The Dynamics and Economy of Recommender Systems

Haifeng Xu

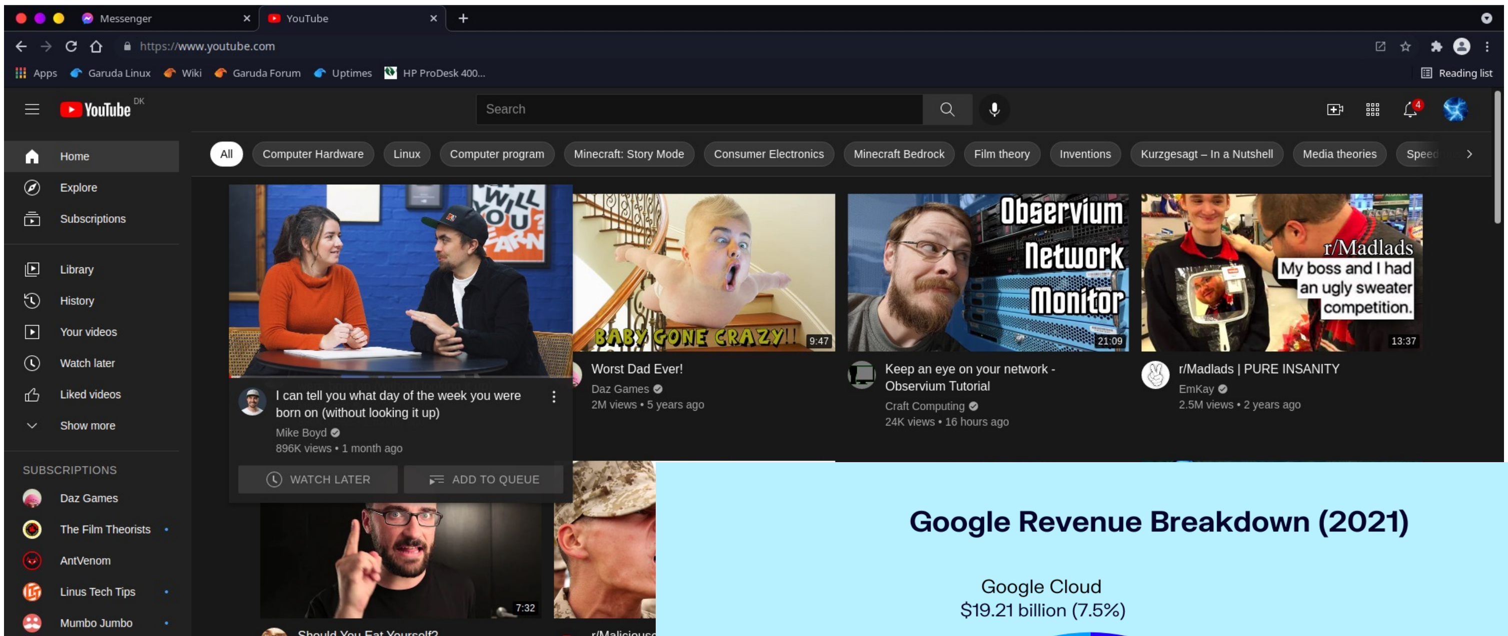
Department of Computer Science
and Data Science Institute

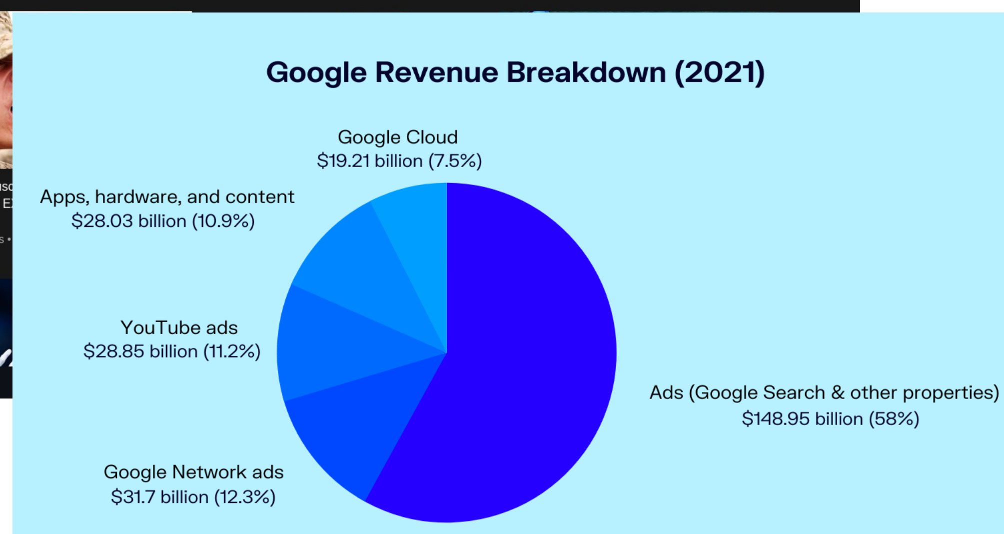
University of Chicago

SIGMA LAB
Strategic IntelliGence in Machine Agents

Recommender System (RS)

➢ An indispensable component of modern information systems





Classic Research Paradigm in RSs

System learning
in
static environments

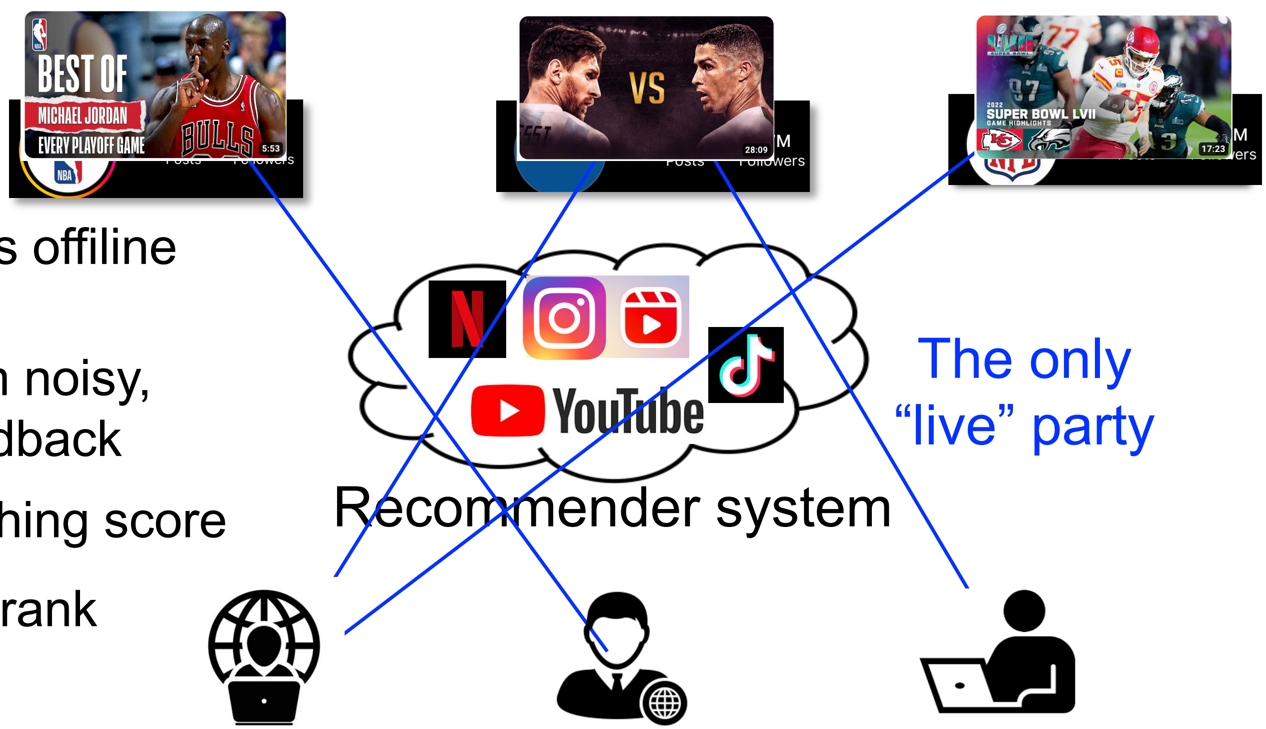
Static

Online vs offiline
learning

Learn from noisy,
bandit feedback

Learn matching score

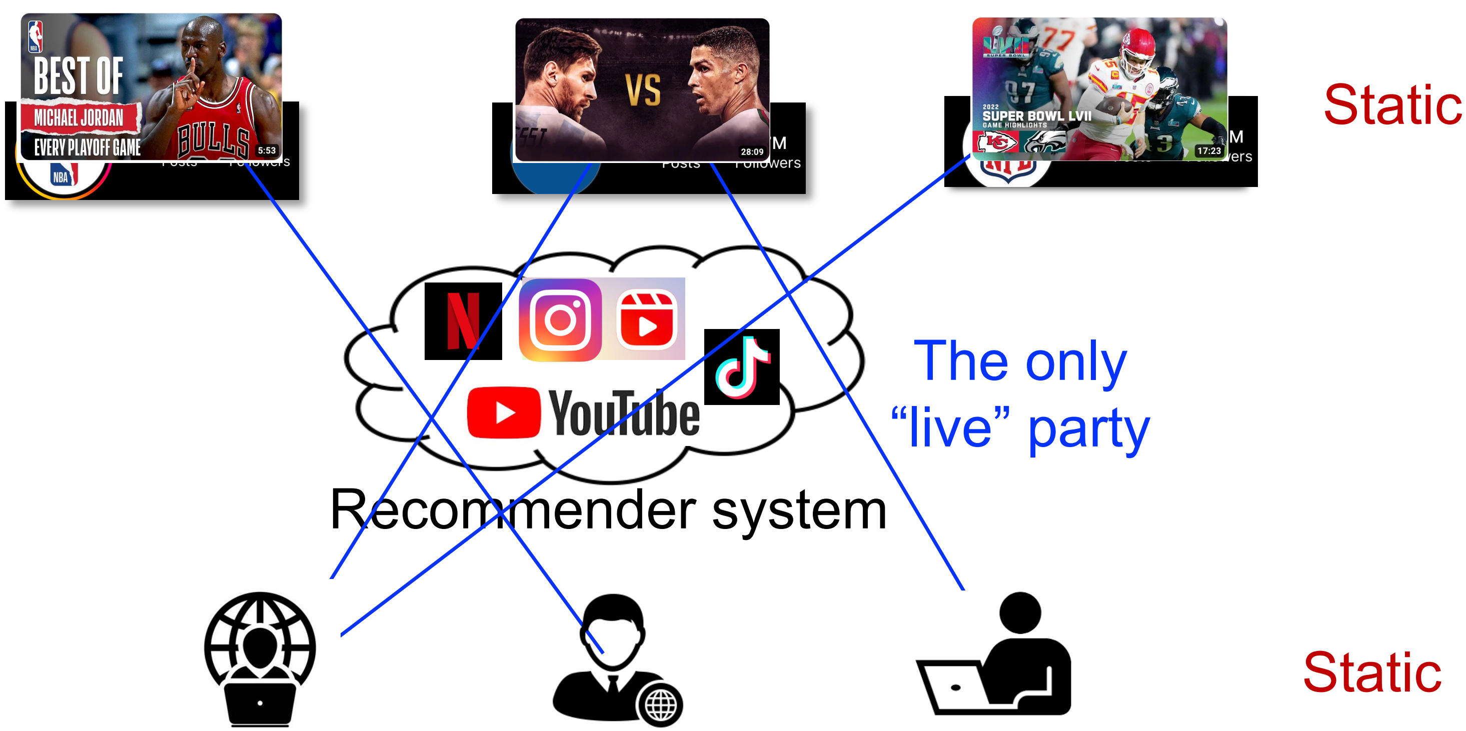
Learn to rank



The only
“live” party

Static

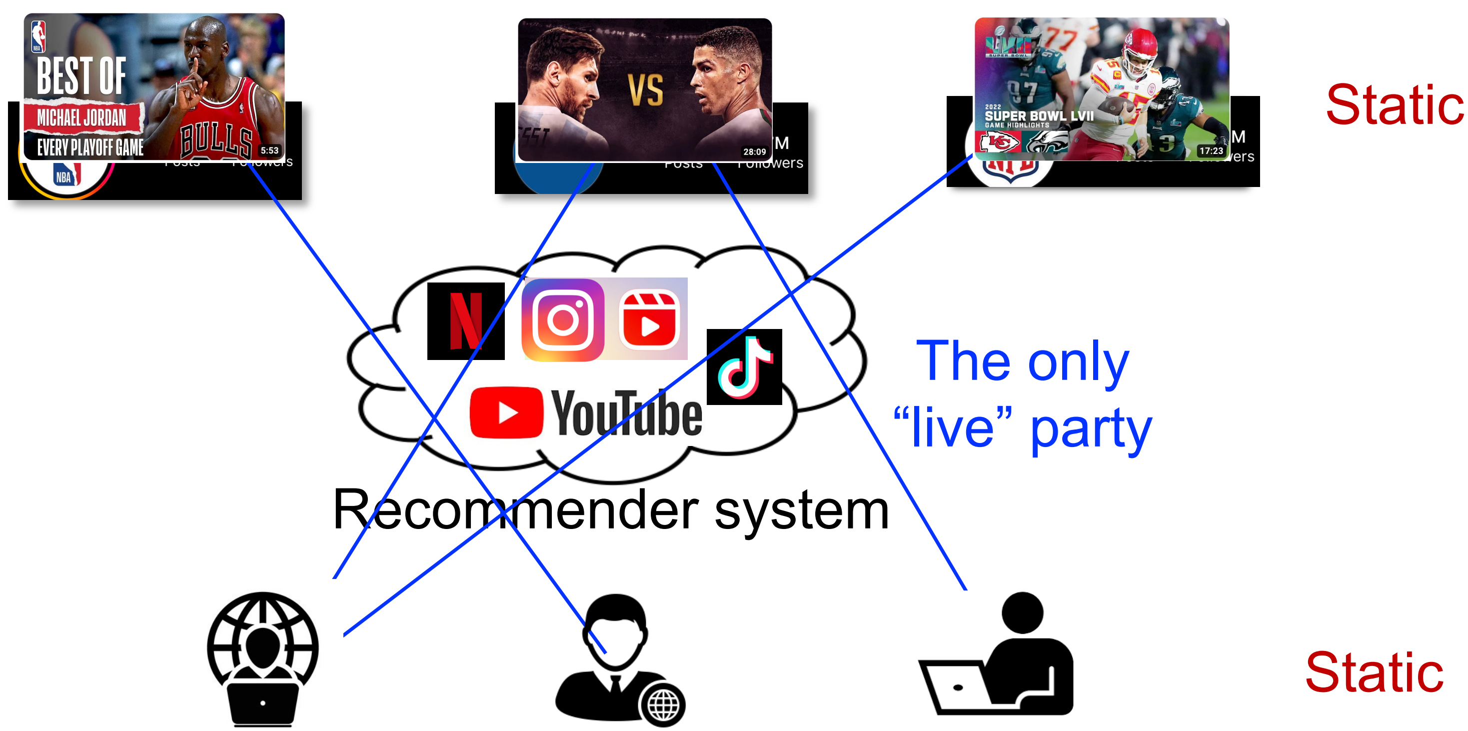
However…Numerous Evidence Supports Dynamic (Often Adaptive) Creator and User Behaviors



SIGMA LAB
Strategic IntelliGence in Machine Agents

However…Numerous Evidence Supports Dynamic (Often Adaptive) Creator and User Behaviors

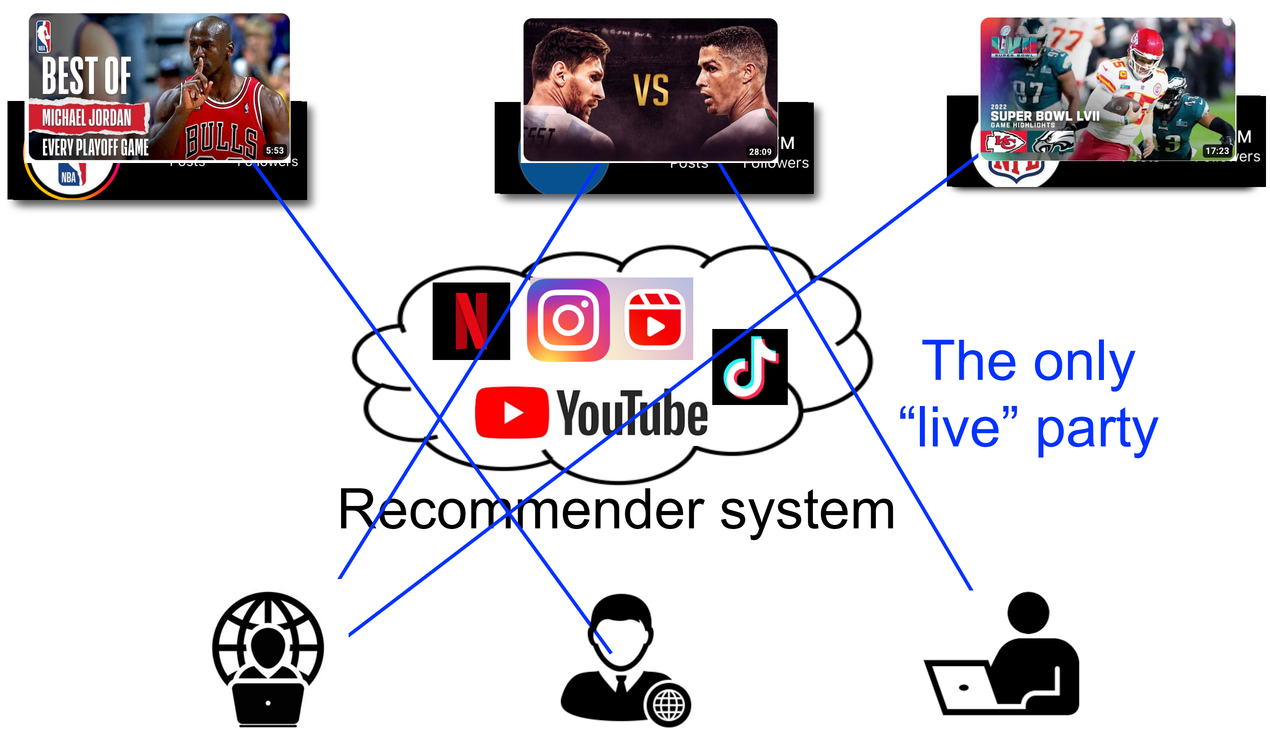
- **Creators** create longer videos after Youtube switches to use view duration to evaluate quality [MC’23]
- **RS users** are explorative at beginning (shown in many behavioral studies); Their feedback becomes more accurate only after sufficient experience



How the YouTube Algorithm Works in 2023: The Complete Guide. Stacey McLachlan, Paige Cooper (2023)

Rethinking this modeling paradigm….

System learning
in
static environments



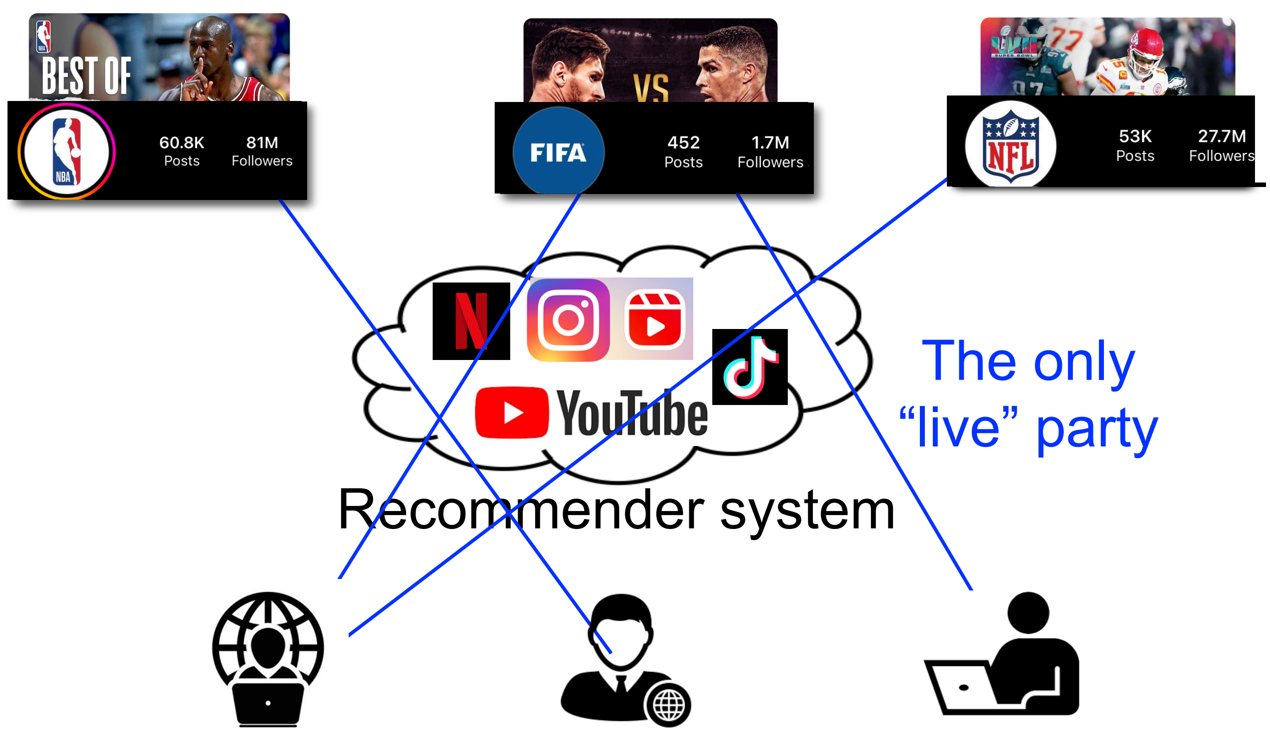
Static

The only
“live” party

Static

Rethinking this modeling paradigm….

System learning
in
static environments



Intelligent
(learning)
creators

The only
“live” party

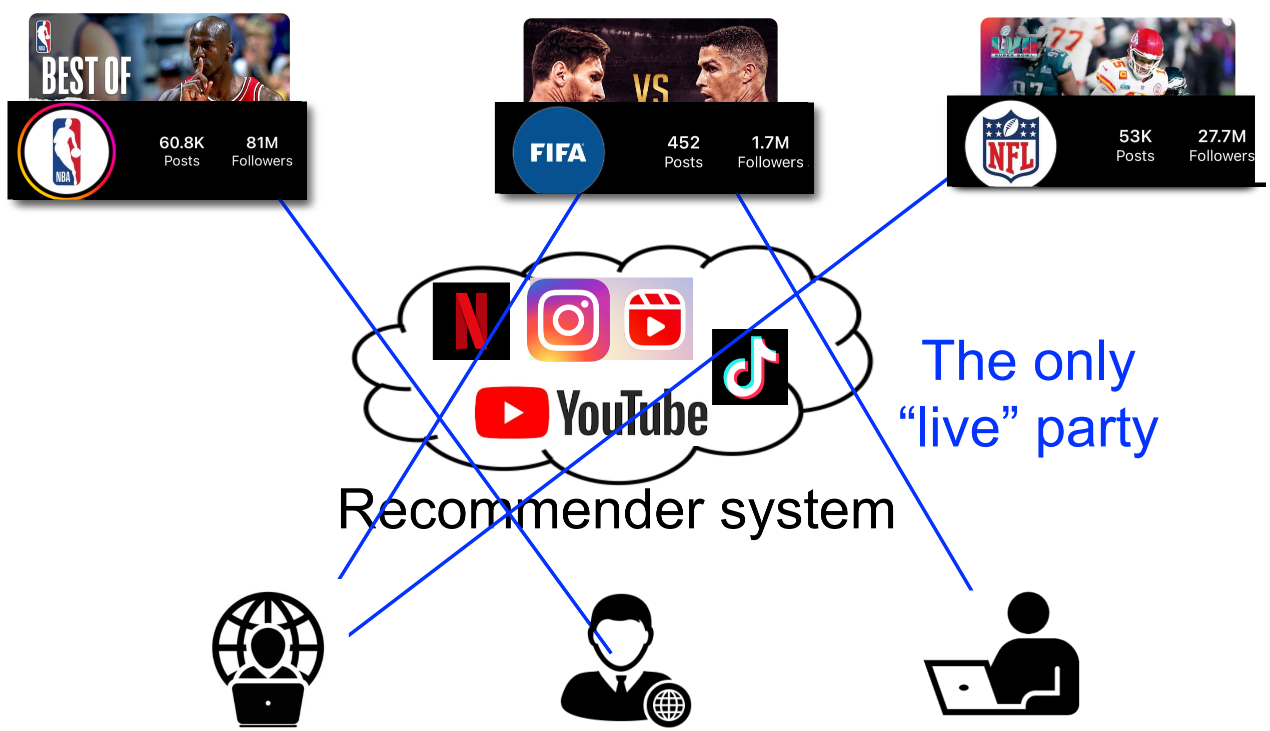
Intelligent
(learning)
users

Theme of This Talk

(multi-agent) economic modeling and optimization of recommender systems

Rethinking this modeling paradigm….

Multi-agent ~~System~~ learning
in
non-stationary ~~static~~ environments



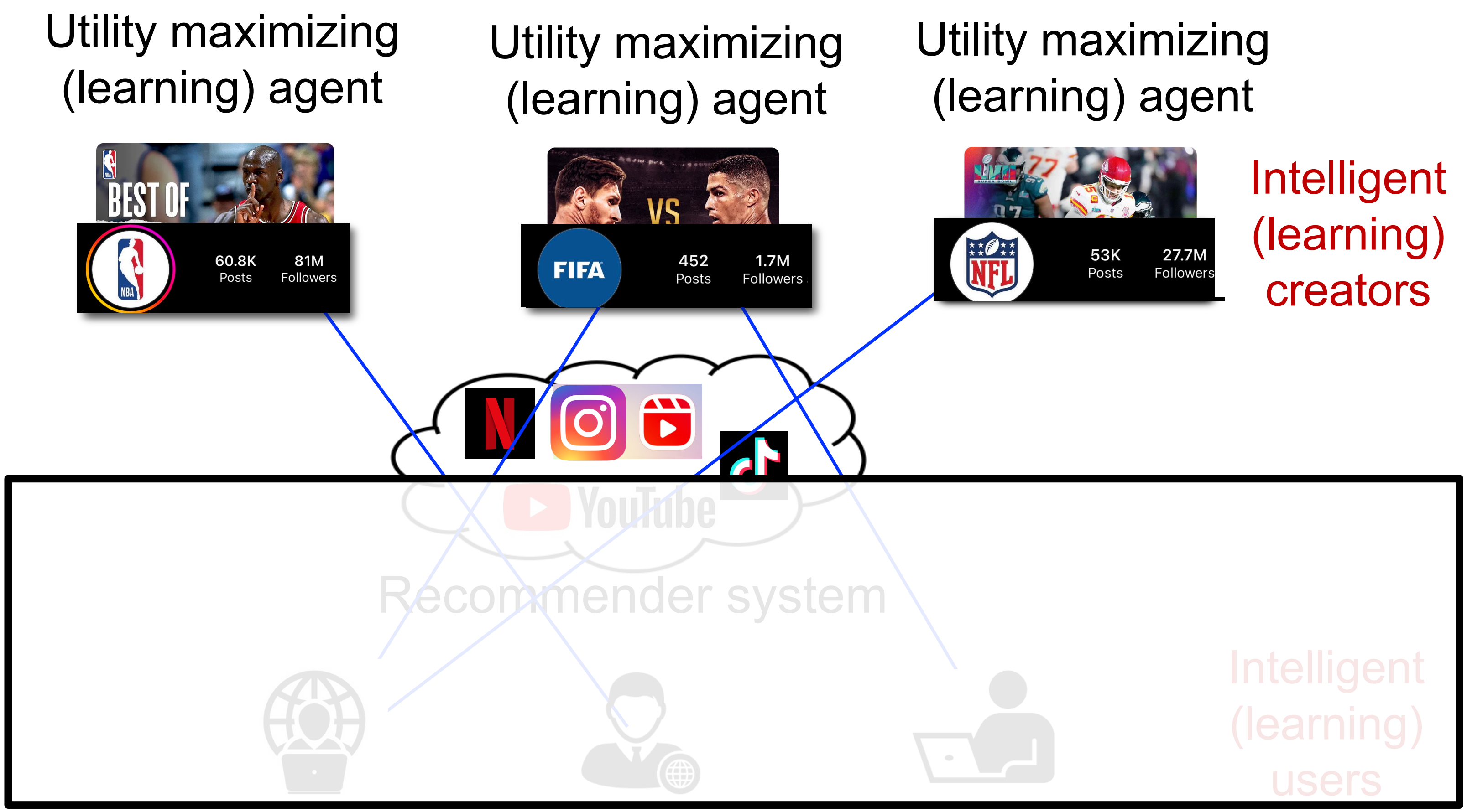
Intelligent (learning) creators

Intelligent (learning) users

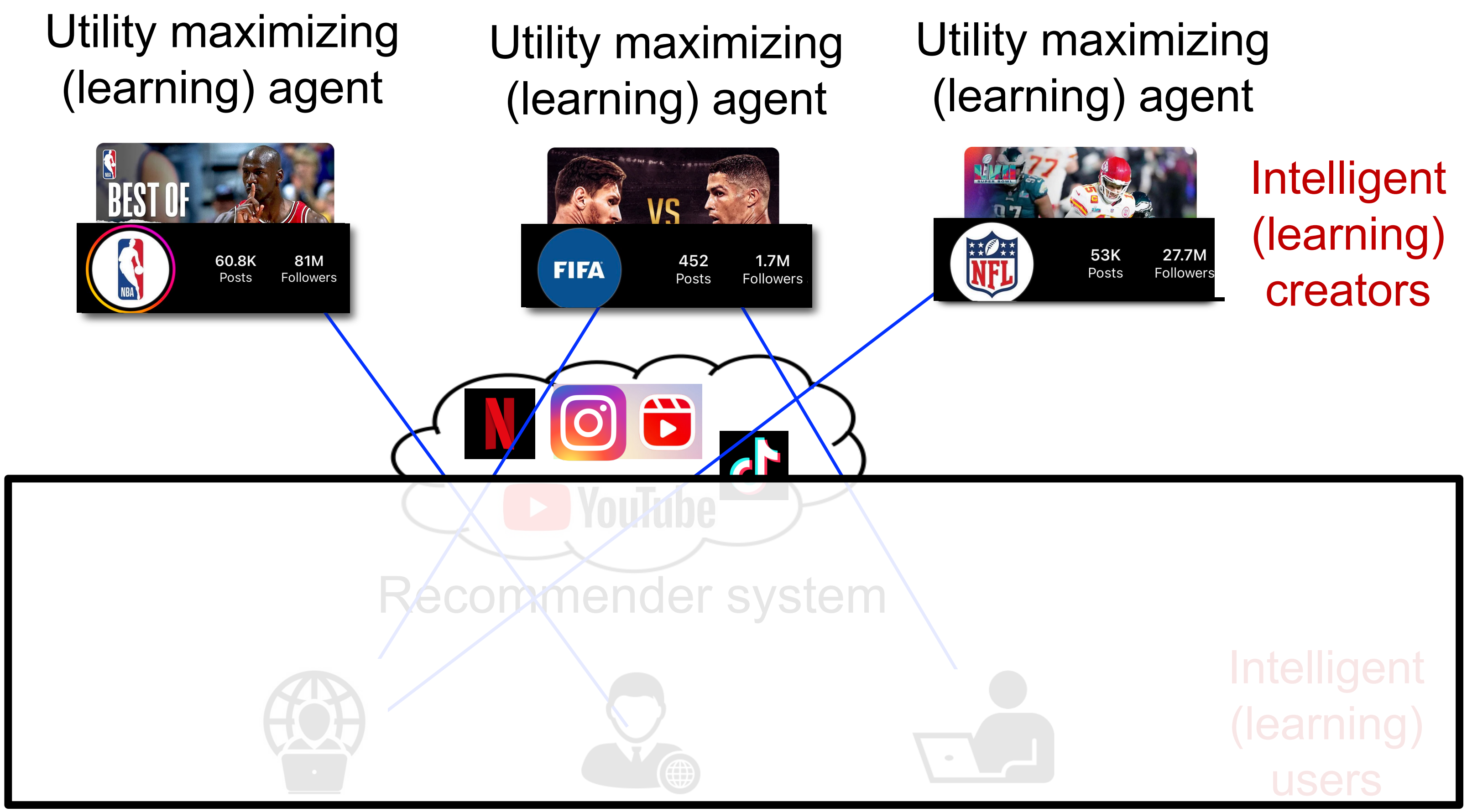
Outline

Part 1: Interacting with Strategic Creators

Part 2: Learning from Learning Users


Utility maximizing (learning) agent
Utility maximizing (learning) agent
Utility maximizing (learning) agent
BEST OF
60.8K Posts
81M Followers
FIFA
452 Posts
1.7M Followers
53K Posts
27.7M Followers
Intelligent (learning) creators
YouTube
Recommender system
Intelligent (learning) users

Q: do creators' learning dynamics converge to good or bad outcome?



Q: do creators' learning dynamics converge to good or bad outcome?

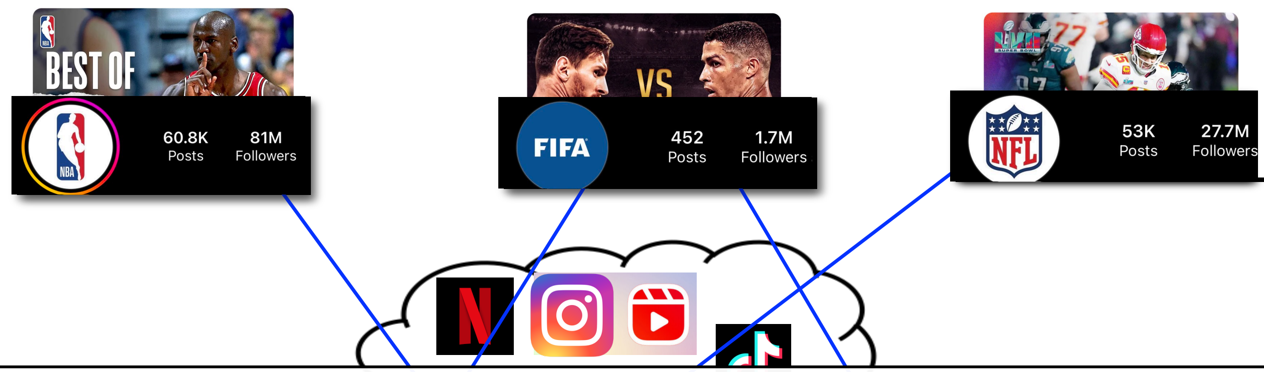
The Competing Content Creation (C3) Game

Utility maximizing (learning) agent

Utility maximizing (learning) agent

Utility maximizing (learning) agent

Intelligent (learning) creators



Reward Generating Environments

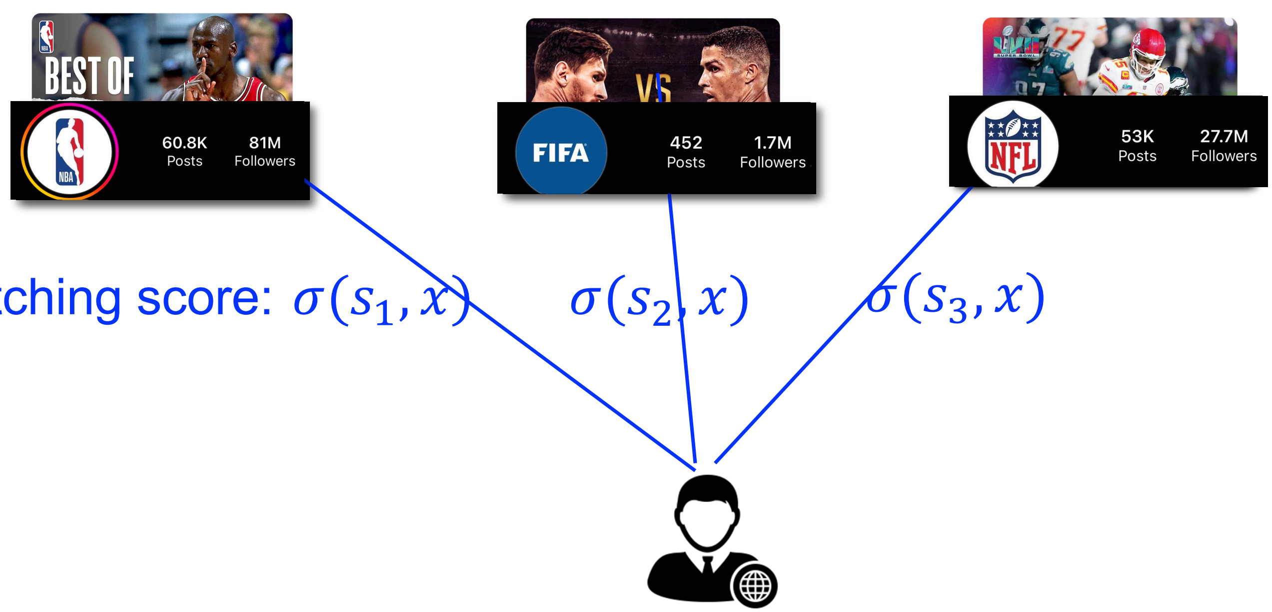
How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

The Competing Content Creation (C3) Game

Contents (actions)

$s_1 \in S_1$ $s_2 \in S_2$ $s_3 \in S_3$



Estimated matching score: $\sigma(s_1, x)$ $\sigma(s_2, x)$ $\sigma(s_3, x)$

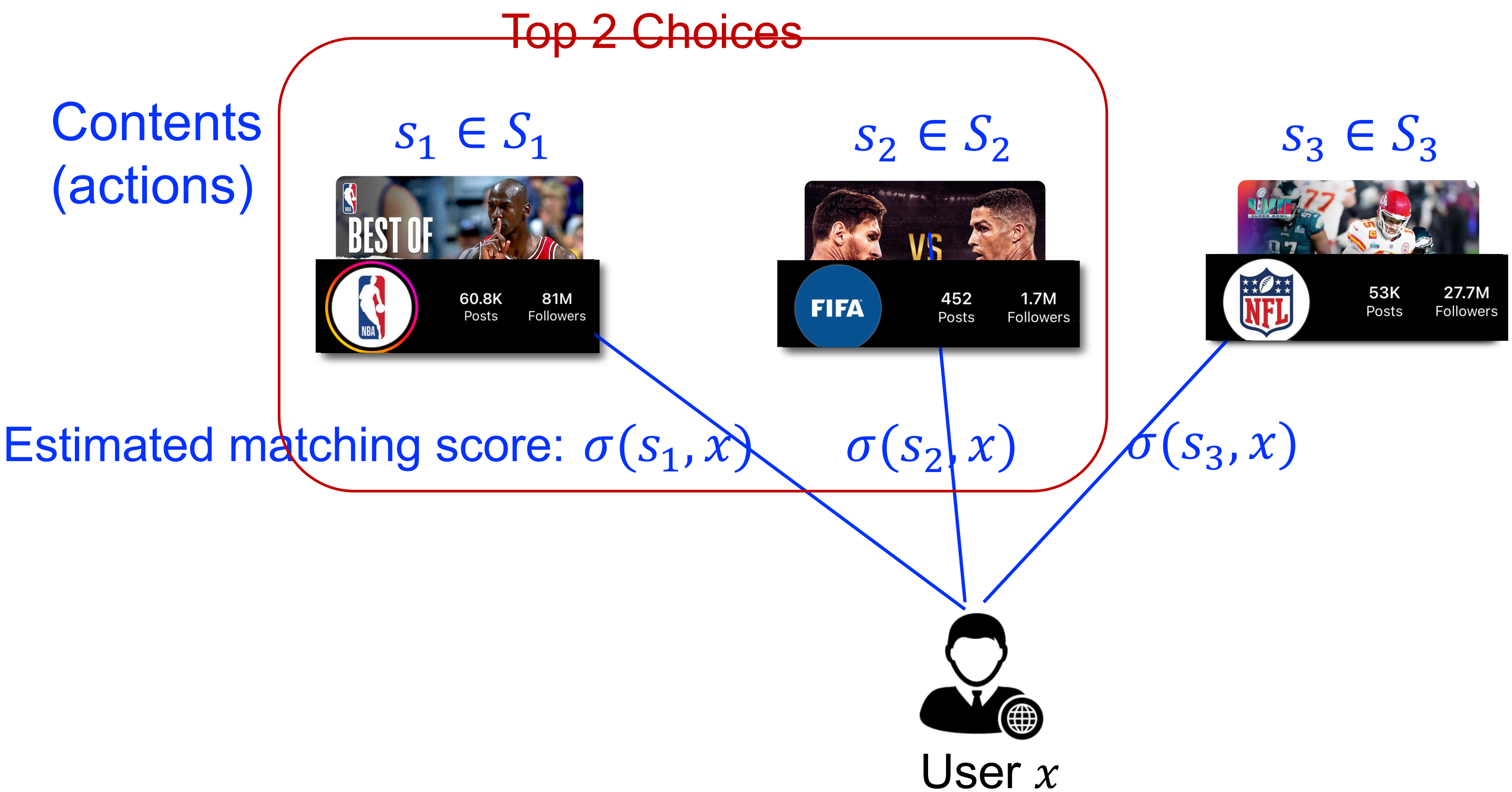
User x

drawn from population/distribution F

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

The Competing Content Creation (C3) Game



drawn from population/distribution F

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

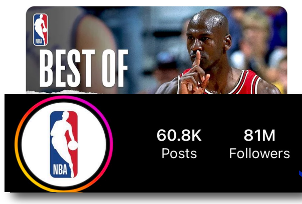
Q: do creators' learning dynamics converge to good or bad outcome?

The Competing Content Creation (C3) Game

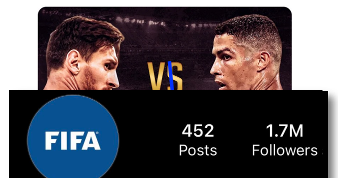
Top 2 Choices

Contents (actions)

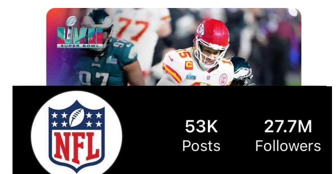
$s_1 \in S_1$



$s_2 \in S_2$



$s_3 \in S_3$



Estimated matching score: $\sigma(s_1, x)$ $\sigma(s_2, x)$ $\sigma(s_3, x)$

True matching score $\sigma(s_1, x) + \epsilon_1$ $(<)$ $\sigma(s_2, x) + \epsilon_2$

i.i.d Gumbel noise (a.k.a random utility model [Manski'77])

User x

drawn from population/distribution F

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

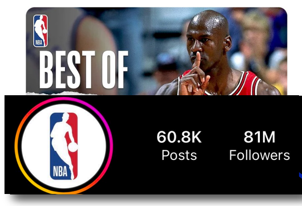
The Competing Content Creation (C3) Game

Top 2 Choices

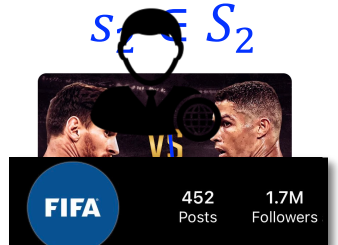
$\sigma(s_2, x) + \epsilon_2$

Contents (actions)

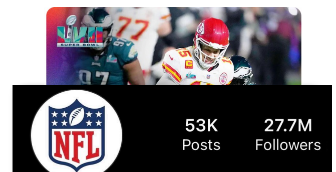
$s_1 \in S_1$



$s_2 \in S_2$



$s_3 \in S_3$



Estimated matching score: $\sigma(s_1, x)$ $\sigma(s_2, x)$ $\sigma(s_3, x)$

True matching score $\sigma(s_1, x) + \epsilon_1$ $(<)$

i.i.d Gumbel noise (a.k.a random utility model [Manski'77])

drawn from population/distribution F

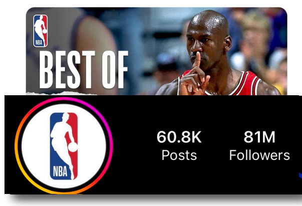
How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

The Competing Content Creation (C3) Game

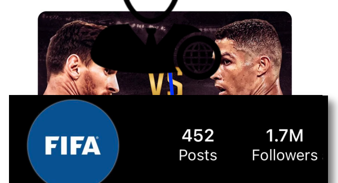
Contents (actions)

$s_1 \in S_1$

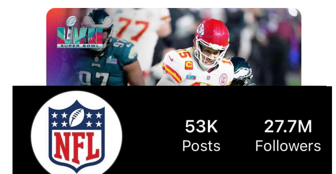


$\sigma(s_2, x) + \epsilon_2$ =Creator 2's utility (= welfare)

$s_2 \in S_2$



$s_3 \in S_3$



Estimated matching score: $\sigma(s_1, x)$ $\sigma(s_2, x)$ $\sigma(s_3, x)$

True matching score $\sigma(s_1, x) + \epsilon_1$ (<)

drawn from population/distribution F

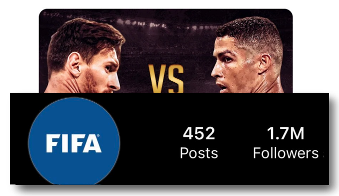
How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

The Competing Content Creation (C3) Game

$\sigma(s_2, x) + \epsilon_2$ =Creator 2's utility (= welfare)

$s_2 \in S_2$



$\sigma(s_2, x)$

How to model each content creator's behavior in the system?

→ Simple – they are just *any* "reasonable" (no-regret) learners who learn to maximize their own users' welfare/happiness

$$\mathbb{E}_{x\sim F}\left[(\sigma(s_2, x) + \epsilon_2) \cdot \mathbb{I}(x \text{ visits the creator})\right]$$

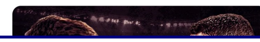
How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

The Competing Content Creation (C3) Game

$\sigma(s_2, x) + \epsilon_2$ =Creator 2's utility (= welfare)

$s_2 \in S_2$



- The goal here is NOT to learn $\sigma(s, x)$ or set S_i's
- Goal is to study convergence property in C3 under (non-stationary) creator learning dynamics, and resultant system welfare

→ Simple – they are just *any* "reasonable" (no-regret) learners who learn to maximize their own users' welfare/happiness

$$\mathbb{E}_{x\sim F}\left[\left(\sigma(s_2, x) + \epsilon_2\right) \cdot \mathbb{I}(x \text{ visits the creator})\right]$$

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

The Competing Content Creation (C3) Game

$\sigma(s_2, x) + \epsilon_2$ =Creator 2's utility (= welfare)

$s_2 \in S_2$

- The goal here is NOT to learn $\sigma(s, x)$ or set S_i's
- Goal is to study convergence property in C3 under (non-stationary) creator learning dynamics, and resultant system welfare
  - We do not directly consider revenue, but RS's revenue is often aligned with total user welfare

→ Simple – they are just *any* "reasonable" (no-regret) learners who learn to maximize their own users' welfare/happiness

$$\mathbb{E}_{x \sim F}\left[(\sigma(s_2, x) + \epsilon_2) \cdot \mathbb{I}(x \text{ visits the creator})\right]$$

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\text{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1 + \beta)\log(K)}$$

K = # of recommendation slots
β^2 = variance of Gumbel noise

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\text{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1+\beta)\log(K)}$$

K = # of recommendation slots
β^2 = variance of Gumbel noise

Remark

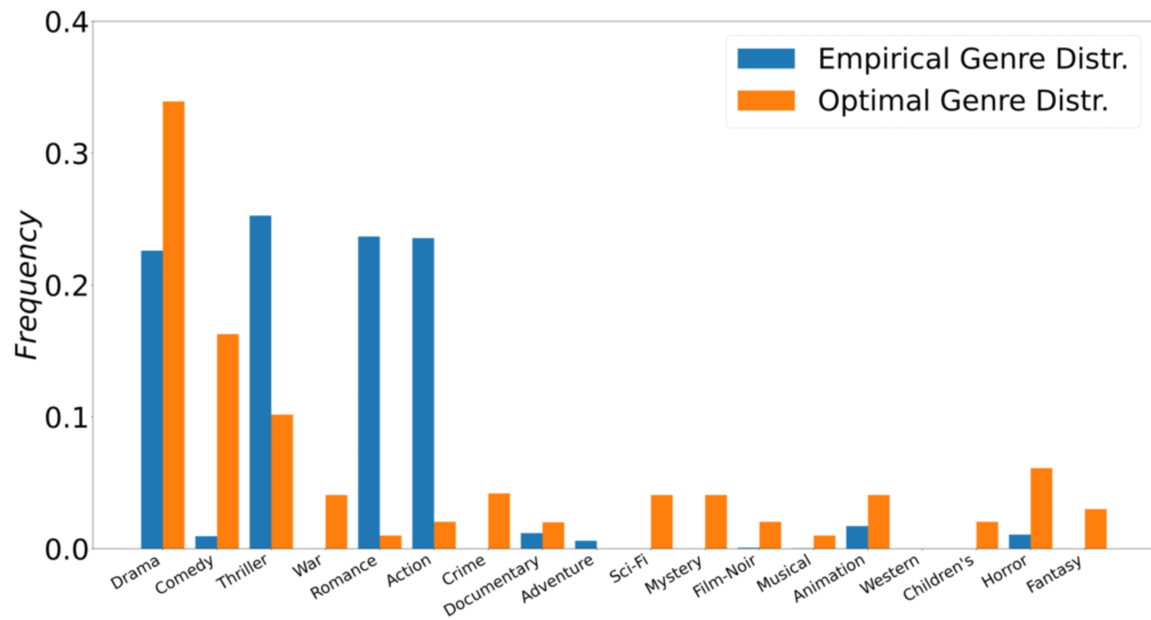
- Also known as the *price of anarchy (PoA)*
  - A very plausible and robust prediction about welfare [Blum et al.'08]
- The bound is an intrinsic property of content competition and user choices
  - Independent of matching score function $\sigma(s, x)$ and #users

Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\text{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1+\beta)\log(K)}$$

Simulation on MovieLens dataset between empirical and ideal content distributions



How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\textit{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1+\beta)\log(K)}$$

This bound is (order-wise) tight

Prop 1. There exists C3 games such that is PoA (even for Nash) satisfies

$$\frac{\textit{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \leq 1 - \frac{1}{2 + 5\beta\log(K)}$$

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\text{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1+\beta)\log(K)}$$

Proof Sketch

- Core insight – C3 is a smooth game [Roughgarden'12]
- Proof turns out to be quite involved
  - Hinges on various analytical properties about the C3 game
  - E.g., total welfare is submodular in the set of contents

Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\textit{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1+\beta)\log(K)}$$

Proof Sketch

- Core insight – C3 is a smooth game [Roughgarden'12]
- Proof turns out to be quite involved
  - Hinges on various analytical properties about the C3 game
  - E.g., total welfare is submodular in the set of contents
- **Fun fact:** smoothness technique for C3 yields (order-wise) tight PoA
  - Before this, only 3 classes of games are known to satisfy this (linear congestion game, second price auction and valid utility games)

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\text{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1+\beta)\log(K)}$$

Economic insights:

- More recommendation slots (K large), more efficient the system is

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\text{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1+\beta)\log(K)}$$

Economic insights:

- More recommendation slots (K large), more efficient the system is
- Setting proper creator incentives matters a lot!

In previous model — Creator's utility ~ True user matching score

$= \sigma(s_2, x) + \epsilon_2$

$\approx$ user engagement

What if — Creator's utility ~ Pr(being matched to user)

$\approx$ user traffic

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\textit{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1+\beta)\log(K)}$$

Economic insights:

- More recommendation slots (K large), more efficient the system is
- Setting proper creator incentives matters a lot!

Prop 2. Suppose creators' utilities are propositional to user traffic in C3 games, then there are C3 games such that

$$\frac{\textit{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \leq \frac{1}{2}$$

What if Creator's utility ~ Pr(being matched to user)

≈ user traffic

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

SIGMA LAB
Strategic IntelliGence in Machine Agents

Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\text{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1+\beta)\log(K)}$$

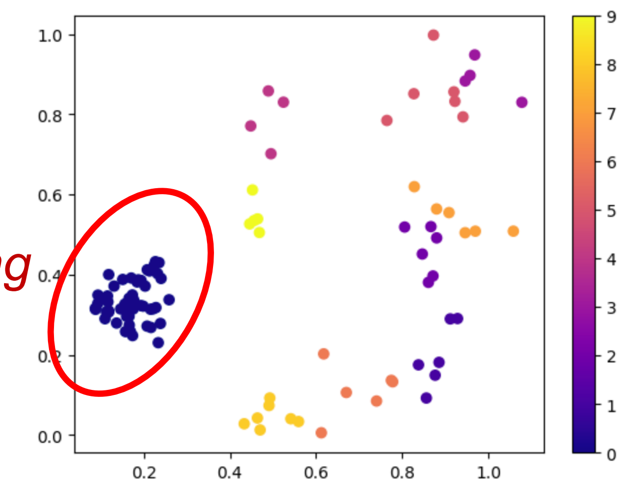
Economic insights:

- More recommendation slots (K large), more efficient the system is
- Setting proper creator incentives matters a lot!

What situation is worrisome?

Trend-vs-Niche!

A dominating user group



How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

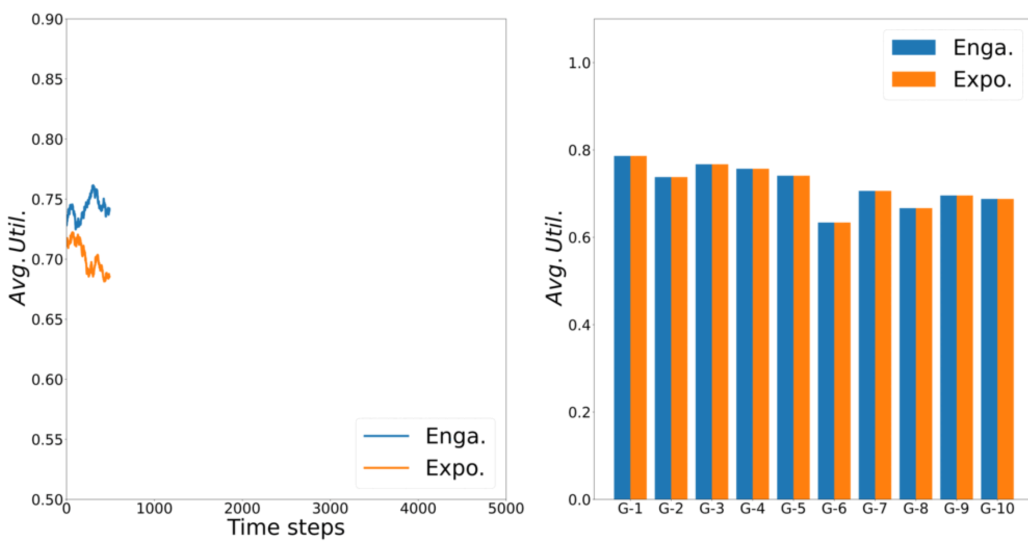
Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\textit{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1+\beta)\log(K)}$$

Economic insights:

- More recommendation slots (K large), more efficient the system is
- Setting proper creator incentives matters a lot!



Simulations of no-regret creators on synthetic data

— user engagement

— user traffic

SIGMA LAB
Strategic IntelliGence in Machine Agents

Q: do creators' learning dynamics converge to good or bad outcome?

Theorem [YLNWX, ICML'23]. In any C3 games, if each creator generates contents via *any* no regret learning algorithms, then w.h.p.

$$\frac{\text{Accumulated total welfare}}{\text{Idealized Maximum Welfare}} \geq 1 - \frac{1}{1 + (1+\beta)\log(K)}$$

Economic insights:

- More recommendation slots (K large), more efficient the system is
- Setting proper creator incentives matters a lot!
- Larger β – users are more explorative – increases efficiency
- In practice, still constant fraction loss since $K \leq 12$

Next Question: how to further improve system's welfare?

How Bad is Top-K Recommendation under Competing Content Creators? Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Incentive Design for Rewarding Creation

Next Question: how to further improve system's welfare?

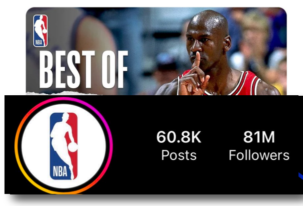
Incentive Design for Rewarding Creation

Previous mechanism's rewards ≈ created welfare

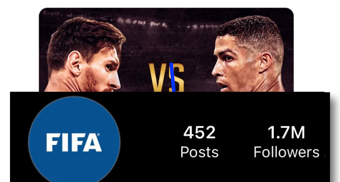
$\sigma(s_2, x) + \epsilon_2$ =Creator 2's utility (= welfare)

Contents (actions)

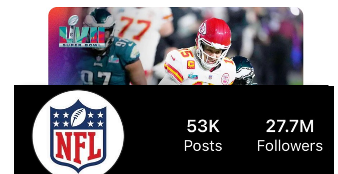
$s_1 \in S_1$



$s_2 \in S_2$



$s_3 \in S_3$



Estimated matching score: $\sigma(s_1, x)$ $\sigma(s_2, x)$ $\sigma(s_3, x)$

True matching score $\sigma(s_1, x) + \epsilon_1 \ (<) \ \sigma(s_2, x) + \epsilon_2$

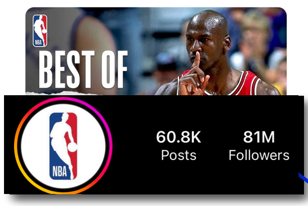
User x

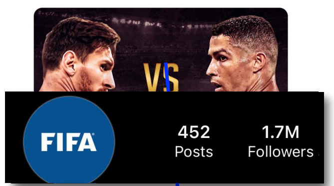
Incentive Design for Rewarding Creation

Q: Can we design/optimize the reward values R to "steer"/incentivize creators' collective behaviors towards better total welfare?

$R(\sigma_1, \sigma_{-1})$ $\quad$ $R(\sigma_2, \sigma_{-2})$ =Creator 2's utility (= welfare)

Contents (actions) $\quad s_1 \in S_1$ $\quad s_2 \in S_2$ $\quad s_3 \in S_3$





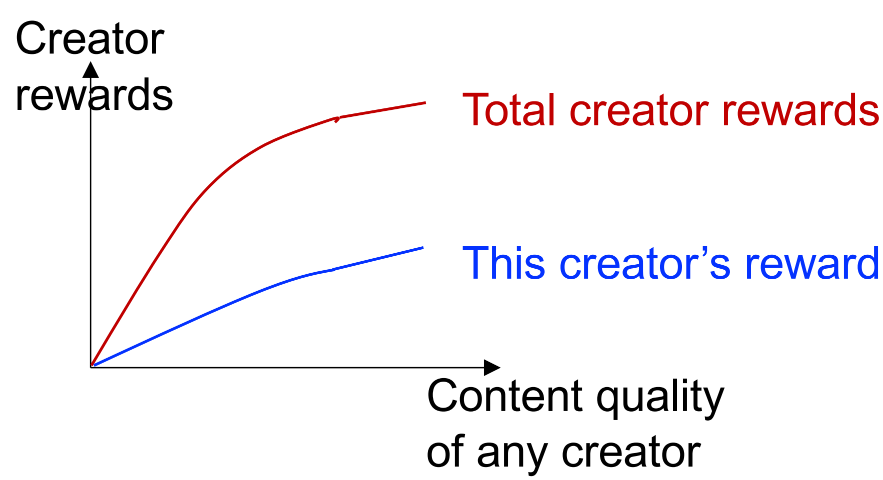
Estimated matching score: $\sigma(s_1, x)$ $\quad \sigma(s_2, x)$ $\quad \sigma(s_3, x)$

True matching score $\quad \sigma(s_1, x) + \epsilon_1 \ (<) \ \sigma(s_2, x) + \epsilon_2$

User x

Why current rewarding mechanism may not be good?

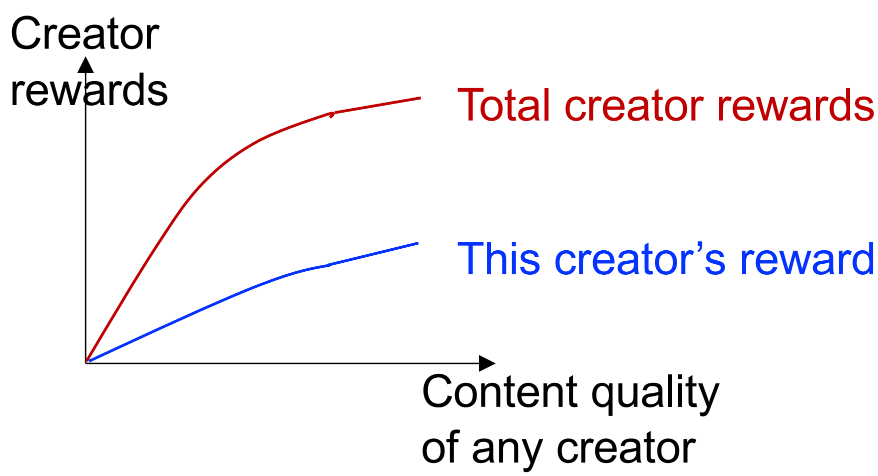
Theorem [Yao et al.'23]. If a rewarding mechanism R are both individual-monotone (better contents get more rewards) and group-monotone, then it necessarily suffer at least $1/K$ fraction of welfare loss at equilibrium



Why current rewarding mechanism may not be good?

Theorem [Yao et al.'23]. If a rewarding mechanism R are both individual-monotone (better contents get more rewards) and group-monotone, then it necessarily suffer at least $1/K$ fraction of welfare loss at equilibrium

- *User engagement* and *user traffic* do satisfy both; so do many natural rewarding mechanisms in real-world

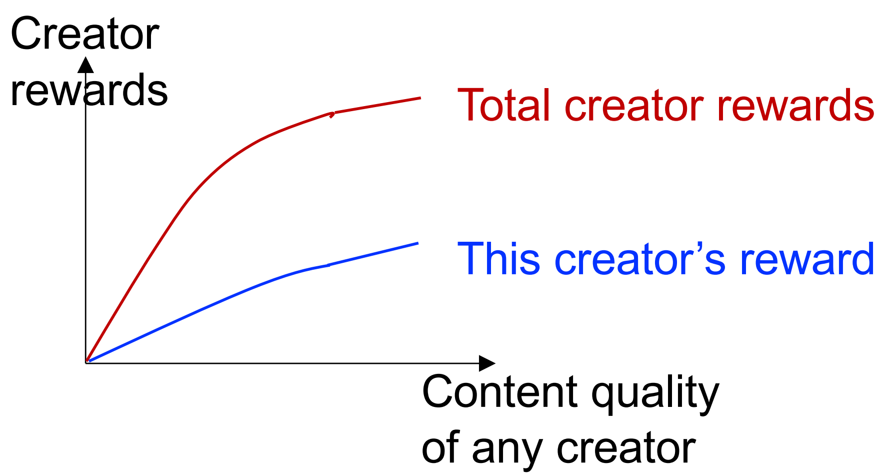


“Rethinking Incentives in Recommender Systems”

Why current rewarding mechanism may not be good?

Theorem [Yao et al.’23]. If a rewarding mechanism R are both individual-monotone (better contents get more rewards) and group-monotone, then it necessarily suffer at least $1/K$ fraction of welfare loss at equilibrium

- *User engagement* and *user traffic* do satisfy both; so do many natural rewarding mechanisms in real-world



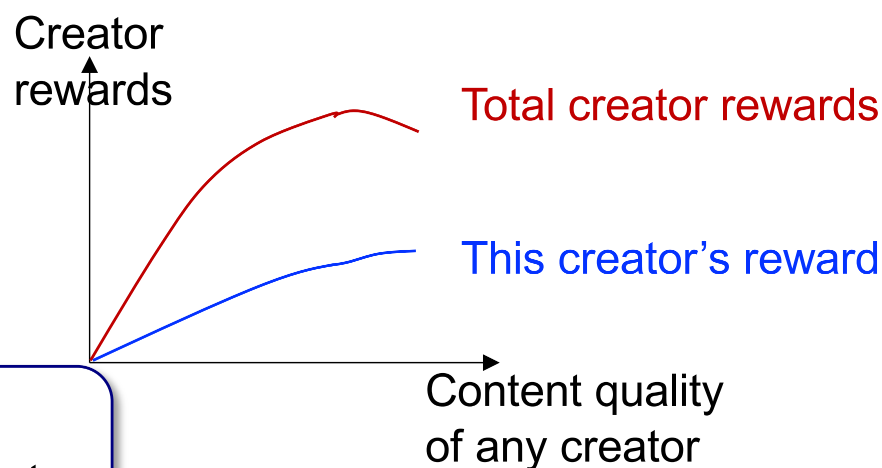
Rethinking Incentives in Recommender Systems: Are Monotone Rewards Always Beneficial? Fan Yao et al., in submission

"Rethinking Incentives in Recommender Systems"

Why current rewarding mechanism may not be good?

Theorem [Yao et al.'23]. If a rewarding mechanism R are both individual-monotone (better contents get more rewards) and group-monotone, then it necessarily suffer at least $1/K$ fraction of welfare loss at equilibrium

- *User engagement* and *user traffic* do satisfy both; so do many natural rewarding mechanisms in real-world
- To overcome this limitation, we drop group-monotonicity



Why reasonable?

Group-monotonicity is generally not satisfied in economic markets!

Rethinking Incentives in Recommender Systems: Are Monotone Rewards Always Beneficial? Fan Yao et al., in submission

“Rethinking Incentives in Recommender Systems”

Our new mechanism. We designed a new rewarding mechanism that drops group-monotoniciy, but provably achieves optimal welfare

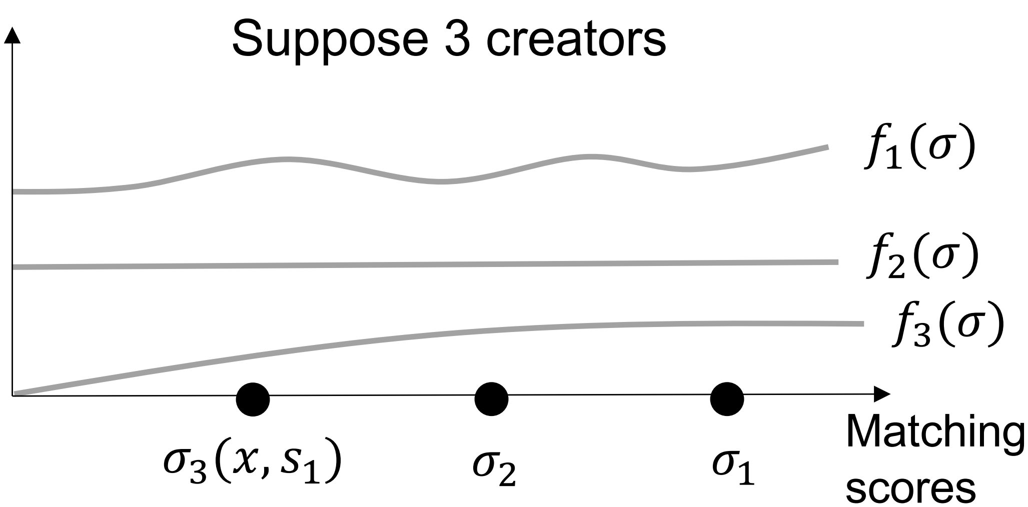
Core idea: reward based on how much you are better than the next

“Rethinking Incentives in Recommender Systems”

Our new mechanism. We designed a new rewarding mechanism that drops group-monotoniciy, but provably achieves optimal welfare

Core idea: reward based on how much you are better than the next

➢ Mechanism is fully described by functions f_1, f_2, f_3



Rethinking Incentives in Recommender Systems: Are Monotone Rewards Always Beneficial? Fan Yao et al., in submission

"Rethinking Incentives in Recommender Systems"

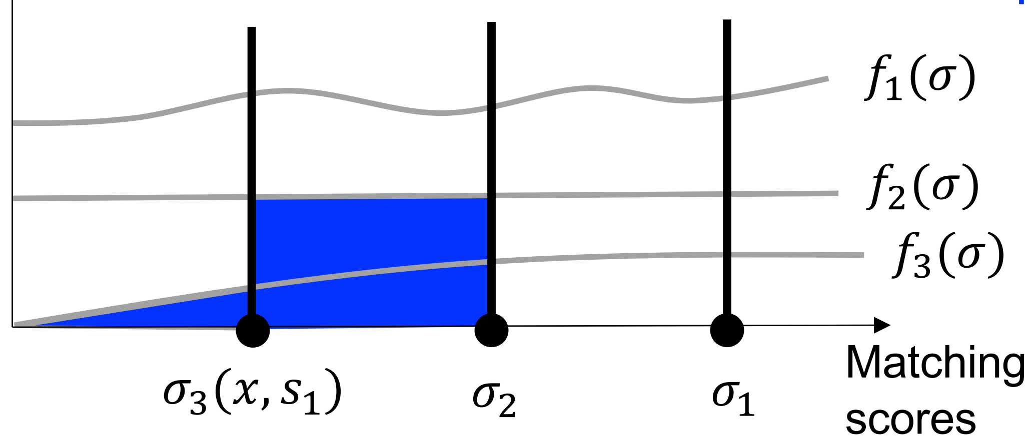
Our new mechanism. We designed a new rewarding mechanism that drops group-monotoniciy, but provably achieves optimal welfare

Core idea: reward based on how much you are better than the next

- Mechanism is fully described by functions f_1, f_2, f_3
- Reward = area of

Suppose 3 creators

Reward for σ_2

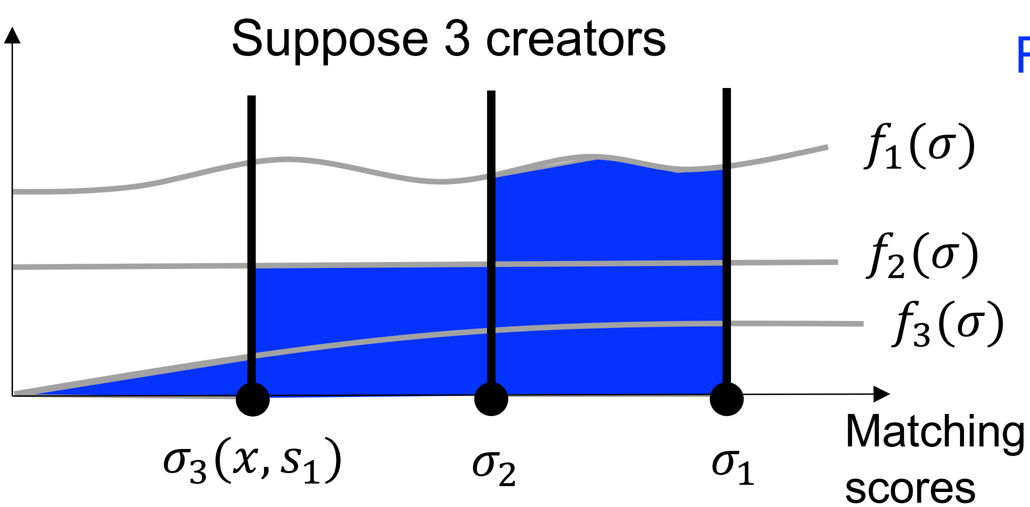


"Rethinking Incentives in Recommender Systems"

Our new mechanism. We designed a new rewarding mechanism that drops group-monotoniciy, but provably achieves optimal welfare

Core idea: reward based on how much you are better than the next

- Mechanism is fully described by functions f_1, f_2, f_3
- Reward = area of



Rethinking Incentives in Recommender Systems: Are Monotone Rewards Always Beneficial? Fan Yao et al., in submission

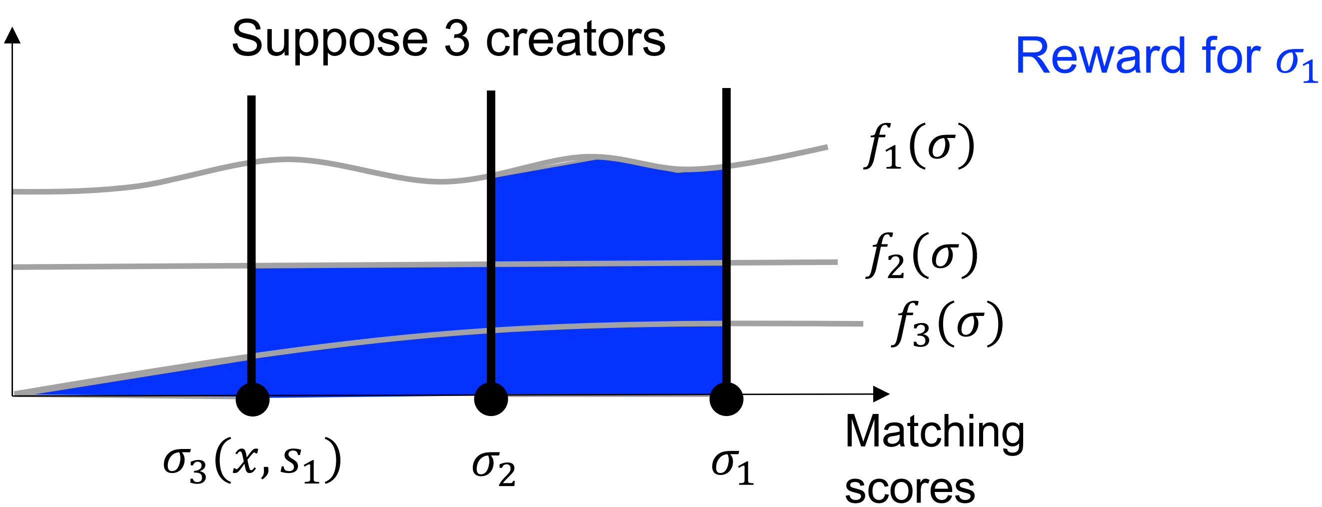
"Rethinking Incentives in Recommender Systems"

Our new mechanism. We designed a new rewarding mechanism that drops group-monotoniciy, but provably achieves optimal welfare

Core idea: reward based on how much you are better than the next

Advantages

✓ σ_1's reward decreases when σ_2 becomes better (i.e., competition reduces rewards)



Rethinking Incentives in Recommender Systems: Are Monotone Rewards Always Beneficial? Fan Yao et al., in submission

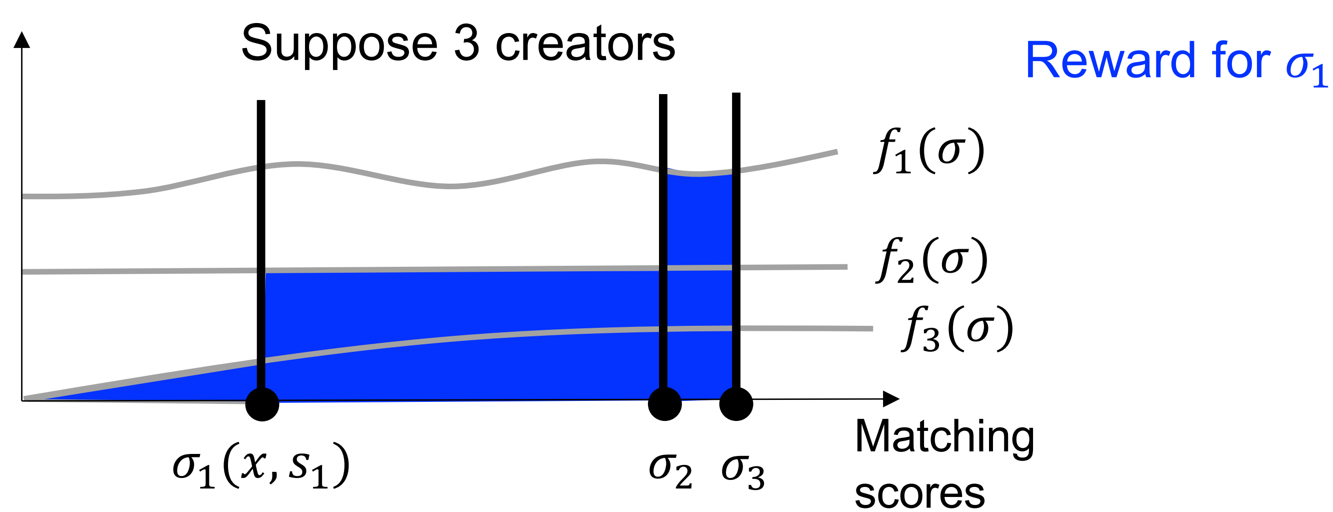
"Rethinking Incentives in Recommender Systems"

Our new mechanism. We designed a new rewarding mechanism that drops group-monotoniciy, but provably achieves optimal welfare

Core idea: reward based on how much you are better than the next

Advantages

✓ σ_1's reward decreases when σ_2 becomes better (i.e., competition reduces rewards)



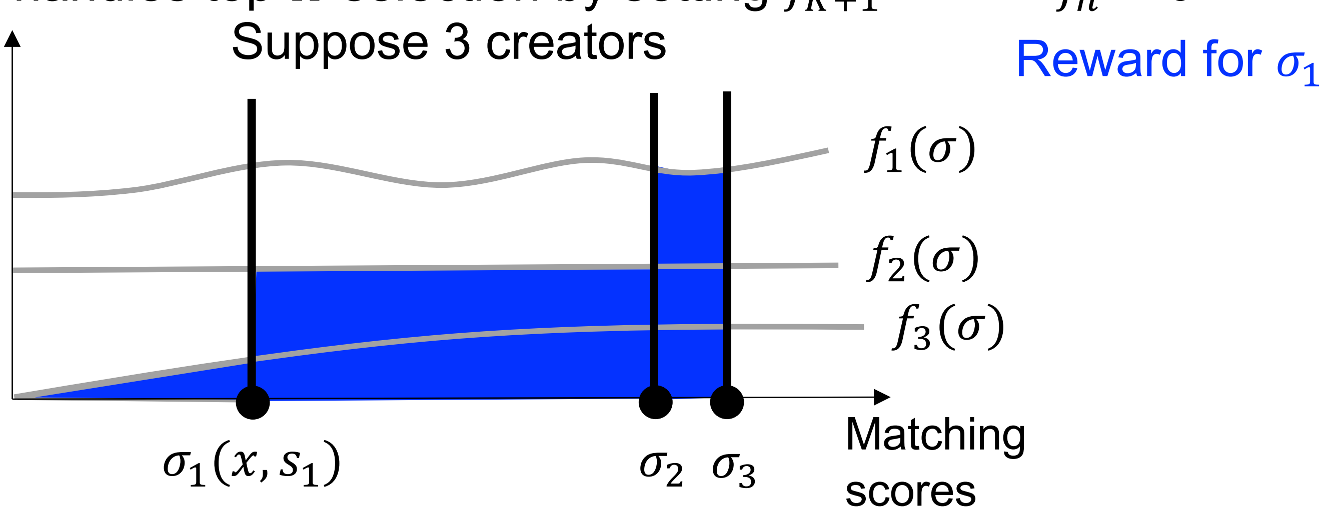
"Rethinking Incentives in Recommender Systems"

Our new mechanism. We designed a new rewarding mechanism that drops group-monotoniciy, but provably achieves optimal welfare

Core idea: reward based on how much you are better than the next

Advantages

- ✓ σ_1's reward decreases when σ_2 becomes better (i.e., competition reduces rewards)
- ✓ Naturally handles top-K selection by setting $f_{K+1} = \cdots = f_n = 0$



Rethinking Incentives in Recommender Systems: Are Monotone Rewards Always Beneficial? Fan Yao et al., in submission

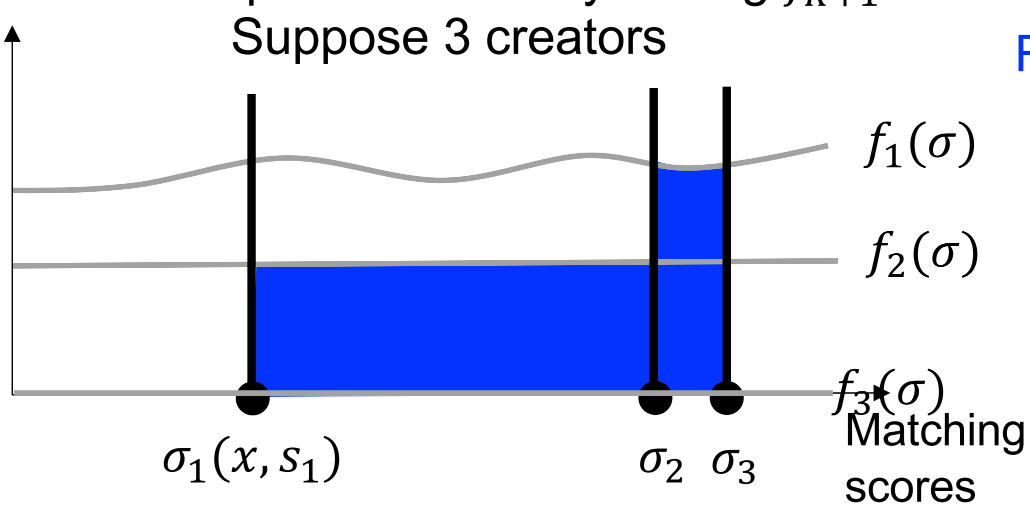
"Rethinking Incentives in Recommender Systems"

Our new mechanism. We designed a new rewarding mechanism that drops group-monotoniciy, but provably achieves optimal welfare

Core idea: reward based on how much you are better than the next

Advantages

- ✓ σ_1's reward decreases when σ_2 becomes better (i.e., competition reduces rewards)
- ✓ Naturally handles top-K selection by setting $f_{K+1} = \cdots = f_n = 0$



Rethinking Incentives in Recommender Systems: Are Monotone Rewards Always Beneficial? Fan Yao et al., in submission

“Rethinking Incentives in Recommender Systems”

Our new mechanism. We designed a new rewarding mechanism that drops group-monotoniciy, but provably achieves optimal welfare

- Proof idea: the reward mechanism above induce a potential game among creators, such that potential function = welfare function

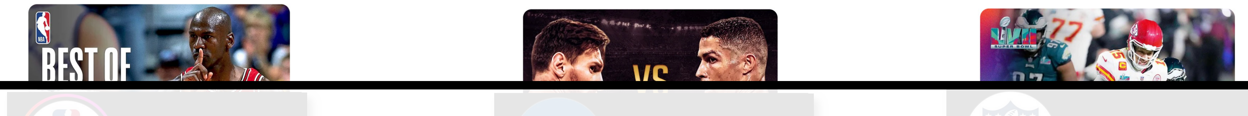
"Rethinking Incentives in Recommender Systems"

Our new mechanism. We designed a new rewarding mechanism that drops group-monotoniciy, but provably achieves optimal welfare

- Proof idea: the reward mechanism above induce a potential game among creators, such that potential function = welfare function
- Project done in collaboration with researchers at Meta
- Under live experiments on Instagram for >1month now
  - Disclaimer: the deployed algorithm is inspired by, but different from the exact design above

Rethinking Incentives in Recommender Systems: Are Monotone Rewards Always Beneficial? Fan Yao et al., in submission

Outline



Intelligent

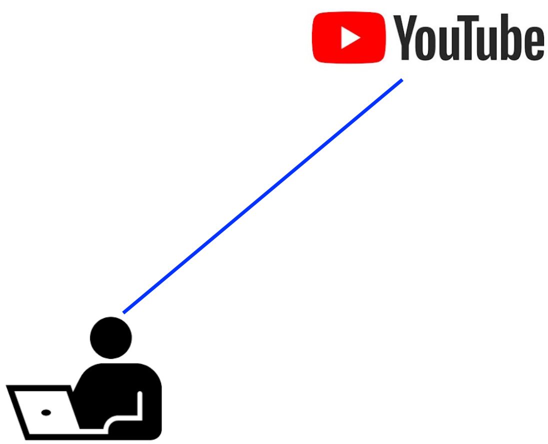
Part 1: Interacting with Strategic Creators

Part 2: Learning from Learning Users

users

Different Research Challenges

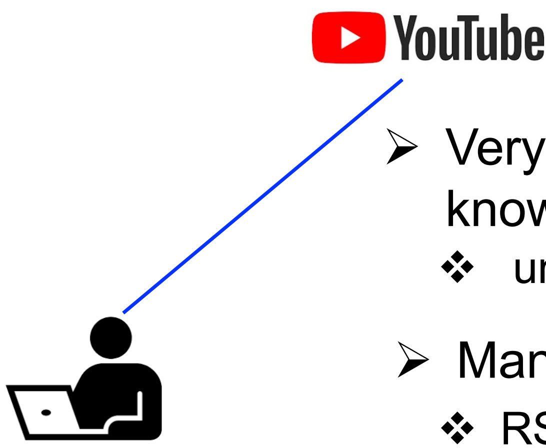
	Creator side	User side
Difficult	Incentives, Strategic behaviors	User preferences
Easy	Contents' embedding	Incentive (typically aligned)



User

Different Research Challenges

	Creator side	User side
Difficult	Incentives, Strategic behaviors	User preferences
Easy	Contents' embedding	Incentive (typically aligned)



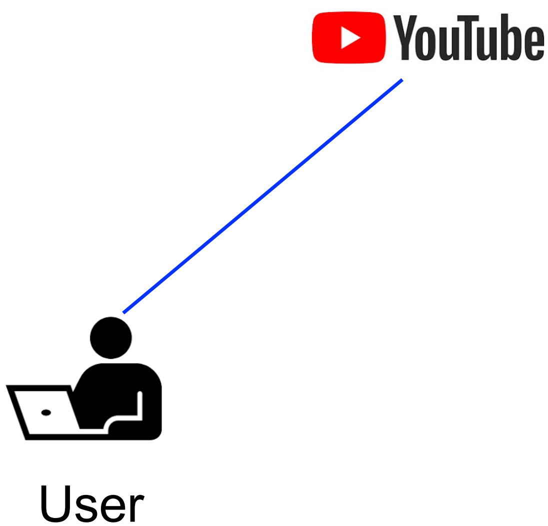
User

- Very often, users themselves even do not know what they like the most
  - uninformative/misleading feedback at beginning
- Many behavioral/marking studies show
  - RS users are explorative at beginning;
  - Their feedback becomes more accurate only after sufficient experience
  - (see more discussions in [Yao et al., ICML22])

Different Research Challenges

Q: how to learn user preferences from evolving/non-stationary behaviors?

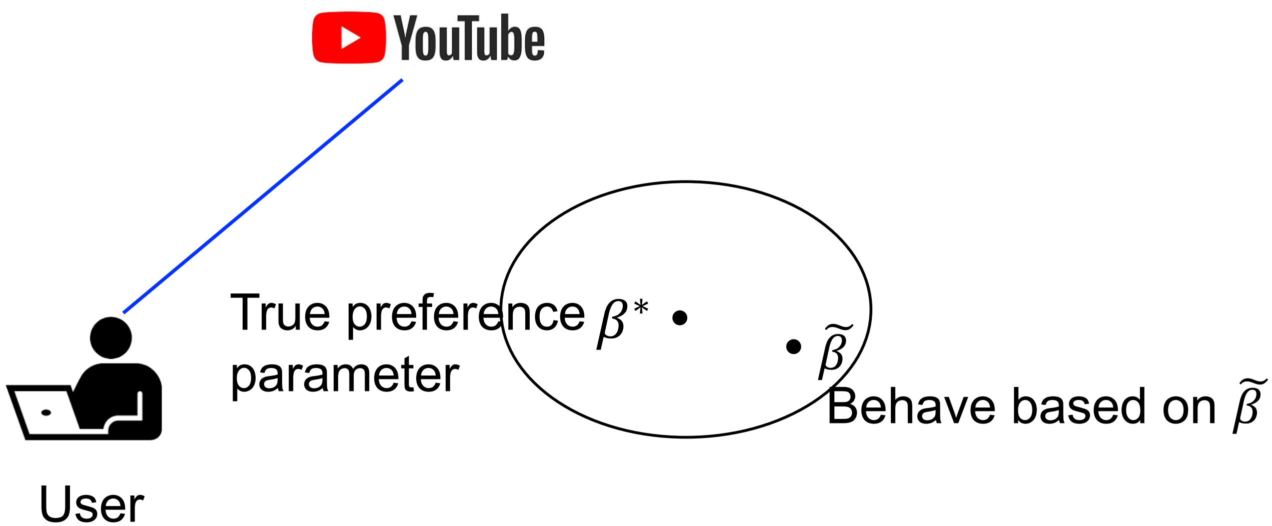
❖ Learning from learning users



Different Research Challenges

Q: how to learn user preferences from evolving/non-stationary behaviors?

❖ Learning from learning users

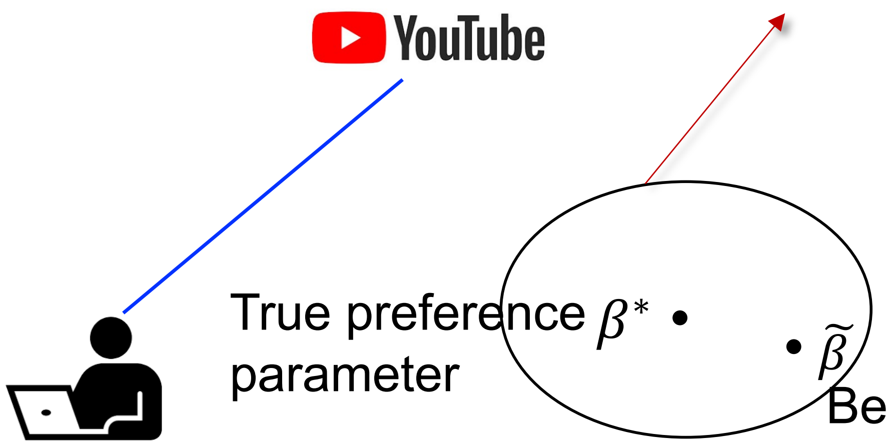


Different Research Challenges

Q: how to learn user preferences from evolving/non-stationary behaviors?

❖ Learning from learning users

✓ Determined by what the user has seen previously



$$(\beta - \beta^*)^T X X^T (\beta - \beta^*) \le t^{\alpha}$$

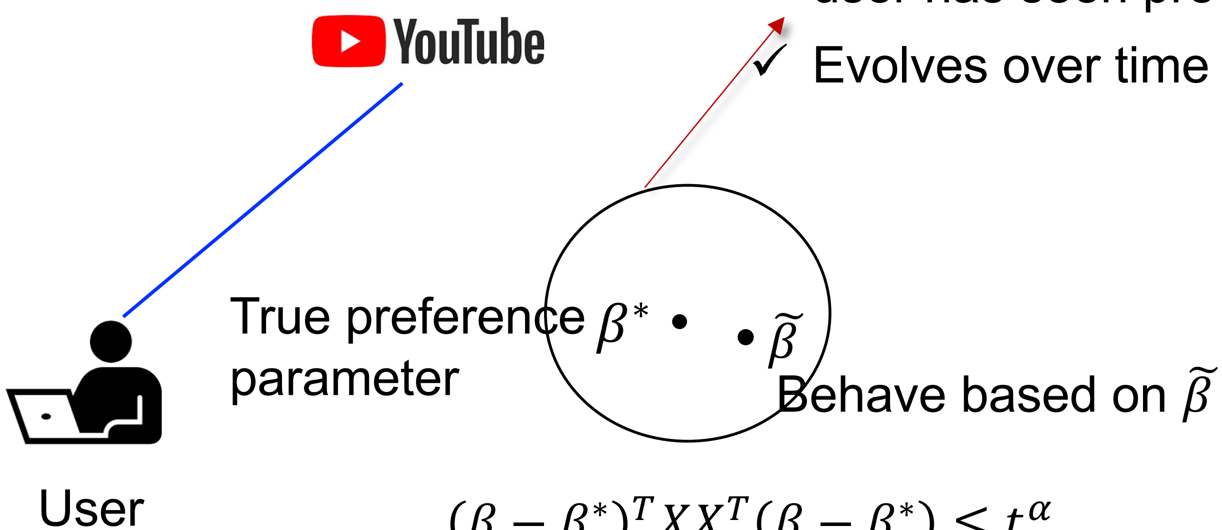
where α = upper bound of user's learning rate

Learning from a Learning User for Optimal Recommendations Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu ICML 2022

Different Research Challenges

Q: how to learn user preferences from evolving/non-stationary behaviors?

- Learning from learning users



- ✓ Determined by what the user has seen previously
- ✓ Evolves over time (so is $\tilde{\beta}$)

$$(\beta - \beta^*)^T X X^T (\beta - \beta^*) \leq t^\alpha$$

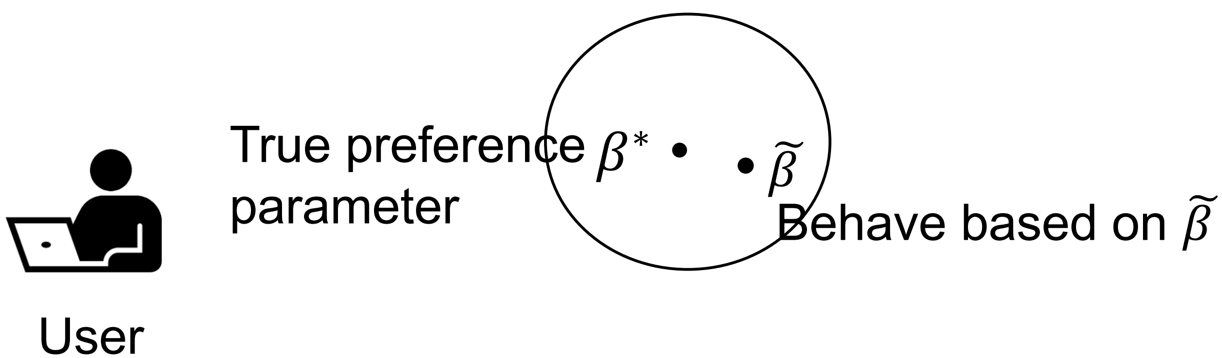
where α = upper bound of user's learning rate

Learning from a Learning User for Optimal Recommendations Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu ICML 2022

Different Research Challenges

Q: how to learn user preferences from evolving/non-stationary behaviors?

Can design algorithm to effectively learn from such non-stationary user feedback (driven by user's own learning)



Different Research Challenges

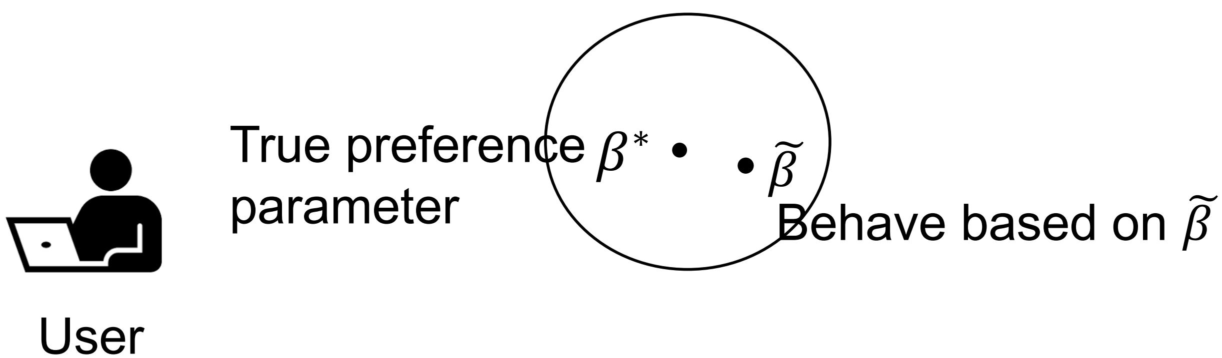
Q: how to learn user preferences from evolving/non-stationary behaviors?

Can design algorithm to effectively learn from such non-stationary user feedback (driven by user's own learning)

Core ideas

✓ Cultivate user's own learning at first with more aggressive exploration

⚠ Challenge: tailor exploration time based on user's learning rate α



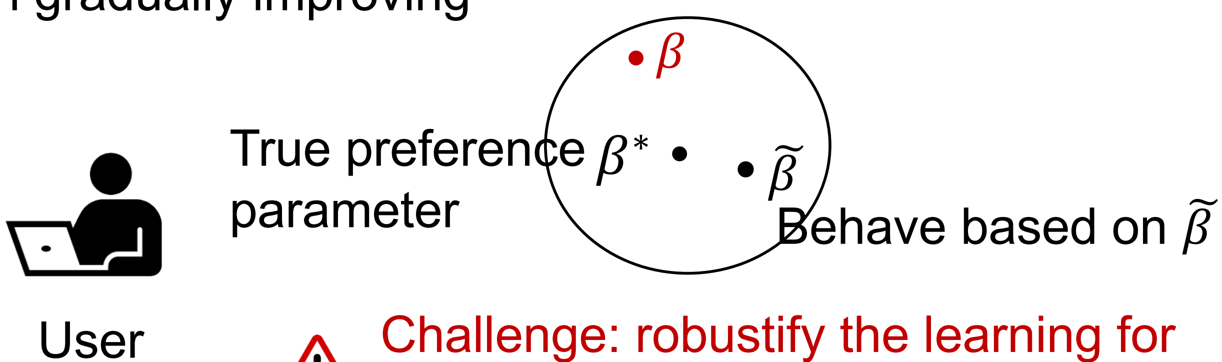
Different Research Challenges

Q: how to learn user preferences from evolving/non-stationary behaviors?

Can design algorithm to effectively learn from such non-stationary user feedback (driven by user's own learning)

Core ideas

- ✓ Cultivate user's own learning at first with more aggressive exploration
- ✓ Robustify the use of user's reward feedback, since it is never perfect though gradually improving



⚠ Challenge: robustify the learning for arbitrary β in the confidence region

Different Research Challenges

Q: how to learn user preferences from evolving/non-stationary behaviors?

Can design algorithm to effectively learn from such non-stationary user feedback (driven by user's own learning)

Overall, it is good news!

Theorem [informal]. There is an algorithm that learns optimal user preferences with regret $O(T^{0.5+\alpha})$ where α is user's own learning rate.

Different Research Challenges

Q: how to learn user preferences from evolving/non-stationary behaviors?

Can design algorithm to effectively learn from such non-stationary user feedback (driven by user's own learning)

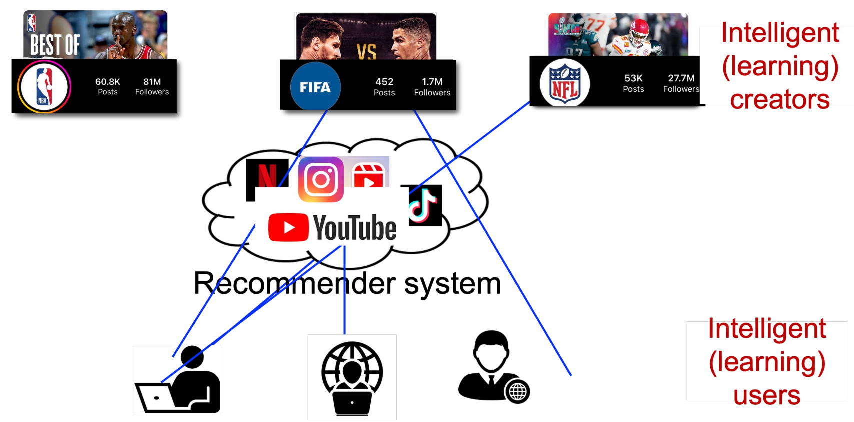
Overall, it is good news!

Theorem [informal]. There is an algorithm that learns optimal user preferences with regret $O(T^{0.5+\alpha})$ where α is user's own learning rate.

- $\alpha = 0$ $\rightarrow$ perfect user, in which case we recover optimal regret for standard setups
- Generally, learning efficiency degrades gracefully as user less efficient

Learning from a Learning User for Optimal Recommendations Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu ICML 2022

Conclusions



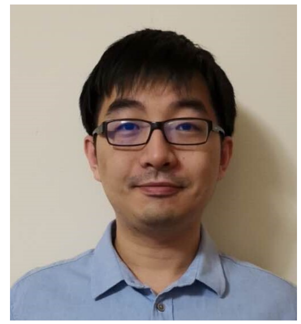
- A framework for economic modeling of contemporary system-creator-user learning + optimization
- Examined some basic questions during system-creator and system-user interactions
- Many open questions
  - What if three parties are learning contemporarily?
  - What if user preference is contextual as well? (e.g., $\theta(x) = \Theta \cdot x$ where x is a search query)

SIGMA LAB
Strategic IntelliGence in Machine Agents

Acknowledgment

Fan Yao

Hongning wang

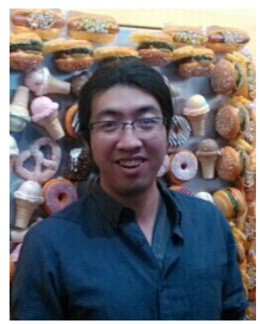


Chuanhao Li

Denis Nekipelov

Karthik Sankararaman

Yiming Liao



Yan Zhu

Qifan Wang

References

Learning from a Learning User for Optimal Recommendations
Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu ICML 2022

How Bad is Top-K Recommendation under Competing Content Creators?
Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang and Haifeng Xu. ICML 2023

Rethinking Incentives in Recommender Systems: Are Monotone Rewards Always Beneficial?
Fan Yao, Chuanhao Li, Karthik Abinav Sankararaman, Yiming Liao, Yan Zhu, Qifan Wang, Hongning Wang and Haifeng Xu, working paper

And many references therein!

Thank You

Questions?

haifengxu@uchicago.edu