Conversion Rate Based Bid Adjustment for Sponsored Search

Benjamin Rey, Ashvin Kannan
Yahoo! Labs, Santa Clara, CA 95054, U.S.A
{benjamir,ashvin}@yahoo-inc.com

ABSTRACT

Advertisers use Sponsored Search to drive traffic to their site at a conversion rate and cost per conversion that provides value to them. However, very often advertisers bid at a constant price on a bundle of keywords, either for lack of enough data to fully optimize their bids at a keyword level, or indirectly by opting into *Advanced Matching* (*AM*) that allows an advertiser to reach a large number of queries while explicitly bidding only on a limited number. Then this single bid price reflects the return the advertiser gets from the full bundle. Under these conditions, the advertiser is competing too aggressively for some keyword auctions and with too low bids for others. In this paper, we propose a solution to improve the fairness of each keyword's bid prices within an *AM* bundle: adjusting the *AM* keyword bid by the ratio of its conversion rate to the conversion rate it would have reached had it been an *Exact Match* (*EM*). First we describe how we measure advertisers' conversion rates despite the opt-in nature of conversion tracking, and illustrate the need for bid adjustment in the context of *AM*. Then we present our approach to predict conversion rates in a robust manner. Our model uses a number of features capturing the quality of the match between the ad and the query. Then we describe how we adjust keyword bid prices to reflect their value to the advertiser thereby improving (1) the auction through fewer incorrectly high bids in the auction, (2) advertiser return through more auctions won by high value keywords and less by low value keywords, and (3) user satisfaction through higher conversion rate. Finally, we present experimental results from our live system.

Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval – retrieval models, search process; H.3.5 [Information Storage and Retrieval]: Online Information Services – commercial services.

General Terms

Algorithms, Measurement, Experimentation

1. INTRODUCTION

All major search engines serve advertisements alongside their organic search results. These ads are sold through a *Generalized Second Price* auction [1] using a *cost-per-click* (CPC) billing model. There have been several attempts [6] from search engines and advertising companies to let the advertiser define bids on different levels, such as *cost-per-million* impressions (CPM) for brand advertising, and *cost-per-action* (CPA) for direct marketing advertising, as well as theoretical research to define hybrid auction mixing CPC, CPM and CPA pricing [7]. However CPC model prevails. Multiple publications have proposed strategies and algorithms to help advertisers optimize their set of keywords and associated bid [8][9]. Yet, very often advertisers bid at a constant price on a bundle of keywords, either for lack of enough data to fully optimize their bids at a the keyword level, or indirectly by opting in for *Advanced Matching*. *AM* is a system proposed by all major search engines which provides advertisers the choice to be eligible to appear not only for user queries exactly matching their bid term, but also for variations of their bid term, such as spelling correction, stemming, minor generalization or specification, etc. In this case, the advertiser optimize her bid price for the bundle of *{query,AM(query)}* and if some of the phrases from the bundle perform better (worse), their bid will be under (over) estimated, limiting the auction effectiveness. In this paper, we propose a solution to improve the fairness of each keyword's bid prices within an *AM* bundle. First we defined a robust conversion rate metric, then we train a model to predict conversion rate given the quality of matching, and adjust the bid of the *AM* keyword by the ratio of its predicted conversion rate to the conversion rate it would have reached had it exactly matched the user query. Finally we present encouraging results from live experimentation of this bid adjustment approach.

2. CONVERSION DATA AND METRIC

When signing up for a sponsored search campaign, advertisers can choose to place conversion trackers on their website that gets triggered when the user performs predefined actions. These data are necessary to optimize the marketplace but they comprise several limitations which we overcome as follows:

- Trackers can be incorrectly set: we filter out advertisers whose conversion gets triggered too easily or too little
- Population of advertiser with tracking isn't representative of the full population: we use covariate shift correction [5] to correct this self-selection bias on the dimensions of volume, average cost per click and query frequency.
- Conversion meaning differs across advertisers and requires normalization: within each advertiser we can study the conversion rate gain/loss of a subset of clicks, such as experiment, type of matching.

For such subset of clicks, using the filters and self selection weighting (*ssw*), we compute the following normalized conversion metric:

$$CVi = \frac{\sum_{a=advertiser} ssw_a * revenue_{a\_subset} \frac{CVR_{a\_subset}}{CVR_{a\_ref}}}{\sum_{a=advertiser} ssw_a * revenue_{a\_subset}}$$

$CVR_{advertiser\_ref}$ *reference conversion rate from a 2 month period*

3. CPA EQUALIZATION

3.1 Advertiser Optimize for Constant CPA

The goals of advertisers using sponsored search may be branding, direct marketing or a combination thereof. In this paper, we assume advertiser is mostly optimizing for conversion rate (*CVR)*:

$$CPA(kw_1) = \frac{bid(kw_1)}{CVR(kw_1)} = CPA(kw_2) = \frac{bid(kw_2)}{CVR(kw_2)}$$

To increase traffic beyond a limited set of keywords, advertisers can opt-in *AM*: the match between the advertiser bid phrase and the user query will be looser. Since *AM* is broader, it will match the user intent with a lower conversion rate. Hence even if overall advertisers are paying a fair price, the *AM* part of their traffic is of lower quality than the other (*EM*).

3.2 Bid Adjustment by CVR Ratio

For a given keyword opted into *AM*, we can determine the bid price the advertiser would have set for it would it be only exactly matched. However, since in general *CVR(EM) > CVR(AM)*, this implies that *bid(EM) > bid*. Increasing bid higher than the advertiser is willing to pay per click is neither practical nor sensible. Instead we can leave the bid for *EM* clicks unchanged and influence only on the bid for clicks from *AM*:

$$bid(AM(kw)) = \frac{CVR(AM(kw))}{CVR(EM(kw))} \times bid(kw) = CVR_{ratio} \times bid(kw)$$

If we adjust bid by CVR_{ratio}, the reference bid price will become the bid of *EM*: next time the advertiser optimizes his bid, he will be estimating the bid price for his keyword exactly matched only. Using such a ratio also has the advantage that the self-definition of conversion per advertiser cancels out and we can compare across advertisers. CVR_{ratio}, is equivalent to *CVi* when $CVR_{advertiser\_ref}$ is computed using only *EM* clicks from the advertisers.

3.3 Impact of Bid Adjustment

The auction runs on the adjusted bid and will have many more implications than changing keyword cost:

(1)**Ranking**: lower conversion rate ads will be ranked lower

(2)**Pricing**: lower price for all ads in auctions with adjusted bids

(3)**Coverage**: some ads will be dropped for too low expected revenue

(4)**Conversion rate**: ads with better (lower) conversion rate will be ranked more (less) prominently and get more (less) traffic

4. CVR RATIO MODELING

To model CVR_{ratio}, we chose to capitalize on several existing scores that assess on different dimensions the quality of the match between the user query and the advertiser ad:

(1) **Quality of match between user query and bid term**[4] using features such as syntactical match, occurrences statistics from user sessions, occurrences statistics in advertiser databases.

(2) **Quality of match between user query and advertiser listing** [3] using features around syntactical match between user query and text from the title, abstract and url from the ad listing.

(3) **Probability of click of the ad given user and query**[2] using features such as click feedback history and word pair text features from the user query and the ad listing.

(4) **Matching technology**: each type of *AM* has different *CVi*.

Given the constraint of robustness and the strong correlation between the features with *CVi*, we chose to experiment with a linear regression weighted by number of clicks, with the features mentioned above and their conjugates. The training set is made of 2 months of Yahoo! Search clicks under conversion tracking.

We evaluated our model offline on two weeks of clicks from Yahoo! Search, starting after the end of the training set and did observe that the lowest predicted quality clicks have a much lower conversion rate. For instance the 10% lowest scored clicks have a *CVi* 25% lower than an average *AM* click.

5. LIVE EXPERIMENT

We ran a number of experiments on a random sample of live search traffic from the Yahoo! US domain. Result metrics are compared to a baseline experiment that was run simultaneously.

Adjusting the bid price only down can have a large immediate revenue impact. An option is to iterate, starting with a fixed smaller budget and give only part of the discount, then wait for the advertisers to have reacted to the CPA gain, then give more discount, until we reach the CPA parity between *AM* and *EM*.

For a fixed budget, we can either spread the limited budget to all *AM* ads by squashing the score by a small value (γ), or shift the score to impact mostly the lowest quality ads:

$$bid\ factor = \min\left(\overline{CVi}^{\gamma} + shif, 1.0\right).$$

The table below shows the metrics from four weeks of live experiment with settings such that the discount is given to the ads with the larger need of it: γ = 1.0, shift = 0.114.

Table 1. Metrics from live experiment

CPAi	-6%	AM vs EM gap: -45%
CVi	+2%	
Click Yield	0%	
Revenue Per Search	-2.5%	
Cost Per Click	-3%	
Coverage	-3%	

The results are convincing: not only did we bring *AM CPA* much closer to *EM CPA*, by reducing the gap by nearly half, but it also substantially improved the user experience, by removing 3% coverage without losing any clicks. A fair amount of the revenue loss from *AM* was compensated by new clicks on *EM*, and overall even though we are losing 2.5% *RPS*, advertisers are getting a 6% better deal, which we can hope they will reinvest. We also ran the experiment with different settings, spreading the budget across all AM, or with a minimal *RPS* impact, however results weren't as significant.

6. CONCLUSIONS AND FUTURE WORK

In this paper we presented a solution to automatically adjust the bid price for *AM* bid terms, to reflect the decrease in direct marketing quality. With such an optimization, advertisers would then only need to optimize the bid price for their keywords being exactly matched. Our simple but robust modeling is giving promising results as shown in the live experimentation, improving not only advertiser CPA and reducing the CPA difference between terms which are *Exactly Matched* and those which are *Advanced Matched*, but also improving user satisfaction. We also presented a method for working within fixed revenue constraints.

7. REFERENCES

[1] B. Edelman et al. Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords American Economic Review, 2007

[2] B. Shaparenko et al. Data driven text features for sponsored search click prediction. In AdKDD Workshop, 2009.

[3] D. Hillard et al. Improving Ad Relevance in Sponsored Search. In WSDM, 2010.

[4] R. Jones et al. Generating query substitutions. In WWW, 2006.

[5] A. Gretton et al. Dataset shift in machine learning. In Covariate Shift and Local Learning by Distribution Matching, pages 131-160, Cambridge,MA, 2008.

[6] D. Fain and J. Pedersen. "Sponsored Search: a Brief History", In Proceedings of the Second Workshop on Sponsored Search Auctions, 2006.

[7] B. Edelman et al. CPC/CPA Hybrid Bidding in a Second Price Auction No. 09-074 Negotiation, Organizations & Markets 2008

[8] J. Feldman et al. Budget optimization in search-based advertising auctions. ACM Conference on EC, 2006.

[9] S. Goel et al. Contract Auctions for Sponsored Search. WINE 2009, LNCS 5929, pp. 196–207, 2009.