Online Platforms and the Fair Exposure Problem Under Homophily

Jakob Schoeffer[1,*], Alexander Ritchie[2,*], Keziah Naggita[3,*], Faidra Monachou[4,*], Jessica Finocchiaro[5,*], Marc Juarez[6]

[1]Karlsruhe Institute of Technology, Germany
[2]University of Michigan, USA
[3]Toyota Technological Institute at Chicago, USA
[4]Stanford University, USA
[5]University of Colorado Boulder, USA
[6]University of Southern California, USA

*Equal contribution, listed in reverse-alphabetical order

Online Platforms and the Fair Exposure Problem Under Homophily

Jakob Schoeffer (jakob.schoeffer@kit.edu), Alexander Ritchie (aritch@umich.edu), Keziah Naggita (knaggita@ttic.edu),
Faidra Monachou (monachou@stanford.edu), Jessica Finocchiaro (jessica.finocchiaro@colorado.edu), Marc Juarez (marc.juarez@usc.edu)

Overview

- **Motivation:** Political extremism, polarization in (social media) networks
- **Introducing the fair exposure problem:** Given limited intervention power of the platform, goal is to enforce balance in the spread of content (e.g., news articles) among two groups of users.

Main Contributions

1. We initiate the **fair exposure problem.**
2. Provide a novel and simple **framework** to study it plus emerging fairness questions in platforms.
3. We show that introducing fairness constraints does **not** automatically imply truly fair outcomes.

Model

M	Finite mass of users
$g \in \{A,B\}$	Group affiliation
π_g	Fraction of users in group g
$s \in \{a,b\}$	Article source affiliated with A,B
$t \in \{1,\dots,T\}$	Discrete time
$\theta_{g,s}$	Fraction of users g shown s at $t=1$
$p_{g,s}$	Prob. of user g liking article s (with A (B) liking a (b) more)
$c_{g,s}$	Cost for reading article
$v_{g,s}$	Valuation for liking read article
$q_g \in (0.5,1)$	Intra-group user replacement

- At time $t > 0$, each user sees an article and decides whether to **click or not**. Users click iff

$$v_{g,s}\, p_{g,s} \geqq c_{g,s}.$$

- At $t+1$, users are **replaced** by same-group users (prob. q_g) or users from the other group ($1-q_g$). If user at t liked article, then replacing user sees the same article—otherwise nothing.

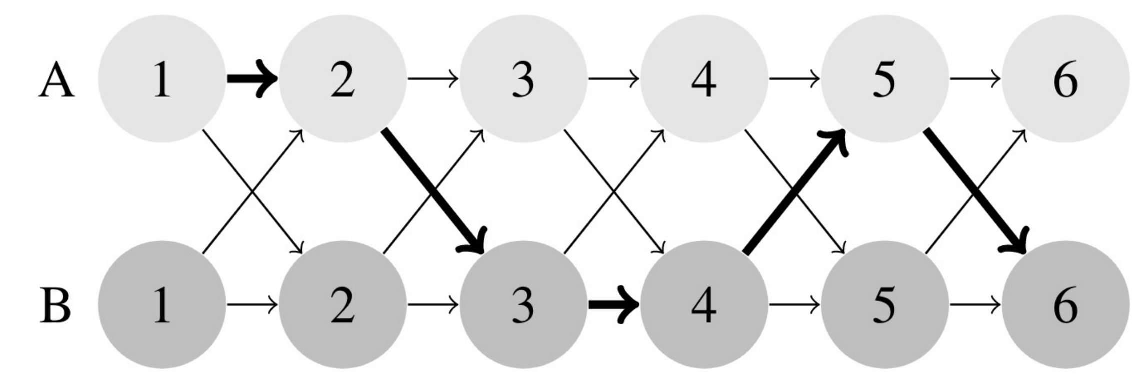

Fig. Exempl. article sharing over time for $T = 6$.

Fairness Constraints

Let $l_{g,s}(t)$ be the **mass** of users at time t belonging to group g who clicked and liked article s.

Constant fair exposure ($e \in [0,1]$):

$$\frac{l_{A,s}(t)}{\pi_A} = \frac{l_{B,s'}(t)}{\pi_B} = e \quad \forall t \leq T,\ \forall s, s' \in \{a,b\}, s \neq s'$$

Approx. fair average exposure ($\underline{\delta} < 1 < \overline{\delta}$):

$$\underline{\delta} \leq \frac{\sum_{t=1}^{T} l_{A,a}(t)}{\sum_{t=1}^{T} l_{B,b}(t)} \leq \overline{\delta} \quad \text{and} \quad \underline{\delta} \leq \frac{\sum_{t=1}^{T} l_{A,b}(t)}{\sum_{t=1}^{T} l_{B,a}(t)} \leq \overline{\delta}$$

- The mass $l_{g,s}(t)$ is a strictly increasing linear function of $\theta_{g,s}$ and $\theta_{g',s}$, except at time $t = 1$.
- We give a non-recursive expression for $l_{g,s}(t)$ using the one-sided $\mathcal{Z}$-transform.

Platform Optimization Problem (LP)

(with approx. fair average exposure constraints (C1) and (C2))

$$\max_{\theta_{A,a},\theta_{B,a} \in [0,1]} \sum_{t=1}^{T} \sum_{g \in \{A,B\}} \sum_{s \in \{a,b\}} l_{g,s}(t)$$

$$\text{s.t. } \underline{\delta} \leq \frac{\sum_{t=1}^{T} l_{A,a}(t)}{\sum_{t=1}^{T} l_{B,b}(t)} \leq \overline{\delta} \qquad \text{(C1)}$$

$$\underline{\delta} \leq \frac{\sum_{t=1}^{T} l_{A,b}(t)}{\sum_{t=1}^{T} l_{B,a}(t)} \leq \overline{\delta}. \qquad \text{(C2)}$$

Theoretical Results (Excerpt)

Proposition (informal): The exclusion of any fairness constraints in LP always results in all users of the same group being shown the same article by the platform at $t = 1$.

Lemma (informal): It is generally not possible to achieve constant fair exposure at every time step unless certain restrictive conditions hold.

Main Takeaways

- From analyzing the optimal solutions to (LP) with (C1) and (C2), we know that introducing fairness constraints does **not** automatically imply that the outcome is truly fair/ balanced.
- Specifically, it can happen that one group is being targeted with only one article (which may not be the group's preferred); whereas the other group sees both articles at unequal rates—thus incurring the "**price of fairness**."

Results from Simulations (Excerpt)

We also use our model to empirically study the effects of **different model parameters** from real-world click data (e.g., Bakshy et al., 2015).

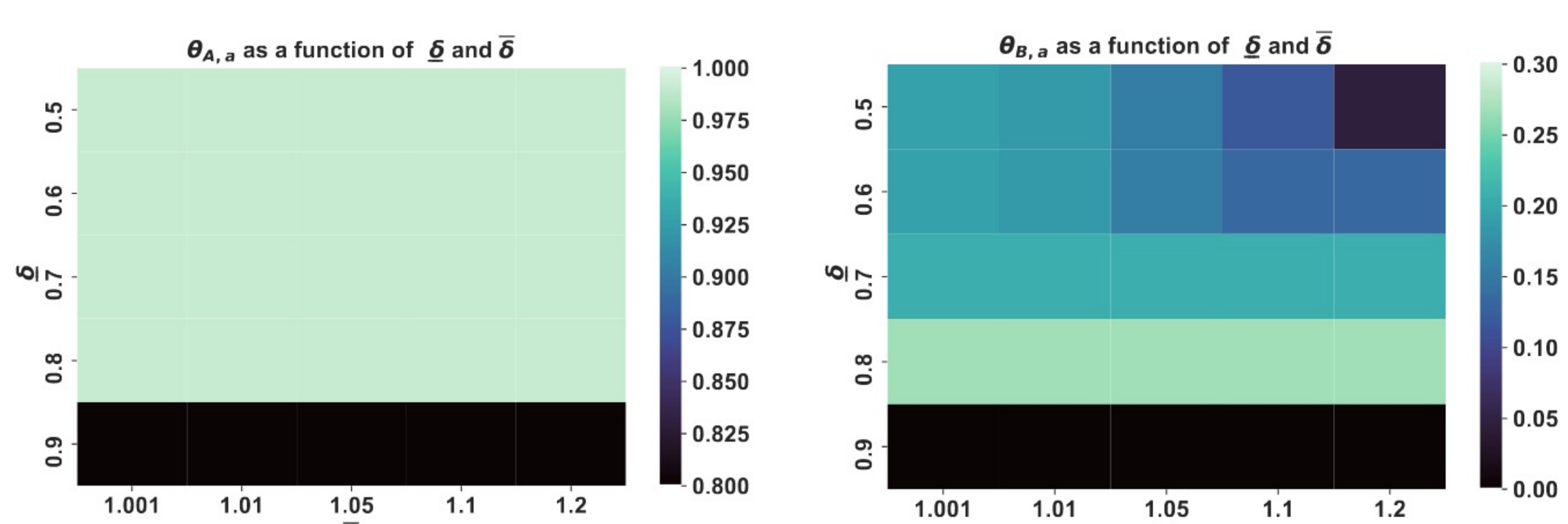

Fig. Calculating $\theta_{g,s}$ and $\theta_{g,s}$ as a function of fairness bounds δ.

Questions? → Please reach out!