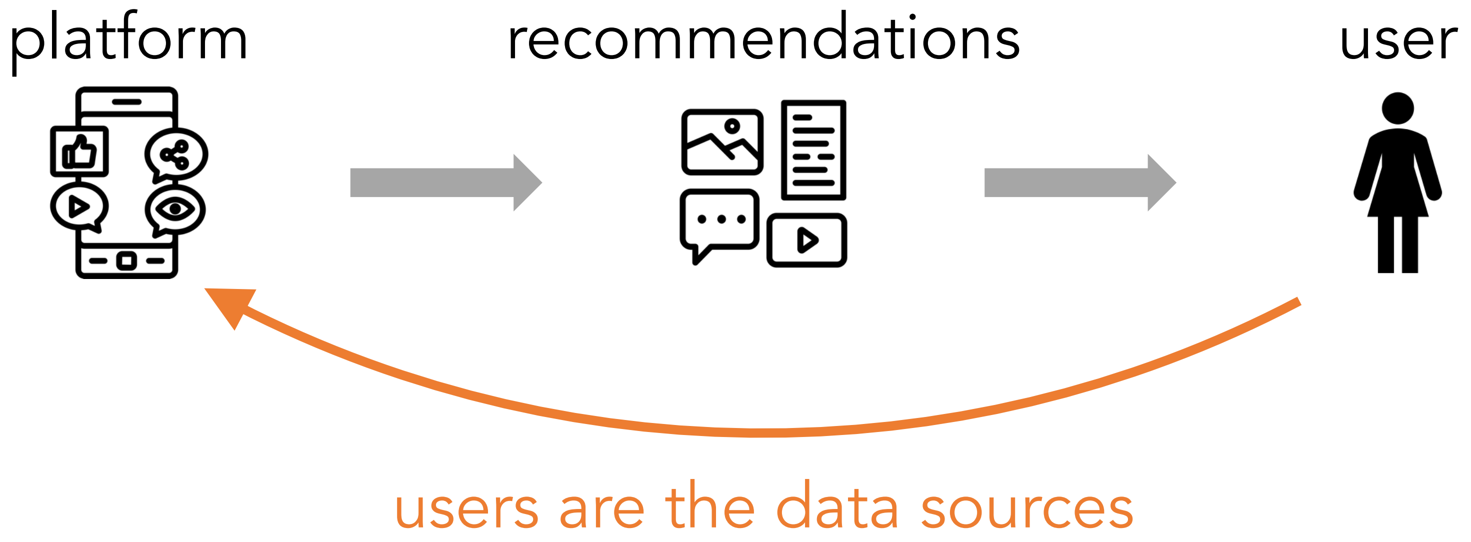
A Game-Theoretic Perspective on Trust in Recommendation

Sarah H. Cen, Andrew Ilyas, and Aleksander Mądry
MIT EECS

Responsible Decision Making in Dynamic Environments, ICML 2022

The role of trust in recommendation



Users are not fixed or truthful. They can **learn, adapt, and strategize**.

Model interactions as an **alternating two-player game**.

Find that **cooperating** can benefit both the user & platform → **trust**!

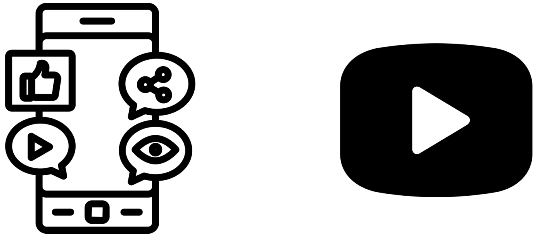
Recommendation

Platform provides (personalized) suggestions to each user.

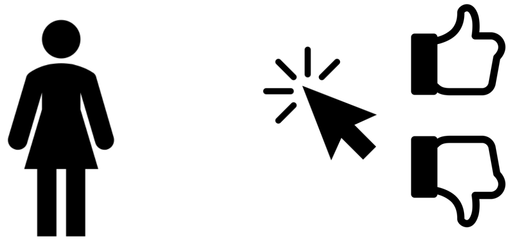
Our focus: trust **between a user and their platform**.

Why do we care about trust?

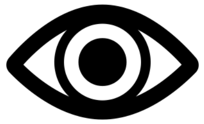
1. Platform recommends a video to user



2. User decides whether to watch & up or down vote



3. Platform observes user's watch & voting behavior



Common assumption: fixed preferences & truthful.

Why do we care about trust?

But humans (not just platforms) are adaptable & strategic.

Poses problem for platforms.

Why? Because users are platforms’ primary data sources.

In reality, the data are not i.i.d., missing uniformly at random, etc.

Punchline: Both users and platforms benefit from trust.

Distrust is a self-defeating cycle

Hiding interests

Protecting privacy

Users don't trust platforms.

Platforms don't trust users.

Trust as **encapsulated interest** (Hardin, 1991).

When two strategic actors interact, trust matters.

Model: Alternating two-player game

We model recommendation as an alternating two-player game:

Platform deploys a recommender system.

User interacts with recommender.

Formally, the game is given by $(\mathcal{F}, \mathcal{B}, U_p, U_u)$, where:

Platform plays recommender $f_t \in \mathcal{F}$

User plays behavior $b_t \in \mathcal{B}$

Receive payoffs $U_p, U_u : \mathcal{F} \times \mathcal{B} \to [-1, 1]$

Model: Alternating two-player game

Truthful strategy:

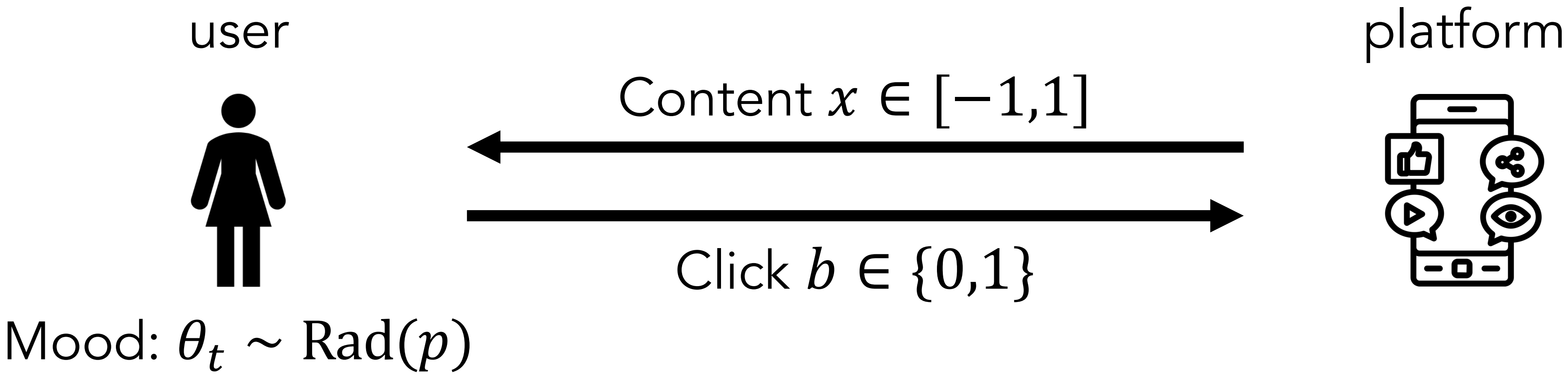
Maximizes payoff w.r.t. platform's most recent action (BR)

Long-term optimal strategy:

Given the platform's strategy, maximizes the long-term payoff.

If a user **trusts** their platform's strategy s_p, then their optimal long-term strategy to s_p is to be truthful at every time step.

Example 1: Multi-modal user



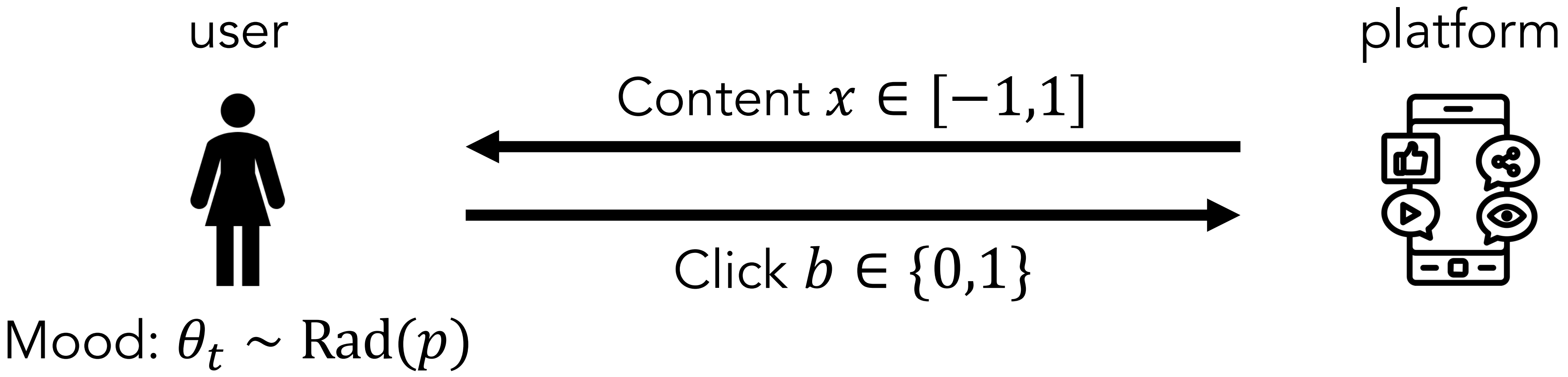
$$U_u(x,b) = b \cdot \mathbf{1}\{\theta_t = x\}$$

User gets +1 if content matches their current mood, 0 otherwise

$$U_p(x,b) = b$$

Platform gets 1 if user clicks, 0 otherwise

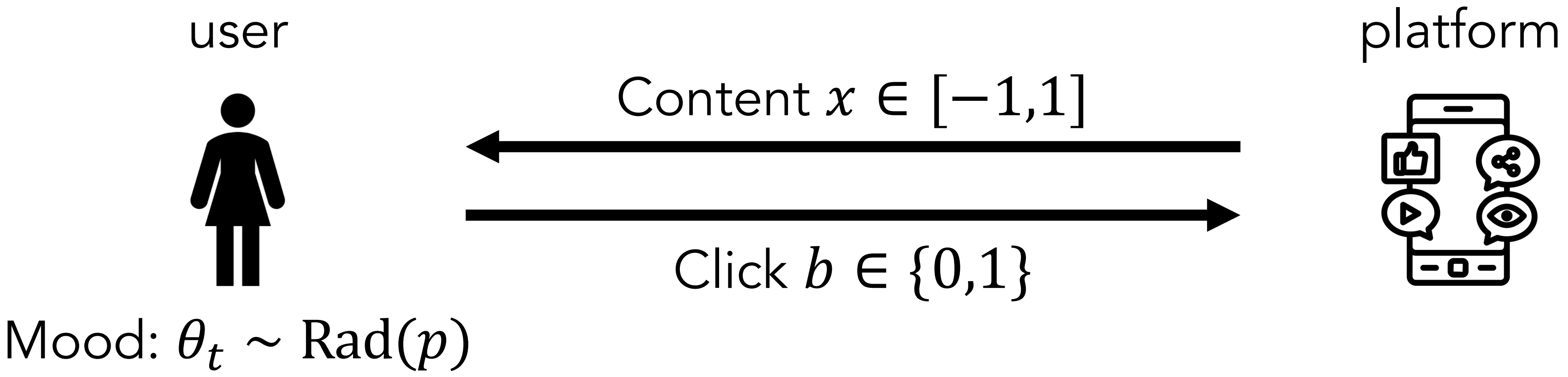
Example 1: Multi-modal user



Naive platform strategy: Use ERM to learn a parameter $\hat{\theta}$, recommend x =clip($\hat{\theta}$ + noise, -1, 1)

User is not incentivized to be truthful: $\hat{\theta}$ diverges (caters to majority mood) or $\hat{\theta} = p$ (reflects "average mood")

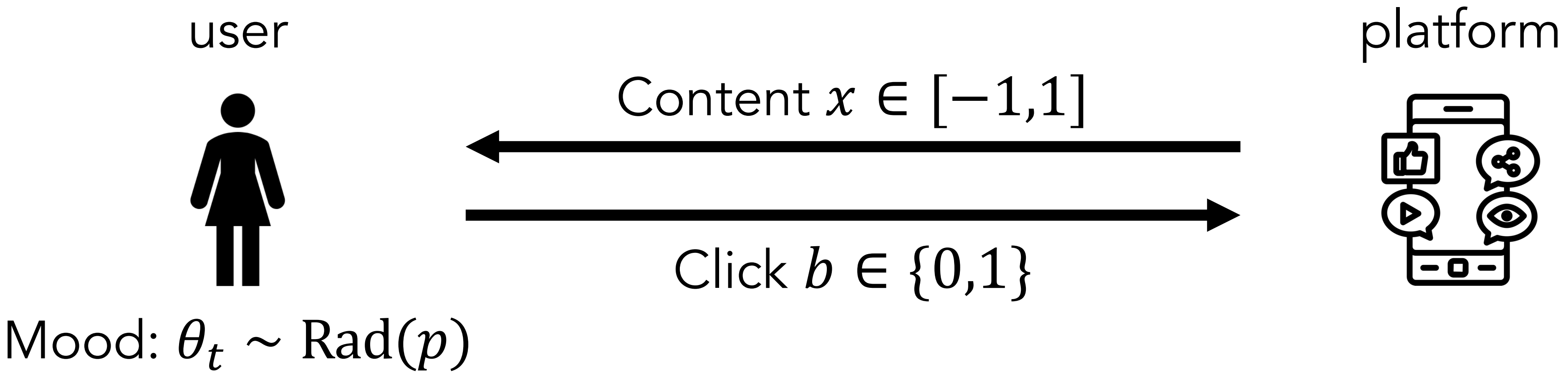
Example 1: Multi-modal user



Naive platform strategy: Use ERM to learn a parameter $\hat{\theta}$, recommend x =clip($\hat{\theta}$ + noise, -1, 1)

Result: User will only visit the platform when in their dominant mood (platform misses out on clicks)

Example 1: Multi-modal user



It's beneficial to cooperate & earn the user's trust:
Solicit mood θ_t from user (e.g., allowing them to filter)

Earning the user's trust by giving them agency:
Platform can always suggest content that the user will enjoy

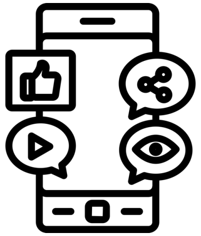
Example 2: Privacy-conscious user

user

Model $\hat{\theta}$, content $x \in \mathbb{R}^d$

Time spent $t \in \mathbb{R}$

platform



Interest:

$y \sim \mathcal{N}(\theta^\top x, 1)$

User features θ

Content features x

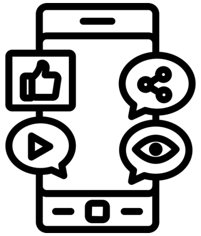
Example 2: Privacy-conscious user

user

Model $\hat{\theta}$, content $x \in \mathbb{R}^d$

Time spent $t \in \mathbb{R}$

platform



Interest:
$y \sim \mathcal{N}(\theta^\top x, 1)$

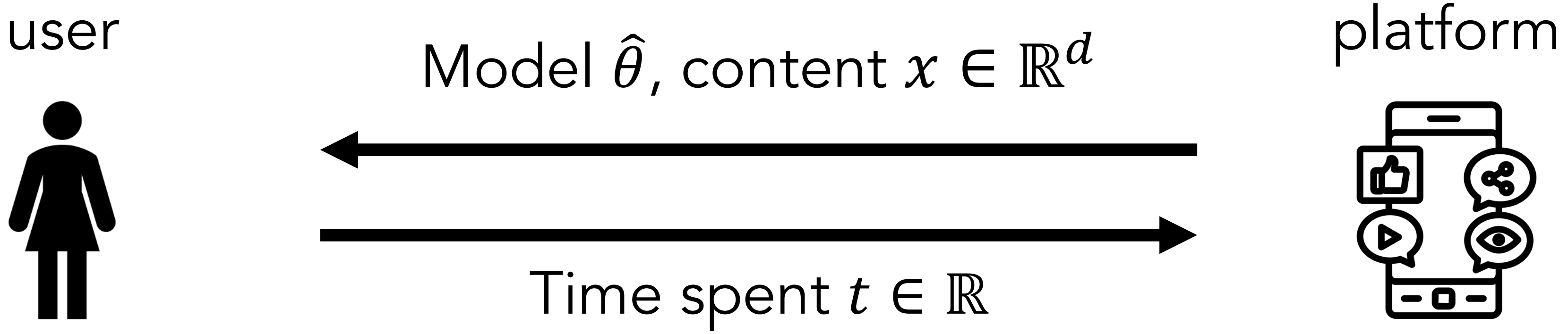
$$U_u(x,t) = (y - t)^2 + \log(|\theta_p - \hat{\theta}_p|)$$

Reward for watching interesting content, but penalty for revealing private feature

$$U_p(x,t) = t$$

Reward for user watching for longer

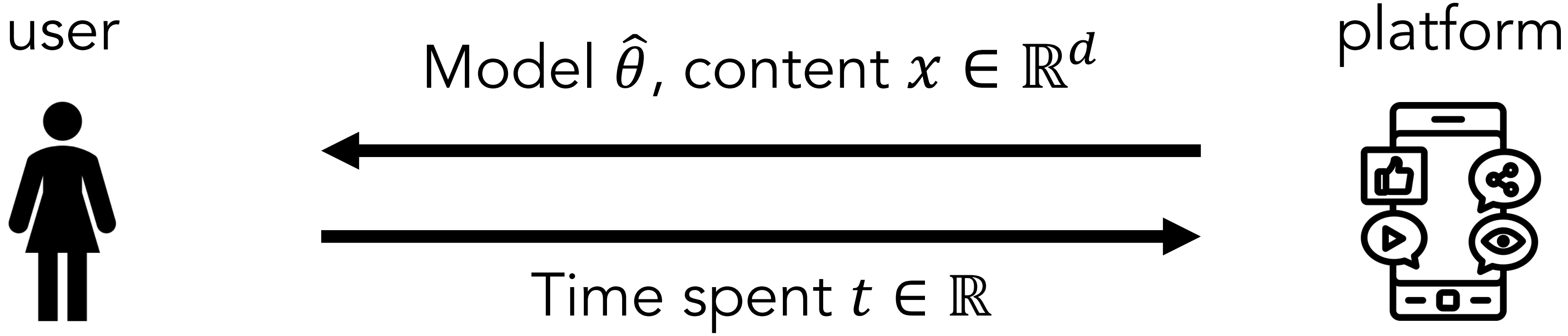
Example 2: Privacy-conscious user



Naive platform strategy: Learn a user model $\hat{\theta}$, and use bandit algorithm to suggest content

User is not incentivized to be truthful: $\hat{\theta}_p \approx \theta_p$ (platform learns private feature), so user reward diverges to $-\infty$

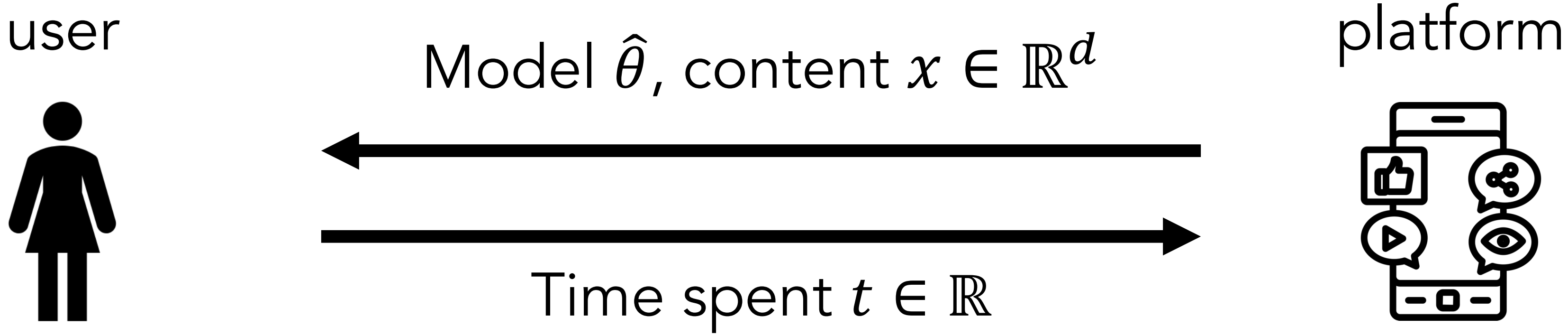
Example 2: Privacy-conscious user



Naive platform strategy: Learn a user model $\hat{\theta}$,
use bandit algorithm to suggest content

Result: User avoids "feature-revealing content" by spending little time on content that for which x_p is large

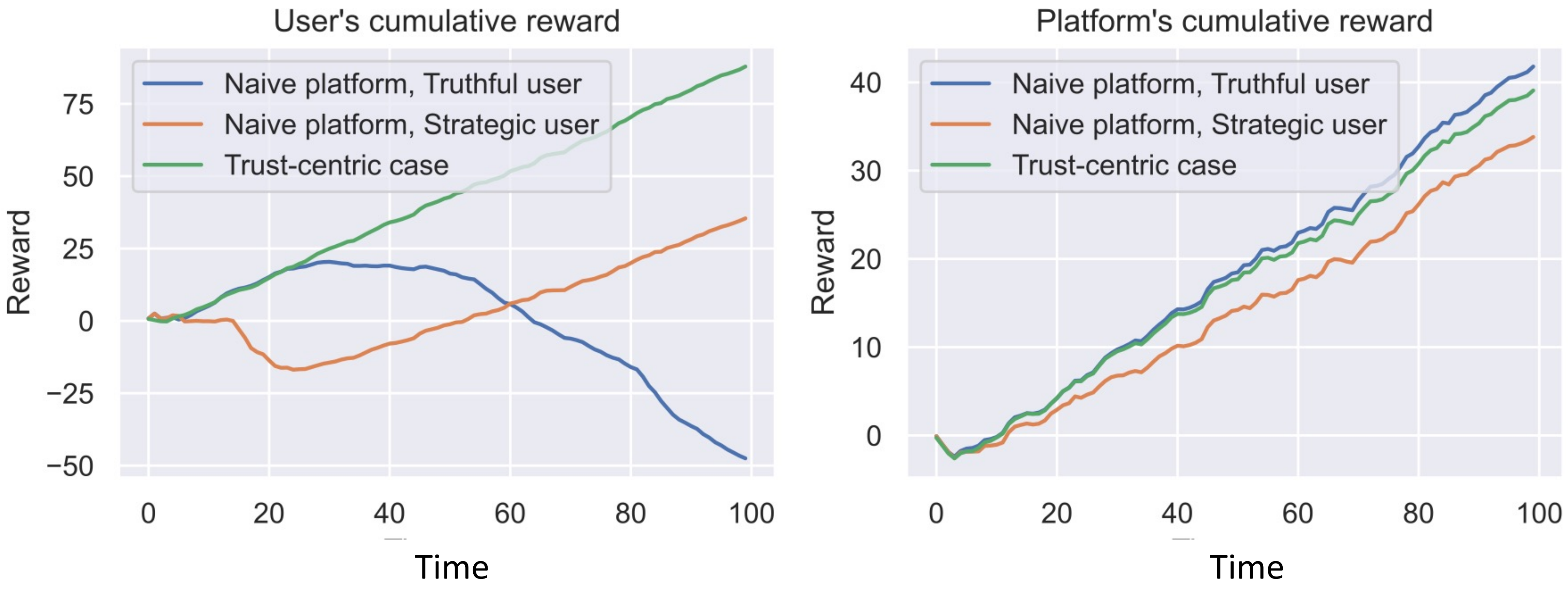
Example 2: Privacy-conscious user



The platform can accommodate the user's privacy concerns: only recommend content with $x_p = 0$ to the user

Cooperating helps platform learn as much as it can: The platform can't infer θ_p anyways, but learns the rest of θ

Example 2: Privacy-conscious user



Trust improves both **platform** and **user** reward!

Takeaways

In recommendation, users are platforms' primary data sources.

Need to account for users' ability to adapt and strategize.

Building trust can benefit both the user and platform.

We model recommendation as alternating two-player game.

Provide formalization of trust → can study effect of cooperation.

Lots of future work: cost of distrust, user studies, better algorithms, & more!

Thank you!

@cen_sarah