Regret Minimization for Reserve Prices in Second-Price Auctions

Claudio Gentile
Universita' dell'Insubria, Italy

SODA 2013
January 7th, 2013

Joint with:
N. Cesa-Bianchi, Y. Mansour
Motivation and goal

Surrounding context:

- Revenue maximization in Second Price auctions
- Widely used for selling ad slots
- Reserve price is main tool allowing seller to influence auction
- Optimizing reserve price in sponsored search has big impact on seller’s profit

This paper:

- **NOT** searching for optimal revenue maximization truthful mechanism, but maximize the seller’s revenue in SP auction with reserve price
- Seller has limited info about actual auction: e-Bay, AdSense, …
  Auctioneer ≠ seller: Seller only observes his/her revenue
- **No distributional assumptions** on bids **BUT** bids are i.i.d.
  (good proxy when seller conducts large no of auctions and bidders rarely participate in many of them)
Sequential setting

At $t$-th repetition of auction:

1. Seller sets a reserve price $p_t$

2. Bids $B_{t,1}, B_{t,2}, \ldots$ arrive [unobserved by seller]

3. Seller gets revenue $R_t = R_t(p_t)$ [observed]
Sequential setting

At $t$-th repetition of auction:

1. Seller sets a reserve price $p_t$
2. Bids $B_{t,1}, B_{t,2}, \ldots$ arrive [unobserved by seller]
3. Seller gets revenue $R_t = R_t(p_t)$ [observed]
Sequential setting

At $t$-th repetition of auction:

1. Seller sets a reserve price $p_t$
2. Bids $B_{t,1}, B_{t,2}, \ldots$ arrive [unobserved by seller]
3. Seller gets revenue $R_t = R_t(p_t)$ [observed]
Preliminaries and basic assumptions

- At $t$-th repetition, bids $B_{t,1}, B_{t,2}, \ldots B_{t,M}$ drawn i.i.d. from same fixed (but unknown and arbitrary and bounded range) distribution

- Seller only observes his/her revenue $R_t(p_t)$ (price at which item is sold, zero if item not sold)

- No. of bids $M$ not observed by seller: $M$ changes over time, but follows known prior distribution

- Seller’s policy is to set reserve prices $p_t$ based on past revenues $R_1(p_1), R_2(p_2) \ldots$

- Want to bound with high probability regret over $T$ repetitions

$$\max_{p \in [0,1]} \sum_{t=1}^{T} (\mu(p) - \mu(p_t)) \text{ want } = o(T)$$

$$\mu(p) = \text{expected revenue of price } p$$
Remarks and related work/1

• Identification of buyer’s utility given auction outcome well studied in Economics [AH02,HT03,...]

• When bid distrib. has monotone hazard rate, optimal reserve price in second-price auction is unique and independent of no. of bidders

• No such assumptions here (actually, no assumptions at all on bid distrib.): at each repetition best price depends on (time-varying) no. of bidders

• Seller does not observe no. of bidders $M \implies p^* = \operatorname{argmax}_{p \in [0,1]} \mu(p)$
  is w.r.t. (known) prior over no. of bidders $M$

• Other refs. related to optimizing reserve price: [D+07,W+08,JLB07...] (discrete bids, nonstationary behavior, hidden bids, ...)

• Item pricing (posted price auction) in stochastic setting with concave $\mu(\cdot)$ [KL03,B+12...]
Remarks and related work/2

- An instance of (one-dim) stochastic bandit optimization:
  - By playing price ("action") $p_t$ we observe noisy measurement of $\mu(p_t)$
  - A continuum of actions

- $\sqrt{T}$-like regret obtained under various assumptions on $\mu(\cdot)$ (Lipschitz, unimodality, well-behaved derivatives, bounded no. of maxima, ...) [AOS07, B+11, YM11,...]

- Is it possible to get $\sqrt{T}$-like regret with no assumptions on bid distribution involved in $\mu(\cdot)$?

  YES: Exploit structure of problem!
Exploiting structure of the problem

Main ideas:

• If reserve price $p_t$ is low seller likely to observe second-highest bid $B_t^{(2)}$

• Collect a sample of these order statistics and construct approximation to bid distribution, from which optimal price can be computed

• Yet, when reserve price too low seller incurs potentially high regret

• Tradeoff between exploration (set low price) and exploitation (set best price based on available data)
Idea of the algorithm

- Set $p_t = 0$ for $t = 1 \ldots T^{1/2}$ rounds, get $B_1^{(2)} \ldots B_{T^{1/2}}^{(2)}$

- Build approximate $\mu(\cdot)$ and confidence interval of size $\sqrt{\frac{1}{T^{1/2}} \log \frac{1}{\delta}}$ for $p^*$

- Set $p_t$ to lowest one in previous interval for next $T^{3/4}$ rounds and iterate $\log \log T$ times (next $T^{7/8}$, next $T^{15/16}$ ... )
Idea of the algorithm

- Set $p_t = 0$ for $t = 1 \ldots T^{1/2}$ rounds, get $B_1^{(2)} \ldots B_{T^{1/2}}^{(2)}$

- Build approximate $\mu(\cdot)$ and confidence interval of size $\sqrt{\frac{1}{T^{1/2}} \log \frac{1}{\delta}}$ for $p^*$

- Set $p_t$ to lowest one in previous interval for next $T^{3/4}$ rounds and iterate $\log \log T$ times (next $T^{7/8}$, next $T^{15/16} \ldots$)
Idea of the algorithm

• Set $p_t = 0$ for $t = 1 \ldots T^{1/2}$ rounds, get $B_1^{(2)} \ldots B_{T^{1/2}}^{(2)}$

• Build approximate $\mu(\cdot)$ and confidence interval of size $\sqrt{\frac{1}{T^{1/2}} \log \frac{1}{\delta}}$ for $p^*$

• set $p_t$ to lowest one in previous interval for next $T^{3/4}$ rounds and iterate $\log \log T$ times (next $T^{7/8}$, next $T^{15/16}$ ... )
Idea of the algorithm

- Set $p_t = 0$ for $t = 1 \ldots T^{1/2}$ rounds, get $B_1^{(2)} \ldots B_{T^{1/2}}^{(2)}$

- Build approximate $\mu(\cdot)$ and confidence interval of size $\sqrt{\frac{1}{T^{1/2}} \log \frac{1}{\delta}}$ for $p^*$

- set $p_t$ to lowest one in previous interval for next $T^{3/4}$ rounds and iterate $\log \log T$ times (next $T^{7/8}$, next $T^{15/16}$ ... )
• Set $p_t = 0$ for $t = 1 \ldots T^{1/2}$ rounds, get $B_1^{(2)} \ldots B_{T^{1/2}}^{(2)}$ \\

• Build approximate $\mu(\cdot)$ and confidence interval of size $\sqrt{\frac{1}{T^{1/2}} \log \frac{1}{\delta}}$ for $p^*$ \\

• Set $p_t$ to lowest one in previous interval for next $T^{3/4}$ rounds 
  and iterate $\log \log T$ times (next $T^{7/8}$, next $T^{15/16}$ ... )
Results/1: Upper bounds

If $\mu(p^*) > 0 \implies \text{w. p. } > 1 - \delta$:

- **Constant (and known) no. $M$** of bidders:
  
  \[
  \text{regret} = O\left(\frac{\sqrt{T \log(1/\delta)(\log \log T)}}{\mu(p^*)}\right)
  \]

- **Random $M$ case**:
  - Extra conditions on prior on $M$:
    
    \[
    \text{regret} = O\left(\frac{\sqrt{T \log(1/\delta)(\log \log T)}}{\mu^2(p^*)}\right)
    \]
  - General prior on $M$:
    
    \[
    \text{regret} = O\left(\frac{E[M]\sqrt{T \log(1/\delta)(\log \log T)}}{\mu^2(p^*)}\right)
    \]
Results/1: Upper bounds

If $\mu(p^*) > 0 \implies \text{w. p. } > 1 - \delta$:

- Constant (and known) no. $M$ of bidders:

  $$\text{regret} = O \left( \frac{\sqrt{T \log(1/\delta)(\log \log T)}}{\mu(p^*)} \right)$$

- Random $M$ case:
  - Extra conditions on prior on $M$:

    $$\text{regret} = O \left( \frac{\sqrt{T \log(1/\delta)(\log \log T')}}{\mu^2(p^*)} \right)$$

  - General prior on $M$:

    $$\text{regret} = O \left( \frac{E[M] \sqrt{T \log(1/\delta)(\log \log T)}}{\mu^2(p^*)} \right)$$
Results/1: Upper bounds

If $\mu(p^*) > 0 \implies \text{w. p. } > 1 - \delta$:

- Constant (and known) no. $M$ of bidders:
  
  $$\text{regret} = O\left(\frac{\sqrt{T \log(1/\delta)} (\log \log T)}{\mu(p^*)}\right)$$

- Random $M$ case:
  - Extra conditions on prior on $M$:
    
    $$\text{regret} = O\left(\frac{\sqrt{T \log(1/\delta)} (\log \log T)}{\mu^2(p^*)}\right)$$

  - General prior on $M$:
    
    $$\text{regret} = O\left(\frac{E[M] \sqrt{T \log(1/\delta)} (\log \log T)}{\mu^2(p^*)}\right)$$
Results/2: Sketch of regret analysis

\[ T \]

\[ T_1, T_2, T_3 \]

\( T_i = \text{length of } i\text{th stage} \)

Regret bound:
\[
T_1 \cdot 1 + \sum_{\text{stage } i} T_i \sqrt{\frac{1}{T_{i-1}}} \log \frac{1}{\delta}
\]

\( T_1 = \sqrt{T}, \ T_2 = T^{3/4}, \ T_3 = T^{7/8}, \ldots, T_i \sqrt{\frac{1}{T_{i-1}}} = \sqrt{T}, \ldots \)

Regret bound:
\[
\sqrt{T} + \sum_{\text{stage } i} \sqrt{T} \sqrt{\log \frac{1}{\delta}}
\]

No. of stages = \( O(\log \log T) \)

Regret bound:
\[
O\left( \sqrt{T \log \frac{1}{\delta}} (\log \log T) \right)
\]
Results/2: Sketch of regret analysis

Regret bound: \( T_1 \cdot 1 + \sum_{\text{stage } i} T_i \sqrt{\frac{1}{T_{i-1}}} \log \frac{1}{\delta} \)

\( T_1 = \sqrt{T}, \ T_2 = T^{3/4}, \ T_3 = T^{7/8}, \ldots, T_i \sqrt{\frac{1}{T_{i-1}}} = \sqrt{T}, \ldots \)

Regret bound: \( \sqrt{T} + \sum_{\text{stage } i} \sqrt{T} \sqrt{\log \frac{1}{\delta}} \)

No. of stages = \( O(\log \log T) \)

Regret bound: \( O \left( \sqrt{T \log \frac{1}{\delta} (\log \log T)} \right) \)
Results/3: Remarks, Extensions, and Lower bounds

Efficiency/Simplicity:
Algo easily computes $p_t$ and confidence intervals as it operates on discretized version of $\mu(\cdot)$ (just sweep over available data in current stage)

Actual regret (stronger):
\[
\max_{p \in [0,1]} \sum_{t=1}^{T} (R_t(p) - R_t(p_t))
\]

Same results as previous slide (via uniform convergence)

Lower bound:
Even if $M = 2$ and seller observes actual bids $B_{t,1}, B_{t,2}$ at each round
\[
\text{regret} = \Omega(\sqrt{T}) \quad T \to \infty
\]
with constant prob. (same with actual regret)
Conclusions and work in progress

• $\sqrt{T}$ regret bounds for reserve price in SP auctions under reasonable conditions:
  – Only observe seller’s revenue (auctioneer $\neq$ seller)
  – i.i.d. assumption on bid distrib. (but no assumptions on it)
  – prior knowledge of distrib. of no. of bidders realistic when log data from auctioneer are available

• Algorithm trades off exploration/exploitation

Open & work in progress:

• Can inverse dependence on $\mu(p^*)$ be avoided in upper bounds? Not inherent to our problem but inherent to our alg.

• GSP auction with multiple items with different qualities