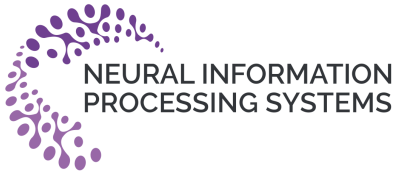


Product Ranking for Revenue Maximization with Multiple Purchases

Renzhe Xu[1], Xingxuan Zhang[1], Bo Li[1], Yafeng Zhang[2], Xiaolong Chen[2], Peng Cui[1]

[1]Tsinghua University

[2]Meituan Inc.

Problem setting

- **Basic setting**
  - **Product ranking** is the core problem for revenue-maximizing online retailers
- **Related works**
  - Most existing works suppose each consumer purchases **at most one product**
- **In this paper**,
  - We propose a more realistic consumer choice model to characterize consumer behaviors under **multiple-purchase** settings

Method

- We study the optimal product ranking policy to maximize online retailers' revenue in both offline and online settings
  - Achieve $\tilde{O}(\sqrt{T})$ regret on the revenue in online settings

Algorithm 1: MPB-UCB (Non-contextual)

Input: Products revenue $\boldsymbol{r}$ and hyper-parameter ϵ_Q.
Initialization: $\tilde{\lambda}_{0,k} = 1$ for $k \in [N]$, $\tilde{q}_0 = 1 - \epsilon_Q$, $\tilde{w}_0 = 1 - \epsilon_Q$.
for *t=1:T* **do**
  Let $\boldsymbol{\sigma}_t$ be the optimal offline policy from Theorem 4.1 with $\boldsymbol{\lambda} = \tilde{\boldsymbol{\lambda}}_{t-1}$, $q = \tilde{q}_{t-1}$, and $s = \tilde{w}_{t-1}/\tilde{q}_{t-1}$.
  Offer ranking policy $\boldsymbol{\sigma}_t$ and observe Φ_t, $\boldsymbol{\Gamma}_t$, $\boldsymbol{\eta}_t$, $\boldsymbol{\mu}_t$.
  Update statistics $C_{t,k}$, $c_{t,k}$, D_t^Q, d_t^Q, D_t^W, and d_t^W.
  Calculate $\tilde{\lambda}_{t,k}$, $\tilde{q}_t$, and $\tilde{w}_t$ by Equations (6) and (8).
end

Algorithm 1: MPB-UCB (Non-contextual)

Algorithm 2: MPB-UCB (Contextual)

Input: Products revenue $\boldsymbol{r}$ and hyper-parameter ϵ_Q, α_Λ, α_Q, and α_W.
Initialization: $\hat{\beta}_0^\Lambda = \mathbf{0}$, $\hat{\beta}_0^Q = \mathbf{0}$, and $\hat{\beta}_0^W = \mathbf{0}$.
for *t=1:T* **do**
  Observe consumer features x_t and $y_{t,k}$ for $k \in [N]$. Let $z_t = \mathrm{vec}(x_t x_t^\top)$.
  Calculate $\tilde{\lambda}_{t,k}$, $\tilde{q}_t$, and $\tilde{w}_t$ according to Equation (16).
  Let $\boldsymbol{\sigma}_t$ be the optimal ranking policy from Theorem 4.1 with $\boldsymbol{\lambda} = \tilde{\boldsymbol{\lambda}}_t$, $q = \tilde{q}_t$, and $s = \tilde{w}_t/\tilde{q}_t$.
  Offer ranking policy $\boldsymbol{\sigma}_t$ and observe Φ_t, $\boldsymbol{\Gamma}_t$, $\boldsymbol{\eta}_t$, $\boldsymbol{\mu}_t$.
  Calculate statistics Σ_t^Λ, ρ_t^Λ, Σ_t^Q, ρ_t^Q, Σ_t^W, and ρ_t^W.
  Calculate estimated parameters $\hat{\beta}_t^\Lambda$, $\hat{\beta}_t^Q$, and $\hat{\beta}_t^W$ according to Equations (12), (13) and (14).
end

Algorithm 2: MPB-UCB (Contextual)

Experiments

- We conduct experiments on both synthetic data and semi-synthetic.

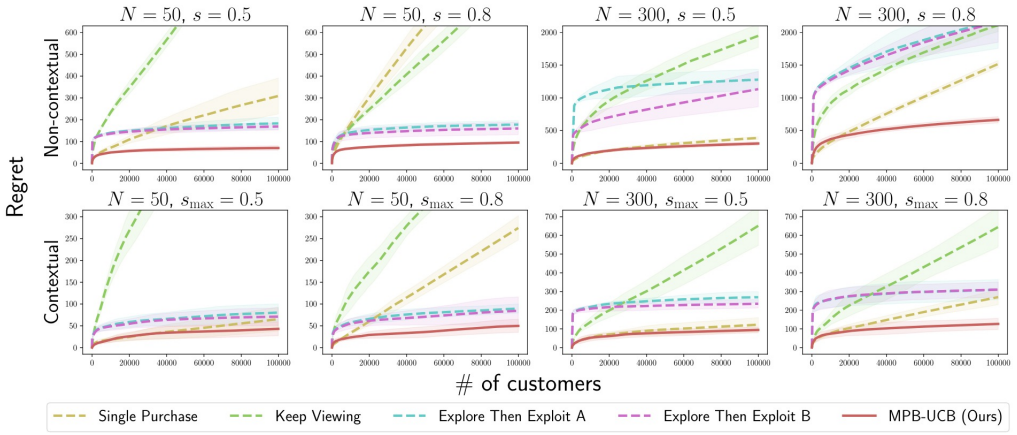


Figure 1: Results on the synthetic data

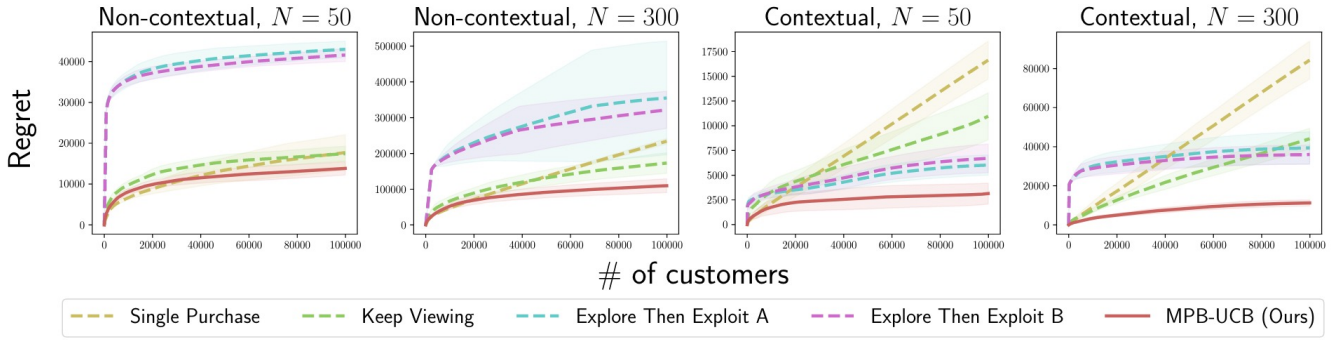
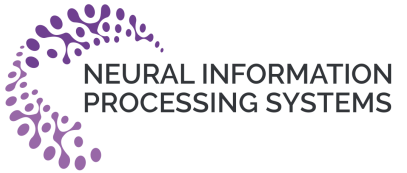


Figure 2: Results on the semi-synthetic data

Meituan



Thanks!

Paper is available at https://arxiv.org/abs/2210.08268

Code is available at https://github.com/windxrz/MPB-UCB