A Novel Ex-Post Truthful Mechanism for Multi-Slot Sponsored Search Auctions

(Extended Abstract)

Debmalya Mandal
Indian Institute of Science, Bangalore, India
dmandal@csa.iisc.ernet.in

Y. Narahari
Indian Institute of Science, Bangalore, India
hari@csa.iisc.ernet.in

ABSTRACT

In this paper, we advance the state-of-the-art in designing ex-post truthful multi-armed bandit (MAB) mechanisms for multi-slot sponsored search auctions (SSA) through two different contributions. First, we prove two important impossibility results which rule out the possibility of an ex-post monotone MAB allocation rule having sublinear regret with time when the click through rates (CTR) of the advertisements (ads) are affected by *ad-dependent* externality or *position-dependent* externality. The above impossibility results motivate our second contribution: when the CTRs are affected by only *position-dependent* externality and follow *click-precedence* property, we design a novel ex-post truthful mechanism for multi-slot SSAs with sublinear regret. The ex-post monotone allocation rule in the proposed mechanism non-trivially generalizes the NewCB allocation rule presented by Babaioff, Sharma, and Slivkins[2]. We derive regret bounds for this allocation rule. When a strong property such as ex-post truthfulness is required, our allocation rule performs as well as the A-VCG mechanism presented by Gatti, Lazaric, and Trovò[4] and in the special case of identical slots, our allocation rule in fact outperforms the A-VCG mechanism and has a regret of $O\left(\sqrt{T}\right)$ with time.

Categories and Subject Descriptors

500 [**Information systems**]: Sponsored search advertising

General Terms

Economics, Algorithms

Keywords

Mechanism Design; Sponsored Search; Multi-Armed Bandit

1. INTRODUCTION

Sponsored search auction (SSA) provides an environment where a mechanism design problem is inherently coupled with a learning problem. As a first step, the auctioneer tries to elicit the true valuations of the advertisers to design a truthful search auction. Besides, the auctioneer has to determine each ad's *click-through-rate* (CTR), which is the probability that his ad will be clicked by a user. This is because the payment policy followed in SSA is the *pay-per-click* scheme which requires an advertiser to make a payment to the search engine only if a user clicks his ad. Since CTRs are not known, the search engine has to combine effective learning algorithms for the estimation of the CTRs with incentive compatible mechanisms. This introduces the classical *explore-exploit* situation. The search engine tries to explore different ads to get better estimates of the CTRs and may lose some revenue in this process. However, this additional information can help the search engine find a better allocation of the ads in future. *Multi-Armed Bandit* (MAB) problems effectively characterize such explore-exploit situations faced in sequential decision problems. The performance of any MAB algorithm is measured in terms of regret which is the expected loss of reward when compared to the optimal algorithm which pulls the best arms every time.

The single play MAB (SPMAB) problem, where the decision maker pulls one arm every round, is a well studied problem. The UCB1 (Upper Confidence Bound) algorithm proposed by Auer, Cesa-Bianchi, and Fischer [1] has distribution independent regret of $O\left(\sqrt{nT\log T}\right)$ for this problem. However, UCB1 algorithm cannot be applied directly to design a regret minimizing SPMAB mechanism because the advertisers are strategic and may misreport their valuations. In fact, Babaioff, Sharma, and Slivkins [3] have shown that any ex-post truthful and normalized MAB mechanism must suffer a regret of at least $\Omega\left(n^{1/3}T^{2/3}\right)$. This increased dependence by a factor of $T^{1/6}$ on time horizon T is termed as the *price of truthfulness*. The multiple play MAB problems (MPMAB) more realistically capture the problem of SSA where multiple slots are up for auction. Gatti, Lazaric, and Trovò [4] have come up with an MPMAB mechanism for multi-slot SSA. They have extended the notions of click probabilities from single slot to the multiple slot case and considered two types of externalities. The *position-dependent* externality captures the fact that the click probability of an ad i depends on its position in the allocation. On the other hand, *ad-dependent* externality refers to the fact that this probability also depends on the ads displayed before ad i in the allocation vector. When k ads can be displayed every time, the A-VCG mechanism proposed in [4] has a regret of $O\left(n^{\frac{1}{3}}k^{\frac{2}{3}}T^{\frac{2}{3}}\right)$ and $O\left(nk^{\frac{2}{3}}T^{\frac{2}{3}}\right)$ respectively in the presence of position-dependent externality and ad-dependent externality.

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

2. RESULTS

In this paper, we advance the state-of-the-art in designing ex-post truthful multi-armed bandit (MAB) mechanisms for multi-slot sponsored search auctions. As a first step, we take up the challenge of designing ex-post monotone allocation rule for the same. An allocation rule is ex-post monotone if it is monotone for every possible click realization. On the other hand, stochastic monotonicity requires monotonicity in expectation over the clicks. Our work is principally motivated by the transformation procedure introduced by Babaioff, Kleinberg and Slivkins [2] which takes as input any ex-post (stochastic) monotone allocation rule, computes payment implicitly and outputs an ex-post (stochastically) truthful mechanism having same regret as the input allocation rule. Note that ex-post monotonicity is a stricter notion than stochastic monotonicity. Moreover, it obviates any modeling of agents' interactions with nature's randomness like random click events. Our contributions in designing ex-post truthful MAB mechanisms for multi-slot SSA are twofold :

Impossibility of Sublinear Regret for Ex-Post Monotone Allocations

First, we prove that it is impossible to design an ex-post monotone MAB allocation rule having sublinear regret with time when the CTRs of the ads are affected by ad-dependent externality. The proof depends on the following observation from the definition of ex-post monotone allocation rule. Suppose, A is an ex-post monotone allocation rule in the presence of ad-dependent externality on the CTRs. Suppose the allocation at time t, A_t allocates ad i at the k-th slot. Now if ad i increases his bid, then to maintain ex-post monotonicity, not only should he get the k th slot at time t, but also the order of the ads allocated before him must not change. In other words, A_t remains unchanged if ad i increases his bid.

There are two routes to get around this impossibility result – (1) design allocation rules which are stochastically monotone or (2) assume externalities which are weaker than ad-dependent externality. We take the second route and assume that the CTRs are affected by only position-dependent externality. We show that even if we relax the ad-dependent externality to position-dependent externality, then also it is impossible to design an ex-post monotone allocation rule having sublinear regret with time.

Ex-Post Truthful Mechanism for Multi-Slot SSA

Motivated by the above impossibility results, we consider the case where the clicks follow a special property called *click-precedence* property, which states that if an ad is clicked at time t when displayed at slot j, then it will also be clicked at time t if displayed at slots higher than j. Under this reasonable assumption, we design an ex-post monotone MAB allocation rule, which is a nontrivial generalization of the NewCB allocation rule presented in [2].

Our allocation rule maintains k lower confidence bounds $\{L_i^j\}_{j=1}^k$ and k upper confidence bounds $\{U_i^j\}_{j=1}^k$ for each ad i. L_i^j and U_i^j are respectively the lower and upper confidence bounds of the click rewards of ad i obtained from slot j. Additionally, it maintains k different activation sets S_{act}^j for $j = 1, \ldots, k$. S_{act}^j is the set of ads that can be assigned to slot j. Ideally, we would like the activation sets to be nested, since any ad which can be assigned slot j, can also be assigned to slots $j+1, \ldots, k$. This will not always be true. Only when sufficient number of samples of the click realization table have been observed for all the ads, we will have all the activation sets nested.

In every round our allocation rule decides in two steps which agents are to be allocated to k slots. At first, it designates k different agents for k slots in a specific way. In the second step, it checks whether the *designated* agents are present in the corresponding activation sets. If j-th designated agent is present in S_{act}^j, then that agent is allocated slot j. The ads for the remaining empty slots are selected from the k-th activation set S_{act}^k.

We prove that our allocation rule is ex-post monotone when the click realization follows the click-precedence property. The proof depends on the following property of our allocation rule – for each ad i, L_i^j and U_i^j are updated in predetermined rounds. Therefore, once the click realization is fixed, the sequence of click rewards used to update L_i^j and U_i^j are actually the same. Next, we derive regret bounds for this allocation rule and under the special case of subset selection (that is slots are identical), we prove that our allocation rule has logarithmic distribution-dependent regret and distribution-independent regret of $O\left(k\sqrt{nT\log T}\right)$. When the number of slots k is 1, the regret of our allocation rule precisely coincides with the regret of the NewCB algorithm. However, it is possible to design an ex-post monotone allocation rule with sublinear regret when clicks follow click-precedence property and the slots are non-identical. In fact, we prove that the allocation rule of the A-VCG mechanism presented in [4] is indeed ex-post monotone in this setting. However, in the special case of subset selection, our algorithm outperforms the A-VCG mechanism.

3. FUTURE WORK

In this paper, we have looked into the design of ex-post monotone MAB allocation rule for multi-slot SSA. One major direction for future work is to design *stochastically monotone* MAB allocation rule in the presence of ad-dependent externality. This problem is complicated by the fact that the problem of finding the best allocation of agents is not only intractable but $\mathcal{APX}$-hard.

4. REFERENCES

[1] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-Time Analysis of the Multi-Armed Bandit Problem. *Machine learning*, 47(2):235–256, 2002.

[2] M. Babaioff, R. Kleinberg, and A. Slivkins. Truthful Mechanisms With Implicit Payment Computation. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC-10)*, pages 43–52, 2010.

[3] M. Babaioff, Y. Sharma, and A. Slivkins. Characterizing Truthful Multi-Armed Bandit Mechanisms. In *Proceedings of the 10th ACM Conference on Electronic Commerce (EC-09)*, pages 79–88, 2009.

[4] N. Gatti, A. Lazaric, and F. Trovò. A Truthful Learning Mechanism for Contextual Multi-Slot Sponsored Search Auctions With Externalities. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC-12)*, pages 605–622, 2012.