Ad Allocation for Browse Sessions

Anand Bhalgat[1] and Sreenivas Gollapudi[2]

[1] Dept of Comp and Info Science, University of Pennsylvania
[2] Microsoft Research Search Labs

Abstract. A user's session of information need often goes well beyond his search query and first click on the search result page and therefore is characterized by both search and browse activities on the web. In such settings, the effectiveness of an ad (measured as CtoC ratio, as well as #(conversions) per unit payment) could change based on what pages the user visits and the ads he encounters earlier in the session. *We assume that an advertiser's welfare is solely derived from conversions.*
Our first contribution is to show that the effectiveness of an ad depends upon the past events in the session, namely past exposure to self as well as to competitors. To this end, we analyze logs of user activity over a period of one month from Microsoft AdCenter Delivery Engine. We then propose a new bidding language that allows the advertiser to specify his valuation of a user's click as a function of these externalities, and study the improvement in prediction of conversion events with the new bidding language. We also study theoretical aspects of the allocation problem under new bidding language and conduct an extensive empirical analysis to measure effectiveness of our proposed allocation schemes.

1 Introduction

The increasing amount of time a user spends online conducting e-commerce transactions has led to a widespread use of online advertising by merchants to attract the potential customer to their sites and/or products. Often, this shopping experience of a user extends beyond his query to a search engine and includes visiting multiple web sites learning more about the product. A *browse session* is a contiguous sequence of webpages visited by a user; and two consecutive browse sessions are separated by a period of user's inactivity.

There has been work on understanding externalities in context of interplay between advertisements on the same page [2, 10, 7, 9, 8], however they neglect an important aspect that the user is not an independent entity on each page, and events in a browse session affect effectiveness of ads shown later in the session. Thus an ad allocation scheme needs to consider the session as a whole, rather than running independent auctions on each page. We initiate the study of understanding externalities and ad allocation for a browse session.

Now we describe the problem in detail: the most prevalent model of payment in search and contextual advertisements is *pay-per-click*, where advertisers bid for an ad position on a the page and they pay their bid value on a user's click. The advertiser's real welfare is derived from the sale of the good or service (i.e.

a conversion), and the additional traffic (or the awareness about the product) generated by a click can contribute to its increase. *The advertiser would want to bid for a <user, page> based on his perceived probability that the given user's click on that page would lead to a conversion.* In other words, the advertiser's welfare and payment are in different "currencies", and if events in the browse session affect his *CtoC ratio* or his *welfare (measured in #(conversions)) per unit payment*, then he would want to change his bid accordingly.

Contributions of this study: We analyzed the entire set of logs of user activity over a period of one month obtained from Microsoft AdCenter Delivery Engine to study the effect of the following two events in a user's browse session on an advertiser's CtoC ratio as well as his welfare per unit payment, namely (a) how many times the ad has been repeated already in the session, and (b) how many competing ads have been shown earlier in the session. We observed that these events affect the CtoC ratio (and the welfare per unit payment) negatively by up to 50%. *While our findings about the externalities from competing ads agree with previous studies in other contexts such as TV advertising which show that competitive advertising has a negative effect on the focal brand [6], our observations for the repeated exposure of ads are contrary to perceptions in other media (such as TV) in which it is considered beneficial to the advertiser [12].*

We model the prior on a user's browse session by a browse graph, and propose a natural language that allows advertisers to express their values of a click as a function of two main externality events. We perform an exhaustive set of experiments to show that *the model can be used to predict the conversion events in the session with better accuracy.* We study theoretical aspects of the allocation problem under new bidding language, and perform an empirical analysis of some natural heuristics on data. Our bidding language is simple, and can be considered as each advertiser specifying his *discount factors* for each externality event. E.g. an advertiser can ask to reduce his bid by a factor of $\texttt{disc\_self}(j)+\texttt{disc\_comp}(k)$ if he is already shown in the session j times and k competing ads have been in the past. Further, *our techniques can also be used internally by the ad allocation engines without exposing the details to advertiser, where the discount factors are computed by the engine, and it scales advertisers' bids with the discount factors.*

Related Work: There has been work on understanding externalities in online advertising [2, 10, 7, 9, 8]. One model of externality that has been studied is the effect of cascade models of user's browsing on the click through rates of ads [2, 10]. Gomes *et al* [9] consider the role of information and position externalities in a similar cascade model. Ghosh *et al* [8] consider a special case where the each advertiser expresses a two bids for a user, for exclusive and non-exclusive display. Modeling externalities into an auction has also been studied using richer bidding languages [4, 11, 5].

Organization: We illustrate the experiments performed to establish the sources of externality in Section 2. We give a brief overview of our bidding language, and empirical analysis of some heuristics for ad allocation in Section 3. The details of the bidding language, its relevance to the conversion events in the data (in terms of accuracy of prediction) and theoretical aspects of the allocation problem with the new bidding language are deferred to the full version of the paper [3].

2 Existence of Externalities

We begin by performing a set of experiments to establish the existence of externalities in a user's browse session.

Data Sets: We used the entire set of user activity logs over a period of one month (June 2011) obtained from Microsoft AdCenter Delivery Engine. These logs consisted information about the user query, the set of competing ads, their bids, ads shown, and the click as well as the conversion information. The pages in consideration were essentially the sponsored search properties as well those sites enrolled in the Microsoft publisher network. We associated all requests coming from same IP (anonymized) to a single user. From this data, we extracted the session information of a user. We defined a *session* as the set of contiguous requests by a user such that two consecutive requests are no more than 10 minutes apart. We labeled an advertiser as a *valid advertiser* if his ad impressions got at least 1000 clicks in the month. *Our experiments are restricted to the set of valid advertisers for whom the conversion data is available.* We further note that, the data is for search and contextual ads, where the payment model in use is *pay-per-click*. We define *two advertisers as competitors* if there are at least 10000 sessions in the month in which both are assigned an impression.

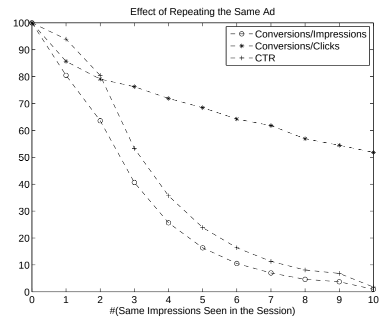
Advertiser's Welfare Model: We assume that an advertiser's welfare is solely derived from a conversion even as the payment model is pay-per-click; and for any fixed advertiser, the welfare derived from a conversion remains same. Given advertiser $\mathbb{A}_i$ and a page p, let $b_{i,p}$ be his bid for a click on that page, and let $b_{i,avg}$ be his average bid over all pages across all sessions. We define his *relative bid* for page p as $\frac{b_{i,p}}{b_{i,avg}}$; the *relative payment* for an advertiser for a click on a page is the value of his relative bid for that page. As an advertiser's welfare is same for each conversion, he would want to pay the same (or similar) amount per unit conversion, and we would expect his relative bid on a given page to be directly proportional to the "anticipated" CtoC ratio from a click on the page.

Existence of Externality: We performed two sets of experiments to establish each type of externality. We considered two types of externality events: (a) the current ad is the advertiser's j^{th} ad in the browse session, and (b) the ad is shown after showing i competitors' ads in the session, where i and j are parameters. Let E be the externality event in consideration.

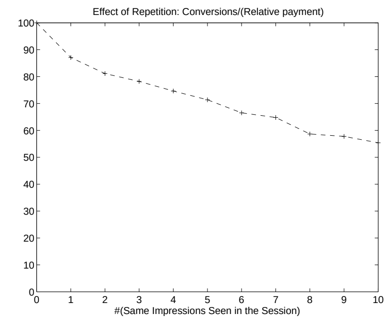
Effect on CtoC ratio: We measure the change in the CtoC ratio over all advertisers as a result of externality event E. Its value is computed as follows: let `conversions`(E) and `clicks`(E) be the number of conversions and clicks summed over all advertisers when event E is true. Then the value of CtoC ratio under event E, denote by `CtoC`(E), is defined to be

$$\mathtt{CtoC}(E) = \frac{\mathtt{conversions}(E)}{\mathtt{clicks}(E)}$$

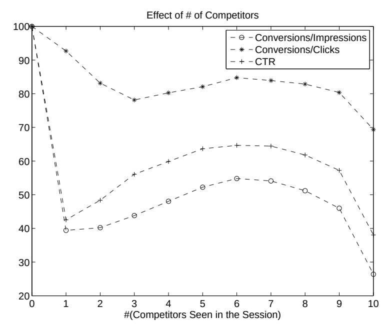
We note that the experiment ignores changes in advertisers' bids from their corresponding average bid values.



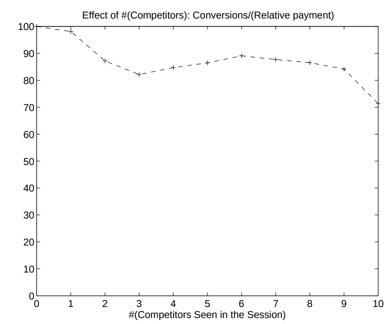
(a) Effect of repeated exposure: Click parameters



(b) Effect of repeated exposure: Efficiency of payment



(c) The effect of preceding competitors: Click parameters



(d) The effect of preceding competitors: Efficiency of payment

Fig. 1. Different externalities that are present in a user's browse session

Effect on Advertiser's Welfare per Unit Payment: We measure the *efficiency-of-payment* for advertisers when the event E is true. It is defined as follows: let `conversions`(E) and `payment`(E) be the number of conversions and the total relative payment summed over all advertisers when event E is true, then the efficiency-of-payment under event E, denote by `EoP`(E), is defined to be

$$\mathtt{EoP}(E) = \frac{\mathtt{conversions}(E)}{\mathtt{payment}(E)}$$

As this experiment scales the clicks by advertisers' bid values, it removes the *effects of external parameters and events* on CtoC ratio such as bad quality of impressions (or less relevant users), as *the relative-bid value of an advertiser is a good indicator of importance* of a (user on a) page to the advertiser. If both experiments show a similar quantitative behavior for the externality event in consideration, then it establishes the externality for the event. Now we illustrate our experimental findings.

Effect of Repeating the Same Ad in the Browse Session – Figure 1(a) plots the values of click parameters over all advertisers based on the prior exposure to the same ad in the current browse session. We measure the prior exposure in terms of the number of times the same ad is shown previously in the session. All values in the plot are *relative* to the maximum possible value of the corresponding click parameter, which happens when the ad is shown for the first time. We observe that the CtoC ratio decreases almost linearly as the ad is repeated multiple times; for instance, it drops to 52% of its maximum value when the same ad is repeated 10 times in the session. Figure 1(b) plots the *efficiency-of-payment* for this externality; we observe that its value also drops linearly and it is around 55% of its maximum value when the ad is repeated 10 times. Thus these two quantities show a similar quantitative behavior.

Effect of Competitors – Next we analyze how important it is for an advertiser that his ad is shown before his competitors in a browse session. Toward this end, we analyzed the effect on an advertiser's click parameters when competing ads are shown on earlier pages in the the current browse session. Figures 1(c) and 1(d) measure the CtoC ratio and the efficiency-of-payment for advertisers as a function of number of preceding competitors in the session. We note that both parameters show a similar quantitative behavior and their values decrease as more competing ads are shown previously in the session. The values of the CtoC ratio and the efficiency-of-payment drop to 70% and 80% respectively of their maximum values as the ad is shown after 10 competing ads in the session. In other experiments, we observed that most advertisers are affected negatively by prior competing ads. In fact, 30% advertisers have their CtoC ratio dropped by more than 50% by preceding competitors, and overall, around 80% advertisers are affected negatively.

3 The Bidding Language and Ad Allocation

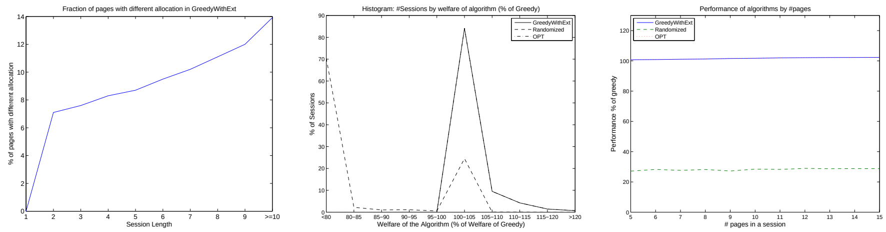
In this section, we give a brief overview of a richer bidding language which enables advertisers to adjust their bids based on the events in the browse session, and study the effectiveness of some natural heuristics for the ad allocation problem. The details of the prior on the user's browse session, the improvement in accuracy of prediction of conversion events with the bidding language and theoretical aspects of the ad allocation problem are deferred to the full version of the paper.

Bidding Language: Advertiser $\mathbb{A}_i$ specifies the *set of his competitors*, and two *discount factors*: $\texttt{disc\_self}_i : N \to R$ and $\texttt{disc\_comp}_i : N \to R$. Let b be the valuation of a click for $\mathbb{A}_i$ on page u. Let j be the number of times advertiser $\mathbb{A}_i$'s ad is shown in the session so far and k be the number of competing ads shown in the session so far, then we have

$$\texttt{ext\_self}(i, j) = \texttt{disc\_self}_i(j) \quad \text{and} \quad \texttt{ext\_comp}(i, k) = \texttt{disc\_comp}_i(k)$$

The total externality is the sum of both externalities. $\mathbb{A}_i$'s *effective valuation* of node u with externality effect is $b(1 + \texttt{ext\_self}(i, j) + \texttt{ext\_comp}(i, k))$.

Expressive Power: We discuss some examples to illustrate the expressive power of the simplified bidding language.



(a) Fraction of pages with different allocation in GreedyWithExt and Greedy a function of the length of the session

(b) Histogram of sessions for various algorithms (% of Greedy).

(c) Performance of various algorithms as a function of number of pages (% ofGreedy)

Fig. 2. Performance of various allocation algorithms

If advertiser $\mathbb{A}_i$ wants his jth repetition in the session to be discounted by $10 \times j\%$, then it can be expressed by setting $\mathtt{disc\_self}_i(j) = -0.1 \times j$.

If advertiser $\mathbb{A}_i$ has two different bids for node u, his bid is b_1 when he is the first among his competitors in the session, and b_2 otherwise, then we can set $w(i,u) = b_1$, $\mathtt{disc\_comp}_i(0) = 0$ and $\mathtt{disc\_comp}_i(> 0) = \frac{b_2 - b_1}{b_1}$. This example is similar to the setting considered in Ghosh *et al* [8], where as advertiser specifies two bids, one for the exclusive display on a page and the other for the non-exclusive display.

The simplified bidding language can also be used as a tool by the ad serving engine, where it computes the discount factors for advertisers to scale their bids, so that the advertisers have better value for their money.

Heuristics for Ad Allocation: As the externalities observed in data are (mostly) negative, the allocation problem is hard to approximate. Hence, we study the performance of some natural heuristics for the allocation problem on real data, using the experimental setup described in Section 2.
a) Greedy allocation with past externality (GreedyWithExt:) Assign the page to the advertiser with maximum (`effectivebid`) $\times$ `CTR` value, where `effective bid` is his bid considering the externality in the session.
b) Randomized allocation (Random:) Chosen advertiser randomly with probability proportional to his (`effectivebid`) $\times$ `CTR` value.
c) Greedy Allocation (Greedy) Assign each page to the advertiser with the maximum (`bid` $\times$ `CTR`) value. This is the optimal allocation in absence of externality effects.
d) Optimal allocation (OPT) (Computed using a dynamic program.)

Observations: The experimental results are given in Figure 2. We note some salient observations:

(1) The first plot (Figure 2(b)) – We measure performances of three algorithms with respect to the Greedy algorithm for sessions with at least 5 pages. For

each algorithm, we classify sessions into ten buckets based on the relative welfare compared to the GREEDY algorithm. We observe that for every session, the performance of GREEDYWITHEXT is at least as good as GREEDY with at least 5% better for 5% sessions, and the performance of GREEDYWITHEXT is indistinguishable from OPT.

(2) The second plot (Figure 2(c)) – We classify sessions based on their length. For a session type, we measure the ratio of the total welfare of the algorithm in consideration summed over these sessions compared to the total welfare of the greedy algorithm for these sessions. As we can see, when the number of pages in a session is small, OPT and GREEDYWITHEXT are not substantially better than GREEDY. This is because, there is less externality in a small session, and the allocation remains almost same even by ignoring it. As session-length increases to 15 pages, the externality effect becomes significant and GREEDYWITHEXT performs 3% better than GREEDY. Furthermore, we note that there is no noticeable difference between the performance of GREEDYWITHEXT and OPT. This suggests that GREEDYWITHEXT works well in practice.

(3) The third plot (Figure 2(a))– We classify sessions based on the number of pages in the session, and measure the fraction of pages that have different allocation in GREEDYWITHEXT and GREEDY algorithms for a given session-type. The difference in allocation increases with an increase in the session length, with 15% difference for sessions with length ≥ 10 pages.

References

1. G. Aggarwal, J. Feldman, and S. Muthukrishnan. Bidding to the top: Vcg and equilibria of position-based auctions. In *WAOA*, pages 15–28, 2006.
2. G. Aggarwal, J. Feldman, S. Muthukrishnan, and M. Pál. Sponsored search auctions with markovian users. In *WINE*, pages 621–628, 2008.
3. A. Bhalgat and S. Gollapudi. Ad allocation in browse sessions. *http://www.seas.upenn.edu/~bhalgat/browse_session.pdf*, 2012.
4. V. Conitzer and T. Sandholm. Expressive negotiation in settings with externalities. In *AAAI*, pages 255–260, 2005.
5. F. Constantin, M. Rao, C.-C. Huang, and D. C. Parkes. On expressing value externalities in position auctions. In *AAAI*, 2011.
6. P. J. Danaher. The effect of competitive advertising interference on sales for packaged goods. *Journal of Marketing Research*, 45:211–225, April 2008.
7. A. Ghosh and M. Mahdian. Externalities in online advertising. In *WWW*, pages 161–168, 2008.
8. A. Ghosh and A. Sayedi. Expressive auctions for externalities in online advertising. In *WWW*, pages 371–380, 2010.
9. R. Gomes, N. Immorlica, and E. Markakis. Externalities in keyword auctions: An empirical and theoretical assessment. In *WINE*, 2009.
10. D. Kempe and M. Mahdian. A cascade model for externalities in sponsored search. In *WINE*, pages 585–596, 2008.
11. S. Lahaie, D. C. Parkes, and D. M. Pennock. An expressive auction design for online display advertising. In *AAAI*, pages 108–113, 2008.
12. S. Verma. Impact of repetitive and contextual advertisements on consumer behavior: An exploratory study. In *IACSIT*, pages 221–225, 2009.