Context Transfer in Search Advertising

Hila Becker‡[1], Andrei Broder†, Evgeniy Gabrilovich†, Vanja Josifovski†, Bo Pang†

‡ 1214 Amsterdam Avenue, New York, NY 10027, USA
† Yahoo! Research, 701 First Avenue, Sunnyvale, CA 94089, USA
hila@cs.columbia.edu | {broder | gabr | vanjaj | bopang}@yahoo-inc.com

ABSTRACT

We define and study the process of *context transfer* in search advertising, which is the transition of a user from the context of Web search to the context of the *landing page* that follows an ad-click. We conclude that in the vast majority of cases, the user is shown one of three types of pages, which can be accurately distinguished using automatic text classification.

Categories and Subject Descriptors

H.3.m [**Information Storage and Retrieval**]: Miscellaneous

General Terms

Experimentation, Measurement

Keywords

online advertising, landing page taxonomy

1. INTRODUCTION

Unbeknownst to most users, when a query is submitted to a search engine two distinct searches are performed: the *organic* or *algorithmic* search that returns relevant Web pages and related data (maps, images, etc.), and the *sponsored search* that returns paid advertisements. The standard approach to textual Web advertising is based on modeling the user's needs and interests to find suitable ads. In particular, in Web search, numerous studies have focused on classifying the query intent [1, 4, 3, 9] and on retrieving most relevant ads [2, 7, 8]. However, surprisingly little research has been devoted to what actually happens *after* an ad is clicked.

To this end, we define and study the process of *context transfer*, that is, the user's transition from her previous activity (to wit, Web search) to the different contexts found on the landing page after clicking on an ad. Arguably, a careful choice of the type of context transfer is among the most important factors explaining subsequent *conversion*, that is the occurrence of an activity that was the aim of the advertisement (e.g., goods purchase, registration, further browsing).

After reviewing a comprehensive sample of several hundred ads and corresponding landing pages, we propose a taxonomy of ad landing pages, where three major classes combined account for over 88% of the ads in our sample dataset. We then use standard machine learning techniques to build a classifier capable of automatically mapping landing pages onto the classes of this taxonomy.

[1]The research described herein was conducted while the first author was a summer intern at Yahoo! Research.

SIGIR'09, July 19–23, 2009, Boston, Massachusetts, USA.
ACM 978-1-60558-483-6/09/07.

2. TAXONOMY OF LANDING PAGES

In this section, we discuss our proposed landing page taxonomy. We compiled a sample set of sponsored search landing pages by issuing 200 unique queries to a Web search engine, and randomly selecting the landing page of an advertisement from the sponsored search results. triggered by issuing 200 unique queries to a Web search engine. The queries were randomly sampled out of the 800 labeled queries used for the 2005 KDD Cup [5]. We computed the frequency of each query from Web search query logs. Finally, we used stratified sampling, dividing the queries into deciles according to frequency, and sampling 20 queries uniformly from each decile. Thus, this dataset was constructed to represent ads that are *shown* for both popular and rare queries.

We inspected each landing page in isolation, noting its structure, appearance and functionality. We observed several distinct, non-overlapping classes that sponsored search landing pages fall into. Each class represents a different context transfer technique that transitions the user from the search engine result page to the advertiser's landing page. It is interesting to note how much or how little context the advertiser preserves by using each class of landing pages.

1. Homepage This is the top-level page of the advertiser's Website. Many advertisers choose to display their home page as a landing page for their ads, often regardless of the query that triggered the ad. This approach is commonly used by either smaller, less experienced advertisers or well known brand-name advertisers that display their homepage when bidding on brand keywords. Unless the user searched for the advertiser's brand name, using the homepage as a landing page does not make for a strong context transfer. For instance, consider a search for the word "Toyota." If Toyota is the advertiser, directing the searcher to Toyota's homepage will likely satisfy the user's information need. On the other hand, any other advertiser that does not have a Website dedicated to Toyota cars, e.g., a Website that provides price quotes for all car makes, would lose some of the context by showing a generic homepage, which does not immediately satisfy the search query (even though the relevant content may be found on the advertiser's Web site by following hyperlinks).

2. Search transfer Landing pages of this type are dynam-

ically generated search results on the advertiser's site. Here, the advertiser uses the original Web search query to perform a new search within its own site, and displays the results as the ad's landing page. For example, given a query "California Zinfandel," an online wine store would return a landing page that dynamically displays their own inventory search results for this query. In such landing pages, context transfer is strong only if the query used to generate the search results corresponds to products, services, or information that the Website offers. However, many advertisers that use this technique do not design their campaigns carefully enough to ensure that all phrases they bid on yield meaningful search results, in which case the context is completely lost.

3. Category browse A Category browse landing page leads the user to a sub-section of the Website that is generally related to the query. This page is not at the top level of the Website (homepage) but could be navigated to from other pages on the site. In our example of an online wine store advertising for the query "California Zinfandel," a Category browse landing page might describe the Zinfandel section of the Web site. This is a technique that advertisers can use if the bid phrase refers to either a general class of products or services, or a specific one. If the user is looking for a general class of products, choosing a Category browse landing page would bring them one step closer to the product they are searching for. If the user is looking for a specific product, while the advertiser only carries different but related products, showing a category page allows the advertiser to present such related offerings.[1]

4. Other These are standalone pages that appear to be disconnected from the rest of the Web site. These pages generally do not have many outgoing links and there is no way to reach them from the home page. Examples of this class include standalone form pages, where the sole purpose of the page is to gather information from the user, and promotion pages, which supply promotional information about a product or service. These pages are similar to print ads in a newspaper, and often include phrases such as "try it now," "limited time," and "special offer."

The distribution of landing page types.

We labeled each landing page in our sample dataset according to the landing-page classes. The following table presents the distribution of labels. Note that the first three classes combined account for over 88% of the ads in our data.

Homepage	Search transfer	Category browse	Other
25%	26%	37.5%	11.5%

3. LANDING PAGE CLASSIFIER

To facilitate future research efforts that study context transfer in the proposed framework, we need to obtain a larger set of landing pages and label them according to the taxonomy. For this, we trained a sufficiently accurate classifier using standard machine learning techniques.

To train the landing-page classifier, we used our sample set of landing pages, labeled with the taxonomy classes defined in Section 2. For each landing page in this set, our classifier assigns one of four labels: Homepage, Category browse, Search transfer, or Other. The features we used include the bag-of-words representation of the visible landing-page text with *tf.idf* weights, as well as a number of frequently observed HTML patterns from the landing-page HTML source. We trained a Support Vector Machine model using Weka's SMO implementation [6] on the reduced feature space induced from a supervised attribute selection technique, aiming to optimize the accuracy of the most frequent classes that accounted for more than 88% of the data. With 10-fold cross validation on the training data, our classifier accurately predicted the class label for 83% of the examples.

[1]A small number of pages in our dataset described a single specific product. For convenience, we include them in the Category browse class.

Class	Precision	Recall	F-Measure
Homepage	0.917	0.786	0.846
Search transfer	0.862	0.926	0.893
Category browse	0.645	0.87	0.741
Other	0.5	0.25	0.333

Table 1: Performances of landing page type classifier

We constructed a separate test set of 100 landing pages, by sampling from activity logs collected from a browser toolbar plug-in. These 100 pages were then manually labeled. We ran the classifier over this test data and it correctly predicted the class label for 80% of the examples. Table 1 presents a breakdown of the performance by class.

4. CONCLUSIONS

In this paper we presented a study of context transfer in sponsored search advertising. By analyzing several hundred examples, we found that the majority of ad landing pages fall into three distinct classes: *Homepage*, *Category browse*, and *Search transfer*. We then built a machine learning classifier, capable of automatically mapping landing pages onto these classes. In future work, we plan to use this classifier to study the effectiveness (e.g., conversion rates) of different context transfer techniques.

5. REFERENCES

[1] A. Broder. A taxonomy of web search. *SIGIR Forum*, 36, 2002.

[2] A. Broder, P. Ciccolo, M. Fontoura, E. Gabrilovich, V. Josifovski, and L. Riedel. Search advertising using Web relevance feedback. In *CIKM'08*, 2008.

[3] D. Downey, S. Dumais, D. Liebling, and E. Horvitz. Understanding the relationship between searchers' queries and information goals. In *CIKM*, 2008.

[4] U. Lee, Z. Liu, and J. Cho. Automatic identification of user goals in web search. In *WWW*, 2005.

[5] Y. Li, Z. Zheng, and H. Dai. KDD CUP-2005 report: Facing a great challenge. In *SIGKDD Explorations*. 2005.

[6] J. C. Platt. Fast training of support vector machines using sequential minimal optimization. In *Advances in kernel methods: support vector learning*. MIT Press, 1999.

[7] F. Radlinski, A. Broder, P. Ciccolo, E. Gabrilovich, V. Josifovski, and L. Riedel. Optimizing relevance and revenue in ad search: A query substitution approach. In *SIGIR'08*, 2008.

[8] B. Ribeiro-Neto, M. Cristo, P. B. Golgher, and E. S. de Moura. Impedance coupling in content-targeted advertising. In *SIGIR'05*, 2005.

[9] D. Rose and D. Levinson. Understanding user goals in web search. In *WWW*, 2004.