Online Second Price Auction with Semi-bandit Feedback Under the Non-Stationary Setting
ID: 4617

Haoyu Zhao, Tsinghua University

hyzhao.me, zhaohy@mails.tsinghua.edu.cn

Joint work with Prof. Wei Chen @ MSRA

清華大學
Tsinghua University

CONTENTS

1

Introduction

Online second-price auction with semi-bandit feedback

We will use the second price auction,
and the reserve price in this round is p.

My value $v_1 \geq p$, and I have a change to win. I will report this value

My value $v_2 < p$, and I will never get the item. I will not report my value.

Online second-price auction with semi-bandit feedback

1. There are m bidders in total.

2. The private value distribution for the bidders is D_t at time t, which is known.

3. At the beginning of round t, we can release our reserve price p_t.

4. The environment draws a private value sample $X_t \sim D_t$.

5. We observe all $X_{i,t}$ with $X_{i,t} \geq p_t$.

6. Our revenue can be computed by the second price auction protocol.

Non-stationary

The distribution D_t are changing through time.

2 measurements:

switching: $S = 1 + \sum_{t=2}^{T} 1(D_t \neq D_{t-1})$

total variation: $V = \sum_{t=2}^{T} \left|\left|D_t - D_{t-1}\right|\right|_{TV}$

Non-stationary (pseudo) regret:

$Reg = E[\sum_{t=1}^{T} \left(R(D_t, p_t^*) - R(D_t, p_t)\right)]$

2

Related Works

Non-stationary MAB

There are K arms, all of them follows a joint distribution D_t

The distribution D_t are changing through time. $X_t \sim D_t$

2 measurements:

switching: $S = 1 + \sum_{t=2}^{T} 1(D_t \neq D_{t-1})$

variation: $V = \sum_{t=2}^{T} \left|\left|E[X_t] - E[X_{t-1}]\right|\right|_\infty$

Non-stationary (pseudo) regret:

$Reg = E[\sum_{t=1}^{T} (\mu_t^* - \mu_t)]$

Non-stationary MAB

The "leading term" in the regret bound is always $O(S^{\gamma_1}T^{\gamma_2}), O(V^{\gamma_3}T^{\gamma_4})$ under different measurement.

$$\gamma_1 + \gamma_2 \geq 1, \gamma_3 + \gamma_4 \geq 1$$

This is because if we allow the distribution to change arbitrarily, we cannot get "sublinear regret".

We hope that we can get regret bound $\gamma_1 + \gamma_2 = 1, \gamma_3 + \gamma_4 = 1$.

If possible, we want the exponential term for S, V are large, since $S, V \leq T$

Non-stationary MAB

We love parameter-free algorithm: we do not know S, V. There are many algorithm that need to know S, V to tune parameters.

There are parameter-free algorithms for MAB, the regret is

$$O(\min\left(\sqrt{ST}, V^{\frac{1}{3}}T^{\frac{2}{3}}\right))$$

Which are nearly optimal (the factor on S,V cannot be larger)

3

Model & Techniques

Online second-price auction – stationary case

Our general technique: Reduce the online second-price auction into a multi-armed bandit problem.

The second-price auction has *one-sided Lipschitz property*: if $p \geq p'$, then $\mathrm{R}(D, p) - \mathrm{R}(D, p') \leq p - p'$.

Because of the Lipschitz property, we can *discretize the reserve prices*.

Online second-price auction – stationary case

Our general technique: Reduce the online second-price auction into a multi-armed bandit problem.

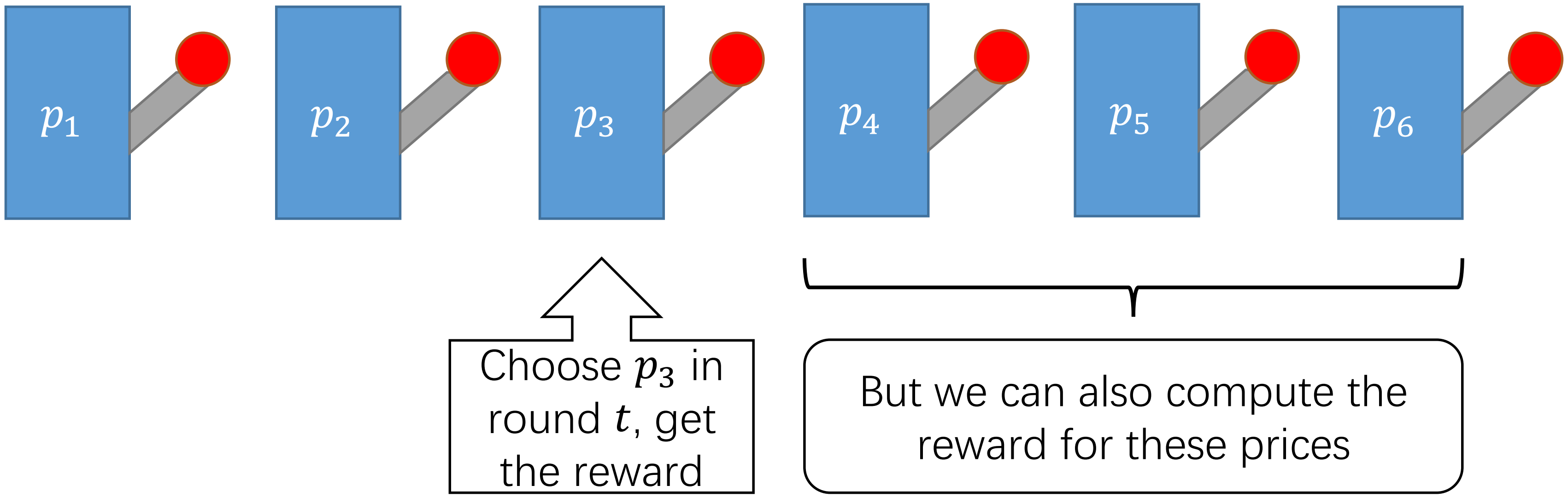
Discretize the reserve prices $p_1 = 0, p_2 = \frac{1}{K}, \dots, p_K = \frac{K-1}{K}$.

However, to reduce the discretization error, we need to set $K = poly(T)$. Recall that the regret lower bound for multi-armed bandit is $\tilde{\boldsymbol{O}}(\sqrt{KT})$, in the stationary case, we cannot get regret to $\tilde{\boldsymbol{O}}(\sqrt{T})$.

Online second-price auction – stationary case

Is there any hope? Yes, we have richer feedback instead of the bandit feedback.

We call this: one-sided feedback



Online second-price auction – stationary case

The algorithm for the stationary case is simple:

1. Start from price p_1, after the reserve price p_1 is "too bad", we eliminate p_1, and start to use p_2.
2. After p_2 is too bad, we eliminate p_2 and use p_3.
3. Repeat the process

The algorithm works really well. In the stationary case, the regret is $\tilde{O}(\sqrt{T})$, which is tight.

Online second-price auction – non-stationary case

The algorithm for the stationary case is simple:

1. Start from price p_1, after the reserve price p_1 is "too bad", we eliminate p_1, and start to use p_2.
2. After p_2 is too bad, we eliminate p_2 and use p_3.
3. Repeat the process

However, in the non-stationary case, the distributions are changing.

An eliminated reserve price may perform really well after some time.

Solution: Recheck the eliminated arms with "some probability", if we detect the non-stationary, we restart the algorithm.

Online second-price auction – non-stationary case

The algorithm for the non-stationary case is build on the stationary case algorithm:

1. Start from price p_1, after the reserve price p_1 is "too bad", we eliminate p_1, and start to use p_2.
2. After p_2 is too bad, we eliminate p_2 and use p_3.
3. Repeat the process

4. In each round, we may use the eliminated reserve price with some probability
5. We check if the non-stationarity happens, if yes, we restart the algorithm.

How to detect the non-stationary?

We check if the mean changes a lot. For example,

$$|\mu[t_1, t_2] - \mu[t_3, t_4]| \geq c(t_2 - t_1) + c(t_4 - t_3)$$

$c()$ stands for the confidence radius.

4

Our Results

Online second-price auction – non-stationary case

In the switching case, the expected (pseudo) regret is $O(\sqrt{ST})$ when omitting log factors

In the dynamic case, the expected (pseudo) regret is $O(V^{\frac{1}{3}}T^{\frac{2}{3}})$ when omitting log factors

Both of them are nearly optimal (the regret bound matches when omitting log factors)

Online second-price auction – non-stationary case

To our best knowledge, this work is the first to theoretically consider the non-stationarity in mechanism design without full-information.

Detecting non-stationarity with full-information is rather easy.

Using the techniques in this work, we also generalize our result to the bandit setting, where we can only get the revenue in each round. However, we need to assume that the bidders are i.i.d and the number is at least 2.

Online second-price auction – non-stationary case

Some further directions:

The mechanism in our case is simple, and we do not assume that the agents are **strategic.** Considering the dynamic mechanism design with strategic agents are also interesting! Even in the full information case!

Thanks ! Any questions?

This ppt has already been uploaded to hyzhao.me/talks

Also welcome to see my poster. You can also ask me questions through emails.

清華大學
Tsinghua University