Online Pricing for Revenue Maximization with Unknown Time Discounting Valuations

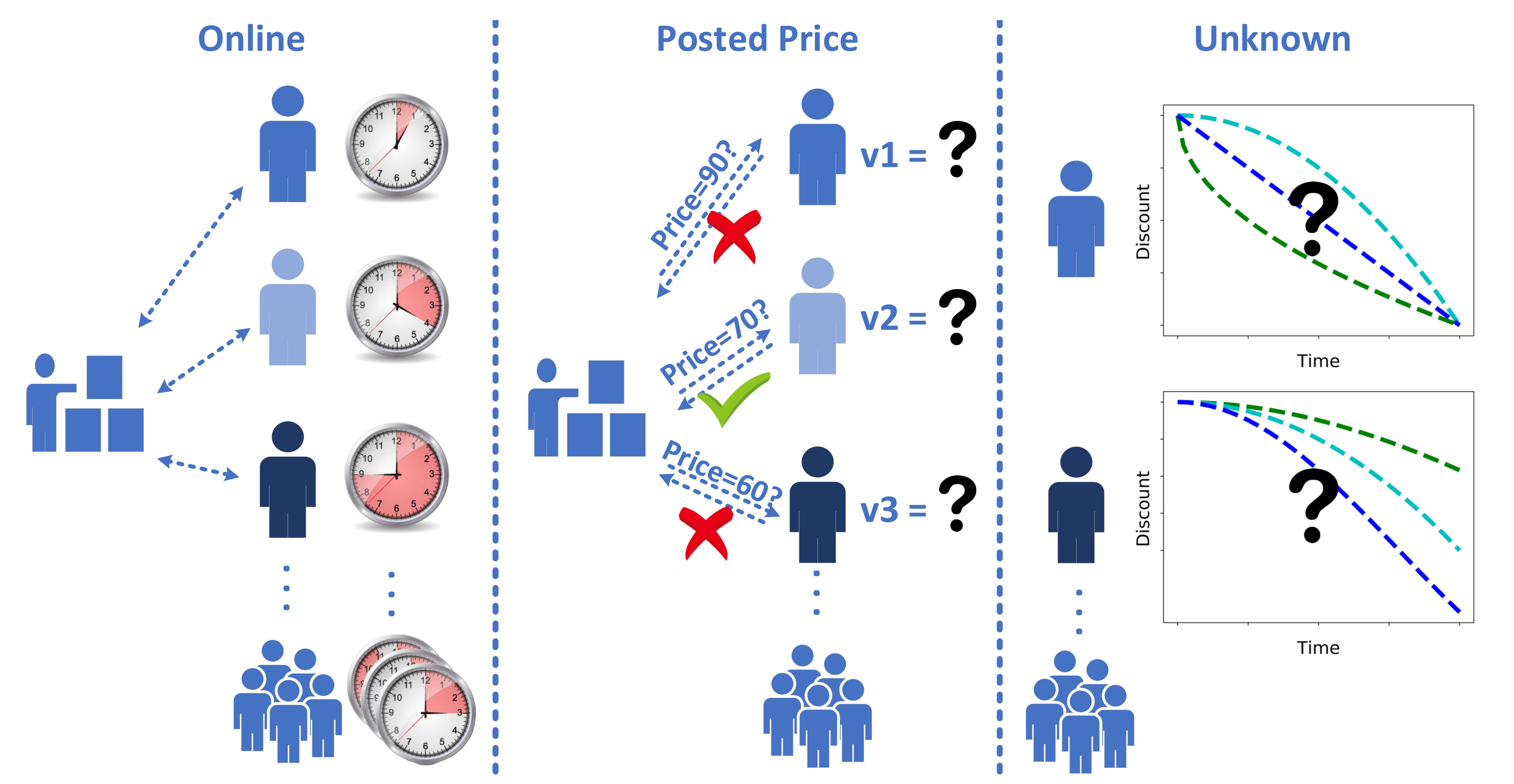
Weichao Mao, Zhenzhe Zheng, Fan Wu, Guihai Chen
Department of Computer Science, Shanghai Jiao Tong University, Shanghai, China
{maoweichao,zhengzhenzhe}@sjtu.edu.cn, {fwu,gchen}@cs.sjtu.edu.cn

1. Introduction

Online pricing mechanisms have been widely applied to resource allocation in multi-agent systems. Most existing online pricing mechanisms assume buyers have fixed valuations over the time horizon. However, in many emerging applications, buyer valuations are discounting over time. For example, advertisers' willingness-to-pay usually decays over the time horizon in current mobile ad auctions, where ad space is sold in different time slots [Mehta et al., 2017]. In crowd-sensed data marketplaces [Zheng et al., 2017], sensor data become less informative over time, leading to the decrease of buyer valuations. Herein we model the revenue maximization problem with discounting buyer valuations as non-stationary multi-armed bandit optimization, and present our pricing mechanism namely Biased-UCB.

2. Design Objective

Design a revenue-maximized online posted price mechanism for digital goods that can handle unknown time discounting buyer valuations.



- Online: Buyers show up one at a time, and the seller has no information about future buyers when dealing with the current buyer.
- Posted price: The seller never learns the valuation of any buyer, but only observes whether the buyer accepts his price.
- Unknown: Buyers have heterogeneous discounting functions, and the seller does not know any of them.

3. Preliminaries

We divide the horizon into T time slots, with one buyer showing up at each slot. Seller offers price p_t to the buyer at slot t. The discounted valuation of buyer i is $v_i(t) = v_i \cdot d_i(t)$, where $v_i \in [1, \bar{v}]$ is her original valuation and $d_i(t) : [T] \to (0, 1]$ is her discounting function.
The seller chooses prices from a discrete candidate price vector $\hat{\boldsymbol{p}} = (\hat{p}_0, \hat{p}_1, \cdots, \hat{p}_H)$, where $\hat{p}_k = (1+\beta)^k$, $\beta > 0$ and $H = \lfloor \log_{1+\beta} \bar{v} \rfloor$. Buyer i at slot t accepts the price if and only if $v_i(t) \geq p_t$.

4. Mechanism

Our mechanism Biased-UCB basically follows the Upper Confidence Bound (UCB) framework from multi-armed bandit (MAB) problems, but with several modifications on the weight function and its update rule.
Weight Function: The classical UCB1 algorithm from [Auer et al., 2002] keeps a record of the average reward u_i of each candidate price, and uses the number of trials n_i to denote the uncertainty of the price. The weight of candidate price $\hat{p}_i$ is defined as follows by a combination of its empirical estimation and uncertainty. At each slot t, the candidate price with highest weight is offered to the buyer.

$$\underbrace{u_i}_{\text{Estimation}} + \underbrace{\sqrt{\frac{c \cdot \ln(n-1)}{n_i}}}_{\text{Uncertainty}}$$

Average Reward ← u_i; n_i → Number of Trials

$$n_i = \sum_{t=0}^{n-1} n_{i,t} \qquad u_i = \frac{\sum_{t=0}^{n-1} r_{i,t}}{n_i}$$

In our problem, however, since the buyer valuations are varying, we are confronted with a non-stationary MAB problem. In this case, we value the recent information more than the historical records a long time ago, and introduce an attenuation parameter γ $(0 < \gamma \leq 1)$ to make the value of information decay over time. The weight function in Biased-UCB is as follows:

$$\hat{u}_i + \sqrt{\frac{c \cdot \ln(n-1)}{\hat{n}_i}} \qquad \hat{n}_i = \sum_{t=0}^{n-1} n_{i,t} \cdot \gamma^{n-t} \qquad \hat{u}_i = \frac{\sum_{t=0}^{n-1} r_{i,t} \cdot \gamma^{n-t}}{\hat{n}_i}$$

γ^{n-t} → Attenuation Factor

Update Rule: Since buyer valuations are generally discounting, we design our update rule to be biased—it always encourages lower prices and suppresses higher prices. Instead of only updating the weight of one particular price $\hat{p}_k$ as in UCB1, we update the weights of all the prices no higher than (or no lower than) this offered price, based on the buyer's response to this price. In the example below, when $\hat{p}_2$ is accepted by the buyer, the weights of $\hat{p}_0$, $\hat{p}_1$ and $\hat{p}_2$ will be all updated by their corresponding expected reward.

Price	$\hat{p}_0$	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$
Result	✓	✓	✓		
Reward	$(1+\beta)^0$	$(1+\beta)^1$	$(1+\beta)^2$		

Price	$\hat{p}_0$	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$
Result			✗	✗	✗
Reward			0	0	0

5. Theoretical Bound

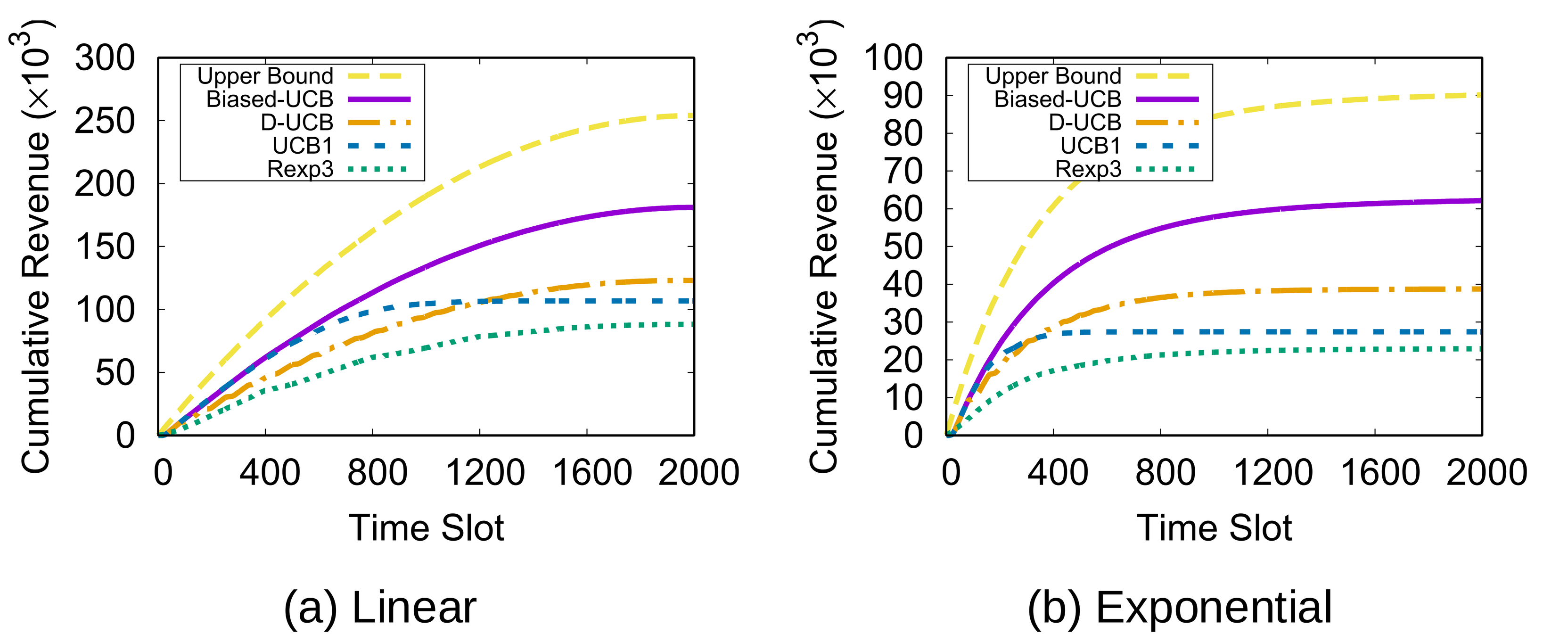
The worst-case competitive ratio of Biased-UCB towards the *ex ante* optimal strategy is lower bounded by the following expression. A typical value of this expression is 0.1. This bound seems relatively weak, but a non-trivial example shows the worst-case revenue of any deterministic mechanism is upper bounded by 0.278 of optimal revenue.

$$\frac{1}{\eta} \cdot \left[1 - \beta r - \frac{2\delta(H+1)}{T}\right] \cdot \left(1 - \frac{1}{1+\beta r}\right)$$

- $\frac{1}{\eta}$: Difference Between $d_i(t)$
- $\left[1 - \beta r - \frac{2\delta(H+1)}{T}\right]$: Abandoned Segments
- $\left(1 - \frac{1}{1+\beta r}\right)$: Abandoned Slots in Remaining Segments

6. Evaluation

The evaluation results of Biased-UCB on the iPinYou dataset [Zhang et al., 2014] with linear and exponential discounting functions are as follows. Four benchmarks include UCB1, D-UCB and Rexp3 modified from existing MAB algorithms, and the revenue upper bound.



(a) Linear (b) Exponential