ICML | 2022

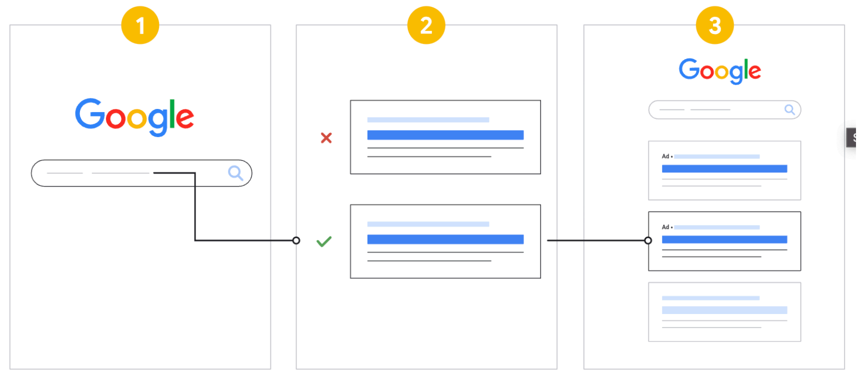
No-Regret Learning in Partially-Informed Auctions

Wenshuo Guo[1], Michael I. Jordan[1], Ellen Vitercik[2]

[1]University of California, Berkeley
[2]Stanford University

Learning in auctions: buyer’s perspective

- Under the canonical mechanism design model, buyers choose whether or not to buy items for sale based on their true values for those items
    - —> Assumes that the buyers know their values



- However, in practice..
    - **Information asymmetry:** sellers might hide information about the item for sell
      [Gershkov, ’09]..
    - **Privacy:** items may contain sensitive information, e.g. user queries are the items in ads auctions
      [Juels, ’01], [Guha et al, ’11], [Epasto et al, ’21]..

How should a buyer determine their purchase strategy with only incomplete item information?

Model: auction with partial item information

- One seller and one buyer, interact over T rounds
- Items are drawn from a distribution $\mathscr{P}$ over an abstract set $\mathscr{X} \subseteq \mathbb{R}^d$
- Buyer has an unknown value $v^*(x) \in [0, H]$ for each item $x \in \mathscr{X}$
- At each round
  - Item x_t is sampled from $\mathscr{P}$
  - Seller publishes **"masked" information** $h(x_t)$, and a price p_t
  - Buyer decides buy or not buy: decision $b_t = s_t(h(x_t), p_t) \in \{0,1\}$
  - Buyer obtains a utility $u_t = (v^*(x_t) - p_t) \cdot b_t$
    - If there is a purchase ($b_t = 1$): buyer observes x_t

Example: advertising auction

How should the buyer selects their strategy at each round to maximize utility?

- Seller: the platform
- Buyer: an advertiser
- Item x_t describes a user visiting the platform: features that uniquely identify each user

- On each round t:
  - The advertiser has value $v^*(x_t)$ for showing the user an ad
  - To protect user privacy, the platform does not reveal x_t to the advertiser, but some summary $h(x_t)$
    - E.g. $h(\cdot)$ can be a SimHash function mapping features from $\mathbb{R}^d$ to $\mathbb{R}^p$ $(p < d)$ [Epasto et al, '21]

Buyer's regret

- Regret w.r.t. an oracle myopic buyer s^*, who has a perfect knowledge of:
  - the item distribution $\mathscr{P}$
  - the masking function h

[Definition 2.1] The buyer's (expected) regret w.r.t. the optimal strategy s^* is:

$$R_T = \mathbb{E}\left[\sum_{t=1}^{T}\left(v^*(x_t) - p_t\right)s^*\left(h(x_t), p_t\right) - \left(v^*(x_t) - p_t\right)s_t\left(h(x_t), p_t\right)\right].$$

SimHash masking functions

- Masking function $h : [0,1]^d \to \{0,1\}^\ell$
- $h_w(x) = (sgn(w_1 \cdot x), sgn(w_2 \cdot x) \dots sgn(w_\ell \cdot x))$
- Buyer receives $h_w(x_t)$ and p_t in each round
- Items' distribution is known, prices are adversarial

[Theorem 3.5] (SimHash) With probability at least $1 - \delta$, the regret of Algorithm 1 is $R_T = \mathcal{O}\left(\sqrt{Td\ell \log(T\ell/\delta)}\right)$.

Algorithm overview

- Recall that s^* maximizes the expected utility at every round:

[Proposition 2.3] The strategy s^* that maximizes the above expected utility is:

$$s^*(h(x), p, h, \mathscr{P}) = 1\left(\mathbb{E}_{x\sim\mathscr{P}}\left[v^*(x) \mid h(x)\right] > p\right).$$

  - —> Observation: the optimal s^* is **a fixed thresholding rule** given $h(x) \in \{1,2,...n\}$

- Algorithm: explore-then-commit (ETC)
  - Exploration: learn the expected value for each group
  - Exploitation: use the learnt threshold to make purchase decisions

General strategy with stochastic prices

- A general masking function $h : \mathcal{X} \to [n]$
- Prices p_t are drawn from some unknown fixed distribution

[Theorem 4.3] (general h) There exists an algorithm (Algorithm 2) that achieves a regret rate with probability at least $1 - \delta$ that is:

$$R_T = \mathcal{O}\left(\sqrt{T(n \log T + \log(\frac{1}{\delta}))}\right).$$

[Theorem 4.4] (computational complexity) Algorithm 2 can be computed in polynomial time, with a per-round complexity that is $\tilde{\mathcal{O}}(n + \sqrt{T})$.

Further results

Item distribution	**Prices**	**Masking function h**	**Regret**
Known	Adversarial	SimHash $h : [0,1]^d \to \{0,1\}^\ell$	$\mathcal{O}(\sqrt{Td\ell \log(T\ell/\delta)})$ (Theorem 3.5)
Unknown	Stochastic	Arbitrary $h : \mathcal{X} \to [n]$	$\mathcal{O}(\sqrt{T(n \log T/n + \log 1/\delta)})$ (Theorem 4.3)
Unknown	Adversarial	Arbitrary $h : \mathcal{X} \to [n]$	$\mathcal{O}(T^{2/3} n^{1/3})$ (Remark 4.5)

Table 1. Summary of regret bounds which hold with probability at least $1 - \delta$.

- Known item distribution + SimHash masking function: exponential improvement
- Stochastic prices: Exp4. VC —adaptive algorithm
- Adversarial prices: ETC (explore-then-commit)

Thank you!