Regret Minimization against Strategic Buyers

Mehryar Mohri
Courant Institute & Google Research

Andrés Muñoz Medina
Google Research


6
?
5
2
4
2


3
2
2

Motivation

- Online advertisement:
  - revenue of modern search engine and popular online sites.
  - billions of transactions every day.
  - key role of revenue optimization algorithms.

Motivation

- Second-price auctions with reserve:
  - widely adopted mechanism in Ad Exchanges.
  - many transactions admit a single bidder → posted-price auctions.
  - study of posted-price auctions with strategic buyers.

Related Work

- Revenue optimization in second-price auctions [Cui et al. 2011; He et al., 2013; Cesa-Bianchi et al., 2013; MM and Muñoz 2014].
- Revenue optimization in generalized second-price auctions (GSP) [MM and Muñoz, 2015; Varian, 2007; Lucier et al., 2014; Sun et al, 2014; Rudolph et al. 2016; Charles et al., 2016; Roughgarden and Wang, 2016].
- Dynamic pricing [Kanoria and Nazerzadeh 2014, Bikhchandani and McCardle 2012; den Boer, 2015; Chen et al. 2015].
- Pricing with strategic and patient buyers [Feldman et al., 2016].
- Preference reconstruction [Blum et al. 2014].
- Learning optimal auctions [Huang et al., 2015; Morgenstern and Roughgarden, 2015].

This Talk

- Scenarios:
  - Fixed valuation.
  - Random valuation.

Setup

- Repeated posted-price auctions:
  - good repeatedly offered for sale by a seller to a single buyer over T rounds.
  - buyer holds private valuation $v \in [0, 1]$.
  - seller offers price p_t and buyer accepts, $a_t = 1$, or rejects, $a_t = 0$, at each round $t \in [T]$.

Setup

- Seller: pricing algorithm $\mathcal{A}$.
  - total revenue: $\sum_{t=1}^{T} a_t p_t$.
  - regret: $\mathrm{Reg}_T(\mathcal{A}) = vT - \sum_{t=1}^{T} a_t p_t$.
- Buyer: discounting factor $\gamma \in (0, 1]$.
  - surplus: $\mathrm{Sur}_T(\mathcal{A}) = \sum_{t=1}^{T} \gamma^{t-1} a_t (v - p_t)$.

Strategic Setting

[Amin et al., 2013]

- Seller announces his algorithm $\mathcal{A}$.
- Buyer acts strategically: he seeks to maximize his surplus $\mathrm{Sur}_T(\mathcal{A})$.
- Seller seeks to minimize his strategic regret, that is regret $\mathrm{Reg}_T(\mathcal{A})$ against strategic buyer.
- Question: can we design algorithms for minimizing strategic regret?

Truthful Setting

[Kleinberg and Leighton, 2003]

- Fast Search (FS) algorithm:
  - keeps track of feasible interval $[a, b]$, starting with $[0, 1]$ and parameter $\epsilon = \frac{1}{2}$.
  - in each phase, offers prices $a + \epsilon, a + 2\epsilon, \ldots$ until a price is rejected.
  - if price $a + k\epsilon$ is rejected, new phase with interval $[a + (k-1)\epsilon, a + k\epsilon]$ and parameter ϵ^2.
  - until size of the interval less than $\frac{1}{T}$.

Truthful Setting

[Kleinberg and Leighton, 2003]

- Fast Search (FS) algorithm:
  - ✦ at most $\lceil \log_2 \log_2 T) \rceil + 1$ phases.
  - ✦ regret in $O(\log \log T)$.
  - ✦ lower bound: $\Omega(\log \log T)$.

Example

Monotone algorithms

[Amin et al., 2013]

- Algorithm:
  - offer price $p_t = \beta^t$ $(\beta < 1)$ until it is accepted.
  - offer accepted price thereafter.
  - idea: slow enough decrease inconvenient for the buyer.
- Strategic regret in $O\left(\sqrt{\frac{T}{1-\gamma}}\right)$.

Monotone algorithms

[MM and Muñoz, 2014]

- Theorem: the strategic regret of any monotone decreasing convex algorithm is in $\Omega(\sqrt{T})$.

Proof idea

- Fix monotone function.
- Choose $v \in [\frac{1}{2}, 1]$ at random.
- Let $\kappa = \inf\{t \colon p_t < v\}$, then

$$\mathbb{E}[\kappa]\,\mathbb{E}[v - p_\kappa] \geq c.$$

- Tradeoff optimized for $p_t - p_{t+1} \sim \dfrac{1}{\sqrt{T}}$.

Lower Bound

[Amin et al. 2013; Kleinberg and Leighton, 2003; MM and Muñoz, 2014]

- Theorem: for any pricing algorithm $\mathcal{A}$, the following lower bound holds:

$$\mathrm{Reg}_T(\mathcal{A}) \geq \max\left(\frac{1}{12(1-\gamma)}, C \log\log T\right)$$

for some universal constant C.

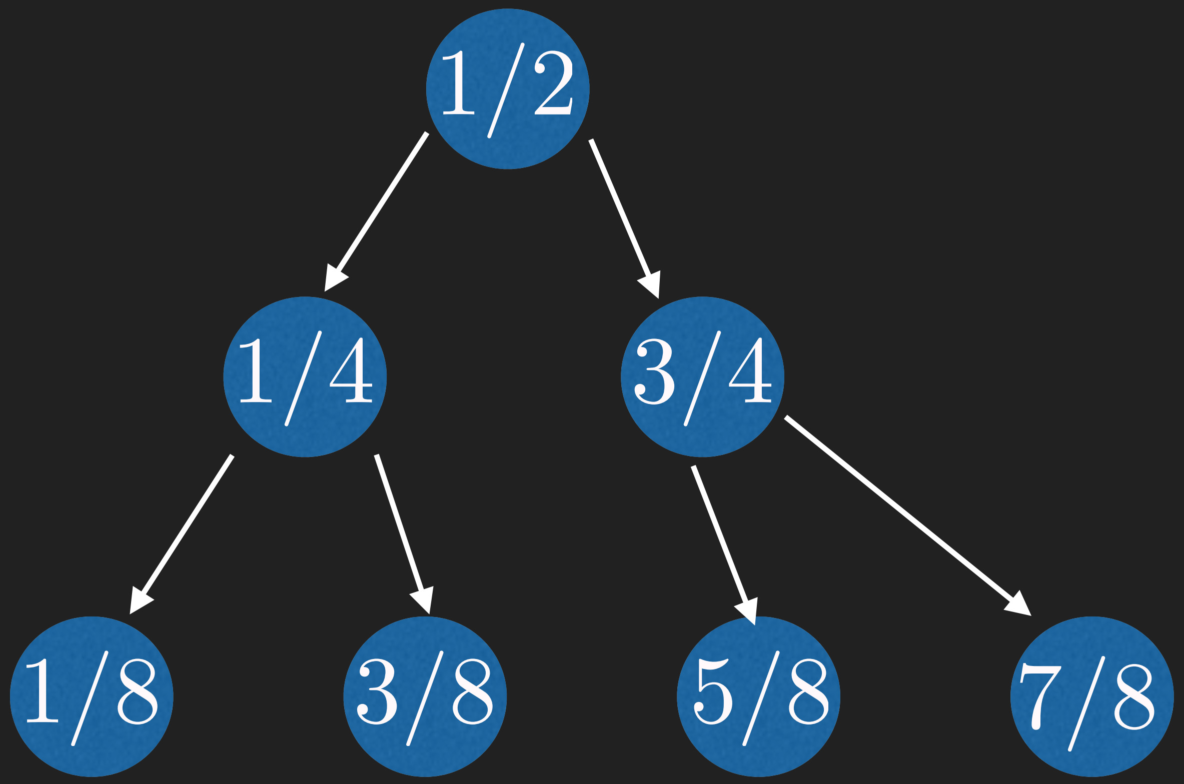
Idea

- Lie $\equiv$ buyer when rejecting $v > p_t$ or accepting when $v < p_t$.

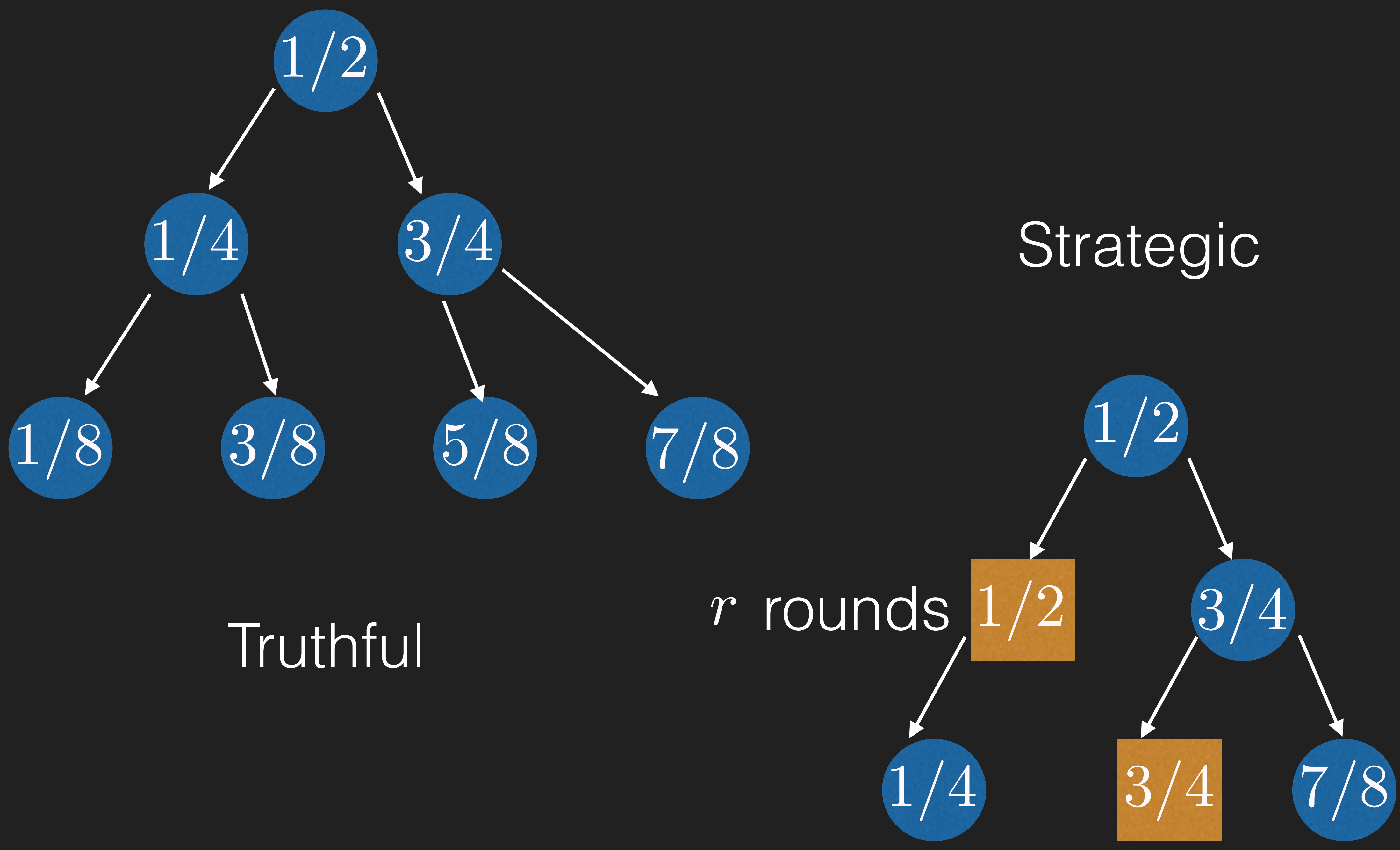
- Can we dissuade the buyer from lying?
  - ✦ buyer's weakness: time (discounted surplus).
  - ✦ penalization: if buyer rejects price, reoffer the price for another $(r-1)$ rounds.
  - ✦ choice of r subject to a trade-off.

Pricing Strategies

- Any deterministic strategy can be represented by a tree.



Meta-Algorithm



PFS Guarantees

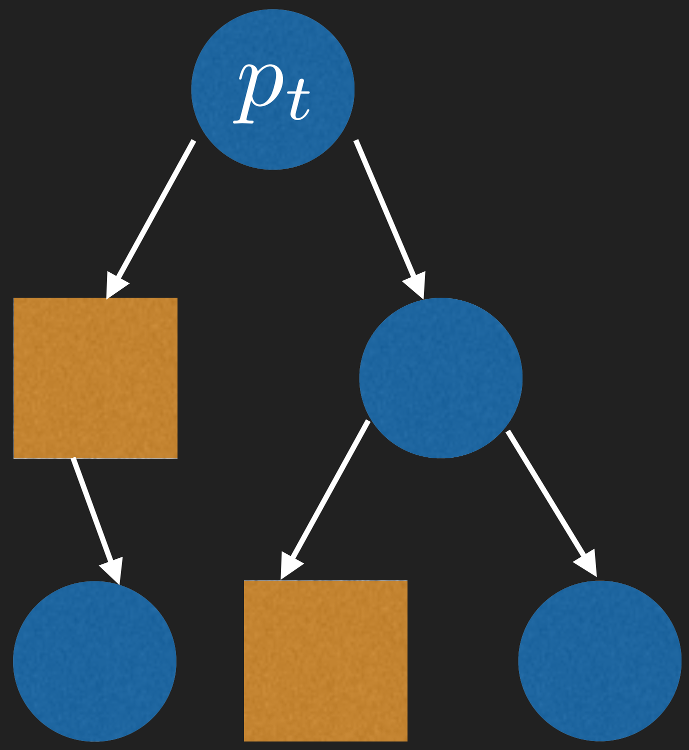
[MM and Muñoz, 2014]

- Theorem: let $\gamma_0 \in (\frac{1}{2}, 1)$; the following strategic regret guarantees hold for Penalized Fast Search (PFS):
  - $\text{Reg}_T(\text{PFS}) = O(\log\log T)$ for $\gamma \in (0, \frac{1}{2}]$;
  - $\text{Reg}_T(\text{PFS}) = O(\log T \log\log T)$ for $\gamma \in (\frac{1}{2}, \gamma_0)$.

Proof Idea

Surplus of rejected path at most

$$\frac{\gamma^{t+r-1}}{1-\gamma}$$



Surplus of accepted path at least

$$\gamma^{t-1}(v - p_t)$$

$$\rightarrow \quad (v - p_t) \leq \frac{\gamma^r}{1-\gamma}$$

Horizon-Indep. Regret

[Drutsa, 2017]

- Extension of PFS via 'exponentiating trick' to horizon-independent algorithm $\widetilde{\mathrm{PFS}}$:
    - length of ith epoch verifies $\log_2 \log_2 T_i = 2^{i-1}$.
    - $\mathrm{Reg}_T(\widetilde{\mathrm{PFS}}) = O(\log\log T)$ for $\gamma \in (0, \frac{1}{2}]$;
    - $\mathrm{Reg}_T(\widetilde{\mathrm{PFS}}) = O(\log T \log\log T)$ for $\gamma \in (\frac{1}{2}, \gamma_0)$.

Further Improvement

[Drutsa, 2017]

- PRRFES algorithm:
  - truthful FES algorithm: modified FS; after rejection, reoffer last accepted price g times.
  - same lie penalization as in PFS.
  - continue to offer until rejection.
  - strategic regret: $\mathrm{Reg}_T(\mathrm{PRRFES}) = O(\log\log T)$ for $\gamma \in (0, \gamma_0]$.

Random valuations

Setup

[Amin et al., 2013]

- Repeated posted-price auctions:
  - good repeatedly offered for sale by a seller to a single buyer over T rounds.
  - buyer receives valuation $v_t \in [0,1]$, $v_t \sim \mathcal{D}$.
  - seller offers price p_t and buyer accepts, $a_t = 1$, or rejects, $a_t = 0$, at each round $t \in [T]$.

Setup

- Seller: pricing algorithm $\mathcal{A}$.
  - ✦ total revenue: $\mathbb{E}\left[\sum_{t=1}^{T} a_t p_t\right]$.
  - ✦ regret: $\mathrm{Reg}_T(\mathcal{A}) = \max_{p \in \mathcal{P}} p\,\mathbb{P}(v > p)T - \mathbb{E}\left[\sum_{t=1}^{T} a_t p_t\right]$.
- Buyer: discounting factor $\gamma \in (0, 1]$.
  - ✦ surplus: $\mathrm{Sur}_T(\mathcal{A}) = \mathbb{E}\left[\sum_{t=1}^{T} \gamma^{t-1} a_t (v - p_t)\right]$.

Strategic buyers

- Simple tree structure for fixed valuation.
- Seller offers price from distribution P_t.
- Surplus of state $s_t = (P_t, H_{t-1}, v_t, p_t)$:

$$S_t(s_t) = \max_{a_t \in \{0,1\}} \gamma^{t-1} a_t (v_t - p_t)$$
$$+ \mathop{\mathbb{E}}_{\substack{(v_{t+1}, p_{t+1}) \\ \sim \mathcal{D} \times f_t(P_t, H_{t-1})}} \Big[S_{t+1}\big(f_t(P_t, H_{t-1}), H_t, v_{t+1}, p_{t+1}\big) \Big].$$

- Solution found in time $\Omega(T^{|\mathcal{P}|})$.

ϵ-strategic buyers

- Stop optimizing if all future surplus is at most ϵ .
- Behave truthfully otherwise,
- Optimize for $\left\lceil \frac{\log\left[\frac{1}{\epsilon(1-\gamma)}\right]}{\log\left(\frac{1}{\gamma}\right)} \right\rceil$ rounds.
- Tractable MDP.

Bandit Formulation

- Only observe reward of price offered.
- Minimize pseudo-regret

$$\mathrm{Reg}_T(\mathcal{A}) = \max_{p \in \mathcal{P}} p\, \mathbb{P}(v > p) T - \mathbb{E}\left[\sum_{t=1}^{T} a_t p_t\right].$$

- Problem: rewards not i.i.d. (strategic buyer).

Regret bound

[MM and Muñoz, 2015]

Theorem: Let $\mathcal{P}$ be a finite set of prices. Let $\mathcal{L}$ be the number of time the buyer lies.
Let $p^* = \operatorname*{argmax}_{p \in \mathcal{P}} p\, \mathbb{P}(v > p)$ and

$\Delta_p = p^*\, \mathbb{P}(v > p^*) - p\, \mathbb{P}(v > p)$. For any $\delta > 0$,

$$\text{Reg}_T \leq \mathbb{E}[\mathcal{L}] + \sum_{p:\ \Delta_p > \delta} \mathbb{E}[T_p(t)]\Delta_p + T\delta$$

where $T_p(t)$ is the number of times price p has been offered up to time t .

R-UCB

- Make UCB robust to lies.
- Use different upper confidence bounds

$$\widehat{\mu}_p(t) = \frac{1}{T_p(t)} \sum_{i=1}^{t} a_i p_i \mathbb{1}_{p_i = p} + \sqrt{\frac{2 \log t}{T_p(t)}} + \frac{Lp}{T_p(t)}$$

Regret analysis

- Proposition: The regret of the R-UCB algorithm is bounded by

$$\mathrm{Reg}_T \leq \mathbb{E}[\mathcal{L}] + \sum_{p:\Delta_p > \delta} 4Lp + \frac{32 \log T}{\Delta_p} + 2\Delta_p + T\delta + \sum_{t=1}^{T} P_t(p, L),$$

where

$$P_t(p, L) := \mathbb{P}\left(\left| \frac{L_t(p)}{T_p(t)} \right| + \left| \frac{L_t(p^*)}{T^*(t)} \right| \geq L\left(\frac{p}{T_p(t)} + \frac{p^*}{T^*(t)} \right) \right).$$

Bound on Lies

- An ϵ-strategic buyer lies at most $\left\lceil \frac{\log(1/\epsilon(1-\gamma))}{\log(1/\gamma)} \right\rceil$.
- Regret of R-UCB in $O\left(\log T + \frac{\mathcal{P}}{1-\gamma}\right)$.
- Extension to continuous set of prices by discretization.
- Regret in $O\left(\sqrt{T} + \frac{T^{1/4}}{1-\gamma}\right)$.

Conclusion

- Analysis of strategic regret.
  - Fixed and random valuation scenarios.
  - Simple algorithms extending truthful scenario.
- Many questions:
  - Can we extend results to other types of buyers?
  - What about if the buyer learns too?
  - Extension to general auctions?

Other Related Questions

- Can the buyer trust the algorithm announced?
  - testing incentive-compatibility [Lahaie, Muñoz, Sivan, and Vassilvitskii, 2017] (Andres's talk).
- Extend analysis to the case where algorithmic details are not known.