Strategic Considerations in Bandit Settings: The Price of Truthfulness in Ad Auctions

Nikhil Devanur & Sham M. Kakade

TTI-C → UPenn

Strategic Considerations in Learning Scenarios

Many technological systems better understood and managed when they are viewed as economic systems rather than just technological ones (e.g. spam, P2P networks, routing, etc).

Strategic Considerations in Learning Scenarios

Many technological systems better understood and managed when they are viewed as economic systems rather than just technological ones (e.g. spam, P2P networks, routing, etc).

This talk: Online Ad Auctions

- huge source of revenue (Google, Yahoo!, Microsoft)
- central issue: exploration/exploitation

Strategic Considerations in Learning Scenarios

Many technological systems better understood and managed when they are viewed as economic systems rather than just technological ones (e.g. spam, P2P networks, routing, etc).

This talk: Online Ad Auctions

- huge source of revenue (Google, Yahoo!, Microsoft)
- central issue: exploration/exploitation

But... now participants have their own selfish interests.

Pay-Per-Impression Auctions

"Classical" Ad Auction:

- An advertiser is charged each time their ad is displayed. e.g. newspapers, billboards, etc...
- Advertisers bid for an impression.

Pay-Per-Impression Auctions

"Classical" Ad Auction:

- An advertiser is charged each time their ad is displayed.
  e.g. newspapers, billboards, etc...
- Advertisers bid for an impression.
- What's wrong with this model for the Internet?

Pay-Per-Impression Auctions

"Classical" Ad Auction:

- An advertiser is charged each time their ad is displayed.
  e.g. newspapers, billboards, etc...
- Advertisers bid for an impression.
- What's wrong with this model for the Internet?
  - There are many different forums for ads:
    search engines (keywords), personal websites, blogs, email. . .

Pay-Per-Impression Auctions

"Classical" Ad Auction:

- An advertiser is charged each time their ad is displayed. e.g. newspapers, billboards, etc...
- Advertisers bid for an impression.
- What's wrong with this model for the Internet?
  - There are many different forums for ads: search engines (keywords), personal websites, blogs, email. . .
  - Advertisers do not know their value for an impression in these highly varied forums.

Pay-Per-Click Ad Auctions

Here, advertisers still bid on a slots, but they are charged *only* if their ad is clicked.

- Advertisers have a better idea of their 'value' for a click.

Pay-Per-Click Ad Auctions

Here, advertisers still bid on a slots, but they are charged *only* if their ad is clicked.

- Advertisers have a better idea of their 'value' for a click.
  - Closer to a desired outcome.
  - Sometimes astronomical prices for certain keywords: "loan consolidation", "mortgage financing", "mesothelioma"

Pay-Per-Click Ad Auctions

Here, advertisers still bid on a slots, but they are charged *only* if their ad is clicked.

- Advertisers have a better idea of their ‘value’ for a click.
  - Closer to a desired outcome.
  - Sometimes astronomical prices for certain keywords: “loan consolidation”, “mortgage financing”, “mesothelioma”
- Mechanism design problem: allocate slots to maximize revenue, which depends on bids and Click-Through-Rates (CTRs):

Pay-Per-Click Ad Auctions

Here, advertisers still bid on a slots, but they are charged *only* if their ad is clicked.

- Advertisers have a better idea of their 'value' for a click.
  - Closer to a desired outcome.
  - Sometimes astronomical prices for certain keywords: "loan consolidation", "mortgage financing", "mesothelioma"
- Mechanism design problem: allocate slots to maximize revenue, which depends on bids and Click-Through-Rates (CTRs):
  - Exploit: use current CTR estimates for immediate profit.
  - Explore: obtain better CTR estimates for future profit.

Strategic Considerations

If we ignored strategic considerations, then we could just use some multi-armed bandit scheme.

Strategic Considerations

If we ignored strategic considerations, then we could just use some multi-armed bandit scheme.

but... consider a one shot auction for a single good:

- The auction's winner should not be charged higher than the second highest bid (else, the winner would lie)

Strategic Considerations

If we ignored strategic considerations, then we could just use some multi-armed bandit scheme.

but... consider a one shot auction for a single good:

- The auction's winner should not be charged higher than the second highest bid (else, the winner would lie)
- English Auction: "open outcry" auction

Strategic Considerations

If we ignored strategic considerations, then we could just use some multi-armed bandit scheme.

but... consider a one shot auction for a single good:

- The auction's winner should not be charged higher than the second highest bid (else, the winner would lie)
- English Auction: "open outcry" auction
- Vickery Auction: "sealed bid" auction. The highest bidder wins and is charged the second highest bid.

Strategic Considerations

If we ignored strategic considerations, then we could just use some multi-armed bandit scheme.

but... consider a one shot auction for a single good:

- The auction's winner should not be charged higher than the second highest bid (else, the winner would lie)
- English Auction: "open outcry" auction
- Vickery Auction: "sealed bid" auction. The highest bidder wins and is charged the second highest bid.
  - Bidding truthfully is a dominant strategy.

Strategic Considerations

If we ignored strategic considerations, then we could just use some multi-armed bandit scheme.

but... consider a one shot auction for a single good:

- The auction's winner should not be charged higher than the second highest bid (else, the winner would lie)
- English Auction: "open outcry" auction
- Vickery Auction: "sealed bid" auction. The highest bidder wins and is charged the second highest bid.
  - Bidding truthfully is a dominant strategy.
  - e.g. Ebay (but all at once).

Strategic Considerations

If we ignored strategic considerations, then we could just use some multi-armed bandit scheme.

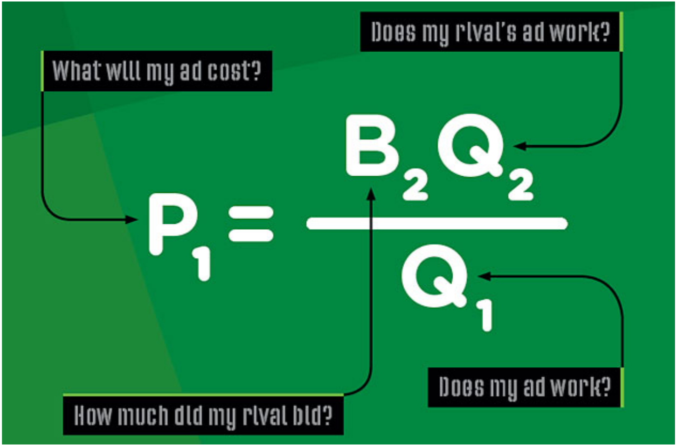
but... consider a one shot auction for a single good:

- The auction's winner should not be charged higher than the second highest bid (else, the winner would lie)
- English Auction: "open outcry" auction
- Vickery Auction: "sealed bid" auction. The highest bidder wins and is charged the second highest bid.
  - Bidding truthfully is a dominant strategy.
  - e.g. Ebay (but all at once).

This talk: What bandit algorithms permit such truthful pricing schemes? What rate can we hope for?

Instantaneous Truthfulness Pricing

Wired [2009] on Google: AdWords Premium → AdWords Select



"Quality score" is central (and depends on estimated CTRs).

Outline

A Repeated PPC Auction

Each round, of a single slot auction with n advertisers, proceeds as:

A Repeated PPC Auction

Each round, of a single slot auction with n advertisers, proceeds as:

1. Each advertiser i at time t:
   - has a true (private) value for a click is v_i^t.
   - bids a value b_i^t.

A Repeated PPC Auction

Each round, of a single slot auction with n advertisers, proceeds as:

1. Each advertiser i at time t:
   - has a true (private) value for a click is v_i^t.
   - bids a value b_i^t.
2. The auction allocates the slot to only one advertiser $i^\star$.

A Repeated PPC Auction

Each round, of a single slot auction with n advertisers, proceeds as:

1. Each advertiser i at time t:
   - has a true (private) value for a click is v_i^t.
   - bids a value b_i^t.
2. The auction allocates the slot to only one advertiser $i^\star$.
3. The auction then observes the click only for $i^\star$
   - Bandit setting: only displayed ad's outcome is observed.

A Repeated PPC Auction

Each round, of a single slot auction with n advertisers, proceeds as:

1. Each advertiser i at time t:
   - has a true (private) value for a click is v_i^t.
   - bids a value b_i^t.
2. The auction allocates the slot to only one advertiser $i^\star$.
3. The auction then observes the click only for $i^\star$
   - Bandit setting: only displayed ad's outcome is observed.
   - Stochastic setting (with no keywords): the probability of i's ad being clicked at time t is an i.i.d., constant probability ρ_i (the CTR).

A Repeated PPC Auction

Each round, of a single slot auction with n advertisers, proceeds as:

1. Each advertiser i at time t:
   - has a true (private) value for a click is v_i^t.
   - bids a value b_i^t.
2. The auction allocates the slot to only one advertiser $i^\star$.
3. The auction then observes the click only for $i^\star$
   - Bandit setting: only displayed ad's outcome is observed.
   - Stochastic setting (with no keywords): the probability of i's ad being clicked at time t is an i.i.d., constant probability ρ_i (the CTR).
4. Advertiser $i^\star$ is charged only if the $i^\star$'s ad was clicked.

Instantaneous Truthfulness and Maximizing Revenue

- Say we have knowledge of ρ_i.

Instantaneous Truthfulness and Maximizing Revenue

- Say we have knowledge of ρ_i.
- We would like to allocate to $i^\star = \mathrm{argmax} \rho_i b_i$.

Instantaneous Truthfulness and Maximizing Revenue

- Say we have knowledge of ρ_i.
- We would like to allocate to $i^\star = \text{argmax} \rho_i b_i$.
- How much should we charge?

Instantaneous Truthfulness and Maximizing Revenue

- Say we have knowledge of ρ_i.
- We would like to allocate to $i^\star = \mathrm{argmax} \rho_i b_i$.
- How much should we charge?
  - Unreasonable to expect a revenue of $\rho_{i^\star} b_{i^\star}$, which involves charging $b_{i^\star}$

Instantaneous Truthfulness and Maximizing Revenue

- Say we have knowledge of ρ_i.
- We would like to allocate to $i^\star = \mathrm{argmax}\rho_i b_i$.
- How much should we charge?
    - Unreasonable to expect a revenue of $\rho_{i^\star} b_{i^\star}$, which involves charging $b_{i^\star}$
    - Instead, we seek the second highest expected revenue.

Instantaneous Truthfulness and Maximizing Revenue

- Say we have knowledge of ρ_i.
- We would like to allocate to $i^\star = \mathrm{argmax} \rho_i b_i$.
- How much should we charge?
  - Unreasonable to expect a revenue of $\rho_{i^\star} b_{i^\star}$, which involves charging $b_{i^\star}$
  - Instead, we seek the second highest expected revenue.
  - Let $\mathrm{smax}_i\{u_i\}$ be the second largest number in the set $\{u_i\}_i$.

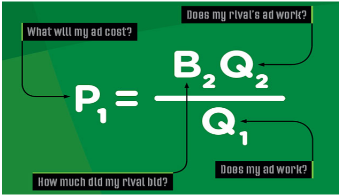
Instantaneous Truthfulness and Maximizing Revenue

- Say we have knowledge of ρ_i.
- We would like to allocate to $i^\star = \mathrm{argmax} \rho_i b_i$.
- How much should we charge?
  - Unreasonable to expect a revenue of $\rho_{i^\star} b_{i^\star}$, which involves charging $b_{i^\star}$
  - Instead, we seek the second highest expected revenue.
  - Let $\mathrm{smax}_i\{u_i\}$ be the second largest number in the set $\{u_i\}_i$.
  - So we seek a revenue of: $\mathrm{smax} \rho_i b_i$

Instantaneous Truthfulness and Maximizing Revenue

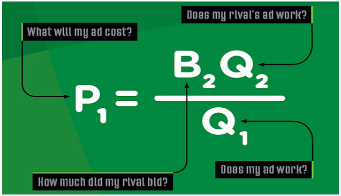
- Say we have knowledge of ρ_i.
- We would like to allocate to $i^\star = \mathrm{argmax} \rho_i b_i$.
- How much should we charge?
  - Unreasonable to expect a revenue of $\rho_{i^\star} b_{i^\star}$, which involves charging $b_{i^\star}$
  - Instead, we seek the second highest expected revenue.
  - Let $\mathrm{smax}_i \{u_i\}$ be the second largest number in the set $\{u_i\}_i$.
  - So we seek a revenue of: $\mathrm{smax} \rho_i b_i$
- It's essentially the best we can do.

Instantaneous Truthfulness Pricing



- Assume $i^{\star} = 1$ and advertiser 2 is second, so: $b_1\rho_1 \geq b_2\rho_2$

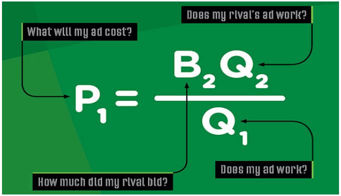
Instantaneous Truthfulness Pricing



- Assume $i^\star = 1$ and advertiser 2 is second, so: $b_1 \rho_1 \geq b_2 \rho_2$
- Hence, the price charged (if clicked) is:

$$p_1 = \frac{b_2 \rho_2}{\rho_1} \leq b_1$$

Instantaneous Truthfulness Pricing

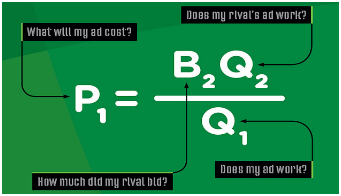


- Assume $i^\star = 1$ and advertiser 2 is second, so: $b_1\rho_1 \geq b_2\rho_2$
- Hence, the price charged (if clicked) is:

$$p_1 = \frac{b_2\rho_2}{\rho_1} \leq b_1$$

- This provides an expected revenue of $b_2\rho_2$.

Instantaneous Truthfulness Pricing

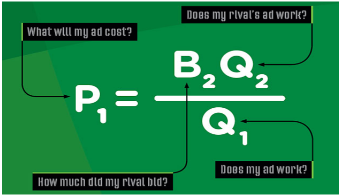


- Assume $i^\star = 1$ and advertiser 2 is second, so: $b_1\rho_1 \geq b_2\rho_2$
- Hence, the price charged (if clicked) is:

$$p_1 = \frac{b_2\rho_2}{\rho_1} \leq b_1$$

- This provides an expected revenue of $b_2\rho_2$.
- This is truthful. (p_1 is independent of $b_{i^\star}$ and it's less than this.)

Instantaneous Truthfulness Pricing



- Assume $i^\star = 1$ and advertiser 2 is second, so: $b_1\rho_1 \geq b_2\rho_2$
- Hence, the price charged (if clicked) is:

$$p_1 = \frac{b_2\rho_2}{\rho_1} \leq b_1$$

- This provides an expected revenue of $b_2\rho_2$.
- This is truthful. (p_1 is independent of $b_{i^\star}$ and it's less than this.)
- (any 'weights' result in an instantaneously truthful auction).

Online Truthfulness

- The auction depends on the advertisers previous bids, so an advertiser could try to manipulate their current bid in order to improve their future utility.

Online Truthfulness

- The auction depends on the advertisers previous bids, so an advertiser could try to manipulate their current bid in order to improve their future utility.
- A truthful auction would prevent this.

Online Truthfulness

- The auction depends on the advertisers previous bids, so an advertiser could try to manipulate their current bid in order to improve their future utility.
- A truthful auction would prevent this.

Definition

An auction is always truthful if for all click sequences, if bidding $v_i^t = b_i^t$ is a dominant strategy for all bidders
e.g. if for all possible bids of other advertisers $\{b_{-i}^t\}$, the utility of i is maximized when i bids $b_i^t = v_i^t$ for all t.

Online Truthfulness

- The auction depends on the advertisers previous bids, so an advertiser could try to manipulate their current bid in order to improve their future utility.
- A truthful auction would prevent this.

Definition

An auction is always truthful if for all click sequences, if bidding $v_i^t = b_i^t$ is a dominant strategy for all bidders
e.g. if for all possible bids of other advertisers $\{b_{-i}^t\}$, the utility of i is maximized when i bids $b_i^t = v_i^t$ for all t.

- How to price clicks? explore/exploit?
- How should we measure success?

Truthful Regret

- Subject to the constraint of being always truthful, the auction seeks to maximize its revenue.

Truthful Regret

- Subject to the constraint of being always truthful, the auction seeks to maximize its revenue.
  - OPT is the revenue of the Vickrey auction that knows ρ_i:

$$\mathrm{OPT}_T = \sum_{t=1}^{T} \mathrm{smax}_i\{\rho_i b_i^t\} .$$

Truthful Regret

- Subject to the constraint of being always truthful, the auction seeks to maximize its revenue.
  - OPT is the revenue of the Vickrey auction that knows ρ_i:

$$\mathrm{OPT}_T = \sum_{t=1}^{T} \mathrm{smax}_i \{\rho_i b_i^t\} \,.$$

  - The expected truthful regret of the auction is:

$$\text{T-Regret} := \mathrm{OPT}_T - \mathbb{E}[\text{Revenue}]$$

What's the difficulty?

Two issues:

What's the difficulty?

Two issues:

- We only have estimates of ρ_i using observations.
  - But even with estimates, it's easy to come up with an instantaneously truthful scheme.

What's the difficulty?

Two issues:

- We only have estimates of ρ_i using observations.
  - But even with estimates, it's easy to come up with an instantaneously truthful scheme.
- It seems like we need to charge using information about the second highest revenue.

What's the difficulty?

Two issues:

- We only have estimates of ρ_i using observations.
    - But even with estimates, it's easy to come up with an instantaneously truthful scheme.
- It seems like we need to charge using information about the second highest revenue.
    - An "optimal" multi-armed bandit mechanism only gleans enough information about the second most profitable arm to distinguish it from the first.

What's the difficulty?

Two issues:

- We only have estimates of ρ_i using observations.
  - But even with estimates, it's easy to come up with an instantaneously truthful scheme.
- It seems like we need to charge using information about the second highest revenue.
  - An "optimal" multi-armed bandit mechanism only gleans enough information about the second most profitable arm to distinguish it from the first.
  - With a naive implementation, the mechanism would end up overcharging the winner (so the winner would lie to force more exploration and be charge a better price).

Outline

1. Pay-Per-Click Model
   - A Bandit Setting
   - Truthful Auctions
2. Price of Truthfulness
   - Achievable Regret
3. Algorithms and Proof Ideas
   - Truthful Pricing

Achievable Truthful Regret

Theorem (Upper Bound)

Let $b_{max} = \max_{i,t} b_i^t$. There exists an always truthful PPC auction with

$$\textit{Truthful-Regret} = O(b_{max} n^{1/3} T^{2/3} \sqrt{\log(nT)}).$$

Achievable Truthful Regret

Theorem (Upper Bound)

Let $b_{max} = \max_{i,t} b_i^t$. There exists an always truthful PPC auction with

$$\textit{Truthful-Regret} = O(b_{max} n^{1/3} T^{2/3} \sqrt{\log(nT)}).$$

- it's 'strawman algo': first explore, then exploit.

Achievable Truthful Regret

Theorem (Upper Bound)

Let $b_{max} = \max_{i,t} b_i^t$. There exists an always truthful PPC auction with

$$\mathit{Truthful\text{-}Regret} = O(b_{max} n^{1/3} T^{2/3} \sqrt{\log(nT)}).$$

- it's 'strawman algo': first explore, then exploit.
- Can we do better? i.e. get $O(\sqrt{T})$
  - try to charge more carefully (not over/under charge)?
  - charge only at end?
  - many different bandit-algorithms to consider?

Lower Bound

Theorem (Lower Bound)

For every non-degenerate, scale invariant and always truthful PPC auction (with $n = 2$), there exists a set of bids bounded in $[0, 1]$ and ρ_i such that T-Regret $= \Omega(T^{2/3})$.

Lower Bound

Theorem (Lower Bound)

For every non-degenerate, scale invariant and always truthful PPC auction (with $n = 2$), there exists a set of bids bounded in $[0, 1]$ and ρ_i such that $T\text{-}Regret = \Omega(T^{2/3})$.

- Truthfulness provides a statistical restriction.
  - usually, $O(\sqrt{T})$ regret.

Lower Bound

Theorem (Lower Bound)

For every non-degenerate, scale invariant and always truthful PPC auction (with $n = 2$), there exists a set of bids bounded in $[0, 1]$ and ρ_i such that T-$\mathcal{Regret} = \Omega(T^{2/3})$.

- Truthfulness provides a statistical restriction.
    - usually, $O(\sqrt{T})$ regret.
- proof idea: need to estimate second highest revenue?

Lower Bound

Theorem (Lower Bound)

For every non-degenerate, scale invariant and always truthful PPC auction (with $n = 2$), there exists a set of bids bounded in $[0, 1]$ and ρ_i such that $T\text{-}Regret = \Omega(T^{2/3})$.

- Truthfulness provides a statistical restriction.
  - usually, $O(\sqrt{T})$ regret.
- proof idea: need to estimate second highest revenue?
  - uses (seminal) characterization of truthful auctions [Myerson, 1981]
  - proof reasons about information gain vs. revenue

Lower Bound

Theorem (Lower Bound)

For every non-degenerate, scale invariant and always truthful PPC auction (with $n = 2$), there exists a set of bids bounded in $[0, 1]$ and ρ_i such that T-Regret $= \Omega(T^{2/3})$.

- Truthfulness provides a statistical restriction.
  - usually, $O(\sqrt{T})$ regret.
- proof idea: need to estimate second highest revenue?
  - uses (seminal) characterization of truthful auctions [Myerson, 1981]
  - proof reasons about information gain vs. revenue
- Concurrent result: Babaioff, Sharma, & Slivkins [2009]

Outline

Algorithm (for the Upper Bound)

- Explore: for the first τ steps:
  - Randomly allocate the slot to an advertiser.
  - Charge nothing.

Algorithm (for the Upper Bound)

- Explore: for the first τ steps:
  - Randomly allocate the slot to an advertiser.
  - Charge nothing.
- Estimate: Let $\hat{\rho}_i^+$ be a (high prob.) empirical upper bound on ρ_i.

Algorithm (for the Upper Bound)

- Explore: for the first τ steps:
  - Randomly allocate the slot to an advertiser.
  - Charge nothing.
- Estimate: Let $\hat{\rho}_i^+$ be a (high prob.) empirical upper bound on ρ_i.
- Exploit: for the remainder, at each time t,
  - Allocate to the bidder i^* which maximizes $\hat{\rho}_i^+ b_i^t$
  - Charge i^*, if clicked,

$$p_i^t = \frac{\mathrm{smax}_i \hat{\rho}_i^+ b_i^t}{\hat{\rho}_{i^*}^+}$$

Algorithm (for the Upper Bound)

- Explore: for the first τ steps:
  - Randomly allocate the slot to an advertiser.
  - Charge nothing.
- Estimate: Let $\hat{\rho}_i^+$ be a (high prob.) empirical upper bound on ρ_i.
- Exploit: for the remainder, at each time t,
  - Allocate to the bidder i^* which maximizes $\hat{\rho}_i^+ b_i^t$
  - Charge i^*, if clicked,

$$p_i^t = \frac{\mathrm{smax}_i \hat{\rho}_i^+ b_i^t}{\hat{\rho}_{i^*}^+}$$

- Proof Idea:
  - truthfulness: exploration is entirely non-adaptive to the bids.
  - optimize τ to show upper bound.

Truthful Pricing

- Let $x^t(\cdot)$ and c^t be the allocation and click vectors: $x_i^t = 1$ iff the allocation is to i at time t, and $c_i^t = 1$ iff i is clicked.

Theorem (Myerson, 1981)

Truthful pricing rule: Fix a click sequence. Let $y_i = \sum_t x_i^t c_i^t$ and let the price $p_i = \sum_t p_i^t$. If an auction x, is truthful then

1. *y_i is monotonically increasing in b_i*
2. *the price p_i charged to i is exactly*

$$p_i(\boldsymbol{b}) = b_i y_i(\boldsymbol{b}) - \int_{z=0}^{b_i} y_i(z, \boldsymbol{b}_{-i}) dz.$$

Truthful Pricing

- Let $x^t(\cdot)$ and c^t be the allocation and click vectors: $x_i^t = 1$ iff the allocation is to i at time t, and $c_i^t = 1$ iff i is clicked.

Theorem (Myerson, 1981)

Truthful pricing rule: Fix a click sequence. Let $y_i = \sum_t x_i^t c_i^t$ and let the price $p_i = \sum_t p_i^t$. If an auction x, is truthful then

1. *y_i is monotonically increasing in b_i*
2. *the price p_i charged to i is exactly*

$$p_i(\boldsymbol{b}) = b_i y_i(\boldsymbol{b}) - \int_{z=0}^{b_i} y_i(z, \boldsymbol{b}_{-i}) dz.$$

- the last term increases as there is "less competition" .

Structural Restrictions

- Def: Round t is a competitive round for bidder 1 if there exists a high enough bid b_1^t (regardless of b_2^t) such that 1 can win.

Structural Restrictions

- Def: Round t is a competitive round for bidder 1 if there exists a high enough bid b_1^t (regardless of b_2^t) such that 1 can win.
- Def: The allocation x_1^t depends on c_2^τ if there exist b_1, b_2 such that $x_1^t(b_1, b_2, c_2^\tau) \neq x_1^t(b_1, b_2, 1 - c_2^\tau)$.

Structural Restrictions

- Def: Round t is a competitive round for bidder 1 if there exists a high enough bid b_1^t (regardless of b_2^t) such that 1 can win.
- Def: The allocation x_1^t depends on c_2^τ if there exist b_1, b_2 such that $x_1^t(b_1, b_2, c_2^\tau) \neq x_1^t(b_1, b_2, 1 - c_2^\tau)$.
- If x_1^t has functional dependence on c_2^τ $(\tau < t)$, the auction must observe c_2^τ, in which case $x_2^\tau(b_1, b_2) = 1$.

Structural Restrictions

- Def: Round t is a competitive round for bidder 1 if there exists a high enough bid b_1^t (regardless of b_2^t) such that 1 can win.
- Def: The allocation x_1^t depends on c_2^τ if there exist b_1, b_2 such that $x_1^t(b_1, b_2, c_2^\tau) \neq x_1^t(b_1, b_2, 1 - c_2^\tau)$.
- If x_1^t has functional dependence on c_2^τ ($\tau < t$), the auction must observe c_2^τ, in which case $x_2^\tau(b_1, b_2) = 1$.

Lemma (Competitive Pricing)

1. *If τ is a competitive round w.r.t bidder 1, then x_1^t does not depend on c_2^τ.*
2. *If τ is not competitive w.r.t. bidder 1, then $p_1^\tau = p_2^\tau = 0$.*

Structural Restrictions

- Def: Round t is a competitive round for bidder 1 if there exists a high enough bid b_1^t (regardless of b_2^t) such that 1 can win.
- Def: The allocation x_1^t depends on c_2^τ if there exist b_1, b_2 such that $x_1^t(b_1, b_2, c_2^\tau) \neq x_1^t(b_1, b_2, 1 - c_2^\tau)$.
- If x_1^t has functional dependence on c_2^τ ($\tau < t$), the auction must observe c_2^τ, in which case $x_2^\tau(b_1, b_2) = 1$.

Lemma (Competitive Pricing)

1. *If τ is a competitive round w.r.t bidder 1, then x_1^t does not depend on c_2^τ.*
2. *If τ is not competitive w.r.t. bidder 1, then $p_1^\tau = p_2^\tau = 0$.*

Roughly: 1 applies to exploit rounds and 2 for explore rounds.

Conclusions/Open Problems

- Provided a connection between bandit settings and strategic considerations.
- Truthfulness imposes a statistical restriction.
- Many open problems:
  - budgets?
  - multiple-slots? GSP?
  - side-info? (e.g. keywords)
  - Bayes regret/reserve price setting?
- Future: “Pay-Per-Action” auctions??