Online Pricing for Multi-User Multi-Item Markets

Yigit Efe Erginbas Thomas Courtade Kannan Ramchandran Soham Phade

BL SS Berkeley Laboratory for Information and System Sciences



Introduction

- Large-scale online marketplace.
- Unknown user valuations.
- Learning through interactions.
- Accept/reject feedback only.
- **Goal:** achieving maximal revenue.
- Optimizing the *offers* and *pricing*.



Problem Setting

- Set of users \mathcal{N} of size N.
- Set of items \mathcal{I} of size M.

At each round $t \in 1, \ldots, T$:

- A subset of users $\mathcal{D}^t \subseteq \mathcal{N}$ are active.
- A subset of items $\mathcal{E}^t \subseteq \mathcal{I}$ are available for sale.
- Each user u has a different valuation v_{ui}^t for each item i.
- The provider decides on

offers
$$\mathbf{X}^t \in \{0,1\}^{N \times M}$$
 and prices $\mathbf{p}^t \in \mathbb{R}^M_+$

 \circ User u accepts an offered item i if and only if

 $v_{ui}^t \ge p_i^t.$



Goal

The goal is to minimize the revenue regret defined as

$$\mathcal{R}(T, \boldsymbol{\pi}) = \text{OPT} - \sum_{t=1}^{T} \sum_{u \in \mathcal{D}^{t}} \sum_{i \in \mathcal{E}^{t}} x_{ui}^{t} p_{i}^{t} \mathbb{1}\{v_{ui}^{t} \ge p_{i}^{t}\}$$

• OPT denotes the revenue of the optimal algorithm (separately for each different valuation model).

Main Theoretical Results

- Design algorithms to maximize revenue
- Three different valuation models:
 - Fixed valuations
 - Random experiences
 - Random valuations
- Regret upper bounds achieved by our algorithms
- Regret lower bounds for multi-user multi-item online learning

Main Theoretical Results

- Regret upper bounds and lower bounds
- Our algorithms are optimal up to logarithmic factors!

| Model | Regret Upper Bound | Regret Lower Bound |
|--------------------|--|---------------------------------------|
| Fixed Valuations | $O\left(NM\log\log(LT) ight)$ | $\Omega\left(NM\log\log(LT/NM) ight)$ |
| Random Experiences | $\widetilde{O}\left(\sqrt{NMLT} ight)$ | $\Omega\left(\sqrt{NMLT}\right)$ |
| Random Valuations | $\widetilde{O}\left(\sqrt{NMLT} ight)$ | $\Omega\left(\sqrt{NMLT}\right)$ |

• Load parameter:
$$L = \max_{1 \le t \le T} \left\{ \min \left(|\mathcal{D}^t|, |\mathcal{E}^t| \right) \right\}$$

(roughly, the maximum number of simultaneous allocations possible)

Model 1: Fixed Valuations

- The valuations of users **do not change over time**.
- Formally, there exist values v_{ui} such that

$$v_{ui}^t = v_{ui}$$

for all $t \in [T]$.

Model 1: Fixed Valuations

Algorithm for fixed valuations:



Initialize the interval $[a_{ui}, b_{ui}] = [0, 1]$ for each user-item pair. Initialize the search step $\beta_{ui} = 1/2$ for each user-item pair.

At each round *t*:

- 1. Choose \mathbf{X}^t as the optimum offers under valuations b_{ui} .
- 2. For each offer (u, i) in \mathbf{X}^t do:



Model 2: Random Experiences

- The valuations of users are given as their **average historical experience**.
- Formally, the valuations are given as

$$v_{ui}^t = \arg\{z_{ui}^\tau | \tau < t, x_{ui}^\tau = 1, v_{ui}^\tau \ge p_i^\tau\}$$

for all $t \in [T]$.

• Each random variable z_{ui}^t represents the experience of a user with an accepted item.

Model 2: Random Experiences

• Algorithm for random experiences:



Initialize the interval $[a_{ui}, b_{ui}] = [0, 1]$ for each user-item pair. At each round *t*:

- 1. Choose \mathbf{X}^t as the optimum offers under valuations b_{ui} .
- 2. For each offer (u, i) in \mathbf{X}^t do:



Model 3: Random Valuations

- The valuations of users at different rounds are **independently drawn**.
- Formally, each valuation v_{ui}^t is drawn from a distribution with c.d.f. F_{ui} specific to each user and each item.

Model 3: Random Valuations

• **Algorithm** for random valuations:



Set $K = (LT/(NM \log(LT)))^{1/4}$. Quantize prices into $\{1/K, 2/K, \dots, 1\}$. At each round t: 1. Set $b_{uik} = \psi_{uik} + \sqrt{\frac{8\log(NMKT)}{n_{uik}}}$ 2. Find $b_{ui} = \max_k b_{uik}$. 3. Choose \mathbf{X}^t as the optimum offers under valuations b_{ui} . 4. For each offer (u, i) in \mathbf{X}^t do: • Set $k^* = \arg \max_k b_{uik}$. • Offer item i to user u at price k^*/K . • Update $\psi_{uik^*} \leftarrow \frac{n_{uik^*}\psi_{uik^*} + p_{ui}\mathbb{1}\{v_{ui}^t \ge p_{ui}\}}{n_{uik^*} + 1}$.

Numerical Experiments

• Instantaneous Regret under different valuation models.



Numerical Experiments

Regret vs. Time Horizon under different valuation models.



- Numerical results verify that our algorithms can achieve
 - o sub-logarithmic regret under the fixed valuations model,
 - sub-linear regret under random experiences model,
 - sub-linear regret under random valuations model.