Neurocomputing 120 (2013) 524–535

Contents lists available at ScienceDirect

Neurocomputing

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

Position-wise contextual advertising: Placing relevant ads at appropriate positions of a web page

Zongda Wu [a,b], Guandong Xu [c], Chenglang Lu [a,d,*], Enhong Chen [b], Yanchun Zhang [e], Hong Zhang [f]

[a] Wenzhou University, Wenzhou 325035, Zhejiang, China
[b] University of Science and Technology of China, Hefei 230026, Anhui, China
[c] University of Technology, Sydney, Australia
[d] Northwestern Polytechnical University, Xi'an 710072, China
[e] Victoria University, Melbourne, Australia
[f] Institute of Science and Technology Information of Zhejiang Province, Hangzhou 310006, China

ARTICLE INFO

Article history:
Received 4 July 2012
Received in revised form
7 April 2013
Accepted 7 April 2013
Communicated by Tao Mei
Available online 15 June 2013

Keywords:
Wikipedia knowledge
Similarity
Contextual advertising

ABSTRACT

Web advertising, a form of online advertising, which uses the Internet as a medium to post product or service information and attract customers, has become one of the most important marketing channels. As one prevalent type of web advertising, contextual advertising refers to the placement of the most relevant ads at appropriate positions of a web page, so as to provide a better user experience and increase the user's ad-click rate. However, most existing contextual advertising techniques only take into account how to select as relevant ads for a given page as possible, without considering the positional effect of the ad placement on the page, resulting in an unsatisfactory performance in ad local context relevance. In this paper, we address the novel problem of position-wise contextual advertising, i.e., how to select and place relevant ads properly for a target web page. In our proposed approach, the relevant ads are selected based on not only global context relevance but also local context relevance, so that the embedded ads yield contextual relevance to both the whole target page and the insertion positions where the ads are placed. In addition, to improve the accuracy of global and local context relevance measure, the rich wikipedia knowledge is used to enhance the semantic feature representation of pages and ad candidates. Last, we evaluate our approach using a set of ads and pages downloaded from the Internet, and demonstrate the effectiveness of our approach.

1. Introduction

Web advertising is becoming an increasingly important and popular advertising market today. PwC[1] predicts that web advertising will become the second largest advertising medium in America after TV within the next four years, and spending in this area will increase from 24 billion dollars in 2009 to 34 billion dollars in 2014. A large part of web advertising consists of textual ads, which are short text messages usually marked as "sponsored links" or similar. Now, there are two main types of textual web advertising, i.e., sponsored search and contextual advertising [1,2]:

1. *Sponsored search* (also called keyword-targeted advertising), which selects ads based on keywords contained in search queries given by users, is characterized by placing paid textual ads links on the result pages returned by a web search engine (e.g., Google).
2. *Contextual advertising* (also called content-targeted advertising), which judges the context relevance of ads to the page that the user is browsing, refers to the selection of relevant commercial ads for the target page.

One of the important advantages of contextual advertising over the sponsored search is that it can support various types of web sites, which range from individual bloggers and small niche communities to large publishers (e.g., major newspapers). Now, almost all for-profit non-transactional sites, i.e., the sites that do not sell anything directly, rely heavily on the revenues from contextual advertising. Without contextual ads, the Web will lose the most of its market value.

The first major contextual advertising platform was provided by Google in 2003 [3]. Now, almost all popular search engines such as Baidu, Yahoo! and Microsoft Bing provide similar platforms for ad

* Corresponding author at: Wenzhou University, Wenzhou 325035, China.
E-mail addresses: zongda1983@163.com (Z. Wu), guandong.xu@vu.edu.au (G. Xu), chenglang.lu@qq.com (C. Lu), cheneh@ustc.edu.cn (E. Chen).

[1] PricewaterhouseCoopers, a global professional services company —http://www.pwc.com.

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.04.018

publishers and web site owners. As shown in Fig. 1, a contextual advertising platform generally consists of the following four parts [1,2,4–6]:

1. The *advertiser* provides the supply of textual ads, which is usually a company that wants to use the ad platform to promote their products or services, and needs to pay for its textual ads.
2. The *publisher* is the owner of a web site on which textual ads are placed, who typically aims to provide a good user experience and increase the number of ad-clicks, so as to maximize the market revenue.
3. The *ad platform* is a software system of matching ads to pages, which allows the publisher to select appropriate ads based on the context similarity between pages and ads for advertisers that create the ad revenue for the publisher.
4. End *users* consist of customer groups who have potential interest in the ads while browsing the content of a web page, supplied by the publisher.

The most dominant online advertising pricing model is pay-per-click (PPC), where the advertisers pay a certain amount to the publisher and the ad platform for each user's click on the ads. In addition, there are also other types of pricing models for textual ads, including: (1) pay-per-impression (PPI), where the advertisers pay for the number of ads displayed on a web page and (2) pay-per-action (PPA), where the payment made by the advertiser is calculated by each sale originating from the ads. Since most existing contextual advertising approaches are based on the PPC model [1,2,4–6], in this paper, we also use this model for simplicity.

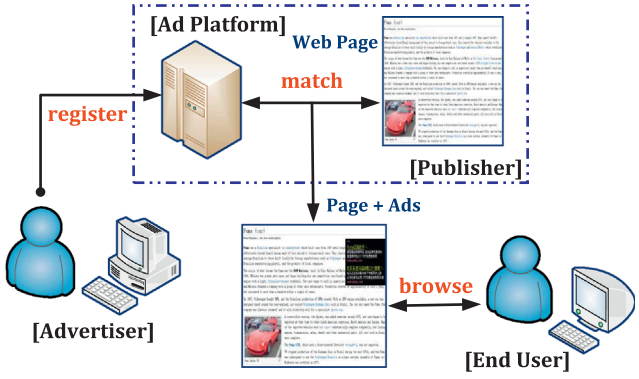

Fig. 1. A contextual advertising platform.

1.1. Problem statement

Under the PPC pricing model, in [2,6,7], it has been pointed out that given a target page $\boldsymbol{p}$, the revenue of the publisher and the ad platform can be estimated as: $\sum_{i=1,2,\ldots,k} P(click|\boldsymbol{p}, a_i) \cdot price(a_i)$, where k is the number of ads displayed on the page $\boldsymbol{p}$ and $price(a_i)$ is the click-price of the current ad a_i. After simplifying the model, the revenue can be maximized by searching for: $\arg\max_i P(click|\boldsymbol{p}, a_i)$. Thus, to maximize the PPC pricing model, we would need to select proper ads to maximize the probability of user's ad-clicks.

In general, the probability of user's ad-clicks is positively related to user's interest in the ad content. Furthermore, we assume that a web user should be interested in the page content that he is currently browsing, i.e., the page content that a user currently browsing could reflect the user's interest to a certain extent. Thus, we believe that the users' ad-click rate can be boosted by increasing the context relevance of an ad to the page position where the ad is placed. A number of studies also have shown similar conclusions, e.g., it was pointed out in [2,6,7] that the ads should be relevant to their surrounding page content; and it was further pointed out in [8–10] that the in-image ads should be not only relevant to the entire page but also locally relevant to their hosting image.

Based on the above observations, to maximize the PPC pricing model (i.e., to increase the users' ad-click rate as much as possible, and thus maximize the revenue), we would need to solve two key problems. The *first problem* is how to select as relevant ads for a given page as possible. As shown in Fig. 2(a) and (b), for example, given a page about "travel in China", embedding the page with an ad about "hotel information" or "tour service" would attract more user attention than randomly chosen ads.

A web page (especially for a long page), however, may cover multiple topics (or multiple subtopics under the same topic), such that an ad, which is relevant to the entire page, may be not locally relevant to the page position where the ad is placed. The *second problem* is the placement of relevant ads at appropriate positions of a web page, making each ad close to the page segments that are really relevant to the ad. For example, as shown in Fig. 2(b) and (c), for the page position next to the segment "Beijing", embedding the position with the ad "Great Wall tour" would increase the interest of users clicking on the ad more than the other ad "Xian hotel", although the two ads are both relevant to the entire page.

Fig. 2. Three examples of placing ads on a web page, where the highlighted areas indicate associated ads.