A Truthful Learning Mechanism for Multi–Slot Sponsored Search Auctions with Externalities

(Extended Abstract)

Nicola Gatti
Politecnico di Milano
ngatti@elet.polimi.it

Alessandro Lazaric
INRIA Lille – Nord Europe
lazaric@inria.fr

Francesco Trovò
Politecnico di Milano
trovo@polimi.it

ABSTRACT

In recent years, effective sponsored search auctions (SSAs) have been designed to incentivize advertisers (advs) to bid their truthful valuations and, at the same time, to assure both the advs and the auctioneer a non–negative utility. Nonetheless, when the click–through–rates (CTRs) of the advs are unknown to the auction, these mechanisms must be paired with a learning algorithm for the estimation of the CTRs. This introduces the critical problem of designing a learning mechanism able to estimate the CTRs as the same time as implementing a truthful mechanism with a revenue loss as small as possible. In this paper, we extend previous results [2, 3] to the general case of multi–slot auctions with position– and ad–dependent externalities with particular attention on the dependency of the regret on the number of slots K and the number of advertisements n.

Categories and Subject Descriptors

I.2 [**Artificial Intelligence**]: Miscellaneous

General Terms

Algorithms, Economics, Theory

Keywords

Sponsored search auction, Learning mechanism

1. NOTATION AND BACKGROUND

We consider a standard model of SSAs. We denote by $\mathcal{N} = \{1, \dots, n\}$ the set of ads. Each ad i is characterized by a *quality* ρ_i, defined as the probability that i is clicked once observed by the user, and by a *value* $v_i \in [0, V]$, that the adv receives once i is clicked (the value is zero if not clicked). While qualities ρ_i are common knowledge, values v_i are private information of the advs. We denote by $\mathcal{K} = \{1, \dots, K\}$ with $K < n$ the set of available slot. An ad–slot allocation rule α is a full bijective mapping from n ads to n slots such that $\alpha(i) = k$ if ad $i \in \mathcal{N}$ is displayed at slot k. For all the non–allocated ads, $\alpha(i)$ takes an arbitrary value from $K+1$ to n so as to preserve the bijectivity of α. We also define the inverse slot–ad allocation rule $\beta = \alpha^{-1}$ such that $\beta(k) = i$ if slot k displays ad i (i.e., $\alpha(i) = k$). We denote by $\mathcal{A}$ and $\mathcal{B}$ the set of all the possible ad–slot and slot–ad mappings respectively. Finally, we define $\mathcal{A}_{-i} = \{\alpha \in \mathcal{A}, \alpha(i) = n\}$ as the set of allocations where ad i is never displayed. We adopt the cascade model [1, 5] to describe the user's behavior. The discount factor $\gamma_k(i)$ is the probability that a user, observing ad i in the slot $k-1$, will observe the ad in the next slot (γ_1 is set to 1 by definition). The cumulative discount factors $\Gamma_k(\beta)$, i.e., the probability with which a user observes the ad displayed at slot k given a slot–ad allocation β, is defined as $\Gamma_k(\beta) = \prod_{l=2}^{k} \gamma_l(\beta(l-1))$ for $2 \leq k \leq K$. With abuse of notation, we use interchangeably $\Gamma_k(\beta)$ and $\Gamma_k(\alpha)$. Given an allocation rule α, $\Gamma_{\alpha(i)}(\alpha)\rho_i$ is the *click through rate* (CTR), representing the probability of ad i to be clicked. Finally, we define the *social welfare* of an allocation α as the cumulative advs' expected values

$$\mathrm{SW}(\alpha) = \mathrm{SW}(\beta) = \sum_{i=1}^{n} \Gamma_{\alpha(i)}(\alpha)\rho_i v_i = \sum_{k=1}^{n} \Gamma_k(\beta)\rho_{\beta(k)} v_{\beta(k)}.$$

At each round, advs submit bids and the auction defines an allocation rule α and payments p_i. The Vickrey–Clark–Groves mechanism (VCG) satisfies a number of interesting properties, notably the incentive compatibility (IC) (i.e., no adv can increase its utility by misreporting its true value and $\hat{v}_i = v_i$), and it allocates ads according to the *efficient* allocation $\alpha^* = \arg\max_{\alpha \in \mathcal{A}} \mathrm{SW}(\alpha)$, and payments are set to 0 if the ad is not clicked and to

$$\tilde{p}_i = \frac{\mathrm{SW}(\alpha^*_{-i}) - \mathrm{SW}_{-i}(\alpha^*)}{\Gamma_{\alpha(i)}(\alpha)\rho_i}, \tag{1}$$

if the ad is clicked, so that $\mathbb{E}[\tilde{p}_i] = p_i = \mathrm{SW}(\alpha^*_{-i}) - \mathrm{SW}_{-i}(\alpha^*)$.

In many practical problems, the qualities ρ_i are not known in advance and must be estimated at the same time as the auction is deployed. This introduces a tradeoff between *exploring* different possible allocations so as to collect information about the quality of the advs and *exploiting* the estimated qualities so as to implement a truthful high–revenue auction. Let $\mathfrak{A}$ be an IC mechanism run over T rounds. At each round t, $\mathfrak{A}$ defines an allocation $\hat{\alpha}_t$ and prescribes an expected payment p_{it} for each ad i. The objective of $\mathfrak{A}$ is to obtain a revenue as close as possible to a VCG mechanism. More precisely, we measure the performance of $\mathfrak{A}$ as its cumulative regret over T rounds:

$$R_T(\mathfrak{A}) = T \sum_{i=1}^{n} p_i - \sum_{t=1}^{T} \sum_{i=1}^{n} p_{it}. \tag{2}$$

The mechanism $\mathfrak{A}$ is a *no–regret* mechanism if its per–round regret decreases to 0 as T increases, i.e., $\lim_{T\to\infty} R_T/T = 0$.

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

2. REGRET BOUNDS

Similar to [3], we define an exploration–exploitation algorithm to approximate the VCG, which we refer to as A–VCG. The algorithm estimates the quality of each adv during a pure exploration phase of length τ when all the payments are set to 0. Then, quality estimates are used to set up a VCG for all the remaining $T-\tau$ rounds. At each round, we can collect K samples (click or not–click events), one from each slot. Let α_t (for $t \leq \tau$) be an arbitrary explorative allocation rule independent from the bids. It is easy to define a sequence of explorative allocations $\{\alpha_t\}_{t=1}^{\tau}$ such that the number of samples collected for each ad i is $S_i = \lfloor K\tau/n \rfloor$. We denote by $c^i_{\alpha_t(i)}(t) \in \{0,1\}$ the click–event at time t for ad i when displayed at slot $\alpha_t(i)$. Depending on the slot we have different CTRs, thus we reweigh each sample by the cumulative discount factor of the slot and we compute the estimated quality $\hat{\rho}_i$ as

$$\hat{\rho}_i = \frac{1}{S_i} \sum_{s=1}^{S_i} \frac{c^i_{\alpha_t(i)}(t)}{\Gamma_{\alpha_t(i)}(\alpha_t)}. \tag{3}$$

Depending on the specific user–model of the auction, we can build a high–probability confidence interval on $|\rho_i - \hat{\rho}_i|$ of size

$$\eta_p := \sqrt{\left(\sum_{k=1}^{K} \frac{1}{\Gamma_k^2}\right) \frac{n}{2K^2\tau} \log \frac{n}{\delta}}\ ; \quad \eta_{pa} := \frac{1}{\Gamma_{\min}} \sqrt{\frac{n}{2K\tau} \log \frac{n}{\delta}},$$

with $\Gamma_{\min} = \min_{\alpha,k} \Gamma_k(\alpha)$, for pos– and pos/ad–dependent externalities respectively. After the exploration phase we define an upper–bound on the quality as $\hat{\rho}_i^+ = \hat{\rho}_i + \eta$. Given $\hat{\rho}_i^+$, we compute the estimated social welfare as $\widehat{\text{SW}}(\alpha) = \sum_{i=1}^{n} \Gamma_{\alpha(i)}(\alpha)\hat{\rho}_i^+ v_i$. The corresponding efficient allocation is denoted by $\hat{\alpha} = \arg\max_{\alpha \in \mathcal{A}} \widehat{\text{SW}}(\alpha)$. Once the exploration phase is over, if ad $i \in \mathcal{N}$ is clicked, then the adv is charged

$$\tilde{p}_i = \frac{\widehat{\text{SW}}(\hat{\alpha}^*_{-i}) - \widehat{\text{SW}}_{-i}(\hat{\alpha}^*)}{\Gamma_{\hat{\alpha}(i)}\hat{\rho}_i^+} \tag{4}$$

which corresponds to an expected payment $\hat{p}_i = \tilde{p}_i \Gamma_{\hat{\alpha}(i)} \rho_i$.

Position–dependent externalities. In case of pos–dependent externalities, the discount coefficients reduce to $\Gamma_k(\alpha) = \Gamma_k$, thus simplifying the computation of both the optimal allocation and the payments. In this case, A–VCG achieves the following regret performance.

THEOREM 1. *In a SSA auction with pos–dependent externalities, by optimizing the parameters τ and δ, the A–VCG is always **truthful** and it achieves a regret*

$$R_T \leq 18^{1/3} V T^{2/3} \Gamma_{\min}^{-2/3} K^{2/3} n^{1/3} (\log(n^2 KT))^{1/3}. \tag{5}$$

where $\Gamma_{\min} = \min_k \Gamma_k \geq 0$.

Remark 1 (Bound). Up to numerical constants and logarithmic factors, the previous bound is $R_T \leq \tilde{O}(T^{2/3}K^{2/3}n^{1/3})$. We first notice that the A–VCG is a zero–regret algorithm since its per–round regret (R_T/T) decreases to 0 as $T^{-1/3}$, thus implying that it asymptotically achieves the same performance as the VCG. Furthermore, the dependence of the regret on n is sub–linear ($n^{1/3}$) and this allows to increase the number of advs without significantly worsening the regret. Finally, according to the bound (5) the regret has a sublinear dependency $\tilde{O}(K^{2/3})$ on the number of slots, meaning that whenever one slot is added to the auction, the performance does not significantly worsen. By analyzing the difference between the payments of the VCG and A–VCG, we notice that during the exploration phase the regret is $O(\tau K)$ (e.g., if all K slots are clicked at each explorative round), while during the exploitation phase the estimation errors sum over all the K slots, thus suggesting a linear dependency on K for this phase as well. Nonetheless, as K increases, the number of samples available per each ad increases as $\tau K/n$, thus improving the accuracy of the quality estimates by $\tilde{O}(K^{-1/2})$. As a result, as K increases, the exploration phase can be shortened (the optimal τ actually decreases as $K^{-1/3}$), thus reducing the regret during the exploration, and still have accurate enough estimations to control the regret of the exploitation phase.

Pos/ad–dependent externalities. In this case the learning problem is more complicated, and the regret is:

THEOREM 2. *In a SSA auction with pos/ad–dependent externalities, by optimizing the parameters τ and δ, the A–VCG is always **truthful** and it achieves a regret*

$$R_T \leq 6^{1/3} \frac{V}{\rho_{\min}} T^{2/3} \Gamma_{\min}^{-2/3} K^{2/3} n (\log(KT))^{1/3}. \tag{6}$$

Remark 1 (Differences with the previous bound). Up to constants and logarithmic factors, the previous bound is $R_T \leq \tilde{O}(T^{2/3}K^{2/3}n)$. We first notice that moving from pos– to pos/ad–dependent externalities does not change the dependency of the regret on the number of rounds T. The main difference is in the dependency on n and on the smallest quality $\rho_{\min}$. While the regret still scales as $K^{2/3}$, it has now a much worse dependency on the number of ads (from $n^{1/3}$ to n). We believe that it is mostly due to an intrinsic difficulty of the position/ad–dependent externalities. The intuition is that now in the computation of the payment for each ad i, the errors in the quality estimates cumulate through the slots (unlike the pos–dependent case where they are scaled by $\Gamma_k - \Gamma_{k+1}$). Nonetheless, this cumulated error should impact only on a portion of the ads (i.e., those which are actually impressed according to the optimal and the estimated optimal allocations). Thus we conjecture that this additional n term is indeed a rough upper–bound on the number of slots K and that we could obtain a regret $\tilde{O}(T^{2/3}K^{4/3}n^{1/3})$, where the dependency on the number of slots becomes super–linear. A more detailed discussion on this conjecture and on the dependency on $\rho_{\min}$ can be found in the extended version of this paper [4].

3. REFERENCES

[1] G. Aggarwal, J. Feldman, S. Muthukrishnan, and M. Pál. Sponsored search auctions with markovian users. In *WINE*, pages 621–628, 2008.

[2] M. Babaioff, Y. Sharma, and A. Slivkins. Characterizing truthful multi-armed bandit mechanisms. *CoRR*, abs/0812.2291, 2008.

[3] N. R. Devanur and S. M. Kakade. The price of truthfulness for pay-per-click auctions. In *ACM EC*, pages 99–106, 2009.

[4] N. Gatti, A. Lazaric, and F. Trovò. A truthful learning mechanism for contextual multi-slot sponsored search auctions with externalities. Technical report, 2012. http://hal.inria.fr/hal-00662549.

[5] D. Kempe and M. Mahdian. A cascade model for externalities in sponsored search. In *WINE*, pages 585–596, 2008.