

On the ranking strategy in adword auctions

Patrick Maillé

Institut Telecom; Telecom Bretagne
2, rue de la Châtaigneraie CS 17607
35576 Cesson Sévigné Cedex, France
Email: patrick.maill@telecom-bretagne.eu

Bruno Tuffin

INRIA Rennes - Bretagne Atlantique
Campus Universitaire de Beaulieu
35042 Rennes Cedex, France
Email: Bruno.Tuffin@inria.fr

Abstract—Search engines get revenue thanks to adword auctions, where commercial links are proposed and charged to advertisers as soon as the link is clicked through. Most search engines have chosen (or switched to) a revenue-based ranking and charging scheme instead of a bid-based one. We investigate here the relevance of that scheme when advertisers' valuation comes from a random distribution. We show that, depending on the search engine's click-through-rate, revenue-based does not always outperform bid-based in terms of revenue to the search engine. As a result, some search engines may have an interest to move to revenue-based ranking while others do not.

I. INTRODUCTION

Search engines are a key piece for web-browsing, allowing to easily find relevant web sites corresponding to given keywords. An important characteristic of their success, in addition to their efficiency, is that this service is free. In order to get a return on investment, search engines propose slots for commercial links to advertisers interested in specific keywords [1]. Those slots are allocated thanks to an auction scheme, hence the terminology *adword auction*.

There are usually a given number K of slots, ordered according to their visibility on the screen. Advertisers submit bids for specific keywords and each time there is a search on that keyword, advertisers are ranked and slots allocated according to a specific criterion: they can for instance be ranked according to their bid value, or by the revenue they can generate (this will be made clearer later). Advertisers may be charged each time their ad is displayed (scheme called *Pay-Per-Impression*), each time the ad is clicked (scheme called *Pay-Per-Click*), or each time the click results in a real sale (called *Pay-Per-Transaction*). The amount of money to be charged each time can be chosen among several possibilities:

- the first-price principle, where advertisers pay their bid,
- the so-called Generalized Second Price (GSP), where they pay
 - the bid of the advertiser just below them in the ranking if the ranking strategy is based on bids;
 - or the price such that the revenue generated corresponds to the (declared) one the advertiser below them if the ranking is based on revenue.
- Another option is to make use of Vickrey-Clarke-Groves (VCG) auctions: each advertiser pays the opportunity cost that its presence introduces to all other advertisers.

For more on adword auction description, the reader can look at [2], [3], [4], [5] and the references therein.

The commonly used strategies are Pay-Per-Click and GSP [2]. Though, there used to be some differences about the ranking strategy procedure used. While bid-based ranking was initially implemented by Yahoo!, Google has introduced a revenue-based ranking, ordering according to the value of the product of the bid by the Click-Through-Rate (CTR), which is the probability that the ad is clicked when displayed. Recently, Yahoo! has changed its ranking policy to a revenue-based one too.

Our goal in this paper is to investigate the best ranking strategy for search engines when Pay-Per-Click and GSP are applied to investigate whether (and when) the move from Yahoo! is relevant. Some papers have studied search engines optimal ranking strategies, see for instance [6], [7]. In [6], a comparison is made between a deterministic allocation based on GSP and a random assignment rule based on VCG. Deterministic allocation will produce a higher revenue for search engines but a lower utility for advertisers; it is argued that under competition, random allocation could eventually be the winner if advertisers have to choose. In [7], a game is applied between two search engines on the ranking policy (choosing between revenue-based and bid-based), given that advertisers choose to submit their ad at only one of them. A conclusion gives a potential reason explaining why Yahoo! has made the move. Though, the paper is based on the assumption that advertisers necessarily choose among the two engines. A practical justification of this could be based on budget constraints, but those are not included in the model of [7]. If advertisers bid truthfully and are sure that bidding is always beneficial, why not sending bids to both engines? We aim at investigating the best ranking strategy of search engines in that case. Actually, here the choice of a search engine does not have an influence on the other. Therefore the analysis can be performed by looking at an engine in isolation. The best strategy will depend on the engine parameters, such that its CTRs which can vary due to different presentations and be more or less attractive to the (heterogeneous) population.

The rest of this paper is organized as follows. Section II reviews the basic model, mostly inspired from [7]. Section III gives a closed-form expression for the average revenue of the search engine in terms of the probability distribution of the advertisers' valuation and the CTRs, for each engine strategy

(bid-based and revenue-based). Section IV illustrates the kind of results that can be obtained by deriving the expression for a specific valuation distribution, namely the uniform one. A numerical comparison of the revenue is then performed in terms of the CTRs. Finally, Section V concludes and gives some directions for future research.

II. MODEL DEFINITION

We assume that a set of n advertisers submit bids to search engines for a given keyword. Bidders (advertisers) are assumed to be truthful about their valuation, and providers are assumed to apply Pay-Per-Click and GSP, so that submitting to all engines at the same time is the best option for advertisers (they will always end-up with a non-negative utility). As a consequence, search engines revenues are independent of the ranking-rule of competitors. The best strategy of an engine therefore only depends on its own choice. As a consequence, we consider a single engine in this paper, and compare the revenue when bid-based or revenue-based mechanisms are implemented, depending on the CTRs of advertisers. As we shall see, two engines may have different optimal strategies when the CTRs of advertisers differ from an engine to another.

Since our goal is to illustrate that phenomenon, we consider only one slot. As a side effect, GSP corresponds to VCG and truthful-bidding (i.e., bidding one's real valuation for a click on one's ad) is the optimal strategy for advertisers. In that context, consider n advertisers with respective bids $(v_i)_{i \in \{1, \dots, n\}}$ and CTRs $(q_i)_{i \in \{1, \dots, n\}}$.

- A *bid-based* Pay-Per-Click GSP mechanism consists in giving the slot to the highest bidder $i_b := \arg \max_i v_i$, and charging him the second-highest bid $\max_{i \neq i_b} v_i$ at each click, hence yielding an expected revenue $q_{i_b} \max_{i \neq i_b} v_i$.
- On the other hand, a *revenue-based* Pay-Per-Click GSP mechanism would give the slot to the bidder with the largest product bid \times CTR, i.e., to the advertiser $i_r := \arg \max_i q_i v_i$. Then bidder i_r would be charged at each click the lowest price p_{i_r} he could have bid while still getting the slot: $p_{i_r} = \frac{1}{q_{i_r}} \max_{i \neq i_r} q_i v_i$. As a result, the expected revenue for the search engine would be $q_{i_r} p_{i_r} = \max_{i \neq i_r} q_i v_i$.

To compare quantitatively the performance of both schemes, we consider a model definition similar to the one in [7]. Specifically, we assume that there are two classes of advertisers:

- high-quality advertisers, with CTR q_h
- low-quality advertisers, with CTR q_l , such that $q_l < q_h$.

The probability of being a high quality one is denoted by α . Advertisers also have different valuation per click v , distributed according to an absolutely continuous pdf F and density f , independent of the quality.

III. AVERAGE REVENUES

Remark that when considering specific instances, revenue-based ranking does not necessarily produce a larger revenue than bid-based one. Consider for instance $n = 2$ advertisers,

where advertiser 1 has valuation $v_1 = 0.7$ and CTR $q_h = 0.6$, and advertiser 2 has valuation $v_2 = 0.5$ and CTR $q_l = 0.1$. In that case, with bid-based ranking rule, advertiser 1 is the winner, and the revenue will be $v_2 q_h = 0.3$. With revenue-based ranking rule, $v_1 q_h > v_2 q_l$ so advertiser 1 is still the winner, but the price per click is $v_2 q_l / q_h$ and the generated revenue is $v_2 q_l = 0.05$, less than with the bid-based strategy. The intuition is that the second-ranked bidder may be a low-quality advertiser, resulting in smaller price per click when using the revenue-based scheme. Note though that this is less likely to occur when the number of bidders increases.

In this section, we intend to quantify the prevalence and the expected impact of such situations, so as to determine the best revenue-making strategy. To do so, we derive closed-form expressions of revenues for both ranking strategies.

Proposition 1: The average revenue under bid-based ranking and charging is

$$R_b = n(n-1)(\alpha q_h + (1-\alpha)q_l) \int x(F(x))^{n-2}(1-F(x))f(x)dx.$$

Proof: Thanks to the independence of valuation and quality, the average revenue under Pay-Per-Click and GSP is the product of the average CTR and the average value of the second highest bid. The average CTR is $\alpha q_h + (1-\alpha)q_l$.

The expected value of the second highest bid is determined from the law of order statistics. It is known that the density value of the k -th statistic $X_{(k)}$, when ordered in a increasing order, within a sample of n independent random variables of density f and cdf F is (see [8])

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} (F(x))^{k-1} (1-F(x))^{n-k} f(x). \quad (1)$$

Then the density of the second highest bid is (for $k = n-1$)

$$\frac{n!}{(n-2)!} (F(x))^{n-2} (1-F(x)) f(x).$$

The proposition immediately follows. ■

Proposition 2: Define

$$\begin{aligned} G(x) &= \alpha F(x/q_h) + (1-\alpha)F(x/q_l) \\ g(x) &= \frac{\alpha}{q_h} f(x/q_h) + \frac{1-\alpha}{q_l} f(x/q_l). \end{aligned}$$

The average revenue under revenue-based ranking and charging is

$$R_r = n(n-1) \int x(G(x))^{n-2} (1-G(x))g(x)dx.$$

Proof: The expected revenue under the revenue-based rule is the expected value of $Q \times V$, with Q the random value of the CTR and V the random valuation. Let G be the cdf of QV . We have

$$\begin{aligned} G(x) &= \mathbb{P}[QV \leq x] \\ &= \alpha F(x/q_h) + (1-\alpha)F(x/q_l). \end{aligned}$$

The corresponding density is

$$g(x) = \frac{\alpha}{q_h} f(x/q_h) + \frac{1-\alpha}{q_l} f(x/q_l).$$

From (1), the density of the second largest value of QV is

$$\frac{n!}{(n-2)!} (G(x))^{n-2} (1-G(x))g(x),$$

hence the proposition. \blacksquare

The expressions in Propositions 1 and 2 are quite different, and difficult to compare in the general case. Nevertheless, we can remark that both are equal when there is no heterogeneity in CTRs, i.e., when $\alpha = 0$ or $\alpha = 1$. This was to be expected, since when all advertisers have the same CTR q the bid-based and revenue-based rankings are equivalent: for any bid profile, we have $i_b = i_r$, and the expected revenue of both schemes are the same: $q \max_{i \neq i_b} v_i = \max_{i \neq i_r} qv_i$.

To study the case when there is heterogeneity among advertisers in terms of CTR, we consider a given distribution of valuations in the next section.

IV. EXAMPLE AND NUMERICAL ILLUSTRATION

Section III has provided general results. We analyze here how computations can be performed for a specific valuation cdf. We consider a uniform distribution over $[0, 1]$, i.e., $F(x) = x$ for $x \in [0, 1]$.

In this case, under the bid-based strategy,

$$\begin{aligned} R_b &= n(n-1)(\alpha q_h + (1-\alpha)q_l) \int_0^1 x^{n-1}(1-x)dx \\ &= n(n-1)(\alpha q_h + (1-\alpha)q_l) \left[\frac{1}{n} - \frac{1}{n+1} \right] \\ &= \frac{n-1}{n+1} (\alpha q_h + (1-\alpha)q_l). \end{aligned}$$

Under the revenue-based strategy, first remark that $\forall x \geq 0$,

$$G(x) = \alpha \min(1, x/q_h) + (1-\alpha) \min(1, x/q_l).$$

Define $\beta = \frac{\alpha}{q_h} + \frac{1-\alpha}{q_l}$ and $\gamma = \frac{\alpha}{q_h}$. The revenue can be expressed by decomposing the interval into three subintervals $[0, q_l]$, $[q_l, q_h]$ and $[q_h, 1]$ as

$$\begin{aligned} R_r &= n(n-1) \int_0^{q_l} x (\beta x)^{n-2} (1-\beta x) \beta dx \\ &\quad + n(n-1) \int_{q_l}^{q_h} x (\gamma x + 1 - \alpha)^{n-2} (\alpha - \gamma x) \gamma dx \\ &\quad + 0. \end{aligned}$$

Remark that the third line is obtained due to a null derivative on this interval. The two integrals can be computed; the first one yields

$$n(n-1)\beta^{n-1}q_l^n \left(\frac{1}{n} - \frac{\beta q_l}{n+1} \right),$$

and the second one gives by multiple integration by part

$$\begin{aligned} &n [(\gamma q_h + 1 - \alpha)^{n-1} (\alpha q_h - \gamma q_h^2) - (\gamma q_l + 1 - \alpha)^{n-1} (\alpha q_l - \gamma q_l^2)] \\ &- \frac{1}{\gamma} [(\gamma q_h + 1 - \alpha)^n (\alpha - 2\gamma q_h) - (\gamma q_l + 1 - \alpha)^n (\alpha - 2\gamma q_l)] \\ &- \frac{2}{(n+1)\gamma} [(\gamma q_h + 1 - \alpha)^{n+1} - (\gamma q_l + 1 - \alpha)^{n+1}]. \end{aligned}$$

Next figures illustrate the gain for specific parameters.

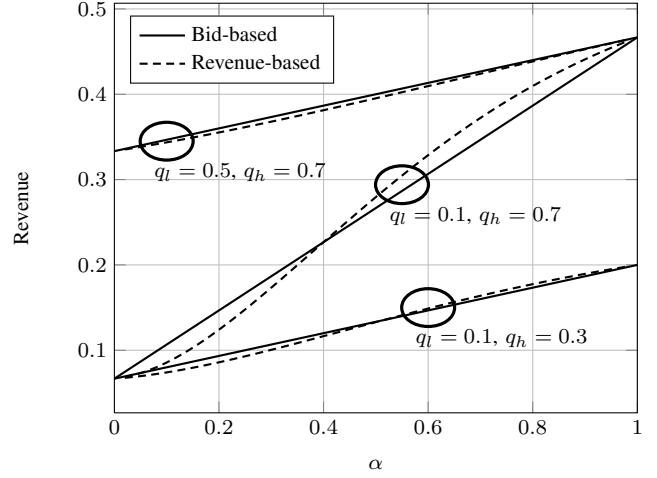


Figure 1: Revenues in terms of α for different values of q_l , q_h , when $n = 5$.

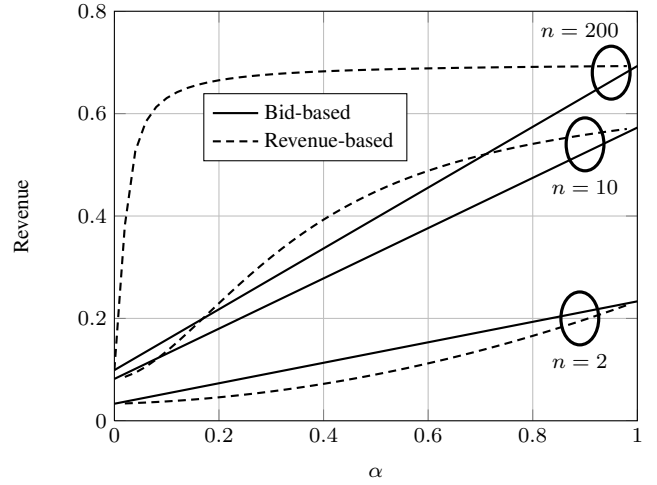


Figure 2: Revenues in terms of α for different values of n when $q_l = 0.1$, $q_h = 0.7$.

Figure 1 displays the revenue in terms of the proportion α of high-quality users for different CTR values. Here one can check that for low values of α , and low values of q_l , the bid-based ranking rule may produce a larger revenue than the revenue-based one. On all the curves, when $\alpha \in \{0, 1\}$, it can be verified that revenues are the same. Revenues also (obviously) increase with the proportion of high-quality advertisers, and with the CTRs.

Figure 2 displays the revenue in terms of the number n of advertisers for different CTR values. For a small number of advertisers the chances to get a larger revenue with bid-based ranking is larger. It is actually the case here when $n = 2$ for every value of α . But as the number of advertisers increases, this is less likely. This is due to the fact that we have more chances to have a high-quality advertiser in second position with the revenue-based rule, while the probability α of having

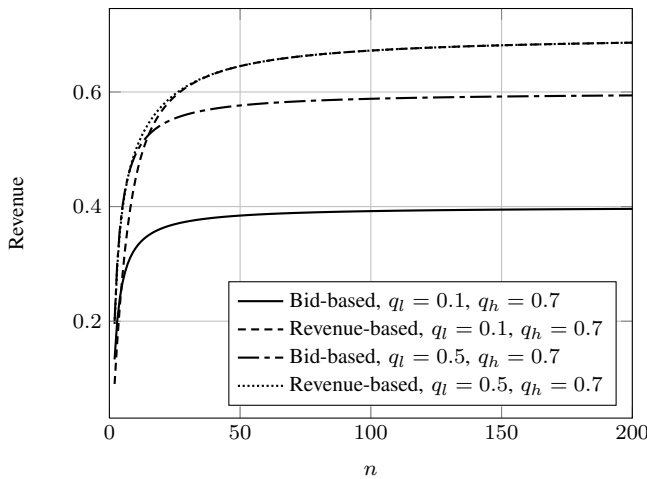


Figure 3: Revenues in terms of n for different values of q_l , q_h , when $\alpha = 0.5$.

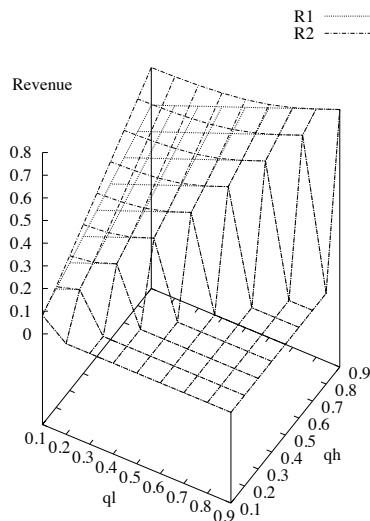


Figure 4: Revenues in terms of q_l , q_h when $n = 10$ and $\alpha = 0.5$.

a high-quality advertiser with the bid-based rule is unchanged. As displayed in Figure 2, for $n = 200$ advertisers the revenue-based mechanism always outperforms the bid-based one.

Figure 3 displays the revenue in terms of the number n of advertisers for different CTR values. As the number of advertisers increase, here again, revenue-based rule yields a larger revenue. Moreover, the average revenue asymptotically depends on the high-quality advertisers parameters with this rule (the two curves tend to coincide) because the second ranked advertisers tend to be a high-quality one. On the other hand, this is not true for bid-based ranking rule because the valuation is independent of the quality parameter.

Figure 4 displays the revenues in terms of q_l , q_h when $n = 10$ and $\alpha = 0.5$. On this 3D curve, we have only considered the cases where $q_h > q_l$ (revenues set to 0 otherwise). As a

remark, bid-based and revenue-based get closer as q_h and q_l get closer, and are the same when $q_l = q_h$. Indeed, those two schemes are exactly the same in that case.

V. CONCLUSIONS

We have investigated in this paper the best strategy between revenue-based and bid-based rankings for search engines. This had already been studied in the literature when engines are in competition, but under the assumption that advertisers choose only one engine, an assumption that has not been justified. If we relax this assumption, then search engines can be considered in isolation (the strategy of the competitors does not have an influence on the revenue). We show here that under randomness over advertisers' valuation and CTRs, the best strategy is not always to apply revenue-based ranking and charging: it actually depends on the various parameters. We derive closed-form expressions for the expected revenues of both revenue-based and bid-based schemes. As an output, a search engine is able to decide, according to the advertisers parameters and reactions to the engine, the best strategy between the two choices.

ACKNOWLEDGMENT

This work was supported by Euro-NF Network of Excellence (<http://euronf.enst.fr/>) and the French research agency through the CAPTURES project (<http://captures.inria.fr/>).

REFERENCES

- [1] Internet Advertising Bureau, <http://www.iab.net>.
- [2] B. Edelman, M. Ostrovsky, and M. Schwarz, "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords," *The American Economic Review*, vol. 97, no. 1, pp. 242–259, 2007.
- [3] P. Jordan and M. Wellman, "Designing an ad auctions game for the trading agent competition," in *IJCAI-09 Workshop on Trading Agent Design and Analysis (TADA)*, Pasadena, California, USA, Jul 2009.
- [4] S. Lahaie, D. Pennock, A. Saberi, and R. Vohra, "Sponsored search auctions," in *Algorithmic Game Theory*, N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, Eds. Cambridge University Press, 2007, ch. 28, pp. 699–716.
- [5] P. Maillé, E. Markakis, M. Naldi, G. Stamoulis, and B. Tuffin, "An overview of research on sponsored search auctions," 2010, available at http://www.irisa.fr/dionysos/pages/_perso/tuffin/Publis/Survey-v4.pdf.
- [6] P. Maillé and B. Tuffin, "On the interest of introducing randomness in ad-word auctions," in *Proc. of the 1st IFIP International Conference on the Network of the Future*, Brisbane, Australia, September 2010.
- [7] D. Liu, J. Chen, and A. Whinston, "Competing keyword auctions," in *Proc. of 4th Workshop on Ad Auctions*, Chicago, IL, USA, Jul 2008.
- [8] H. A. David, *Order Statistics*, 2nd ed. Wiley, 1981.