Meta-learning of Bidding Agent with Knowledge Gradient in a Fully Agent-based Sponsored Search Auction Simulator

Extended Abstract

Donghun Lee
Department of Computer Science
Princeton University
Princeton, New Jersey
d.lee@princeton.edu

Warren B. Powell
Department of Operations Research and Financial Engineering
Princeton University
Princeton, New Jersey
powell@princeton.edu

ABSTRACT

We take a practical approach on learning how to bid in sponsored search auctions, and model the problem of improving real world profit of advertisers in sponsored search auction as a meta-learning problem of configuring adaptive bidding agents. We construct a fully agent-based sponsored search auction simulator that 1) captures the dynamic nature of sponsored search auctions, 2) emulates the interface of Google AdWords platforms, and 3) can be customized and extended by modules. We then present Meta-LQKG algorithm, an agent-based meta-learning algorithm using knowledge gradient, and show the effect of meta-learning with Meta-LQKG on the performance of adaptive bidding agents.

KEYWORDS

Meta-learning; Learning to Bid; Sponsored Search Auction;, Agent-Based Simulation; Knowledge Gradient

ACM Reference Format:
Donghun Lee and Warren B. Powell. 2019. Meta-learning of Bidding Agent with Knowledge Gradient in a Fully Agent-based Sponsored Search Auction Simulator . In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 3 pages.

1 INTRODUCTION

In this paper, we illuminate the fundamental agent-based nature of sponsored search auctions and propose a meta-learning approach to improve the performance of learning agents facing competitors with real-time learning capacity in the auctions.

There are many adaptive approaches using reinforcement learning and artificial intelligence to find the optimal bidding from the advertiser's perspective [16] [17] [13]. However, when it comes to revenue maximization by learning from auctions, the revenue of the search engine is the usual target [10] [3] [5] [12] [14], and there are few works optimizing from the individual bidder's side in the sponsored search auction [2] [8].

Most research on the adaptive bidding algorithms for sponsored search auctions rely on the snapshots of auction history, despite the fact that adaptive bidding agents can participate in the actual sponsored search auctions. On the contrary, there are many real world auctions modeled using agent-based simulation, such as dynamic online auctions [11] [6], FCC spectrum auctions [4], the auctions for electricity [9] and its transmission rights [18].

Our contribution in this paper is to address the gaps in the current literature of learning in online advertisement auctions by:

(1) modeling the adaptive bidding problem from the persepctive of the bidding agent acting as a proxy for the advertiser,
(2) constructing a fully agent-based, modular simulator configured for sponsored search auctions, and
(3) designing an online meta-learning algorithm that can be applied to adaptive bidding algorithms.

2 ALGORITHM AND SIMULATOR

We model the meta-learning problem of finding the suitable hyperparameter $\rho \in \mathcal{P}$ in the learning algorithm for a bidding agent, as a stochastic optimization problem, by extending the language from the Markov decision process that models the underlying bidding problem in the sponsored search auction. The meta-learning problem structure assumes that the bidding agent have a budget of testing N different ρ's, where each ρ is tested for T time steps, as depicted in Figure 1.

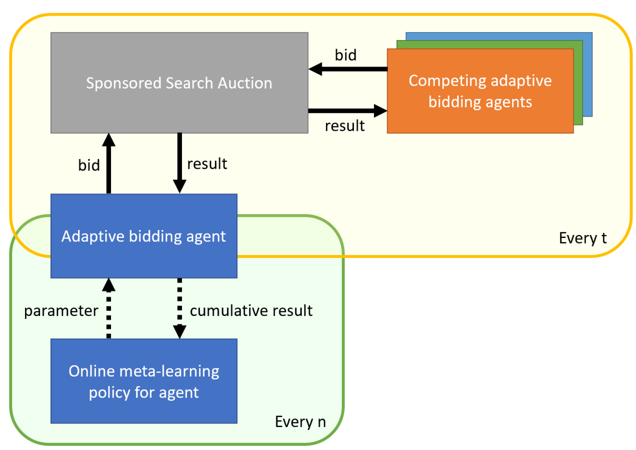
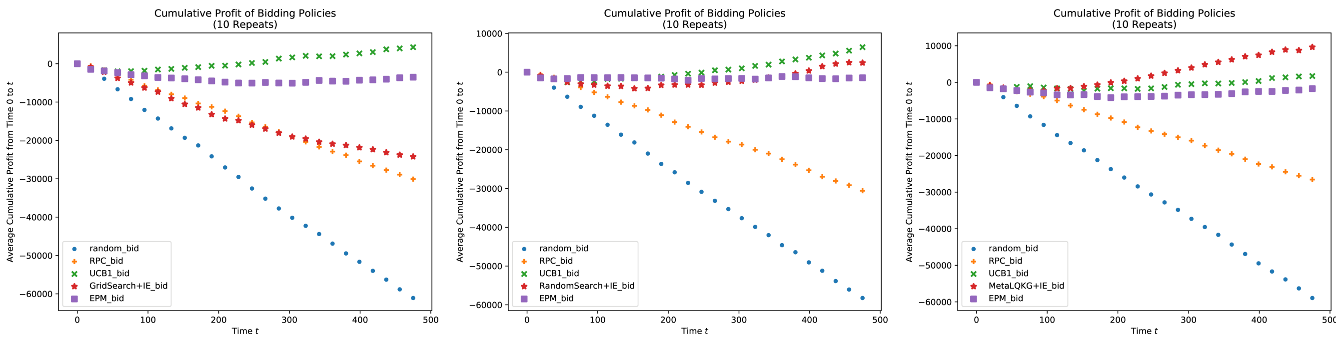


Figure 1: Diagram of meta-learning of a bidding agent in a fully agent-based sponsored search auction simulator.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.



(a) Grid Search **(b) Random Search** **(c) Meta-LQKG**

Figure 2: Plots of the average cumulative profit of all bidding policies in the simulated auctions ($T = 24, N = 20$).

The objective function of the agent-based meta-learning in learning how to bid is defined as:

$$\overbrace{\max_{\pi \in \Pi^{\rho}} \mathbb{E}\left[\sum_{n=0}^{N-1} \underbrace{\sum_{t=0}^{T-1} C\left(S_t^n, X^{\pi}\left(S_t^n \middle| \theta_t^n, \rho^n\right), W_{t+1}^n\right)}_{\text{Bidding problem for } T \text{ steps}} \middle| S_0^0\right]}^{\text{Meta-learning problem to find best } \rho \text{ in } N \text{ trials}}, \tag{1}$$

where during each iteration n the bidding agents bid and learn its parameter θ_t^n from $t = 0$ to $T-1$, and then ρ^{n+1} is determined. The optimization goal is to maximize the expectation of the sum of the advertising profit from every time step t, where $C\left(S_t^n, X^{\pi}\left(S_t^n \middle| \theta_t^n, \rho^n\right), W_{t+1}^n\right)$ is the observed profit from following the bidding policy X^{π} given θ_t^n and ρ^n in time t. S_t^n stands for the state variable at time t in iteration n, which corresponds to $nT + t$ time steps from the initial state S_0^0. We set $S_0^{n+1} = S_T^n$ to model online learning.

We construct Meta-LQKG, a meta-learning algorithm to solve problem (1) by adapting locally quadratic knowledge gradient algorithm [1] to general online learning problems [15]. The details of the Meta-LQKG algorithm is shown in Algorithm 1, where the belief state B^n at iteration n contains the multivariable Normal distribution parameters that model the three quadratic coefficients modeling the "bidding problem" portion of (1).

The fully agent-based simulator of sponsored search auction is designed to interact with a group of bidding agent modules and other modules corresponding to different elements of sponsored search auctions. The Python code and the documentation of the simulator are available in a Github repository [7].

3 RESULTS

We use the finite horizon cumulative profits to compare the practical benefit of using different methods of searching hyperparameter for adaptive bidding policy. We compare the performance of the Meta-LQKG algorithm against two most frequently used methods in practice: grid search and random search. Against a pool of adaptive bidding competitors, we test two types of tunable policies: interval estimation bidding (IE-bid) policy and UCB bidding (UCBtuned-bid) policy, and due to space limitation we present the result from IE-bid.

Algorithm 1 Meta-LQKG: Meta Learning with Locally Quadratic Knowledge Gradient

Require: $\mathcal{P}, N, T, S_0^0, B^0, W, \rho_0, \kappa$
  Randomly select $\rho_0 \in \mathcal{P}$, if ρ_0 not provided.
  for $n = 0, \cdots, N-1$ **do**
    for $t = 0, \cdots, T-1$ **do**
      Let $\hat{C}_{t+1}^n = C\left(S_t^n, X_t^{\pi}\left(S_t^n \middle| \theta_t^n, \rho^n\right), W_{t+1}^n\right)$
      Update $\theta_{t+1}^n \leftarrow \Theta^{\pi}\left(\theta_t^n, W_{t+1}^n \middle| \rho^n\right)$
    end for
    Compute $\hat{g}^{n+1} = \sum_{t=0}^{T-1} \hat{C}_{t+1}^n$
    Let $\kappa = \max \operatorname{diag}(\Sigma^n)$ if κ not provided
    $B^{n+1} \leftarrow$ LQKG-Update $(B^n, \rho^n, \hat{g}^{n+1}, \kappa)$
    $\nu_n^{LQKG}(\rho) \leftarrow$ LQKG-Compute $(B^{n+1}, \kappa, \mathcal{P})$
    $\rho^{n+1} = \arg\max_{\rho \in \mathcal{P}} \nu_{n+1}^{LQKG}(\rho)$
  end for
  $\rho^{*,N} \leftarrow$ LQKG-Select-Best $\left(B^N, \mathcal{P}\right)$

For IE-bid, we set $\mathcal{P}$ as $[-0.3, 0.3]$ separated into 50 elements, as the bid function of IE-bid policy allows both positive and negative values. We repeat 10 independent runs of experiment with different random seeds, and report the evolution trajectory of the sample mean of the cumulative profits averaged over 10 runs in Figure 2. The result suggests that for the IE-bid policy the Meta-LQKG algorithm improves the practical performance of IE-bid policy as shown in Figure 2c compared to Figures 2a and 2b.

4 CONCLUSION

From a practical perspective on deploying adaptive bidding algorithms in sponsored search auction, we model a meta-learning problem of optimally configuring the adaptive bidding agent with the goal of improving the expected finite horizon cumulative profit of the agent. We construct a fully agent-based simulator with customizable modules and present the Meta-LQKG algorithm to tune the bidding algorithms. We demonstrate the potential advantage of using meta-learning approach to tune bidding algorithms to gain comparative advantage over grid search and random search.

REFERENCES

[1] Nana Kwabena Aboagye. 2018. *Knowledge Gradient for Expensive Locally Quadratic Functions and Stochastic Optimization of Aid Allocation.* Ph.D. Dissertation. Princeton University. https://search.proquest.com/docview/2128016297?accountid=188395

[2] Deepak Agarwal, Souvik Ghosh, Kai Wei, and Siyu You. 2014. Budget pacing for targeted online advertisements at LinkedIn. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '14*. https://doi.org/10.1145/2623330.2623366

[3] Kareem Amin, Afshin Rostamizadeh, and Umar Syed. 2013. Learning Prices for Repeated Auctions with Strategic Buyers. In *Advances in Neural Information Processing Systems 26*, C J C Burges, L Bottou, M Welling, Z Ghahramani, and K Q Weinberger (Eds.). Curran Associates, Inc., 1169–1177. http://papers.nips.cc/paper/5010-learning-prices-for-repeated-auctions-with-strategic-buyers.pdf

[4] JA Csirik, ML Littman, Satinder Singh, and Peter Stone. 2001. FAucS: An FCC spectrum auction simulator for autonomous bidding agents. In *Electronic Commerce.*

[5] Yash Kanoria and Hamid Nazerzadeh. 2014. Dynamic Reserve Prices for Repeated Auctions: Learning from Bids. (2014). https://doi.org/10.2139/ssrn.2444495

[6] A Lee, M Lyu, and I King. 2001. An agent-based platform for online auctions. In *IC'2001: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON INTERNET COMPUTING, VOLS I AND II.*

[7] Donghun Lee. 2018. Fully Agent-based Sponsored Search Auction Simulator. (2018). https://github.com/donghun2018/adclick-simulator

[8] Donghun Lee, Piotr Zioło, Weidong Han, and Warren B. Powell. 2018. Optimal online learning in bidding for sponsored search auctions. In *2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings.* https://doi.org/10.1109/SSCI.2017.8285393

[9] Zhen Liu, Jianming Yan, Yuren Shi, Kaiwei Zhu, and Gangqing Pu. 2012. Multi-agent based experimental analysis on bidding mechanism in electricity auction markets. *International Journal of Electrical Power and Energy Systems* (2012). https://doi.org/10.1016/j.ijepes.2012.05.056

[10] Brendan Lucier, Renato Paes Leme, and Eva Tardos. 2012. On revenue in the generalized second price auction. In *Proceedings of the 21st international conference on World Wide Web - WWW '12.* https://doi.org/10.1016/S0025-7125(16)31737-0

[11] Hideyuki Mizuta and Ken Steiglitz. 2000. Agent-based simulation of dynamic online auctions. *Winter Simulation Conference Proceedings* (2000). https://doi.org/10.1109/WSC.2000.899168

[12] Mehryar Mohri and Andres Munoz. 2015. Revenue Optimization against Strategic Buyers. In *Advances in Neural Information Processing Systems 28.*

[13] Kan Ren, Weinan Zhang, Ke Chang, Yifei Rong, Yong Yu, and Jun Wang. 2018. Bidding Machine: Learning to Bid for Directly Optimizing Profits in Display Advertising. *IEEE Transactions on Knowledge and Data Engineering* (2018). https://doi.org/10.1109/TKDE.2017.2775228 arXiv:1803.02194

[14] Jiang Rong, Tao Qin, Bo An, and Tie-Yan Liu. 2017. Revenue Maximization for Finitely Repeated Ad Auctions. (2017). https://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14871

[15] Ilya O. Ryzhov, Warren B. Powell, and Peter I. Frazier. 2012. The Knowledge Gradient Algorithm for a General Class of Online Learning Problems. *Operations Research* (2012). https://doi.org/10.1287/opre.1110.0999

[16] Jie Zhang, Yanwu Yang, Xin Li, Rui Qin, and Daniel Zeng. 2014. Dynamic dual adjustment of daily budgets and bids in sponsored search auctions. *Decision Support Systems* (2014). https://doi.org/10.1016/j.dss.2013.08.004

[17] Weinan Zhang, Yifei Rong, Jun Wang, Tianchi Zhu, and Xiaofan Wang. 2016. Feedback Control of Real-Time Display Advertising. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining - WSDM '16.* https://doi.org/10.1145/2835776.2835843 arXiv:1603.01055

[18] N. P. Ziogos and A. C. Tellidou. 2011. An agent-based FTR auction simulator. (2011). https://doi.org/10.1016/j.epsr.2011.03.004