Multi-Task Combinatorial Bandits for Budget Allocation

Ad KDD 2024

amazon ads

Research team

Lin Ge
Amazon

Fuhong Li
Amazon

Yang Xu
North Carolina State University

Kelly Paulson
Amazon

Jianing Chu
Amazon

Rui Song
Amazon

David Cramer
Amazon

Amazon DSP (ADSP)
Advertising across the web and on Amazon

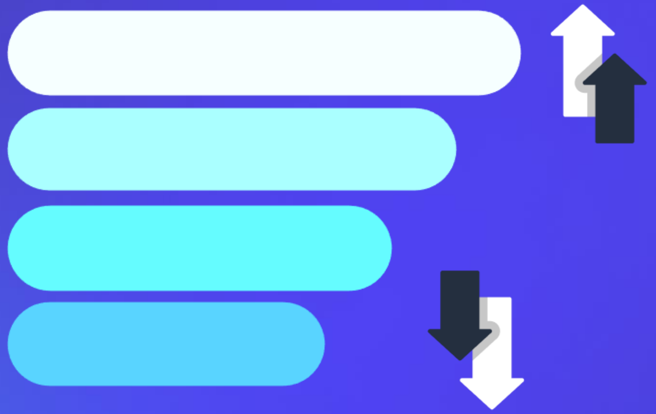
Advertisers use a Demand-side platform (DSP) to automate the process of buying digital ads

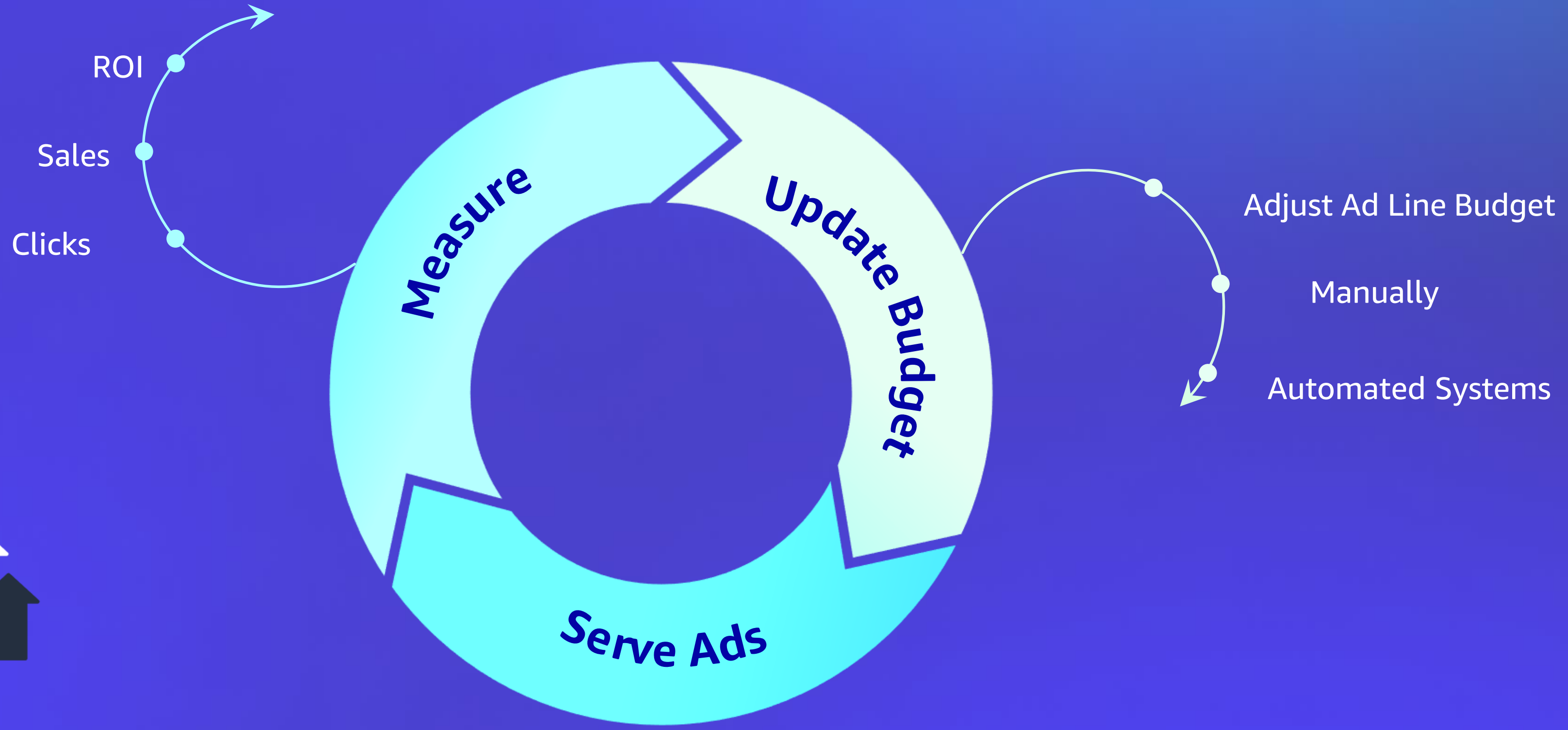
DSP Campaigns

Campaigns can have multiple ad lines (e.g. audiences, channels) with different shares of a campaign budget.

Optimization

Advertisers and systems may reallocate budget between ad lines in a campaign to top performers from bottom performers





Advertisers and Agencies can have Multiple Campaigns

Fire TV

Promoting TV devices

Echo

Promoting smart assistants

Fire Tablet

Promoting tablets

Each Campaign has Multiple Ad Lines

Fire TV

Promoting TV devices

Echo

Promoting smart assistants

Fire Tablet

Promoting tablets

Consider Optimization Challenges

1
Time

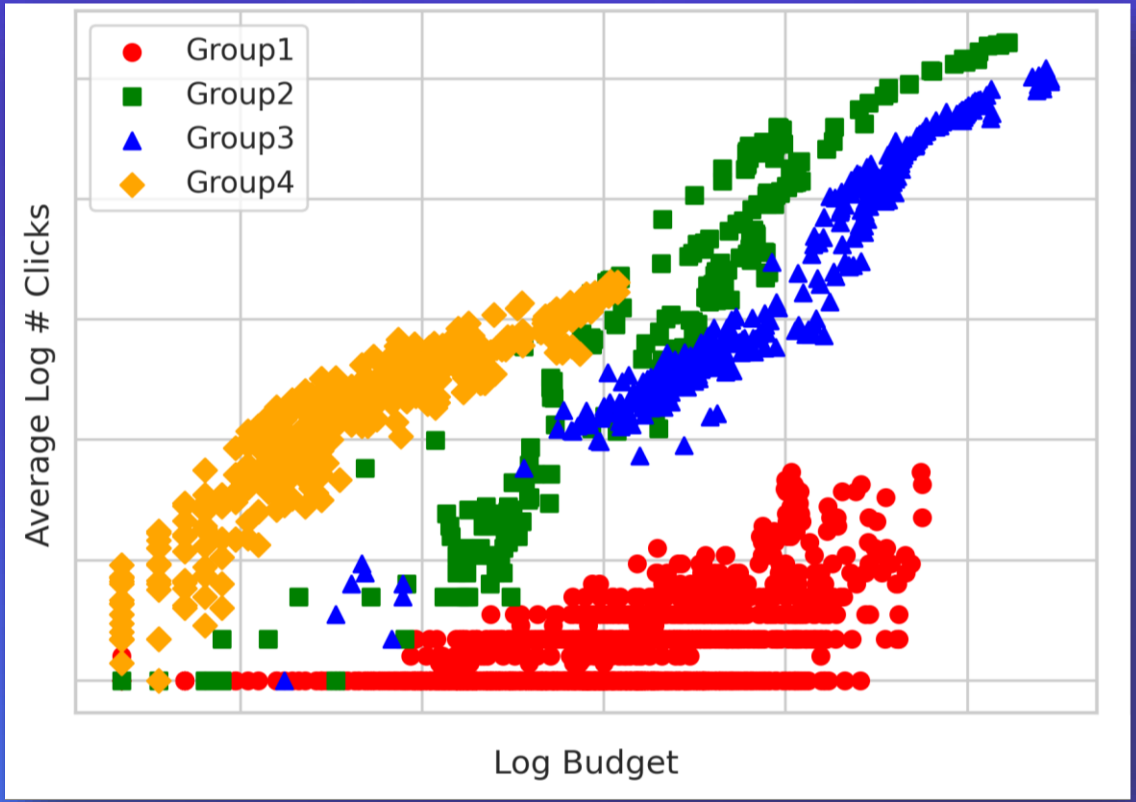
Learning takes time and results in inefficiency at the start of each campaign

2
Measurement

Historic measurement does not always reflect future (seasonality, competitive bidding, privacy, changes within a DSP)

Share Learnings Across Campaigns

Relationship between budget allocation and clicks



The significant influence of ad/campaign attributes on budget-performance relationships

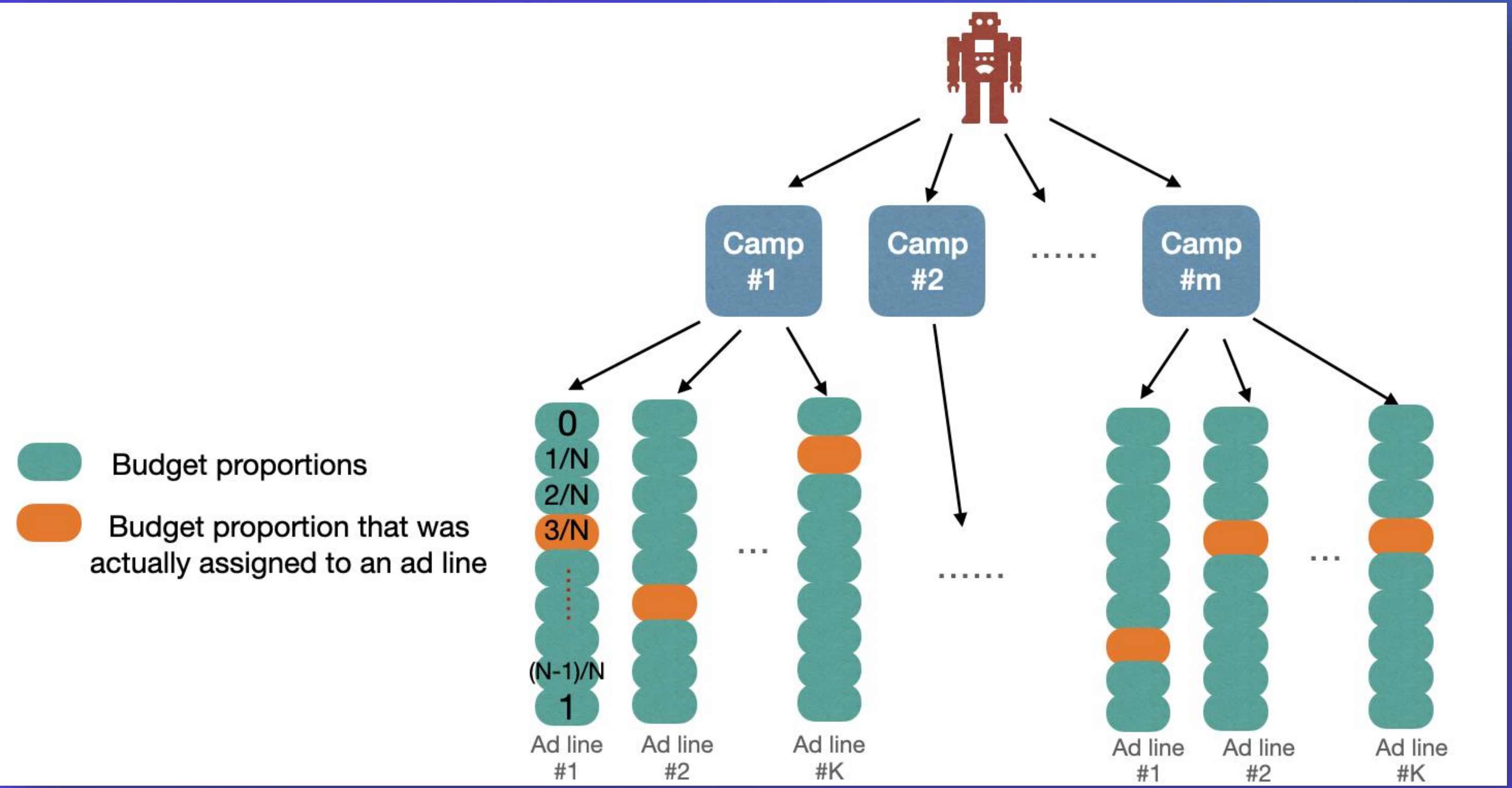
Scatter plots of the log of average number of clicks received and the log of budget allocated. Observations are classified based on advertisers' industry, channel, supply source, and audience.

Combinatorial multi-armed bandit (CMAB)

We formulated this problem as a combinatorial multi-armed bandit (CMAB) problem.

Others have applied bandits to budget allocation

- A. Nuara, F. Trovo, N. Gatti, and M. Restelli, "A combinatorial-bandit algorithm for the online joint bid/budget optimization of pay-per-click advertising campaigns," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32, no. 1, 2018.
- A. Nuara, F. Trovò, N. Gatti, and M. Restelli, "Online joint bid/daily budget optimization of internet advertising campaigns," Artificial Intelligence, vol. 305, p. 103663, 2022.
- J. Zuo and C. Joe-Wong, "Combinatorial multi-armed bandits for resource allocation," in 2021 55th Annual Conference on Information Sciences and Systems (CISS). IEEE, 2021, pp. 1–4.



Problem Formulation

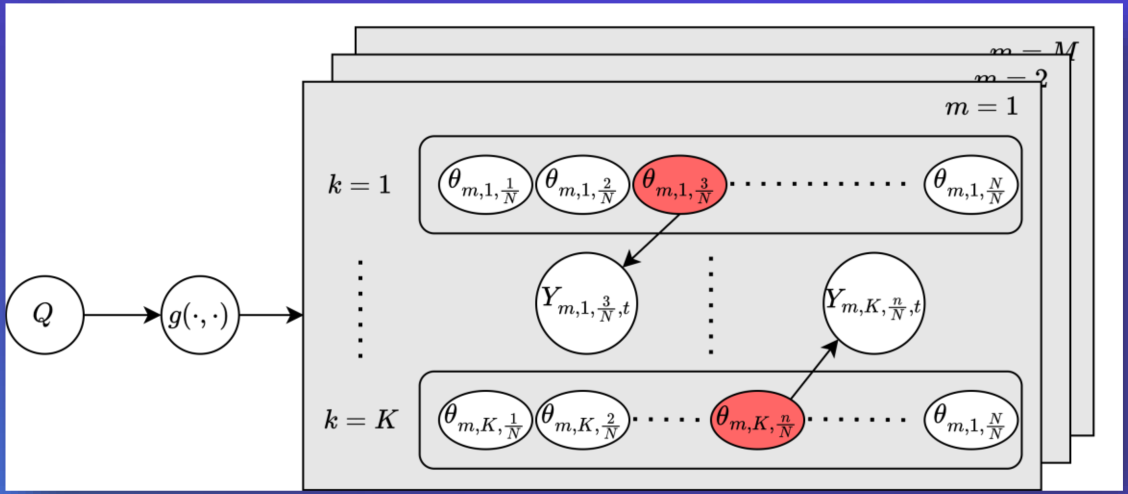
Consider a collection of M advertising campaigns. In each campaign m, there are K_m ad lines designated for daily budget allocation for T_m days with a daily budget of B_m. $x_{m,k}$ contains attribute information on campaign and ad line configurations. WLOG, we consider an action space $\mathcal{A}_d = \{0, \frac{1}{N}, \frac{2}{N}, \cdots, \frac{N-1}{N}, 1\}$, where N denotes the total number of shares we divide the budget equally into. For each campaign m, at its day t,

- allocate $a_{m,k,t} \in \mathcal{A}_d$ proportion of the total budget to ad line k in campaign m,
- observe a random reward $R_{m,k,t}(a_{m,k,t})$.

Let $\theta_{m,k,a} \equiv \mathbb{E}\{R_{m,k}(a)\}$, our goal is:

$$\begin{aligned} &\underset{a_{m,k,t}}{\text{maximize}} \sum_{m=1}^{M} \sum_{t=1}^{T_m} \sum_{k=1}^{K_m} \theta_{m,k,a_{m,k,t}}, \\ &\text{subject to} \sum_{k=1}^{K_m} a_{m,k,t} \leq 1, \quad \sum_{n=0}^{N} I\left(a_{m,k,t} = \frac{n}{N}\right) = 1, \quad \forall m, t. \end{aligned} \tag{1}$$

Bayesian Hierarchical Model



To tackle the optimization problem (1), a key step is the estimation of $\theta_{m,k,a}$, which we propose to accomplish using the following Bayesian hierarchical model:

$$
\begin{aligned}
&\text{(Prior)} && \text{Prior information Q related to } g, \\
&\text{(Generalization)} && \theta_{m,k,a} \mid \mathbf{x}_{m,k}, a = g(\mathbf{x}_{m,k}, a) + \delta_{m,k,a}, \quad \forall m, k, a, \\
&\text{(Observation)} && Y_{m,k,a_{m,k,t},t} = \theta_{m,k,a_{m,k,t}} + \epsilon_{m,k,t}, \\
&\text{(Reward)} && R_{m,t} = \sum_{k \in [K_m]} Y_{m,k,a_{m,k,t},t},
\end{aligned} \tag{2}
$$

- The **generalization layer** shares information across campaigns by utilizing $\mathbf{x}_{m,k}$ and allows inter-arm heterogeneity, while the **observations layer** and the **reward layer** constitute a general CMAB setup.
- $g(\mathbf{x}_{m,k}, a)$ captures the average impact of available features $\mathbf{x}_{m,k}$ and action a on the reward, with Q as the prior belief about g's distribution.
- $\boldsymbol{\delta}_m = [\delta_{m,1,0}, \cdots, \delta_{m,K,\frac{N}{N}}] \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ accounts for the inter-arm heterogeneity conditioned on $\mathbf{x}_{m,k}$ and a.

Algorithm Multi-Task CMAB

Multi-Task CMAB for Budget Allocation (MCMAB) is a TS-type algorithm.

- **Two-step TS (Steps 2-5)**: i) sample $\tilde{g} \sim P(g \mid \mathcal{H})$, which uses features to estimate an informative prior for $\boldsymbol{\theta}_m = (\tilde{\theta}_{m,1}, \tilde{\theta}_{m,2}, \cdots, \tilde{\theta}_{m,K_m})$ to guide the subsequent learning of the distribution of $\boldsymbol{\theta}_m$; ii) sample $\tilde{\boldsymbol{\theta}}_m \sim \mathbb{P}(\boldsymbol{\theta}_m \mid \tilde{g}, \mathcal{H})$.
- **Optimization (Step 6)**: we optimize each campaign $m \in [M]$ separately by solving a Multiple-Choice Knapsack Problem (MCKP) [1], which can utilize dynamic programming to find the optimal solution.

[1] H. Kellerer, U. Pferschy, D. Pisinger, H. Kellerer, U. Pferschy, and D. Pisinger, "The multiple-choice knapsack problem," *Knapsack Problems*, pp. 317–347, 2004.

Algorithm 1: Multi-Task Combinatorial Bandits for Budget Allocation (MCMAB)

Input : Specification of g and the corresponding prior; known parameters (i.e., σ_ϵ, $\boldsymbol{\Sigma}$) of the hierarchical model; $\mathcal{H} = \{\}$

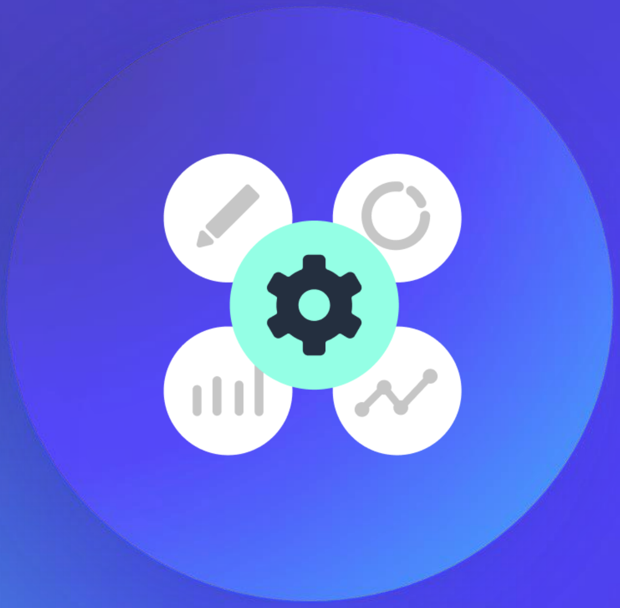
for *every decision point j* **do**

1. Retrieve the campaign index m;
2. Update the posterior for g as $\mathbb{P}(g|\mathcal{H})$, according to the hierarchical model (2);
3. Sample a $\tilde{g} \sim \mathbb{P}(g|\mathcal{H})$;
4. Given $\tilde{g}$, update the posterior for $\boldsymbol{\theta}_m$ as $\mathbb{P}(\boldsymbol{\theta}_m|\tilde{g}, \mathcal{H})$;
5. Sample an utility vector $\tilde{\boldsymbol{\theta}}_m \sim \mathbb{P}(\boldsymbol{\theta}_m|\tilde{g}, \mathcal{H})$;
6. Take action $\boldsymbol{a}_{m,t} = argmax_{\boldsymbol{a}_{m,\cdot} \in \mathcal{S}_m} \sum_{k \in [K_m]} \tilde{\theta}_{m,k,a_{m,k,\cdot}}$;
7. Receive reward $R_{m,t}$;
8. Update the dataset as $\mathcal{H} \leftarrow \mathcal{H} \cup \{(m, \boldsymbol{a}_{m,t}, R_{m,t})\}$

end

Simulation

MCMAB and baseline approaches.



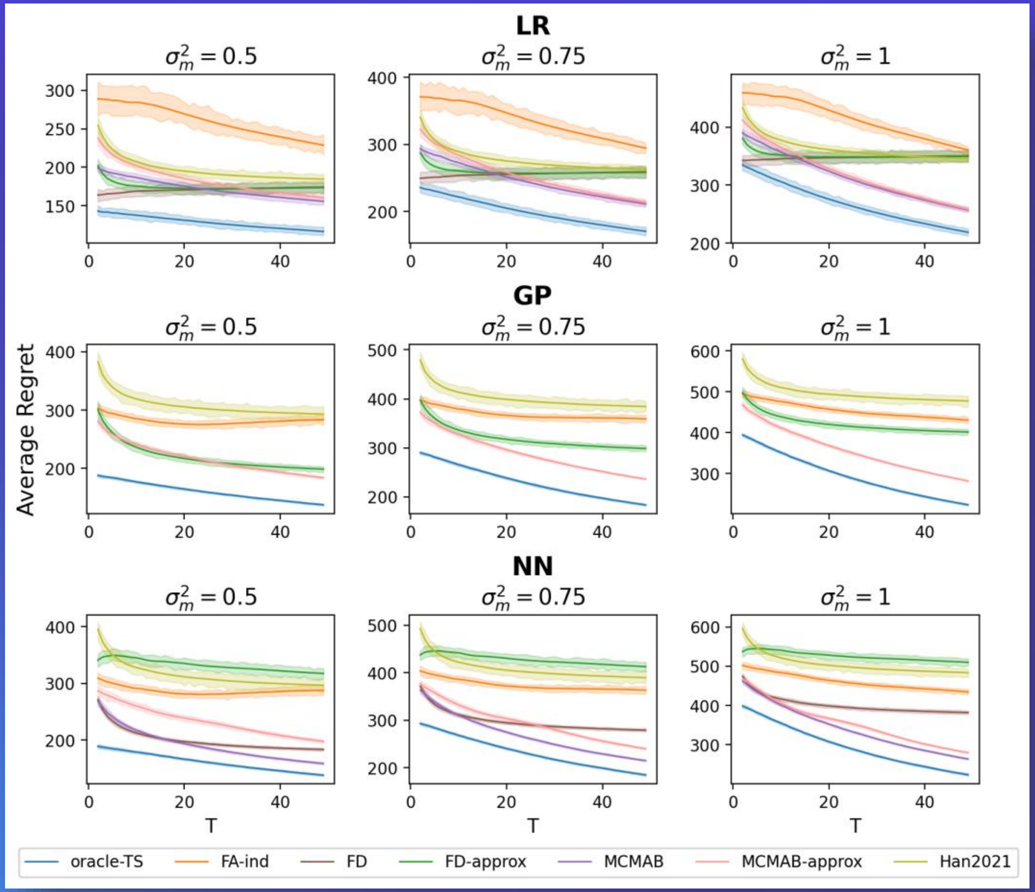
Algorithm	Utilize $\boldsymbol{x}$	Address Heterogeneity	Linear Model Assumption
MCMAB (ours)	✓	✓	× (LR/GP/NN)
FD	✓	×	× (LR/GP/NN)
FA-ind	×	✓	× (no modeling)
Han2021	✓	×	✓

- **Setup:** $M = 50$; $(K_m, T_m) = (5, 50)\ \forall\ m$; $N = 50$; $B_m \sim$ Uniform$(20, 30)$; linear environment for LR; nonlinear environment for GP and NN.
- **Bayes Regret:** $\mathbb{E} \sum_{m=1}^{M} \sum_{t=1}^{T_m} \sum_{k=1}^{K_m} \max_{\boldsymbol{a}_{m,t} \in \mathcal{S}_m} \theta_{m,k,a} - \theta_{m,k,a_{m,k,t}}$.
- **Concurrent Setting:** all M campaigns run concurrently.
- **Sequential Setting:** campaigns only start when the previous campaign ends.

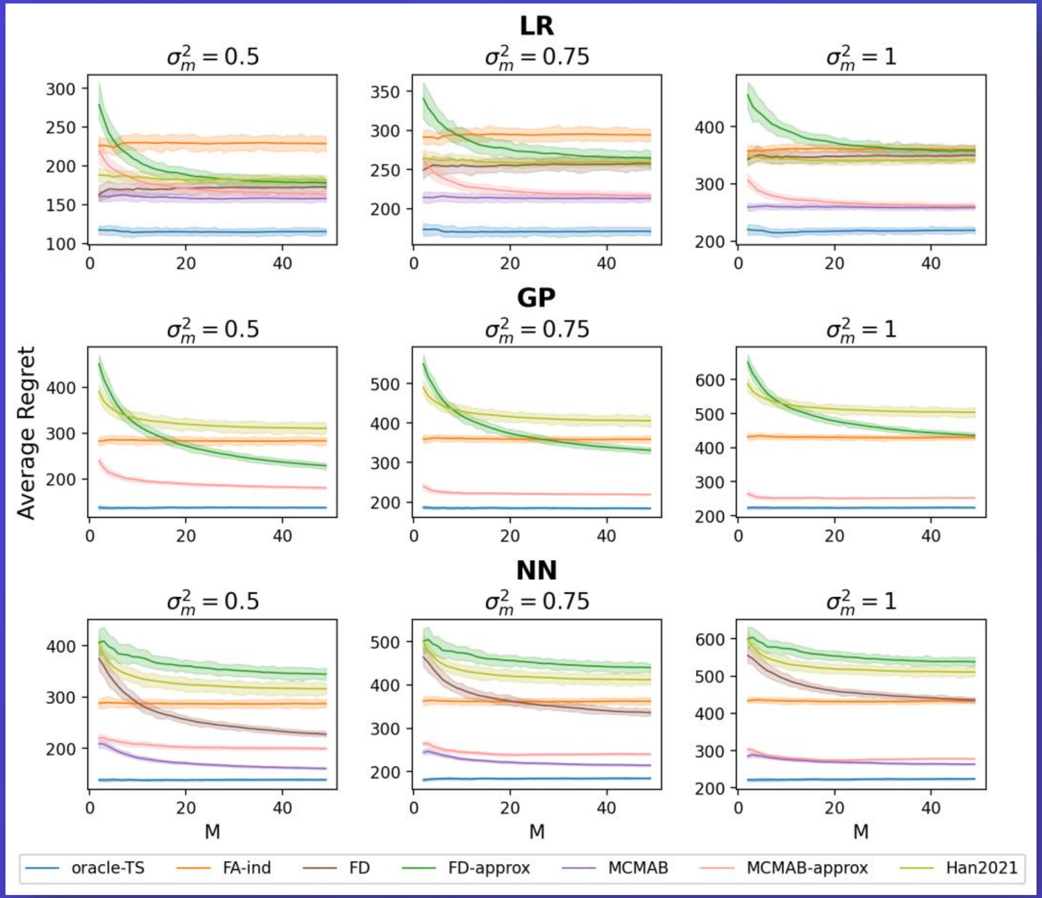
Simulation Results

100 random seeds. Shaded areas indicate the 95% confidence interval.

Concurrent



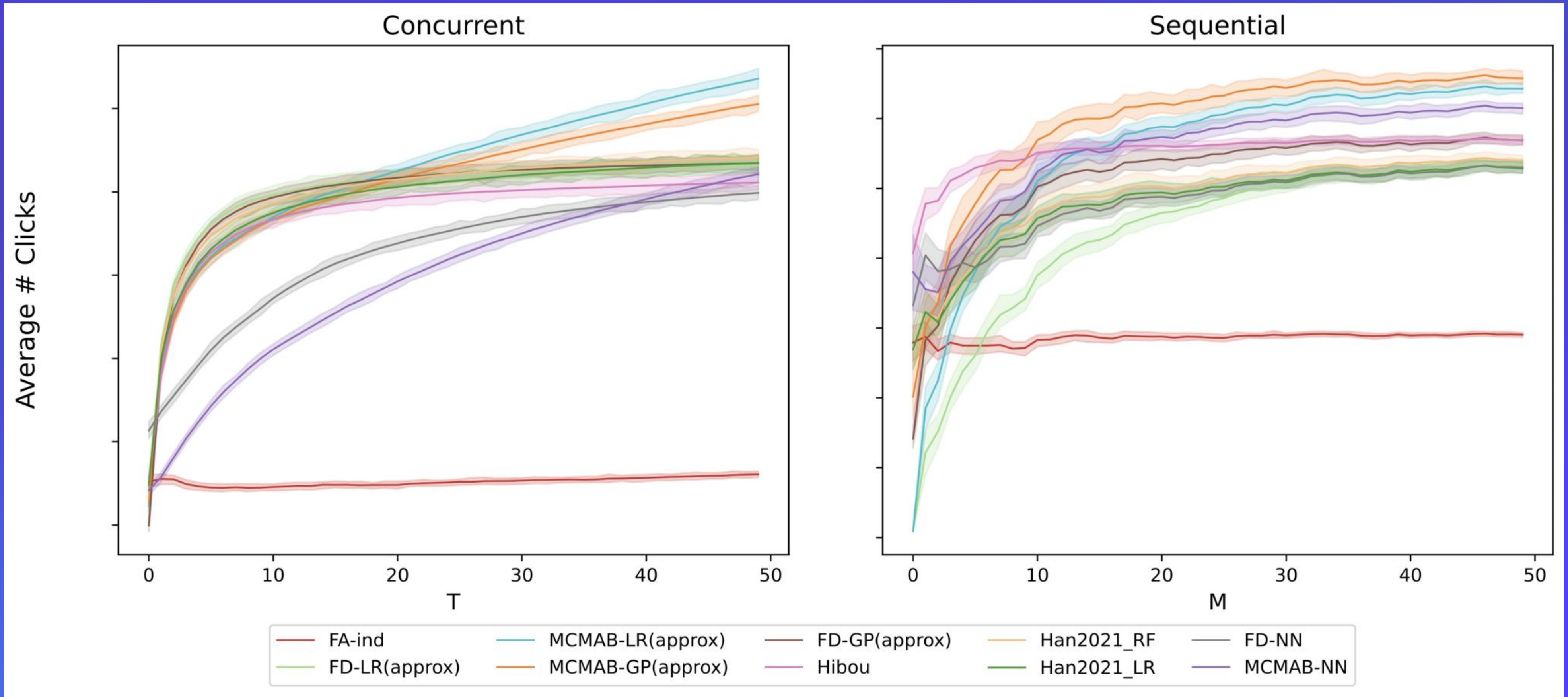
Sequential



Offline Evaluation on Amazon Campaigns

Data: Let $\boldsymbol{\Sigma} = \sigma_m^2 \boldsymbol{I}$, $\{g, \sigma_m, \sigma_\epsilon\}$ are learned from the Amazon Digital Advertisements' campaign data from 2023 Q1.

Setup: $M = 50$; $(K_m = 5, B_m = 300, T_m = 50)\ \forall\ m$; $N = 50$.



100 random seeds. Shaded areas indicate the 95% confidence interval.

Online Evaluation on Amazon Campaigns

13%

Increase in daily clicks after 3 weeks

30 Campaigns, 300 ad lines