Information Sciences 276 (2014) 242–254

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Advertiser-centric approach to understand user click behavior in sponsored search ☆

Sungchul Kim [a], Tao Qin [b], Tie-Yan Liu [b], Hwanjo Yu [a,*]

[a] Data Mining Laboratory, Department of Computer Science and Engineering, POSTECH, Pohang, South Korea
[b] Internet Economics & Computational Advertising group, Microsoft Research Asia, Beijing, China

ARTICLE INFO

Article history:
Received 9 May 2012
Received in revised form 9 January 2014
Accepted 9 February 2014
Available online 20 February 2014

Keywords:
Sponsored search
User study
User model

ABSTRACT

Sponsored search is the major business model of commercial search engines. The number of clicks on ads is a key indicator of success for both advertisers and search engines, and therefore increasing ad clicks is a goal of both of them. Many existing works stand on the view of search engines concerning how to help search engines to earn more revenue by accurately predicting ad clicks. Unlike these works, this paper aims at understanding user clicks on ads from "the view of advertisers", in order to help advertisers to improve their ad quality and therefore advertising effectiveness. To do this, a factor graph model is proposed, which considers two advertiser-controllable factors to understand user click behaviors: (1) the relevance between a query and an ad, which has been well studied in the literature, and (2) the "attractiveness" of the ad, which is a newly-proposed concept. The proposed model can be used to predict user clicks and also to mine a set of attractive words that could be leveraged to improve the quality of the ads. We have verified the effectiveness of the proposed approach using real world datasets, through quantitative evaluations and informative case studies.

1. Introduction

Sponsored search is the prevalent revenue source of major search engines, and search engines and advertisers would like to see more clicks on ads. To obtain more clicks, the first step is to know why users click on an ad. Existing works investigates user click behaviors on ads from the view of search engines [13,35]. However, the analyses on user click behaviors do not directly help advertisers to improve their ads, since many factors used in click prediction cannot be accessible and controlled by advertisers (e.g., position bias, proper bidding price, or externalities). Unlike the existing works, we look at the problem from the perspective of advertisers. In particular, we want to find the factors that explain user click behaviors, which can be exploited by the advertisers to improve their ads and attract more clicks.

It has been widely accepted that the relevance between a query and an ad is an important factor to explain user click behaviors. However there are some cases that relevance might not be able to explain what we have observed. For example, it is easy to find that a less relevant ad attracts more clicks [4,28] than a more relevant ad. Fig. 1 shows two ads for query "arcade games". The first ad is more relevant to the query than the second ad, since it contains query terms in both the title and the description. However, the logs show that the second ad has much more clicks than the first one although the second

☆ A previous version of the manuscript is published in Proc. of the 2011 ACM Conf. Information and Knowledge Management (CIKM) as a full paper.

* Corresponding author. Tel.: +82 1041187006.
E-mail addresses: subright@postech.ac.kr (S. Kim), taoqin@microsoft.com (T. Qin), tyliu@microsoft.com (T.-Y. Liu), hwanjoyu@postech.ac.kr (H. Yu).

http://dx.doi.org/10.1016/j.ins.2014.02.058
0020-0255/

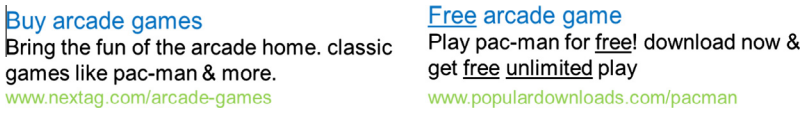

Fig. 1. The comparison of two ads for query "arcade games".

ad is relatively less relevant to the query. Through a user study, we find that people tend to click on the second ad because they think some words in it are attractive, such as "free" and "unlimited". Some further examples of the attractive words provided by the participants of our user study include "sale", "deal" and "save". This indicates the existence of a new factor beyond relevance which affects user click behaviors. We call this factor "**attractiveness**", which refers to how likely an ad (or a word in the ad) will attract users' attention and further get their clicks.

To understand how relevance and ad attractiveness affect user clicks, we model user clicks using a factor graph [19] which consists of several variables and their dependencies represented as factor nodes. Our proposed graph describes four types of dependencies: (1) the attractiveness of a word in an ad depends on a set of attractiveness features extracted for the word; (2) the attractiveness of an ad depends on the attractiveness of the words in it; (3) the relevance between an ad and a query depends on a set of relevance features; and (4) user clicks depend on both relevance and attractiveness of the ad. The model can be used to infer the relevance score and the attractiveness score of a given word or ad. Then clicks are predicted based on the scores. In particular, the inferred attractiveness scores of words can also be used as references for advertisers to effectively improve their ads and attract more clicks.

We have conducted experiments on real-world data, i.e., sponsored search logs obtained from a commercial search engine. According to our experiment results, while the relevance between query and ads plays an important role in modeling the user click behaviors, the attractiveness features further enhance the modeling quality and thus result in higher ads selection accuracy. This implies that advertisers should also pay attention to the attractiveness of their ads as well as improving relevance. Our interpretation on this finding is as follows. In web search, users' goal is to find some useful information. As a result, they hope to see pages relevant to their queries (e.g., pages containing many query keywords) and will click on such pages if their information needs are satisfied. In contrast, ads are just by-products of search and they are not the motivation of the users search. In this situation, whether the ads can attract users' attention become also important, given that all the displayed ads are already relevant to the query.

There are several main contributions of this work.

- We propose the concept of attractiveness to explain user click behaviors in sponsored search.
- We model the effect of relevance and ad attractiveness on user clicks using a factor graph. The model can be used to estimate the attractiveness score of a given word or ad, which can be used by advertisers to create better ads.
- The proposed model, by additionally adopting the attractiveness features, enhances the ad selection accuracy.

This paper is organized as follows. Section 2 provides briefly review the background knowledge of sponsored search and provide an overview on related work. Section 3 introduces the two factors of user click behaviors from advertisers' view. Section 4 describes the factor graph model. Section 5 presents the experimental results to verify the effectiveness of the proposed model and several examples of attractive words and ads. Finally, Section 6 concludes our study.

2. Related work

This section briefly introduces sponsored search and discusses related work in different kind of viewpoint.

2.1. Sponsored search advertising

Search engines provide sponsored search results along with organic web search results, in response to a given query issued by a user. The displayed ads in the sponsored search results are usually ranked according to their estimated click-through rates (CTR) and their bids. If a user clicks an ad, the advertiser will pay the search engine a certain amount of money for the click, and the amount is determined by the auction mechanism (such as generalized second price auctions [9]) used in the search engine. This revenue model is called the pay-per-click model. Fain and Pedersen [10] provide a brief review of the field of sponsored search, and Broder [2] give an excellent introduction to online advertising (including sponsored search).

2.2. Search engine centric approach

The pay-per-click model determines that clicks are very important to both search engines and advertisers. As aforementioned, for search engines, clicks mean revenue; for advertisers, clicks mean potential customers and purchases. There have been many attempts on increasing clicks in sponsored search, and most works are designed only for search engines.

Some works [1,5,32] focus on query expansion, which can lead to more ad impressions and thus (very likely) more clicks since it can provides ads more related to user intent. However, they are designed or have aims only for search engines.