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.

General idea

- **Product ranking** is the core problem for revenue-maximizing online retailers
- 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
  - We study the optimal product ranking policy to maximize online retailers' revenue in both offline and online settings

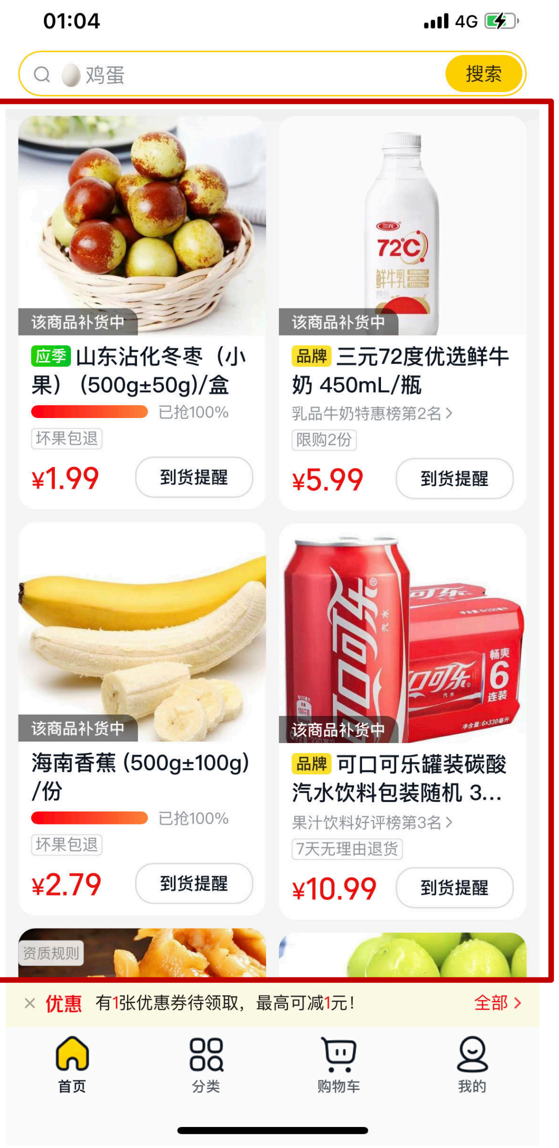
Problem setting

- Consider online retailer platforms
  - The platform recommends a list of products to consumers
  - Consumers purchase products according to a choice model

- Characterize consumer choice model with multiple purchases
  - Each Consumer views the list sequentially
  - Attention span and purchase budget
    - maximal number of products that the consumer is willing to view / purchase
    - They are random and obey geometric distributions

- Target: the total revenue achieved by the online retailer

Product list



Proposed method --- offline setting

- Optimal ranking policy when given a consumer's characteristics
  - Sort products in descending order according to the following score

$$\frac{\lambda_i r_i}{1 - q + q(1 - s)\lambda_i}$$

  - r_i: the revenue of product i
  - λ_i: the purchase probability for product i
  - q, s: the geometric distribution parameters *w.r.t.* attention span and purchase budget

- Special case --- $s = 0$
  - The consumer will purchase at most one product
  - The result becomes the same as the ranking policy in [1], which considers the single-purchase setting

[1] Cao, Junyu, and Wei Sun. "Dynamic learning of sequential choice bandit problem under marketing fatigue." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. No. 01. 2019.

Proposed method --- online setting

- Online Learning of the ranking policy
  - The online retailer has no prior knowledge about consumers' characteristics
  - We consider two settings
    - Non-contextual setting: all consumers share the same parameters
    - Contextual setting: consumers have personalized behaviors
  - We develop the Multiple-Purchase-with-Budget UCB (MPB-UCB) algorithms
    - Achieve $\tilde{O}(\sqrt{T})$ regret on the revenue

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

- Conduct experiments on both synthetic data and semi-synthetic data
- We plot the regret curve for different settings
- MPB-UCB (Ours) achieves the best performance
  - MPB-UCB beats Single Purchase and Keep Viewing
    - Baselines consider different consumer choice models
  - MPB-UCB beats explore-then-exploit-based methods
    - We have a better exploration-exploitation trade-off

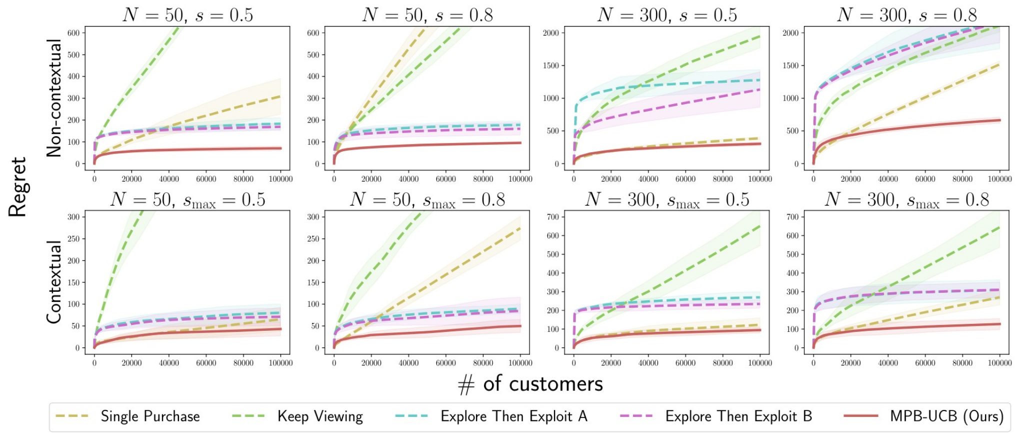


Figure 1: Results on the synthetic data

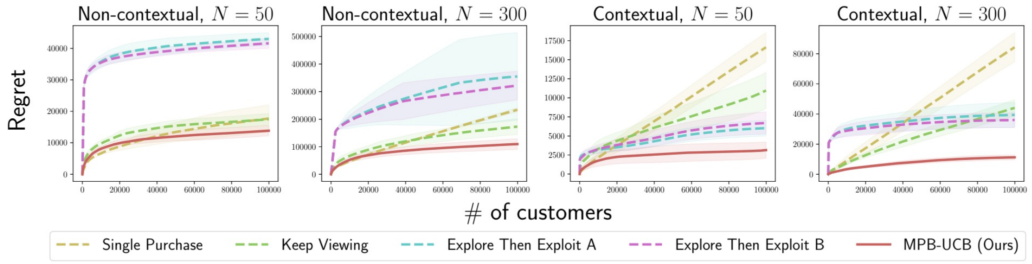


Figure 2: Results on the semi-synthetic data

Thanks!

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

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