Bid keyword suggestion in sponsored search based on competitiveness and relevance ☆

Ying Zhang [a,*], Weinan Zhang [b], Bin Gao [c], Xiaojie Yuan [a], Tie-Yan Liu [c]

[a] College of Software, Nankai University, Tianjin 300071, PR China
[b] Department of Computer Science, University College London, London WC1E 6BT, UK
[c] Microsoft Research, No. 5 Danling Street, Beijing 100080, PR China

ARTICLE INFO

ABSTRACT

In sponsored search, many advertisers have not achieved their expected performances while the search engine also has a large room to improve their revenue. Specifically, due to the improper keyword bidding, many advertisers cannot survive the competitive ad auctions to get their desired ad impressions; meanwhile, a significant portion of search queries have no ads displayed in their search result pages, even if many of them have commercial values. We propose recommending a group of relevant yet less-competitive keywords to an advertiser. Hence, the advertiser can get the chance to win some (originally empty) ad slots and accumulate a number of impressions. At the same time, the revenue of the search engine can also be boosted since many empty ad shots are filled. Mathematically, we model the problem as a mixed integer programming problem, which maximizes the advertiser revenue and the relevance of the recommended keywords, while minimizing the keyword competitiveness, subject to the bid and budget constraints. By solving the problem, we can offer an optimal group of keywords and their optimal bid prices to an advertiser. Simulation results have shown the proposed method is highly effective in increasing ad impressions, expected clicks, advertiser revenue, and search engine revenue.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Sponsored search is one of the major types of online advertising and is also the key source of revenue for search engine companies. In sponsored search, a set of ads are displayed along with organic search results when answering a query. Although displayed simultaneously and in similar forms, sponsored search results are actually generated by a quite different mechanism from that of organic search. While the organic search results are produced according to the relevance of each web page to the query, the sponsored search results are generated according to an auction process (Aggarwal, Goel, & Motwani, 2006; Varian, 2007).

Before the auction happens, each advertiser is asked to participate in a bid process, in which he/she creates a group of ad copies and bids on some keywords for the ad group with their match types. The match type might be exact match or broad match (a.k.a. advanced match). When a query is issued, the search engine will first retrieve some candidate ads whose

---

bid keywords match the query. Then the search engine will run an auction on these candidate ads by considering both the ad quality and the bid prices of them (Feng, Bhargava, & Pennock, 2003). Those ads with the highest quality scores and bid prices will win the auction and be displayed on the search result page. If such an ad is clicked by a user, its advertiser will be charged by the search engine. Mainstream search engines adopt the generalized second price (GSP) (Edelman, Ostrovsky, & Schwarz, 2005) auction mechanism, which means that the advertiser's cost of a click depends on the bid price and the relevance score of the next ad in the ranking list of the auction.

As can be seen above, an advertiser should carefully consider which keywords to bid and what price to set for each of these keywords, in order to win the auction. However, data statistics show that not many advertisers are good at dealing with this (see Section 3.1.2). For example, too many advertisers bid on a small number of popular keywords, and thus as a result those advertisers with low bid prices will fail due to the hot competition in the auction process and do not have as many ad impressions as expected. This is bad for these advertisers since they have missed their campaign goals. This is also bad for search engines since the real contributions of these advertisers to the revenue of search engine will be much lower than their budgets. On the other hand, however, data statistics also show that a significant portion of search queries have no ads (or no enough ads) displayed in their search result pages, even if many of them have commercial values (see Section 3.1.1). In other words, many keywords that may potentially lead to ad clicks have been ignored by the advertisers. Again, this is bad for search engines since many potentially valuable ad slots have been wasted. This is also bad for advertisers since they have missed many advertising opportunities.

It is clear that if we can effectively solve the aforementioned problem, we can improve the satisfaction of advertisers and increase the revenue of search engines simultaneously. A straightforward way to solve this problem is to suggest appropriate keywords to the advertisers to bid. Actually, the ad platforms in many search engines already provide this type of services, such as Keyword Group Detection of Microsoft AdCenter,[1] Overture (Yahoo!) Keyword Selector Tool,[2] and the Keyword Tool of Google AdWords.[3] In the research community, there are also many papers on bid keyword optimization (Abhishek & Hosanagar, 2007; Bartz, Murthi, & Sebastian, 2006; Kitts & Leblanc, 2004; Chen, Xue, & Yu, 2008; Fuxman, Tsaparas, Achan, & Agrawal, 2008; Kiritchenko & Jiline, 2008). However, most of these existing works are based on the semantic similarity between keywords and/or the co-occurrence of bid keywords across advertisers. As a consequence, it is easy to understand that these methodologies will not effectively solve the aforementioned problem, and sometimes make the situation even worse: they will tend to suggest popular keywords to the advertisers and the competition on these popular keywords will become tougher and tougher.

In this paper, we propose a novel keyword suggestion technology, which can alleviate the aforementioned problem. In particular, instead of suggesting popular keywords, we recommend a group of relevant yet less-competitive keywords to the advertisers, by optimizing the expected advertiser revenue. Here, less-competitive keywords correspond to the keywords that have not been intensively bid by advertisers. This idea is inspired by the long tail theory (Anderson, 2008). That is, the aggregated popularity of a large numbers of less-competitive (sometimes tail) items can make a large fraction of the total popularity. As pointed out by many previous works, for search engines, the query volume just follows a long tail distribution. To validate this phenomenon in the data analysis experiment, we used a search log dataset that records the submitted queries from a commercial search engine,[4] which was collected in half a month (from 01-April-2011 to 15-April-2011). It contains about 3.3 billion user queries in volume, in which there are 623 million unique queries. The distribution of the query volume is shown in a log–log scale in Fig. 1. From the figure, we can see that the query volume follows a long tail distribution obviously. This suggests that the aggregated volume of less-competitive queries may take a large part of the total query volume. Therefore, if an advertiser bids on a package of less-competitive keywords, he/she may win a large number of ad auctions and accumulate high ad impressions (and potentially high return on investment, i.e., ROI). Meanwhile, the search engine can also get extra revenue since many empty ad slots related to the suggested keywords can be utilized.

Mathematically, we formulate the keyword recommendation problem as a mixed integer optimization problem. Given a target ad group, we first collect a set of candidate keywords whose relevance score to the ad group can exceed the floor relevance score in sponsored search. Then we perform selection among these candidate keywords and try to give an optimal bid price for each selected keyword, by maximizing the revenue of ad group with the budget control constraint. In this process, more relevant and less competitive candidate keywords will have higher possibility to be selected. This constrained optimization problem can be solved by binary integer programming and sequential quadratic programming (SQP) in an alternate manner. Simulation results on the sponsored search log obtained from a commercial search engine,[5] show that our proposed technology can effectively increase the ad impressions of advertisers with a low average cost per click, and it can effectively help advertisers obtain increased revenue. At the same time, our method can significantly reduce the empty ad slots and enlarge the revenue of search engine.

To sum up, the contributions of our work are listed as below. (i) We perform a comprehensive study on keyword bidding in sponsored search, and point out a problem with the current sponsored search services that may make both advertisers and search engines unsatisfied. (ii) We propose a bid keyword suggestion method based on competitiveness and relevance, and

---

[1] http://adlab.msn.com/Keyword-Group-Detection.
[2] http://inventory.overture.com/d/searchinventory/suggestion.
[3] https://adwords.google.com/select/KeywordToolExternal.
[4] http://www.bing.com.
[5] http://www.bing.com.