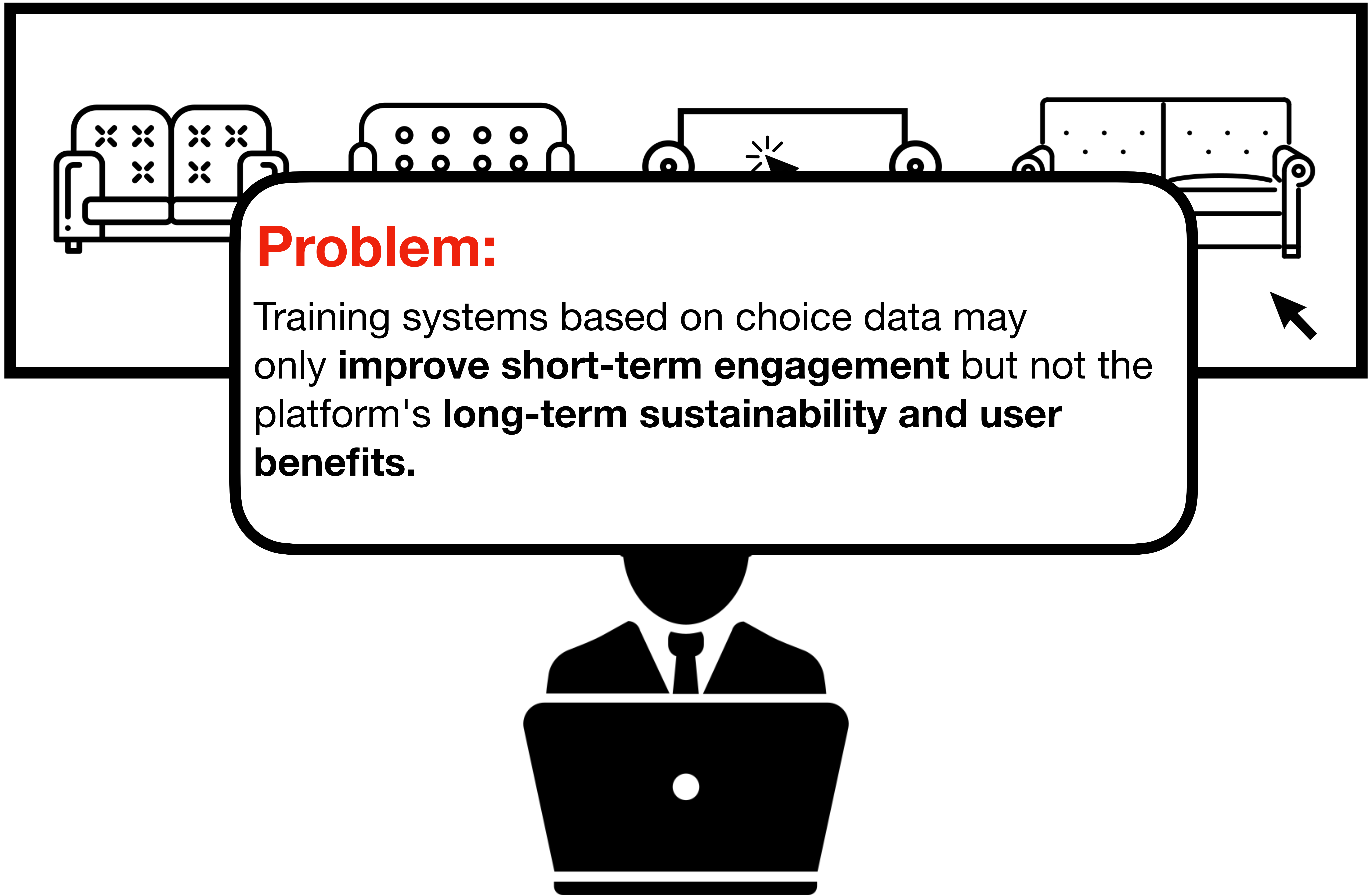
Ranking with Long-Term Constraints

Joint work with:
Zhichong Fang, Sarah Dean, and Thorsten Joachims

Recommendation Platform

problem

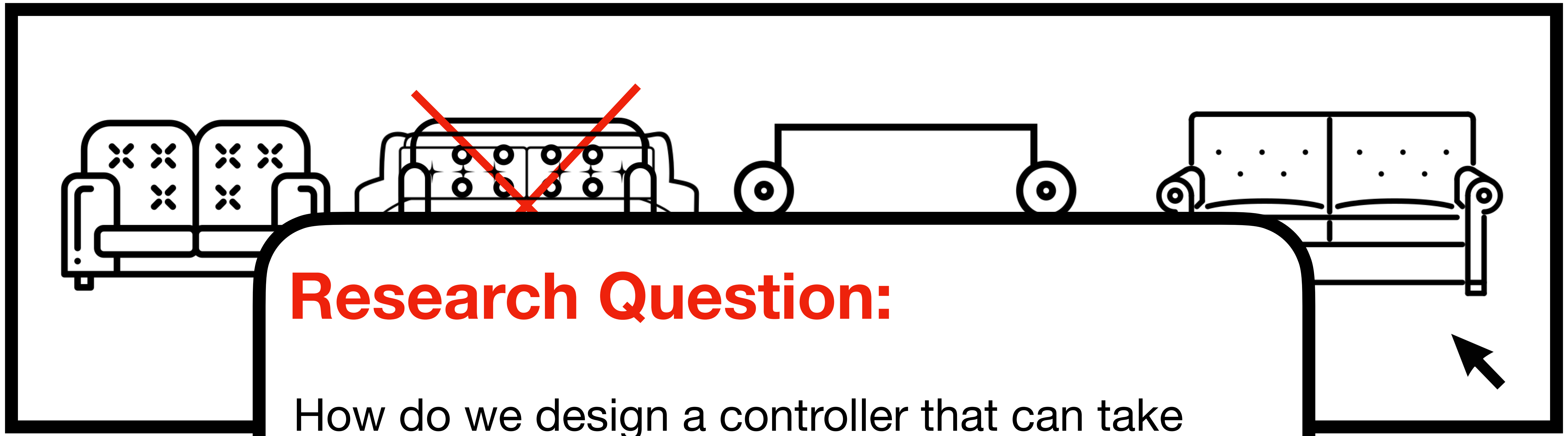


Problem:

Training systems based on choice data may only **improve short-term engagement** but not the platform's **long-term sustainability and user benefits.**

Recommendation Platform

macro-level and micro-level



Research Question:

How do we design a controller that can take complex macro-level goals and convert them into rankings with the least impact on micro-level goals?

Micro-level Controls of Platf

Metrics: engagement through clicks, purchases, etc.

Interventions: ranking, push-notifications, etc.

cro-level Controls of Platform

Metrics: user satisfaction, supplier pool size, etc.

Interventions: exposure location, novelty, etc.

Π

Controller

Macro/Micro Control Problem

objective overview

Context - x_t

Ranking - a_t

Utility - $u(a_t|x_t)$

Exposure - $c(a_t|x_t)$

Target - τ

Micro-level Controls of Platform

$$\sum_{t=1}^{T} u(a_t|x_t) \quad \textbf{(1)}$$

Macro-level Controls of Platform

$$\sum_{t=1}^{T} c(a_t|x_t) \geq \tau \quad \textbf{(2)}$$

maximize (1) s.t. (2)

Overall Objective

$$\sum_{t=1}^{T} u(a_t|x_t) - \phi^T\left(\tau - \sum_{t=1}^{T} c(a_t|x_T)\right)_+$$

Hinge loss - $(\ )_+$

Violation Cost - ϕ^T

Macro/Micro Control Problem

objective overview

Context - x_t

Ranking - a_t

Utility - $u(a_t | x_t)$

Micro-level controls of Platform

$$\sum_{t=1}^{T} u(a_t | x_t) \quad (1)$$

Macro-level controls of Platform

$$\sum_{t=1}^{T} c(a_t | x_t) \geq \tau \quad (2)$$

Problem:

The overall objective can only be computed after the final time step T and we can not make intermediate decisions.

$$\sum_{t=1}^{T} u(a_t | x_t) - \phi^T \left(\tau - \sum_{t=1}^{T} c(a_t | x_T) \right)_+$$

Violation Cost - ϕ^T

Macro/Micro Control Problem

objective overview

Micro-level controls of Platform

$$\sum_{t=1}^{T} u(a_t | x_t) \qquad \text{(1)}$$

Context - x_t

Ranking - a_t

Utility - $u(a_t | x_t)$

Exposure - $c(a_t | x_t)$

Target - τ

Macro-level controls of Platform

$$\sum_{t=1}^{T} c(a_t | x_t) \geq \tau \qquad \text{(2)}$$

maximize (1) s.t. (2)

Overall Objective

$$\sum_{t=1}^{T} u(a_t | x_t) - \phi^T \left(\tau - \sum_{t=1}^{T} c(a_t | x_T) \right)_+$$

Violation Cost - ϕ^T

Macro/Micro Control Problem

objective overview

Observe:

At time t we only know the contexts $x_1, \ldots, x_t$, but not the future contexts $x_{t+1}, \ldots, x_T$.

This means we make decisions under **partial information** because we cannot evaluate our constraint.

Research Question:

How do we design controllers for ranking to optimize our macro and micro level goals under partial information?

ntext -

nking -

Utility -

osure -

Target -

maximize (1) s.t. (2)

Overall Objective

$$\sum_{t=1}^{T} u(a_t | x_t) - \phi^T \left(\tau - \sum_{t=1}^{T} c(a_t | x_T) \right)_+$$

Violation Cost - ϕ^T

Controllers for Ranking

Myopic Controller (MC)

Define an intermediate objective at each time t

$$\sum_{t=1}^{T} u\left(a_t \mid x_t\right) - \phi^\top \left(\frac{t}{T}\tau - \sum_{t=1}^{T} c\left(a_t \mid x_t\right)\right)_+$$

Scale the target linearly

$$\sum_{t'=1}^{t} u\left(a_{t'} \mid x_{t'}\right) - \phi^\top \left(\frac{t}{T}\tau - \sum_{t'=1}^{t} c\left(a_{t'} \mid x_{t'}\right)\right)_+$$

Create intermediate objectives that do not depend on future

$$\operatorname*{argmax}_a u\left(a \mid x_t\right) - \phi^\top \left(\frac{t}{T}\tau - s_{t-1} - c\left(a \mid x_t\right)\right)_+$$

Remove terms that do not affect argmax

$$s_{t-1} = \sum_{t'=0}^{t-1} c(a_{t'} \mid x_{t'})$$

s_{t-1} : past exposure

Controllers for Ranking

Problem:

The cost vector ϕ^T still depends on T because it is the final cost violation not the intermediate cost violation.

This is bad because we **suffer the maximum cost for every intermediate constraint violation**.

Define an intermediate objective at each time t

Scale the target linearly

Create intermediate objectives that do not depend on future

$$\underset{a}{\text{argmax}}\, u\left(a \mid x_t\right) - \phi^{\mathsf{T}}\left(\frac{t}{T}\tau - s_{t-1} - c\left(a \mid x_t\right)\right)_+$$

Remove terms that do not affect argmax

$$s_{t-1} = \sum_{t'=0}^{t-1} c(a_{t'} \mid x_{t'})$$

s_{t-1} : past exposure

Controllers for Ranking

Stationary-Controller (SC)

Define an intermediate objective at each time t

Define an intermediate cost vector at each time t

$$\sum_{t=1}^{T} u(a_t \mid x_t) - \phi^\top \left(\tau - \sum_{t=1}^{T} c(a_t \mid x_t)\right)_+$$

ϕ is bounded between 0 and ϕ

$$\min_{0 \le \lambda \le \phi} \frac{1}{T} \sum_{t=1}^{T} u(a_t \mid x_t) - \lambda^\top \left(\frac{1}{T}\tau - \frac{1}{T} \sum_{t=1}^{T} c(a_t \mid x_t)\right)$$

intermediate cost vector - λ is bounded between 0 and ϕ

Algorithm:

Step 1

$$\operatorname*{argmax}_a u(a \mid x_t) - \lambda_{t-1}^\top \left(\frac{1}{T}\tau - c(a \mid x_t)\right)$$

Remove terms that do not affect argmax

Step 2

$$\lambda_t = \Gamma_{[0,\phi]} \left[\lambda_{t-1} + \gamma \left(\frac{1}{T}\tau - c(a \mid x_t)\right)\right]$$

Update our intermediate cost vector

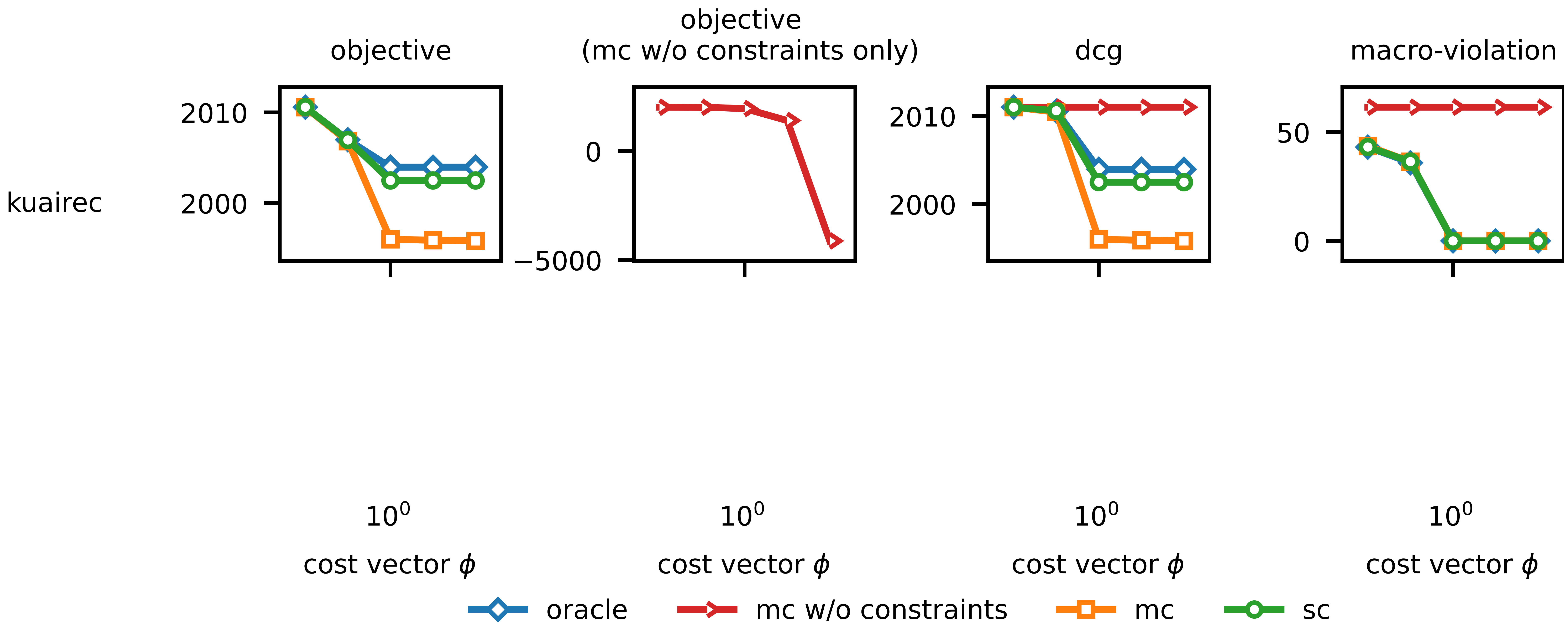
Experimental Results

- **KuiaRec Dataset**
  - Collected from the recommendation logs of the video-sharing mobile app Kuaishou
  - Stationary data **(not-time dependent)**

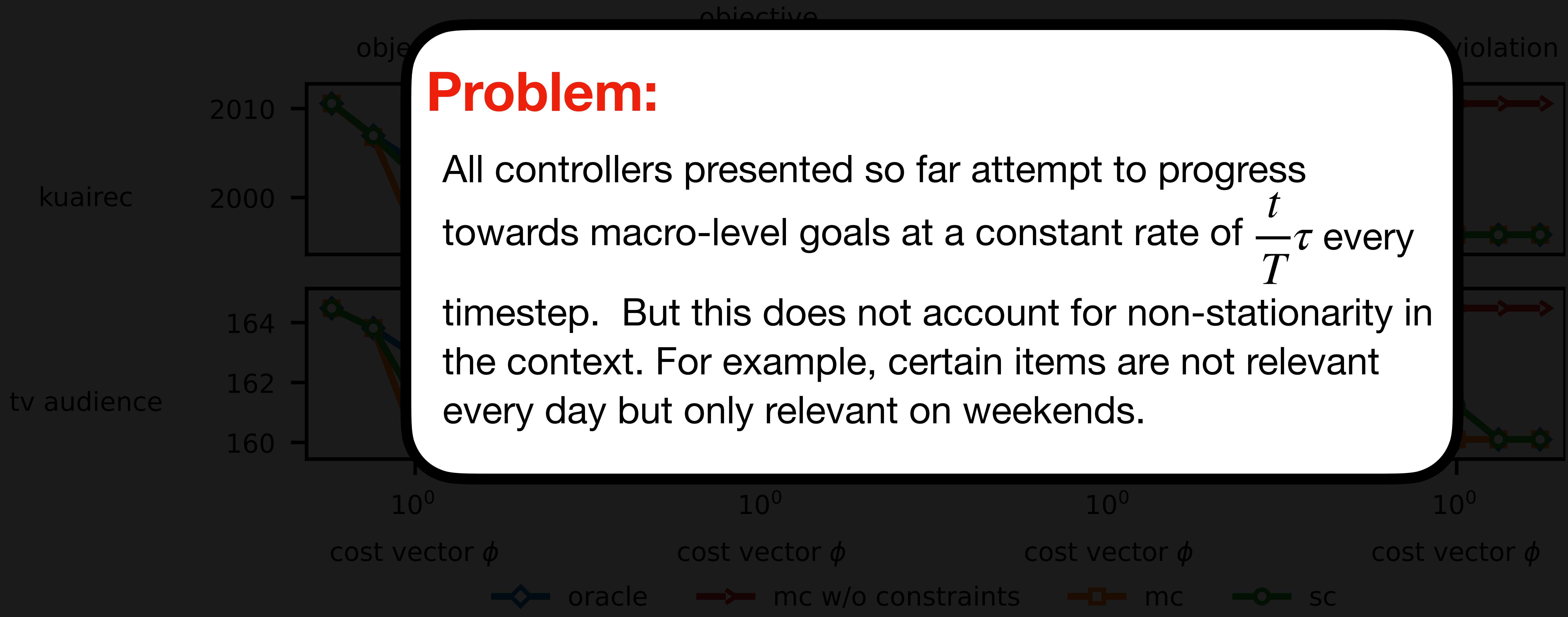
- **Tv Audience Dataset**
  - Television watching behavior of users
  - Non-stationary data **(time dependent)**

- **Metrics**
  - Utility - nDCG
  - Exposure - Reciprocal Rank

Experimental Results



Experimental Results



Problem:

All controllers presented so far attempt to progress towards macro-level goals at a constant rate of $\frac{t}{T}\tau$ every timestep. But this does not account for non-stationarity in the context. For example, certain items are not relevant every day but only relevant on weekends.

Controllers for Ranking

Define an intermediate objective at each time t

Define an intermediate cost vector at each time t

Model future contexts

Predictive Controller (PC)

$$\sum_{t=1}^{T} u\left(a_t \mid x_t\right) - \phi^\top \left(\tau - \sum_{t=1}^{T} c\left(a_t \mid x_t\right)\right)_+$$

$$\min_{0 \le \lambda \le \phi} \frac{1}{T} \sum_{t=1}^{T} u\left(a_t \mid x_t\right) - \lambda^\top \left(\tau - \left[s_{t-1} + c\left(a_t \mid x_t\right) + C\left(A_t \mid C_t\right)\right]\right)_+$$

s_{t-1} : past exposure

$c\left(a_t \mid x_t\right)$: current exposure

$C\left(A_t \mid C_t\right)$: future exposure

$$s_{t-1} = \sum_{t'=0}^{t-1} c(a_{t'} \mid x_{t'})$$

$$C(A_t \mid C_t) = \sum_{t'=t+1}^{T} c(a_{t'} \mid x_{t'})$$

Controllers for Ranking

Define an intermediate objective at each time t

Define an intermediate cost vector at each time t

Model future contexts

Predictive Controller (PC)

$$\min_{0 \le \lambda \le \phi} \frac{1}{T} \sum_{t=1}^{T} u\left(a_t \mid x_t\right) - \lambda^{\top}\left(\tau - \left[s_{t-1} + c\left(a_t \mid x_t\right) + C\left(A_t \mid C_t\right)\right]\right)_+$$

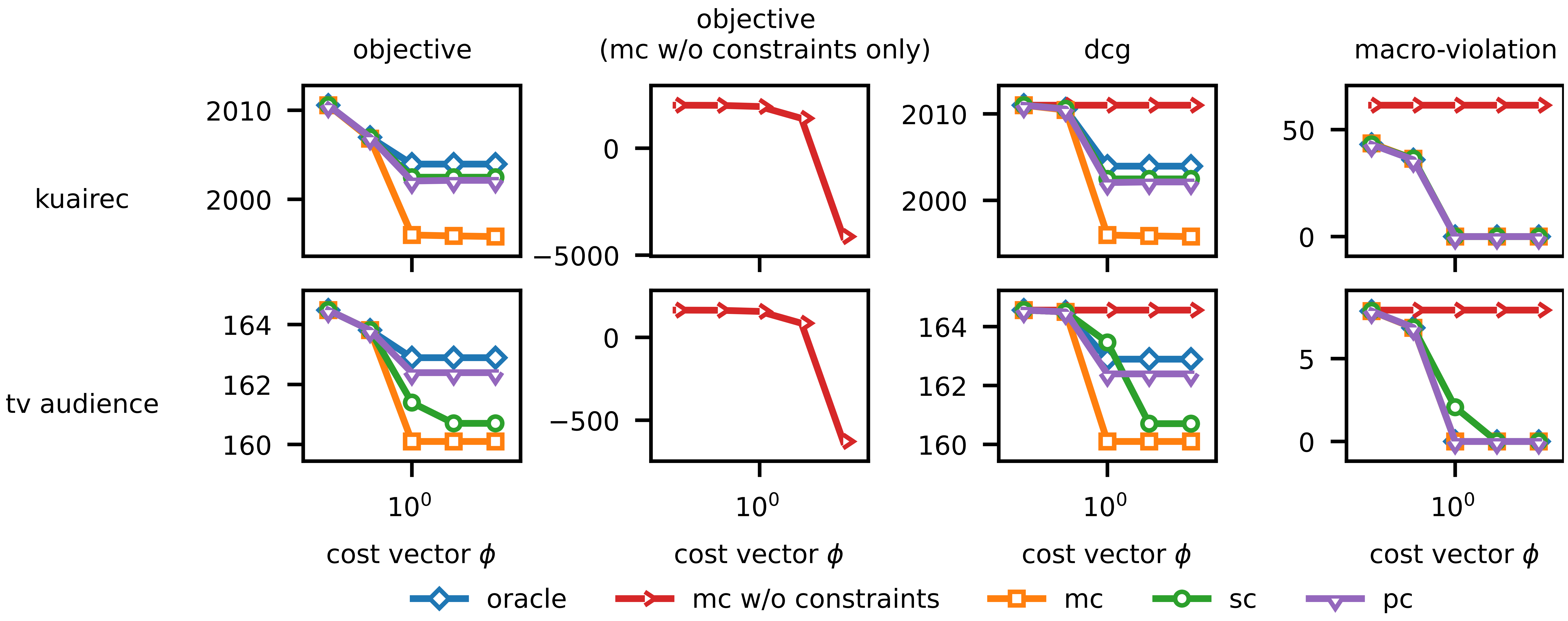
estimate $\hat{C}(A_t \mid C_t)$ using offline data

Algorithm:

Step 1 $\quad \operatorname*{argmax}_a u\left(a \mid x_t\right) - \lambda_{t-1}^{\top} c\left(a \mid x_t\right)$ — Remove terms that do not affect argmax

Step 2 $\quad \lambda_t = \Gamma_{[0,\phi]}\left[\lambda_{t-1} + \gamma\left(\tau - [s_{t-1} + c(a_t \mid x_t) + \hat{C}_t\left(A_t \mid C_t\right)]\right)\right]$ — Update our intermediate cost vector

Experimental Results



Controllers for Ranking

$W \in \mathbb{R}^{m \times n}$	w_{ij}	contribution item j to constraint i
$\Sigma \in \mathbb{R}^{n \times n}$	$\Sigma_{kj} = 1$	ranking matrix
$e \in \mathbb{R}^{n}$		linear exposure utility vector
$u \in \mathbb{R}^{n}$		linear utility vector
$r \in \mathbb{R}^{n}$		relevance vector

Cost-Aware Predictive-Controller

$$\sum_{t=1}^{T} r_t^\top \Sigma_t u - \phi^T \left(\tau - \sum_{t=1}^{T} W \Sigma_t e \right)_+$$

objective: $$\min_{0 \le \lambda \le \phi} \frac{1}{T} \sum_{t=1}^{T} r_t^\top \Sigma_t u - \lambda^\top \left(\tau - \left[s_{t-1} + W \Sigma_t e + \hat{C}(A_t \mid C_t) \right] \right)_+$$

estimate $\hat{C}(A_t \mid C_t)$: $$\arg \max_{(\Sigma^1, \ldots, \Sigma^B) \in \Pi} \Big[\quad \ldots \quad \Big]$$

using offline data

iterative objective:

Step 1 $$\Sigma_t = \arg \max_{\Sigma} \; r_t^\top \Sigma u - \frac{1}{B} \sum_{b=1}^{B} \left(\lambda_{t-1}^b \right)^\top W \Sigma e$$

Break the objective into incremental rankings

Step 2 $$\lambda_t^b = \Gamma_{[0,\phi]} \left[\lambda_{t-1}^b + \gamma \left(\tau - [s_{t-1} + c(a_t \mid x_t) + \hat{C}_t^b] \right) \right], \quad b = 1, \ldots, B$$

Macro/Micro Control Problem

control-loop

Input: target τ

Input: violation cost ϕ

Input: final time T

Input: controller Π

Initialize : state $s_t = \mathbf{0}$

For $t = 1$ **to** T **do**

$x_t = (\mathbf{q}_t, r_t) \sim P_t$ — context x_t draw independently but not identically

$a_t = \Pi(x_t, s_{t-1}, t)$ — a_t is a ranking

$s_t = s_{t-1} + c(a_t | x_t)$ — sum macro-level metrics up to t

Compute Objective — violations are subject to a cost ϕ

$$\sum_{t=1}^{T} u(a_t | x_t) - \phi^T \left(\tau - \sum_{t=1}^{T} c(a_t | x_T) \right)_+$$

Controllers for Ranking

matrix notation

$$\sum_{t=1}^{T} u(a_t|x_t) - \phi^T \left(\tau - \sum_{t=1}^{T} c(a_t|x_T)\right)_+$$

Micro-level Metric (e.g. DCG)

$$u(a_t|x_t) = r_t^\top \Sigma_t u$$

$W \in \mathbb{R}^{m \times n}$	w_{ij}	contribution item j to constraint i
$\Sigma \in \mathbb{R}^{n \times n}$	$\Sigma_{kj} = 1$	ranking matrix
$e \in \mathbb{R}^n$		linear exposure vector
$u \in \mathbb{R}^n$		linear utility vector
$r \in \mathbb{R}^n$		relevance vector

Macro-level Metric (e.g. exposure)

$$c(a_t|x_t) = W\Sigma_t e$$

$$\sum_{t=1}^{T} r_t^\top \Sigma_t u - \phi^T \left(\tau - \sum_{t=1}^{T} W\Sigma_t e)\right)_+$$