Pricing under the Generalized Markov Chain Choice Model: Learning through Large-scale Click Behaviors

Mo Liu

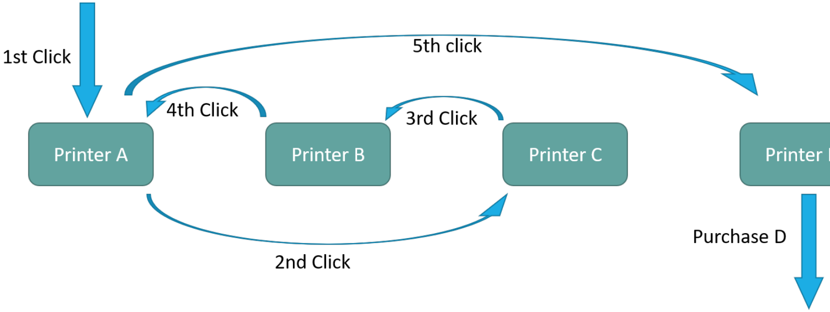
Joint work with Junyu Cao and Zuo-jun Max Shen

Department of Industrial Engineering and Operations Research
UC Berkeley

INFORMS 2022

Motivation: Large-scale click data reveals the preference of customers.

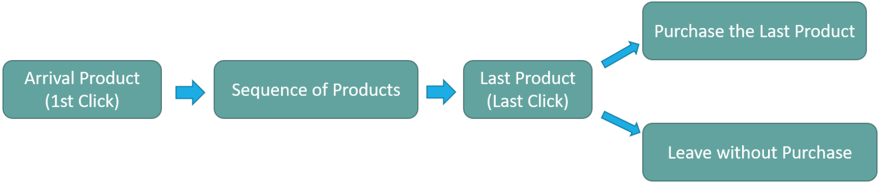
The click behaviors of customers are available to online retailers.
Example: Buy a printer at Amazon.com.



- Click data reveals the search and comparing behaviors of customers before purchasing or leaving the system.

Motivation: click trajectories

- How to use click behaviors to learn the preference of customers?
- Click trajectories:

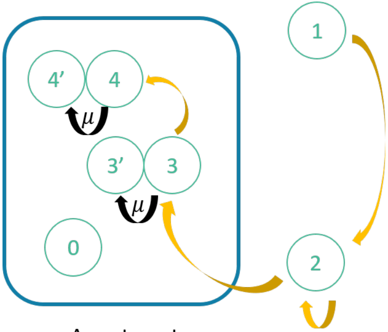


- ☐ Click trajectories have more information than click-through rate on a single product.
- ☐ Click trajectories are random and contain the back-and-forth transitions.

- How to model this random click trajectories among millions of products?
- How to consider the effect of recommendation on the click trajectories?

Generalized Markov Chain Choice Model (GMCCM)

- GMCCM is a choice model, independent of click behaviors.
- Proposed in [Goutam et al., 2019], and [Dong et al., 2019].
  - State i: product i.
  - State 0: no-purchase state.
- State Transition:
  - If the current state is outside the assortment, keep transitioning.
  - If the current state is within the assortment, purchase it and leave the system with probability μ, otherwise keep transitioning.
- Three types of parameters:
  - Transition matrix
  - Arrival probability
  - Instant purchase probability μ, which is a function of price, assortment, and product.



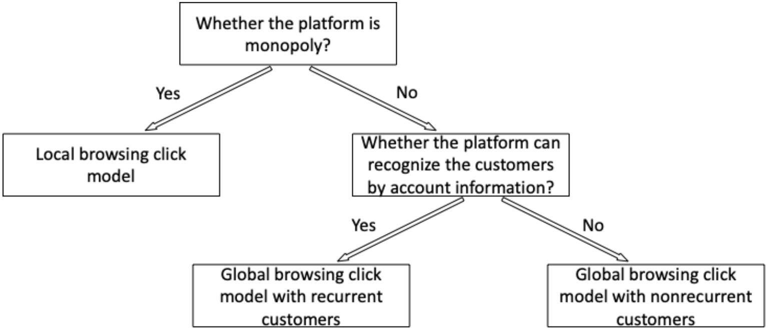
Assortment

Three click models

Assumption 1

A customer clicks product i if she transits to the state of product i in the GMCCM, given that product i is in the assortment S.

Online retailers can only observe the click behavior within their own products.
More assumptions are needed to consider the click behaviors in competitive platforms:



Estimation methods and error bound

We consider the low-rank structure of the transition matrix.

$$
\begin{aligned}
\min_{\boldsymbol{\rho}} \quad & \mathcal{L}(\boldsymbol{\rho}) := \ \ell_{mle}(\boldsymbol{\rho}) + \gamma ||\boldsymbol{\rho}||_* && (1)\\
s.t. \quad & \sum_{j \in [\bar{n}]} \rho_{ij} = 1, \forall i \in [n] && (1a)\\
& \rho_{ij} \geq 0, \forall i \in [n], \forall j \in [\bar{n}] && (1b)\\
& \rho_{00} = 1, \rho_{0j} = 0, \forall j \neq 0, j \in [\bar{n}]. && (1c)
\end{aligned}
$$

☐ $\ell_{mle}(\boldsymbol{\rho})$ is the negative log-likelihood function for each click model.

▶ We use the subgradient projection method.

▶ Result: We prove the Frobenius norm of estimation error grows in $\mathcal{O}(\sqrt{r \ln(|S|)/N_S})$, where $|S|$ is the number of products, N_S is the number of click transition pairs.

☐ In [Kallus and Udell, 2020], the order was $\mathcal{O}(\sqrt{r|S| \ln(|S|)/N_S})$.

Multi-product pricing

Algorithm 1 Optimal Pricing in GMCCM

Initialization: Initialize vector $\mathbf{r}^0 \in \mathbb{R}^n$ randomly
While (True):
  For all Product $i \in [n]$:
    If $i \in S^p$:
      $r_i^{t+1} = \max_{p_i} \{\mathbb{I}_{\{i \in S^r\}} \mu(p_i)(p_i - c_i) + \left(1 - \mu(p_i)\mathbb{I}_{\{i \in S^r\}}\right) \sum_{j \in [n]} \rho_{ij} r_j^t\}$
    Else:
      $r_i^{t+1} = \mathbb{I}_{\{i \in S^r\}} \mu(p_i)(p_i - c_i) + \left(1 - \mu(p_i)\mathbb{I}_{\{i \in S^r\}}\right) \sum_{j \in [n]} \rho_{ij} r_j^t$
  If $||\mathbf{r}^t - \mathbf{r}^{t+1}||_2 \leq \epsilon_r$: **Break**
Set $p_i \leftarrow \mathbb{I}_{\{i \in S^r\}} \mu(p_i)(p_i - c_i) + \left(1 - \mu(p_i)\mathbb{I}_{\{i \in S^r\}}\right) \sum_{j \in [n]} \rho_{ij} r_j^t$, for all $i \in S^p$
Return p_i, for all $i \in S^p$

- Higher click rate of product i does not necessarily mean its higher optimal price.
- The change of optimal prices depends on our defined *optimal stationary revenue*.

Dynamic pricing problem

Algorithm 2 Exploration-free greedy algorithm for dynamic pricing in GMCCM

1: **Initialization:** Initialize the values of prices randomly.
2: **For** $t = 1, ..., T$ **do**
3: Observe the click trajectories and click behavior of customer t; Update the set of click trajectories $\mathbb{C}_S$, N_K, and w_{ii}^S
4: Estimate the parameters $(\hat{\boldsymbol{\rho}}, \hat{\alpha})$ in GMCCM.
5: Run Algorithm 1 to get the optimal price $\mathbf{p}_t$, with parameter $(\hat{\boldsymbol{\rho}}, \hat{\alpha})$.
6: Set price as $\mathbf{p}_t$
7: **End For**

- We show the regret is in the order of $\tilde{\mathcal{O}}(\sqrt{nrT})$, which is smaller than $\tilde{\mathcal{O}}(n\sqrt{T})$.
- The order of T matches the lower bound of regret of the online pricing problem in [Broder and Rusmevichientong, 2012].

References

Broder, J. and Rusmevichientong, P. (2012).
Dynamic pricing under a general parametric choice model.
Operations Research, 60(4):965–980.

Dong, J., Simsek, A. S., and Topaloglu, H. (2019).
Pricing problems under the markov chain choice model.
Production and Operations Management, 28(1):157–175.

Goutam, K., Goyal, V., and Soret, A. (2019).
A generalized markov chain model to capture dynamic preferences and choice overload.
arXiv preprint arXiv:1911.06716.

Kallus, N. and Udell, M. (2020).
Dynamic assortment personalization in high dimensions.
Operations Research, 68(4):1020–1037.

Thank you.