Online Learning Applications

Part 7: Learning in truthful auctions with budget constraints

Formal setting

- Sequence of T **truthful** auctions

Formal setting

- Sequence of T **truthful** auctions (for the ease of exposition, we will focus on second-price auctions, i.e., a single slot)

Formal setting

- Sequence of T **truthful** auctions (for the ease of exposition, we will focus on second-price auctions, i.e., a single slot)
- The bidder has a valuation $v \in [0,1]$ (i.e, the utility when the ad is displayed)
- The bidder has an initial budget B

At each round $t \in [T]$:

1. The bidder chooses $b_t \in [0,1]$
2. m_t is the maximum among the competing bids
3. The bidder utility is $f_t(b_t) = (v - m_t)\mathbf{1}[b_t \geq m_t]$
4. The bidder incurs a cost $c_t(b_t) = m_t\mathbf{1}[b_t \geq m_t]$
5. The budget is decreased by $c_t(b_t)$
6. If the budget is smaller than 1 the bidder interaction stops (this avoids spending more than the budget)

Formal setting

Remark

In the previous slide, we assumed that $q_a = 1$ for each $a \in A$. This can be easily relaxed replacing the valuation v with the expected valuation $q_a v$. Then, $f_t(b_t)$ and $c_t(b_t)$ are the **expected** utility and payment.

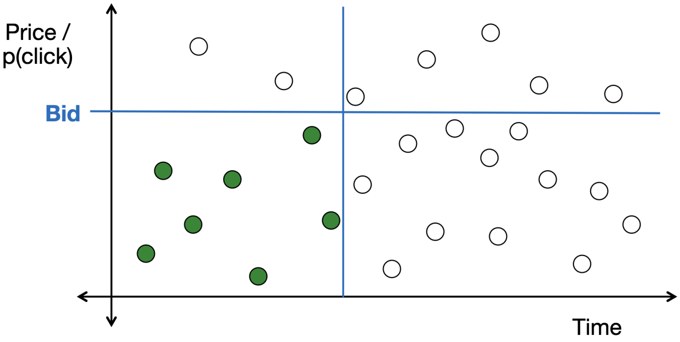
Remark

We assume to observe only $f_t(b_t)$ and $c_t(b_t)$ (instead of m_t):

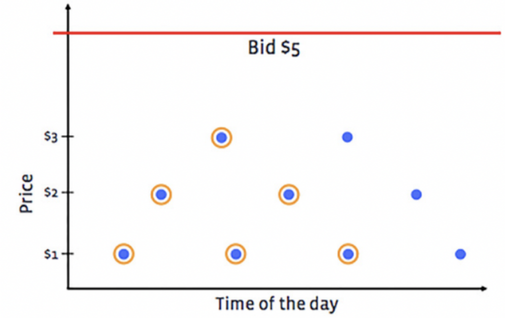
- This is equivalent to **bandit** feedback in MAB
- The bidder knows only if they won the auction and how much they payed.
- The bidder doesn't observe the bids of the other bidders

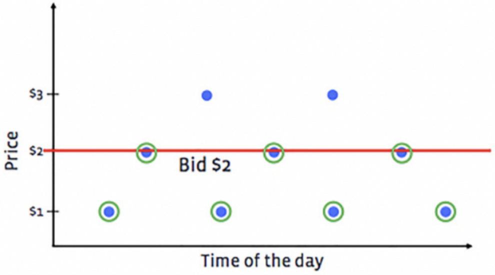
Truthful bidding

Naive way of taking budgets into account: bid as if there was no budget constraint (i.e., truthfully) until budget depletion. At the point stop participating in auctions.



Truthful bidding



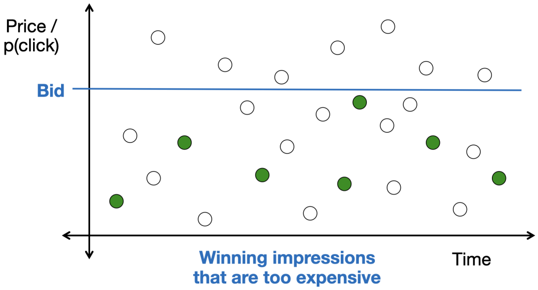


⚠ **Not optimal:** we should bid less aggressively in earlier stages to participate in later auctions.

Picture from Conitzer et al. "Multiplicative Pacing Equilibria in Auction Markets", Operations Research, 2022

Probabilistic pacing

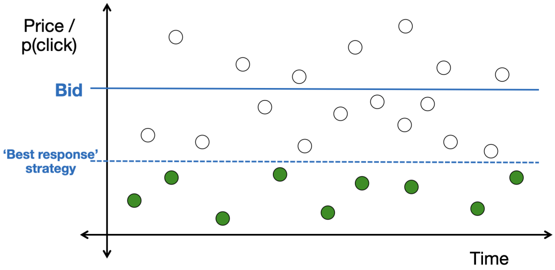
Take part to the auction with some probability.



- **Advantage:** It helps not to deplete the budget
- **Drawback:** It wins auctions that are too expensive

Multiplicative pacing

Scale the bid by some multiplicative factor.



Advantages:

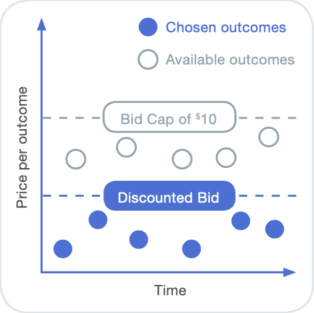
- It helps not to deplete the budget
- It wins the less expensive auctions

Why multiplicative pacing?

Lowest cost bid strategy with standard delivery enabled

This uses “discount” pacing to spend budget on results with the lowest costs and aims to spend budget evenly over the course of your campaign. This option maximizes advertiser value by minimizing your cost per result.

> **Example:** Assume you set a bid cap of $10. As your ad enters each auction, we may “discount” your bid to ensure we can spend your full budget over the duration of the ad set. The mechanism of lowering or “discounting” your bid means that you might not win every auction you could have, but you’ll have a chance to capture more outcomes at more efficient costs over the course of your campaign, instead of using up your budget too quickly on more expensive results. Note that the discounted bid for a particular auction might be higher or lower based on how much budget has been spent and the time elapsed in the campaign.



Source: Meta’s Guide for Advertisers

Environment and baseline

We consider two possible sequences of m_t:

- **Stochastic**: m_t is sampled from a distribution D
- **Adversarial**: no assumptions on m_t

We want to have **no-regret** with respect to:

Baseline

The reward of the best dynamic and feasible sequence of bids when the input sequence is known in advance (i.e., solve the optimal allocation under full information)

Environment and baseline

We consider two possible sequences of m_t:

- **Stochastic**: m_t is sampled from a distribution D
- **Adversarial**: no assumptions on m_t

We want to have **no-regret** with respect to:

Baseline

The reward of the best dynamic and feasible sequence of bids when the input sequence is known in advance (i.e., solve the optimal allocation under full information)

⚠ Stronger than the usual baseline (i.e., the best **fixed** action in insight). We can strengthen the baseline since the auction is truthful and hence the unconstrained problem is trivial.

Multiplicative pacing

idea:

- Use **multiplicative pacing** with a parameter λ
- Bid $b_t = \frac{v}{1+\lambda}$
- Increase the parameter λ when we spent more than the per-round budget $\rho = B/T$ and decrease λ otherwise

Result:

- The bid decreases when we are spending too much (more than the per-round budget) and increases if we are not spending enough (less than the per-round budget)

Pacing strategy

Algorithm: Pacing strategy

input: Budget B, number of rounds T, learning rate η;
initialization: $\rho \leftarrow B/T, \lambda \leftarrow 0$;
for $t = 1, 2, \ldots, T$ **do**
  bid $b_t = \frac{v}{1+\lambda}$;
  observe $f_t(b_t)$ and $c_t(b_t)$;
  $\lambda \leftarrow \Pi_{[0,1/\rho]}(\lambda - \eta(\rho - c_t(b_t)))$;
  $B \leftarrow B - c_t(b_t)$;
  if $B < 1$ **then**
    terminate;

Pacing strategy

Algorithm: Pacing strategy

input: Budget B, number of rounds T, learning rate η;
initialization: $\rho \leftarrow B/T, \lambda \leftarrow 0$;
for $t = 1, 2, \dots, T$ **do**
    bid $b_t = \frac{v}{1+\lambda}$;
    observe $f_t(b_t)$ and $c_t(b_t)$;
    $\lambda \leftarrow \Pi_{[0,1/\rho]}(\lambda - \eta(\rho - c_t(b_t))) \rightarrow$ gradient descent on the loss $\lambda(\rho - c_t(b_t))$;
    $B \leftarrow B - c_t(b_t)$;
    if $B < 1$ **then**
        terminate;

Regret guarantees: stochastic

Theorem [Balseiro and Gur, 2019]

Assume the sequence of m_t is stochastic. The pacing strategy with $\eta = T^{-1/2}$ guarantees regret

$$O(\sqrt{T}),$$

where we ignore the dependency on the other parameters.

Regret guarantees: stochastic

Theorem [Balseiro and Gur, 2019]

Assume the sequence of m_t is stochastic. The pacing strategy with $\eta = T^{-1/2}$ guarantees regret

$$O(\sqrt{T}),$$

where we ignore the dependency on the other parameters.

⚠ The result follows by Lagrangian duality. We will see more details for the case of non-truthful auctions.

Lower bound for adversarial environment

Theorem [Balseiro and Gur, 2019]

No algorithm can achieve strictly more than a $\rho := B/T$ fraction of the optimal utility.

proof Sketch for $\rho = 1/2$.

- Valuation is $v = 1$
- $1 > m_{high} > m_{low} > 0$
- $1 - m_{high}$ is much smaller than $1 - m_{small}$
- $v \approx m_{high} \approx m_{low}$, implying that every time the bidder wins an auction he pays something close to 1.
- The vector of competing bid is:
  - $m^1 = (m_{high}, m_{high}, \dots m_{high}, v, v, \dots, v)$, or
  - $m^2 = (m_{high}, m_{high}, \dots m_{high}, m_{low}, m_{low}, \dots, m_{low})$

$\square$

Lower bound for adversarial environment

proof Sketch for $\rho = 1/2$.

- The bidder doesn't know the sequence of competing bids
- The sequence m^1 and m^2 are equivalent until round $T/2$
- If the bidder spends **more** than half of the budget in the first $T/2$ rounds:
  - Suppose the vector is m^2
  - Can win at most $T/4$ of the last (cheap) auctions
  - The utility in the first $T/2$ auctions is negligible
  - The optimal policy wins all the final $T/2$ auctions
  - The algorithm extracts a $\frac{(T/4)(1-m_{low})}{(T/2)(1-m_{low})} = \frac{1}{2}$ fraction of the optimal utility

$\square$

Lower bound for adversarial environment

proof Sketch for $\rho = 1/2$.

- The bidder doesn't know the sequence of competing bids
- The sequence m^1 and m^2 are equivalent until round $T/2$
- If the bidder spends **less** than half of the budget in the first $T/2$ rounds
  - Suppose the vector is m^1
  - Can win at most $T/4$ of the first auctions
  - In the final $T/2$ auctions the utility is 0 even if the bidder wins
  - The optimal policy wins all the first $T/2$ auctions
  - The algorithm extracts a $\frac{(T/4)(1-m_{high})}{(T/2)(1-m_{high})} = \frac{1}{2}$ fraction of the optimal utility

$\square$

Regret guarantees: adversarial

Theorem [Balseiro and Gur, 2019]

The pacing strategy with $\eta = T^{-1/2}$ guarantees utility at least:

$$\rho OPT - O(T^{-1/2}),$$

where

- OPT is the expected reward of the baseline
- $\rho := B/T$ is the per-round budget
- We ignore the dependency on the other parameters

Regret guarantees: adversarial

Theorem [Balseiro and Gur, 2019]

The pacing strategy with $\eta = T^{-1/2}$ guarantees utility at least:

$$\rho OPT - O(T^{-1/2}),$$

where

- OPT is the expected reward of the baseline
- $\rho := B/T$ is the per-round budget
- We ignore the dependency on the other parameters

- If the environment is well-behaved then we can expect much better performances
- If the environment changes "slightly" the guarantees approaches $O(\sqrt{T})$ regret

References

Santiago R Balseiro and Yonatan Gur. Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Science*, 65(9):3952–3968, 2019.