No-Regret Algorithms in non-Truthful Auctions with Budget and ROI Constraints

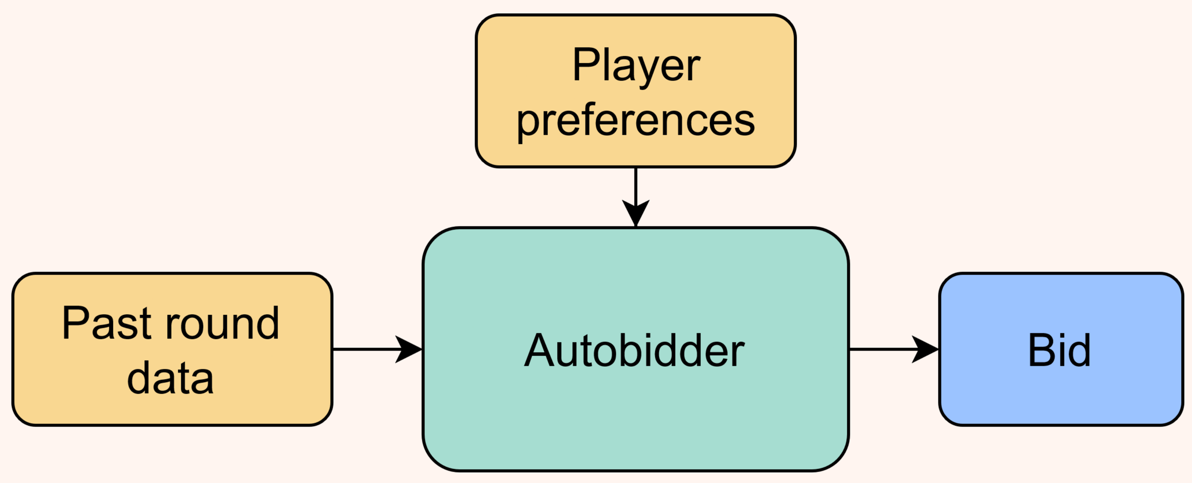
Giannis Fikioris (while SR at Google), Gagan Aggarwal, Mingfei Zhao

Google Reseach & Cornell University

Setup

Automated Bidding

Global digital advertising: **$750bn** in 2024



Automated bidder interacting with stochastic environment

T rounds

Player Model

Budget B & ROI constant γ

Value v_t and highest competing bid d_t

$$\text{maximize value: } \sum_t v_t \mathbf{1}[b_t \geq d_t]$$

$$\text{subject budget: } \sum_t b_t \mathbf{1}[b_t \geq d_t] \leq B$$

$$\text{ROI: } \sum_t b_t \mathbf{1}[b_t \geq d_t] \leq \gamma \sum_t v_t \mathbf{1}[b_t \geq d_t]$$

Stationary environment: $(v_t, d_t) \sim \mathcal{D}$

Learning to bid

Bidding function: $b_t = f(v_t)$

Any $f \implies$ impossible to learn

Class of bidding functions $\mathcal{F}$

[Lucier, Pattathil, Slivkins, Zhang] & [Fikioris, Tardos]:
$\mathcal{F}$ = linear functions
- loss of utility in simple examples:
  - Value $v_t \sim U[0,1]$
  - Competing bid $d_t = 1/2$

This work: $\mathcal{F}$ = Lipschitz functions

Theorem 1 – Full Information

Online learning algorithm that
- exact budget and ROI satisfaction
- total value $\tilde{O}(\sqrt{T})$ suboptimal wrt optimal Lipschitz bidding

When d_t is not revealed after round t:

Theorem 2 – Bandit information

No algorithm $O(T^{2/3})$ suboptimal, even if $v_t = 1$
There exists algorithm that is $\tilde{O}(T^{3/4})$ suboptimal

All results apply to
- first-price/second-price/hybrid auctions
- value maximizers or quasi-linear utilities

Primal / Dual Framework

Maximize / Minimize Lagrangian:

$$\max_f \min_{\lambda,\mu} \left(v_t + \lambda \cdot \left(B/T - f(v_t)\right) + \mu \cdot \left(\gamma\, v_t - f(v_t)\right)\right)\mathbf{1}[f(v_t) \geq d_t]$$

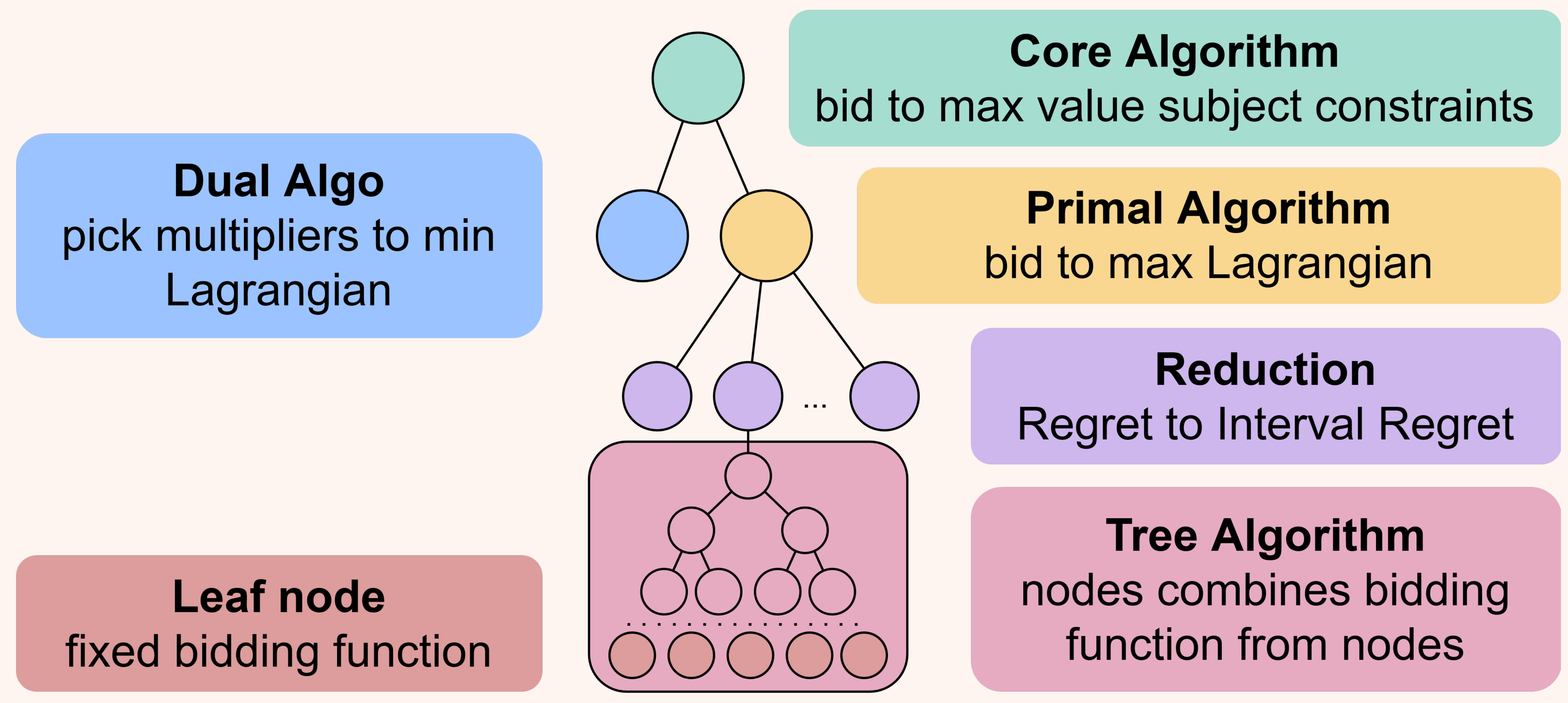
Online algorithms that pick f and λ, μ.

[Castiglioni, Celli, Kroer]: OGD for λ, μ & "right" primal for f
- low error for constrained problem

Full Paper Link

Algorithm

Full information



Safe bidding

Bidding function $f \in \mathcal{F}$ could have negative Lagrangian reward

"Safe" bid for round t: $b_t^\circ = \min\left\{\frac{1+\mu_t}{\lambda_t+\mu_t} v_t, 1\right\}$
- If $\exists d_t$ where $f(v_t)$ has negative reward, b_t° always better

Tree algorithm

$\mathcal{F}_\varepsilon$: ε-cover of Lipschitz functions, $|\mathcal{F}_\varepsilon| = \exp(\Theta(1/\varepsilon))$
- Hedge would get $T^{2/3}$ regret

Arrange $\mathcal{F}_\varepsilon$ into a tree where "similar" functions are "close"
- Leaves are fixed $f \in \mathcal{F}_\varepsilon$
- Non-leaves combine children via Hedge-like algorithm

Similar to [Cesa-Bianchi, Gaillard, Gentile, Gerchinovitz] and [Han, Zhou, Flores, Ordentlich, Weissman]

Range-agnostic Regret

Reward upper bound $U_t \propto \lambda_t, \mu_t$ unknown before round t

Hedge with learning rate $\eta_t \propto \frac{1}{U_t}$ achieves regret $\tilde{O}(U_t\sqrt{t})$
- similar to U_t being known

Interval Regret

Primal/Dual framework requires interval regret

Low regret in every interval $[T_1, T_2]$

Black-box reduction: standard regret $\mathcal{A} \implies$ interval regret
- Hedge over T instances of $\mathcal{A}$, each starting at round t

Exact ROI Satisfaction

Previous algorithm violates ROI
- but whp violation $\leq \tilde{O}(\sqrt{T})$

Black box reduction: when constraint almost violated, bid to

$$\max_f \left(\gamma v_t - f(v_t)\right)\mathbf{1}[f(v_t) \geq d_t]$$

Re-formulation of Lagrangian
- safe bids ensures non-negative

Used for τ rounds, increases slack by $\Theta(\tau) - O(\sqrt{\tau})$

Dynamic instead of static ([Feng, Padmanabhan, Wang])

Bandit Information

$v_t = 1$, no ROI

Base CDF: "All" bids optimal

Perturbation hides optimal bid
- Small enough: hard to find
- Large enough: need to find

$\Omega(T^{2/3})$ regret

Similar to [Kleinberg, Leighton]