user's queries are relevant to the current query she is typing, then after typing a single character, NearestCompletion's MRR is 48% higher relative to the MRR of the standard MostPopularCompletion algorithm. When the context is irrelevant, however, NearestCompletion's MRR is essentially zero. To mitigate this problem, we propose HybridCompletion, which is a hybrid of NearestCompletion with MostPopularCompletion. HybridCompletion is shown to dominate both NearestCompletion and MostPopularCompletion, achieving a total improvement of 31.5% in MRR relative to MostPopularCompletion.

1 Introduction

Query auto completion [3, 20, 12, 2] is one of the most visible features in Web Search today. It is offered by all major search engines and in almost all their search boxes. Query auto completion helps the user formulate her query, while she is typing it. Its main purpose is to predict the user's intended query and thereby save her keystrokes. With the advent of instant as-you-type search results (a la the recently released Google Instant[1]), the importance of correct query prediction is even more acute, because it determines the speed at which the user sees the suitable results for her intended search and the amount of irrelevant results that are displayed to her along the way.

The basic principle that underlies most query auto completion systems is the wisdom of the crowds. The search engine suggests to the user the completions that have been most popular among users in the past (we

*Despite the affiliation of the first author, this paper did not rely on privileged access to Google's resources, data, or technologies. The entire research was done at the Technion.

[1] http://www.google.com/instant/