SIGIR 2020
XI'AN·CHINA

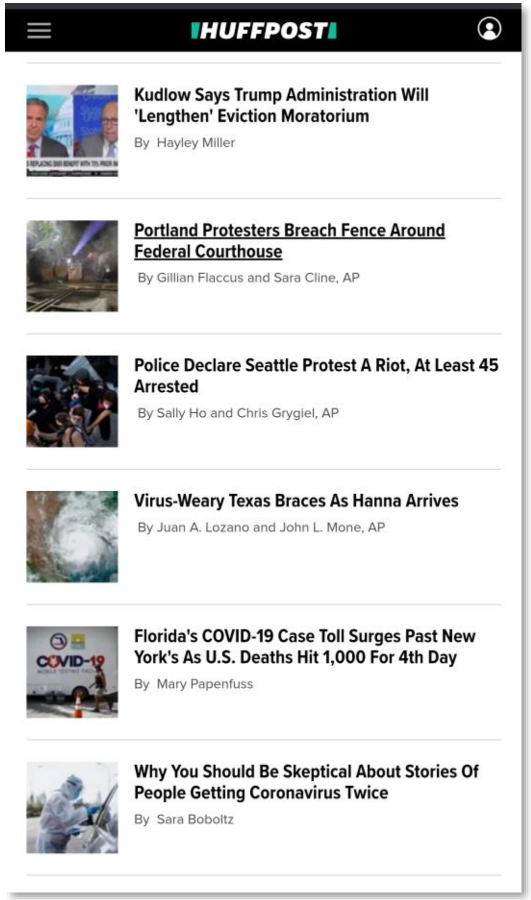
Controlling Fairness and Bias in Dynamic Learning-to-Rank

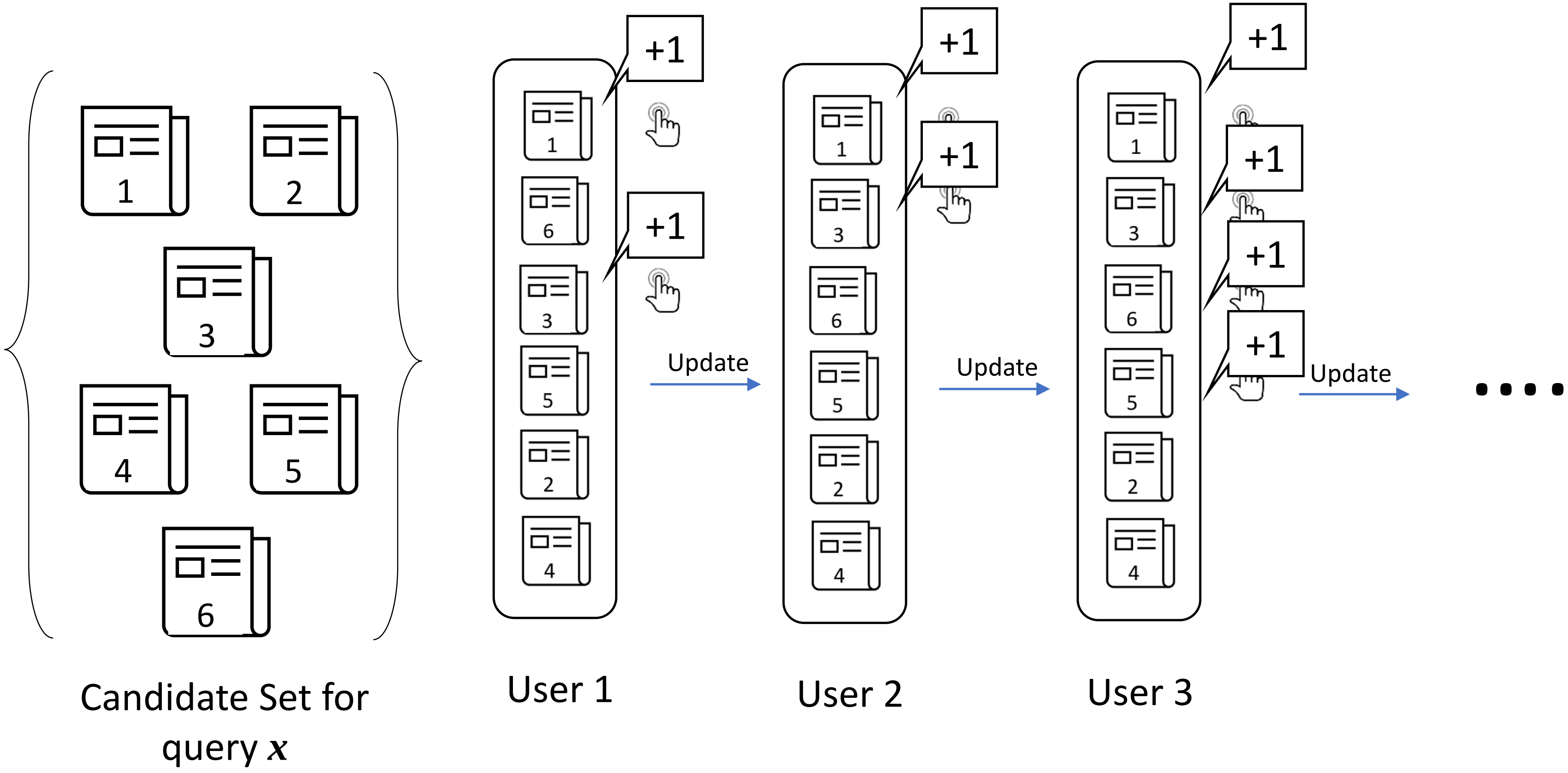
ACM SIGIR 2020

Marco Morik[*†], Ashudeep Singh[*‡], Jessica Hong[‡], Thorsten Joachims[‡]

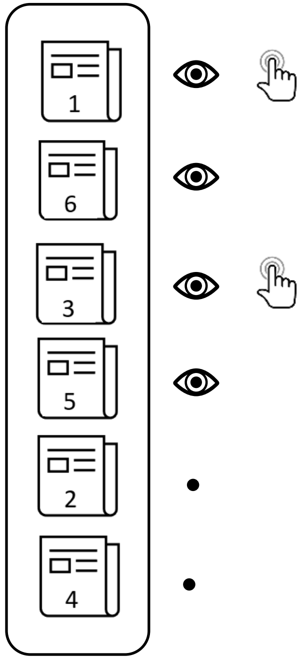
[†] TU Berlin, [‡] Cornell University

Dynamic Learning-to-Rank





Problem 1: Selection Bias due to position

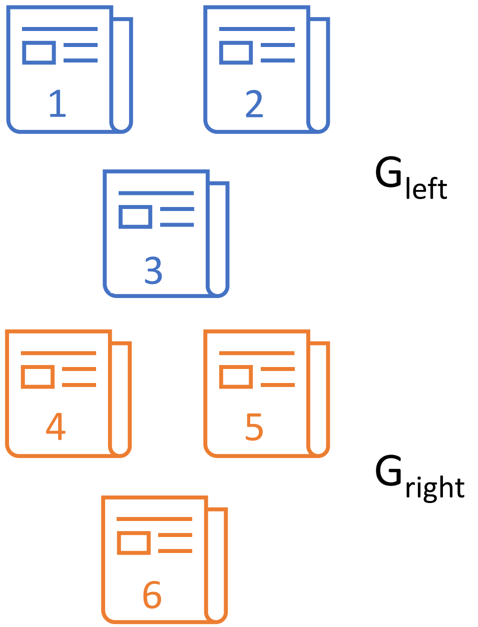


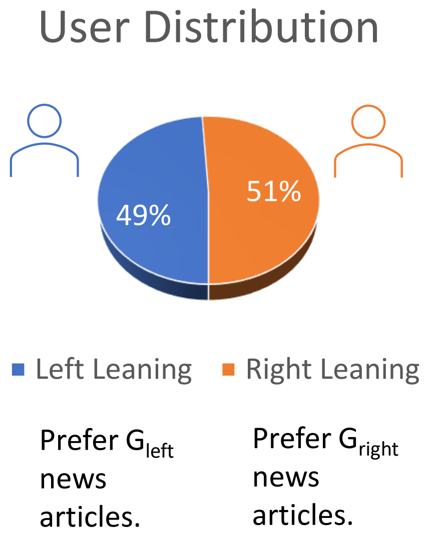
Position Bias

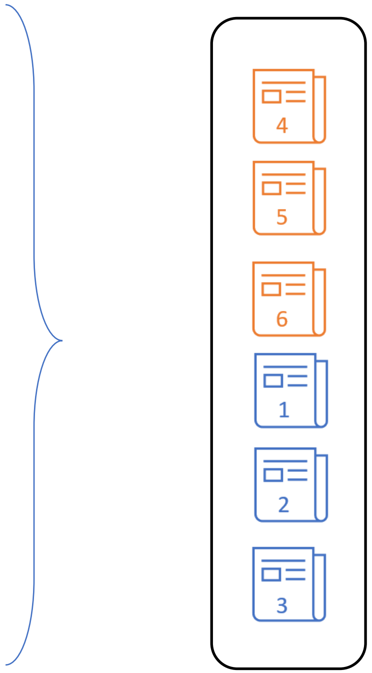
- Number of clicks is a biased estimator of relevance.
  - Lower positions get lower attention.
  - Less attention means fewer clicks.

- Rich-get-richer dynamic: What starts at the bottom has little opportunity to rise in the ranking.

Problem 2: Unfair Exposure







Probability Ranking Principle [Robertson, 1977]:

Rank documents by probability of relevance → y^*.

Maximizes utility for virtually any measure U of ranking quality

$$y^* \coloneqq \operatorname{argmax}_y[\mathrm{U}(y|x)]$$

Ranking by true average relevance leads to unfair rankings.

Position-Based Exposure Model

Definition:

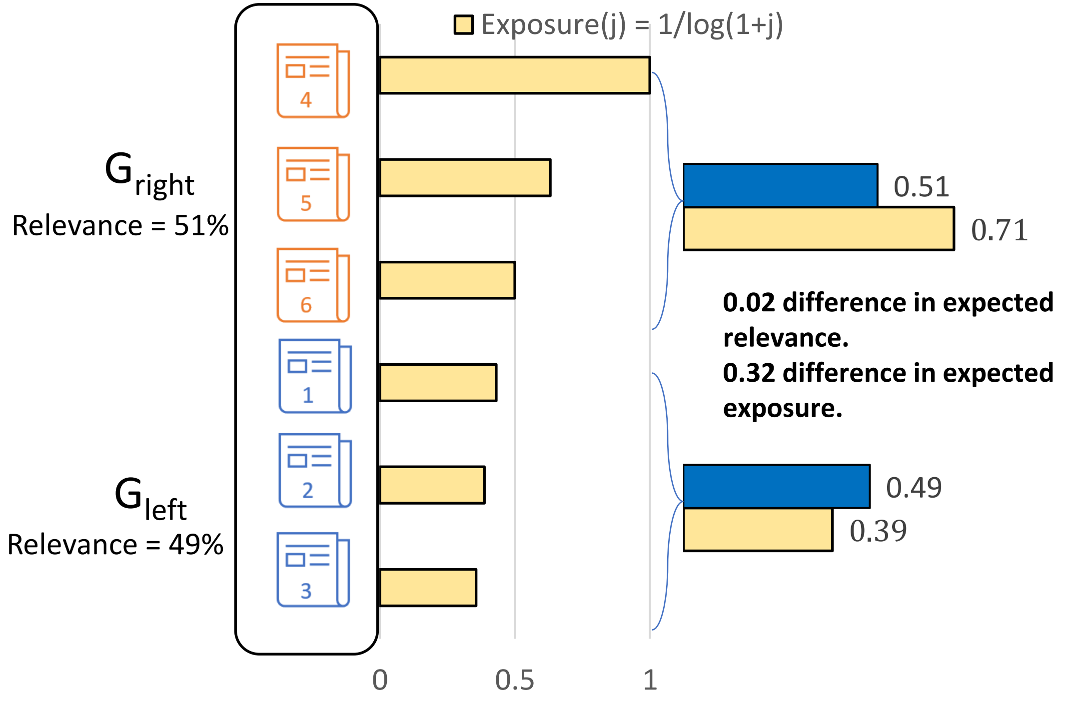
Exposure e_j is the probability a users observes the item at position j.

Exposure of Group:

$$Exp(G|x) = \sum_{j \in G} e_j$$

How to estimate?

- Eye tracking [Joachims et al. 2007]
- Intervention studies [Joachims et al. 2017]
- Intervention harvesting [Agarwal et al. 2019, Fang et al. 2019]



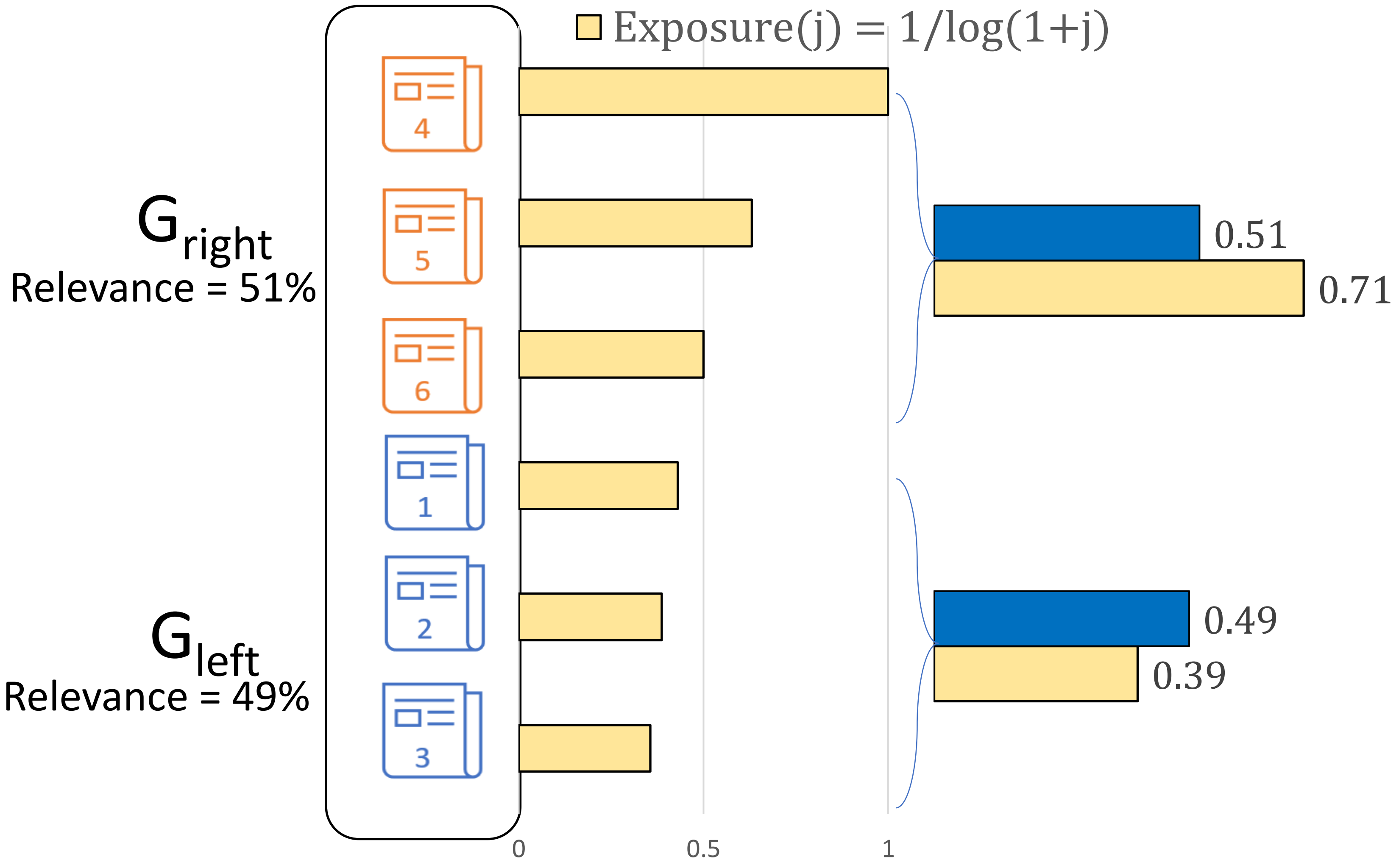
Disparate exposure allocation: A small difference in average relevance, leads to a large difference in average exposure!

Outline

- Exposure Model
- Fairness Notions
- FairCo Algorithm
  - Unbiased Average Relevance estimation
  - Unbiased Relevance estimation for Personalization

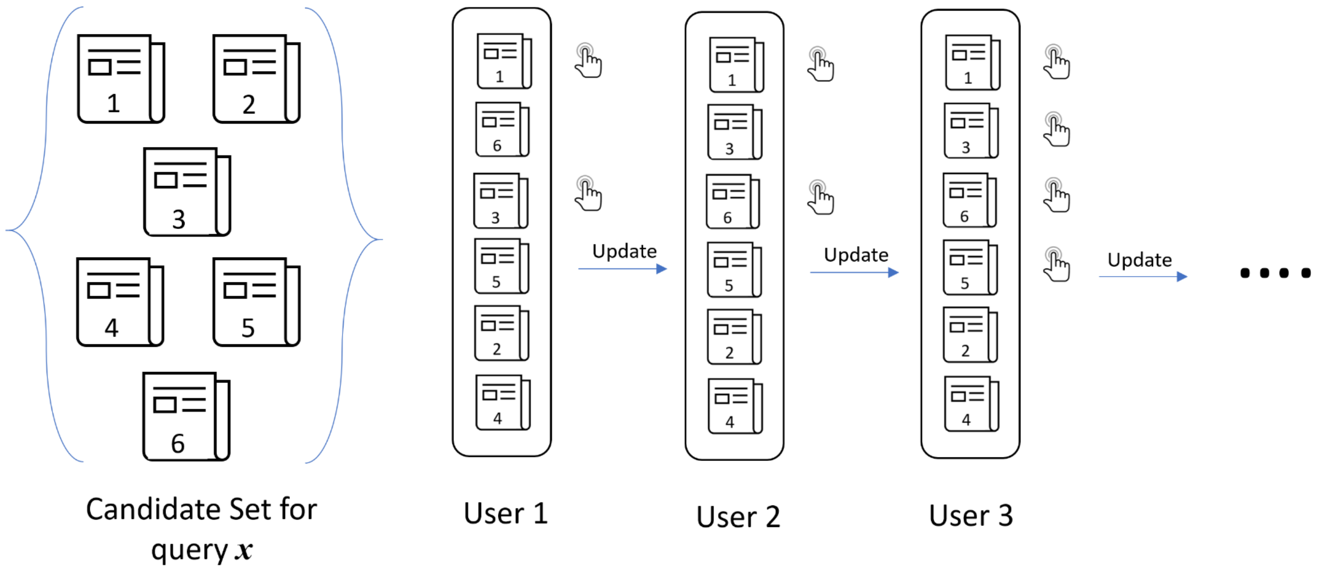
<table>
<tr><th></th><th>Exposure Fairness</th><th>Impact Fairness</th></tr>
<tr><td>Goal</td><td>Exposure is allocated based on relevance of the group.
$$Exp(G|x) = f\left(Rel(G|x)\right)$$</td><td>The expected impact (e.g. clickthrough rate) is allocated based on merit.
$$Imp(G|x) = f(Rel(G|x))$$
For the position bias model,
$$Imp(d|x) = Exp(d|x)Rel(d|x)$$</td></tr>
<tr><td>Constraint</td><td>Make exposure proportional to relevance (per group)
$$\frac{Exp(G_0|x)}{Exp(G_1|x)} = \frac{Rel(G_0|x)}{Rel(G_1|x)}.$$</td><td>Make the expected impact proportional to relevance (per group)
$$\frac{Imp(G_0|x)}{Imp(G_1|x)} = \frac{Rel(G_0|x)}{Rel(G_1|x)}.$$</td></tr>
<tr><td>Disparity Measure</td><td>$$D^E(G_0, G_1) = \left[\frac{Exp(G_0|x)}{Rel(G_0|x)} - \frac{Exp(G_1|x)}{Rel(G_1|x)}\right].$$</td><td>$$D^I(G_0, G_1) = \left[\frac{Imp(G_0|x)}{Rel(G_0|x)} - \frac{Imp(G_1|x)}{Rel(G_1|x)}\right].$$</td></tr>
</table>

[Singh & Joachims. Fairness of Exposure in Rankings. KDD 2018]



Does not satisfy Fairness of Exposure or Fairness of Impact.

Dynamic Learning-to-Rank



Sequentially present rankings to users that

- ❑ Maximize Expected User Utility $\mathbb{E}[U|x]$
- ❑ Ensure Unfairness D_τ goes to 0 with τ.

Fairness Controller (FairCo) LTR Algorithm

FairCo: Ranking at time τ

$$\sigma_\tau = \operatorname{argsort}_{d \in \mathcal{D}} \left(\hat{R}(d|x) + \lambda \, \mathrm{err}_\tau(d) \right)$$

Proportional Controller: Linear feedback control system where correction is proportional to the error.

$\hat{R}(d|x)$: Estimated Conditional Relevance

$\lambda > 0$

$err_\tau(d) = (\tau - 1) \max_{G_i} (\hat{D}^E_\tau(G_i, G(d)))$

- Theorem: When the problem is well posed, FairCo ensures that $\overline{D}_\tau \to 0$ as $\tau \to \infty$ at the rate of $\mathcal{O}\left(\frac{1}{\tau}\right)$.
- Requirements:
  - Estimating Average Relevances $\hat{R}(d)$.
  - Estimating Unbiased Conditional Relevances $\hat{R}(d|x)$ for personalization.

Estimating Average Relevances

- Average number of clicks is not a consistent estimator.
- IPS weighted clicks:

$$\hat{R}^{\mathrm{IPS}}(d) \quad = \quad \frac{1}{\tau}\sum_{t=1}^{\tau}\frac{\mathbf{c}_t(d)}{\mathbf{p}_t(d)}.$$

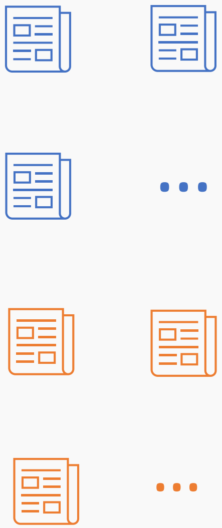
$c_t(d)$: Click on d at time t.
$p_t(d)$: Position bias at the position of d.

[Joachims et al., 2017]

- $\hat{R}^{\mathrm{IPS}}(d)$ is an unbiased estimator of a document's relevance.

Experimental Evaluation

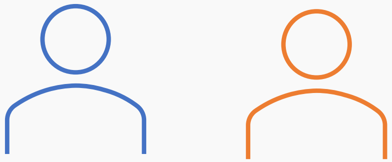
Simulation on **Ad Fontes Media Bias Dataset**



G_{left}
$\rho^d < 0$

G_{right}
$\rho^d \geq 0$

Each news source in the dataset has a polarity assigned $\rho^d \in [-1, 1]$.



Prefer G_{left} news articles.

Prefer G_{right} news articles.

Sample user u_t is drawn with a polarity parameter $\rho^{u_t} \in [-1, 1]$ and an openness parameter $o_t \in (0.05, 0.55)$.

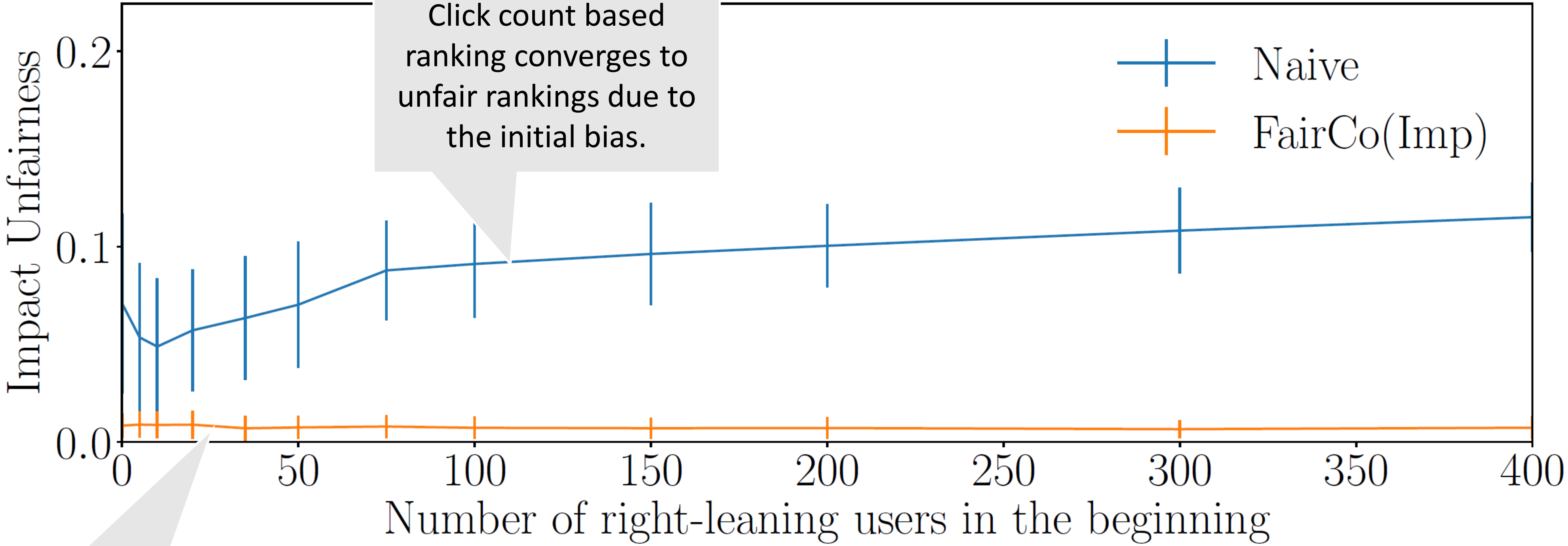
A user's relevance is a function of their polarity and the news article's polarity, and their openness.

$$\mathbf{r}_t(d) \sim \text{Bernoulli}\left[p = exp\left(\frac{-(\rho^{u_t} - \rho^d)^2}{2(o^{u_t})^2}\right)\right]$$

Goal: Present rankings to a sequence of users to maximize their utility while <u>providing fair exposure to the news articles relative to their average relevance over the user population.</u>

Can FairCo break the Rich-get-richer dynamic?

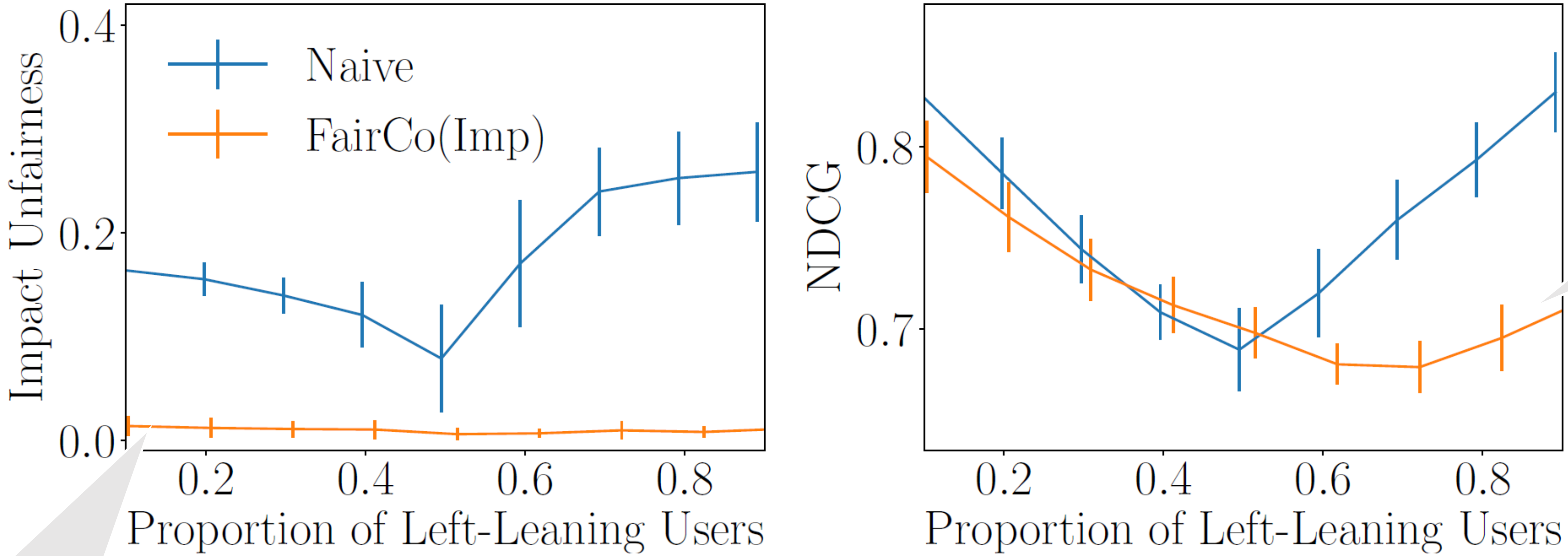
Effect of the initial ranking after 3000 users.



Click count based ranking converges to unfair rankings due to the initial bias.

FairCo keeps the Unfairness low for any amount of head start.

Can FairCo ensure fairness for Minority user groups?



Trades off utility for fairness when there is an imbalance in user distribution.

FairCo converges to fair ranking for all user distributions.

Outline

- Exposure Model
- Fairness Notions
- FairCo Algorithm — Selection Bias ☑ Fairness ☑
  - Unbiased Average Relevance estimation
  - ➡ Unbiased Relevance estimation for Personalization

D-ULTR: Relevance Estimation for Personalized Ranking

$$\mathcal{L}^{\mathbf{c}}(w) = \sum_{t=1}^{\tau} \sum_{d} \hat{R}^{\mathrm{w}}(d|\boldsymbol{x}_t)^2 + \frac{\mathbf{c}_t(d)}{\mathbf{p}_t(d)}(\mathbf{c}_t(d) - 2\hat{R}^{\mathrm{w}}(d|\boldsymbol{x}_t))$$

$\hat{R}^w$: Output of a Neural Network with weights w.
$c_t(d)$: Click on d at time t.
$p_t(d)$: Position bias at position of d.

- To estimate: $\hat{R}^w(d|\boldsymbol{x}_t)$ – Relevance of document d for query $\boldsymbol{x_t}$.
- Train the neural network by minimizing $\mathcal{L}^c(\mathrm{w})$.
- $\mathcal{L}^c(\mathrm{w})$ is unbiased i.e. in expectation it is equal to a full information squared loss (with no position bias).

Evaluation on Movielens dataset

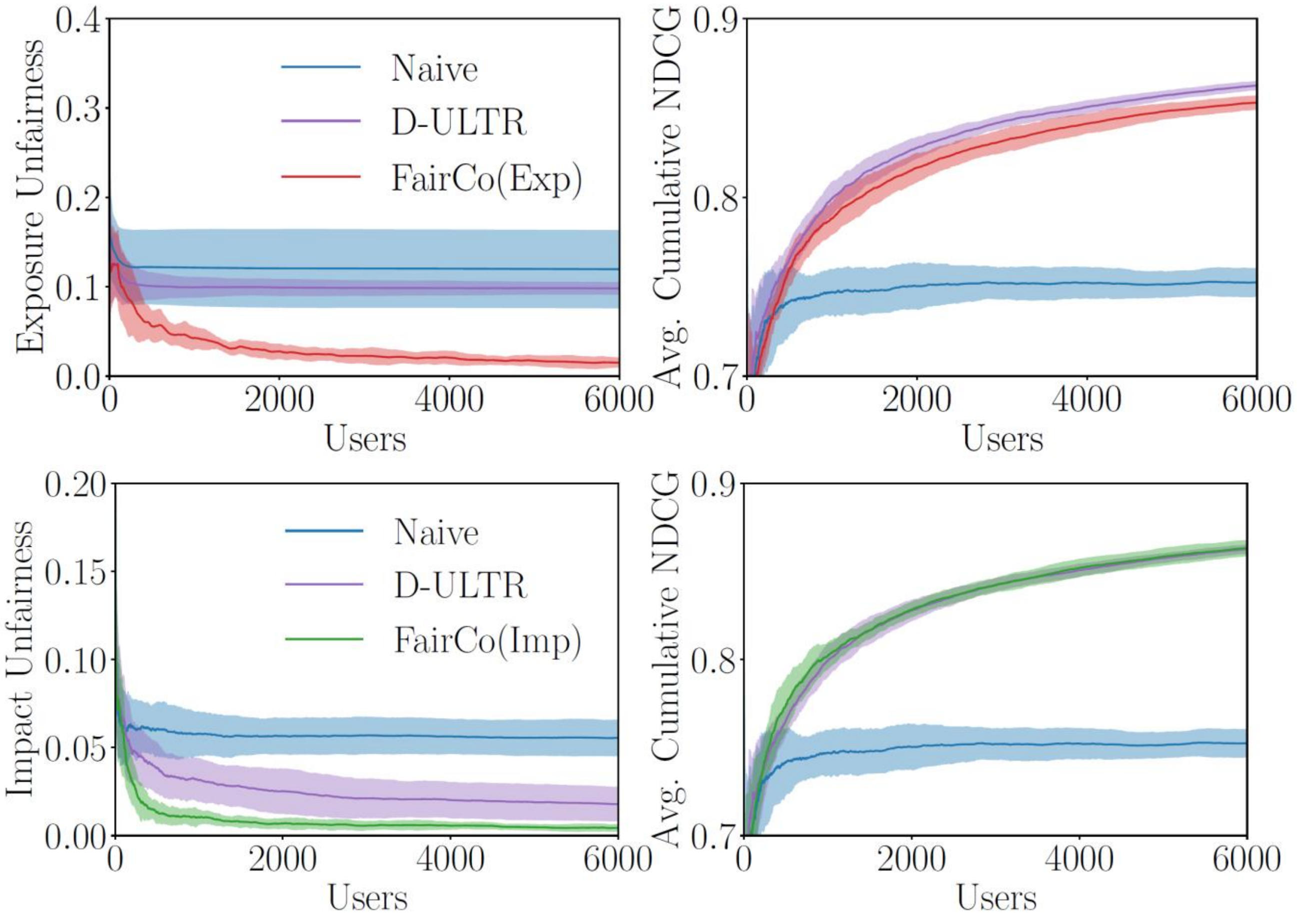
- Completed a subset of Movielens dataset ($10\mathrm{k} \times 100$ ratings matrix) using matrix factorization.
  - Selected 100 movies from top-5 production companies in ML-20M dataset. Groups: *MGM, Warner Bros, Paramount, 20th Century Fox, Columbia.*
  - Selected $10\mathrm{k}$ most active users.
- User features x_t come from this matrix factorization.

Goal: Present ranking to each user u_τ to maximize NDCG while making sure the production companies receive fair share of exposure relative to the average relevance of their movies.

Does FairCo ensure fairness with effective personalization?

Exposure Unfairness

Impact Unfairness



Personalized Rankings achieve high utility (NDCG), while reducing Unfairness to 0 with τ.

Conclusions

- Identified how biased feedback leads to unfairness and suboptimal ranking in Dynamic-LTR.
- Proposed FairCo to adaptively enforce amortized fairness constraints while relevances are being learned.
  - Easy to implement and computationally efficient at serving time.
- The algorithm breaks the rich-get-richer effect in Dynamic-LTR.

Thank you!

Controlling Fairness and Bias in Dynamic Learning-to-Rank

Marco Morik[*†], Ashudeep Singh[*‡], Jessica Hong[‡], Thorsten Joachims[‡]

[†] TU Berlin, [‡] Cornell University

SIGIR 2020
XI'AN·CHINA

ACM SIGIR 2020