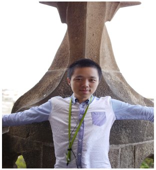
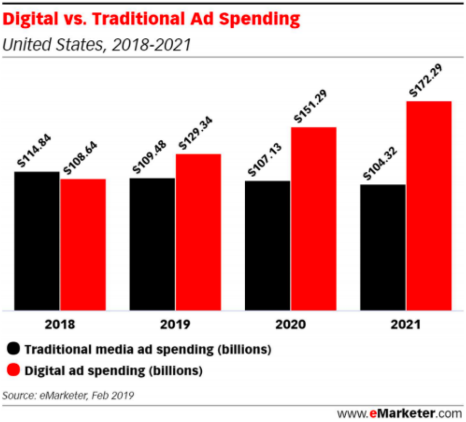
Leveraging The Hints: Adaptive Bidding in Repeated First-Price Auctions


Wei Zhang (MIT), Yanjun Han (MIT), Zhengyuan Zhou (NYU),
Aaron Flores (Yahoo! Research), Tsachy Weissman (Stanford)

Background

- Success of digital ads

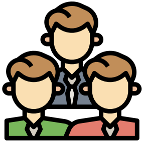


- Ads exchange format: first-price auctions

Bidder's sequential decision model

private source

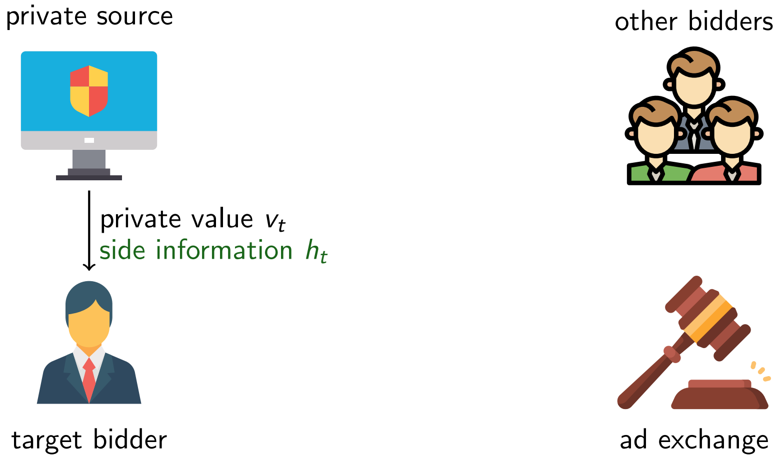
other bidders



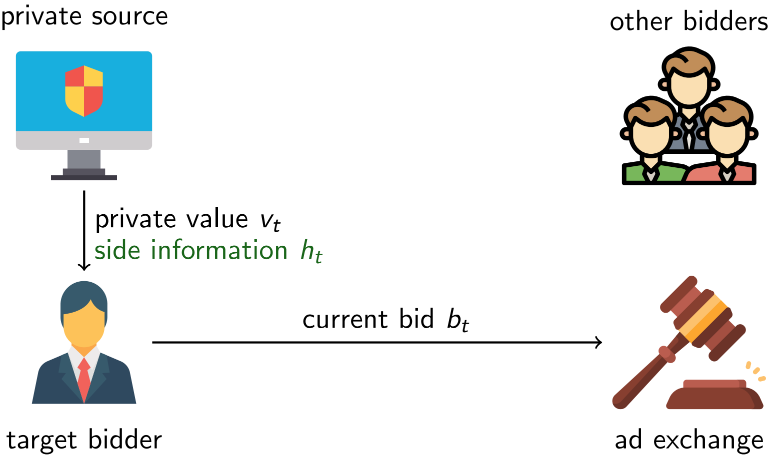
target bidder

ad exchange

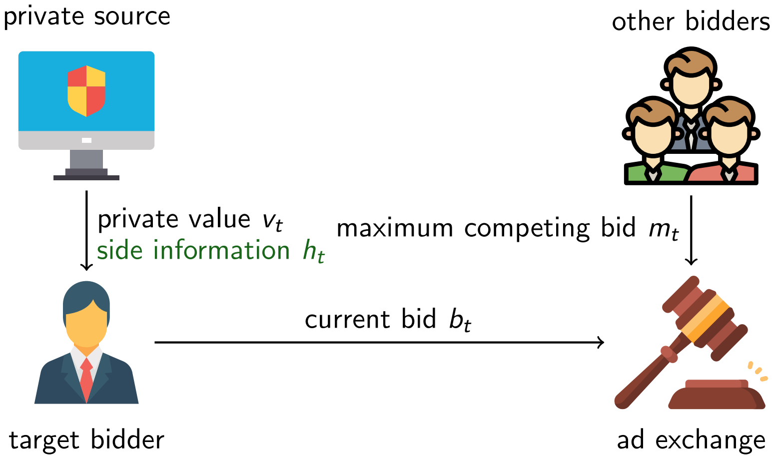
Bidder's sequential decision model



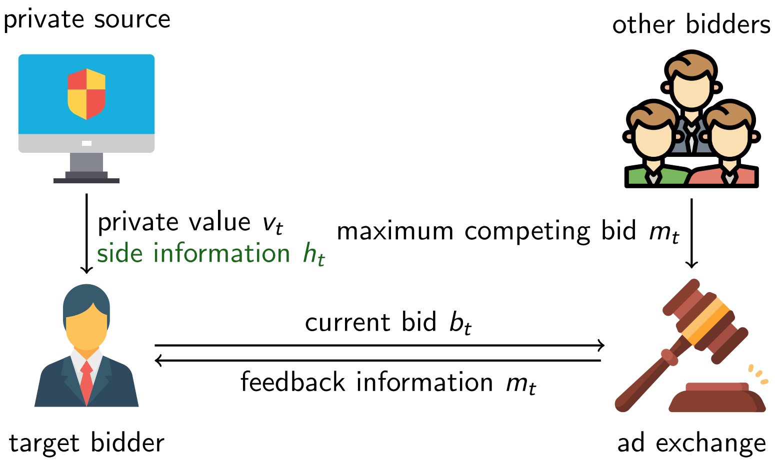
Bidder's sequential decision model



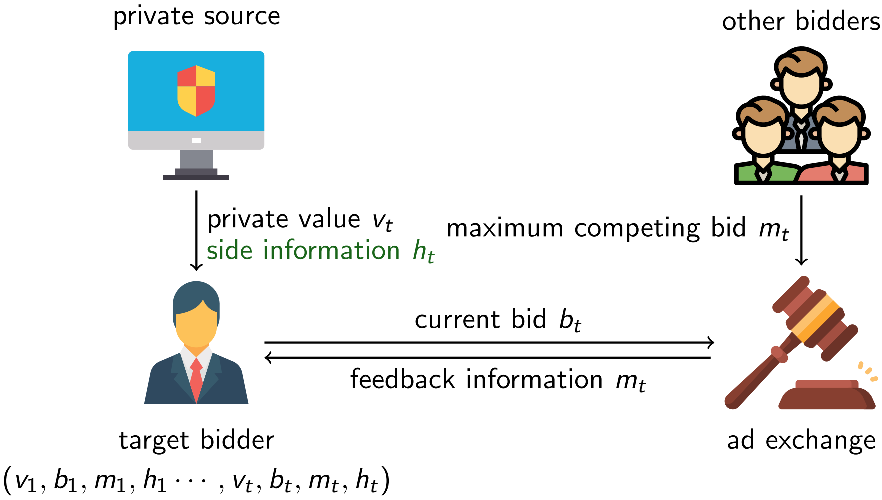
Bidder's sequential decision model



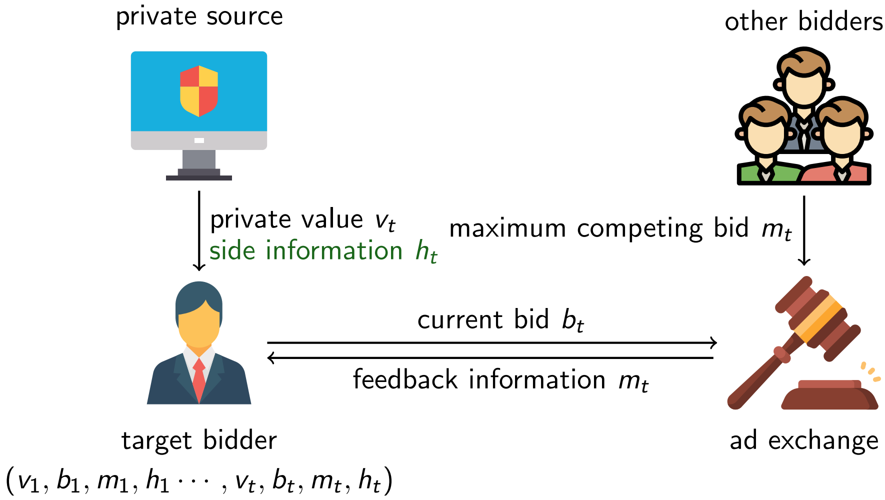
Bidder's sequential decision model



Bidder's sequential decision model



Bidder's sequential decision model



Instantaneous reward: $r(b_t; v_t, m_t) = (v_t - b_t) \cdot \mathbb{1}(b_t \geq m_t)$

Problem formulation

- Private value & others' highest bid: $v_t, m_t \in [0, 1]$

adversarially chosen

Problem formulation

- Private value & others' highest bid: $v_t, m_t \in [0,1]$

  adversarially chosen

- Side information (hint): (h_t, σ_t)

$$\mathbb{E}\left[|h_t - m_t|^q\right] \le \sigma_t^q, \quad L := \sum_{t=1}^{T} \sigma_t$$

  - Hint interval: (h_t, σ_t)
  - Point estimation: h_t

Problem formulation

- Private value & others' highest bid: $v_t, m_t \in [0, 1]$

  adversarially chosen

- Side information (hint): (h_t, σ_t)

$$\mathbb{E}\left[|h_t - m_t|^q\right] \le \sigma_t^q, \quad L := \sum_{t=1}^{T} \sigma_t$$

  - Hint interval: (h_t, σ_t)
  - Point estimation: h_t

Bidder's Target: Regret

$$\mathrm{Reg}(\pi) \triangleq \underbrace{\max_{f \in \mathcal{F}} \mathbb{E}\left[\sum_{t=1}^{T} r(f(v_t); v_t, m_t)\right]}_{\text{oracle's reward}} - \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} r(b_t; v_t, m_t)\right]}_{\text{bidder's reward}}$$

Help from the hint

Theorem

For $L \in [1, T]$, $q \in [1, \infty)$,

- *if $v_t \equiv 1$ and the bidder observes a hint interval at each time t,*

$$\inf_{\pi} \sup_{\{m_t, h_t, \sigma_t\}} \mathrm{Reg}(\pi) = \tilde{\Theta}\left(\sqrt{T^{\frac{1}{q+1}} \cdot L^{\frac{q}{q+1}}}\right).$$

- *If $v_t \equiv 1$ and the bidder observes a point estimation at each time t, then for every $q \in [1, \infty)$,*

$$\inf_{\pi} \sup_{\{m_t, h_t\}} \mathrm{Reg}(\pi) = \tilde{\Theta}\left((T \cdot L)^{\frac{1}{4}}\right).$$

- Smaller regret with good hints (L is small)
- Regret separation on two types of hint

Algorithm for Upper Bound

Multiplicative Weights Update:

$$p_{t,a} = \frac{\exp\left(\eta_t \cdot \sum_{s<t} r_{s,a}\right)}{\sum_{a' \in [K]} \exp\left(\eta_t \cdot \sum_{s<t} r_{s,a'}\right)}, \quad a \in [K], \quad t = 1, \ldots, T,$$

Algorithm for Upper Bound

Multiplicative Weights Update:

$$p_{t,a} = \frac{\exp\left(\eta_t \cdot \sum_{s<t} r_{s,a}\right)}{\sum_{a' \in [K]} \exp\left(\eta_t \cdot \sum_{s<t} r_{s,a'}\right)}, \quad a \in [K], \quad t = 1, \dots, T,$$

Incorporate hints into rewards $r_{s,a}$? No.

Add one extra expert who bids $h_t + \sigma_t^{\frac{q}{q+1}}, t = 1, \dots, T$ to the pool

Large regret with varying private values

Theorem

Let $L \in [1, T]$, $q \in [1, \infty)$, and $v_t \in [0,1]$ for $t = 1, 2, \ldots, T$. If the bidder observes hint intervals

$$\inf_{\pi} \sup_{\{v_t, m_t, h_t, \sigma_t\}} \mathrm{Reg}(\pi) = \widetilde{\Theta}\left(\min\left\{T^{\frac{1}{q+1}} L^{\frac{q}{q+1}}, \sqrt{T}\right\}\right),$$

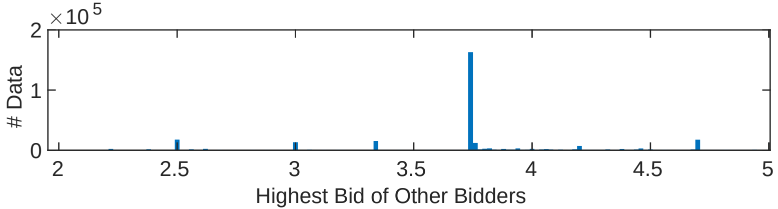
Theorem

Let $L \in [1, T]$, and $v_t \in [0,1]$ for $t = 1, 2, \ldots, T$. If the bidder observes point estimation, then $\forall q \in [1, \infty)$

$$\inf_{\pi} \sup_{\{v_t, m_t, h_t, \sigma_t\}} \mathrm{Reg}(\pi) = \widetilde{\Theta}\left(\sqrt{T}\right),$$

Same as bidding hint itself!

Improved regret with finite support of m_t



Assumption

The support of competing bid is at most K.

Improved regret with finite support of m_t

Theorem

For $q \in [1, \infty)$ and varying private prices, suppose that the minimum-bid-to-win m_t only takes K support values.

- *With hint intervals*

$$\inf_{\pi} \sup_{\{v_t, m_t, h_t, \sigma_t\}} \mathrm{Reg}(\pi) = \tilde{\Theta}\left(\min\left\{ \sqrt{T^{\frac{1}{q+1}} L^{\frac{q}{q+1}} K}, T^{\frac{1}{q+1}} L^{\frac{q}{q+1}}, \sqrt{T} \right\} \right)$$

- *With point estimation*

$$\inf_{\pi} \sup_{\{v_t, m_t, h_t, \sigma_t\}} \mathrm{Reg}(\pi) = \widetilde{\Theta}\left(\min\left\{ \sqrt{T}, \sqrt{\sqrt{LT} \cdot K} \right\} \right)$$

Improved regret with finite support of m_t

Theorem

For $q \in [1, \infty)$ and varying private prices, suppose that the minimum-bid-to-win m_t only takes K support values.

- *With hint intervals*

$$\inf_{\pi} \sup_{\{v_t, m_t, h_t, \sigma_t\}} \mathrm{Reg}(\pi) = \tilde{\Theta}\left(\min\left\{ \sqrt{T^{\frac{1}{q+1}} L^{\frac{q}{q+1}} K}, T^{\frac{1}{q+1}} L^{\frac{q}{q+1}}, \sqrt{T} \right\} \right)$$

- *With point estimation*

$$\inf_{\pi} \sup_{\{v_t, m_t, h_t, \sigma_t\}} \mathrm{Reg}(\pi) = \widetilde{\Theta}\left(\min\left\{ \sqrt{T}, \sqrt{\sqrt{LT} \cdot K} \right\} \right)$$

Adaptive to K and L

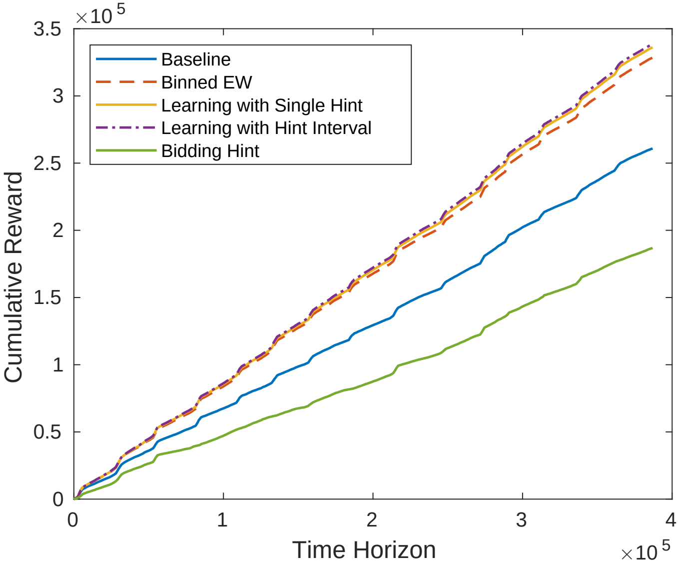
Meta-Algorithm

MWU with hint

Hint itself

OPT ALG w/o hint

Experiments on real dataset



Learning with hint leads to larger cumulative reward.