Learning to Bid Without Knowing your Value
Zhe Feng, Chara Podimata and Vasilis Syrgkanis
Harvard University, Harvard University, MSR New England

Motivation

Standard Assumption in Mechanism Design/ Game Theory:
Bidder knows her value for an item, before the auction

- Small markets, infrequent transactions
- Online ad-auctions

Online ad-auctions Flowchart

1. Bids
2a. Generates
2b. Auctioneer (e.g., Google AdWords, MS BingAds)
3. User Clicks Ad
Allocation Curve: \( x_t() \)
Payment Curve: \( p_t() \)

Value \( v_t \) generated

4. Bidder observes \( v_t \) (if clicked)

5. Bidder’s Utility: \( u_t(b_t) = (v_t - p_t(b_t)) \cdot x_t(b_t) \)

Allocation and Payment Curves in Real Life Ad Auctions

Google AdWords
Microsoft BingAds

Main Result

First Thought:
- Uniform \( \epsilon \)-Discretization of bidding space: \( B = [1/\epsilon] \).
- Apply off-the-shelf EXP3 for the discretized space.
- Regret: \( R(T) \leq O(\sqrt{T}) \).
- Careful! Utility not Lipschitz!

We can do better than that! Instead of \( R(T) \leq O(\sqrt{T|B|}) \), achieve regret of the form: \( R(T) \leq O(T \log |B|) \).

Intuition: Use extra information provided by allocation and payment curves in real life ad auctions!

WIN-EXP Algorithm (Win-Only Feedback)

- Distribution over bidding space: \( \pi_t(b) \in [1/\epsilon] \)
- At each round \( t \in [T] \):
  - Draw bid \( b_t \sim \pi_t \)
  - Observe allocation \( x_t(b) \), \( \forall b \in B \)
  - If you win (e.g., get clicked), observe value, compute reward: \( r_t(b) = v_t - p_t(b) \), \( \forall b \in B \) and estimate:
    \[ \hat{u}_t(b) = \frac{(r_t(b) - \pi_t(b))}{\pi_t(b)} \]
  - Else, estimate: \( \hat{u}_t(b) = \frac{1 - x_t(b)}{1 - x_t(b) \cdot \pi_t(b)} \), \( \forall b \in B \)
  - Update: \( \pi_{t+1}(b) = \pi_t(b) \exp(\eta \hat{u}_t(b)) \), \( \forall b \in B \)

Regret

Theorem. Regret of WINEXP algorithm with win-only feedback is: \( R(T) \leq O(T \log |B|) \).

Applications to Learning in Auctions

Auctions have win-only/outcome-based feedback structure!

- Value-Per-Click Auctions: \( R(T) \leq O(\sqrt{T \log |B|}) \).
- Multi-Slot Auctions & Unit Demand Auctions \( K \) items:
  \( R(T) \leq O(\sqrt{(K + 1)T \log |B|}) \).

Experiments

Extensions

- [Continuous Action Spaces] Proof that Discretization Error of \( \Delta^0 \)-Piecewise Lipschitz utilities is \( O \rightarrow \) WINEXP Algo Regret:
  \[ R(T) \leq O\left(\frac{1}{\Delta^0} \log \left(\max\left\{\frac{1}{\Delta^0}, T\right\}\right)\right) \]
- Applications: Sponsored Search, First-Pay, All-Pay auctions.

Open Questions

1) In what other real-life settings (apart from auctions) does win-only feedback appear?
2) Proof that WINEXP estimates are robust to noise up to a certain degree (experimental results suggest so).
3) Proof that WINEXP estimates can be reconstructed in some cases with fully-bandit feedback (experimental results suggest so too, in some cases of interest).

Main Question Addressed

How can the bidder learn how to bid without knowing their value a priori?

- Use No-Regret Learning!

Minimize: \( R(T) = \sup_{b^*} E[\sum_{t=1}^T u_t(b^*)] - E[\sum_{t=1}^T u_t(b_t)]\)