Fairness implications of relevance-based ranking policies in two-sided platforms

Hannah Li, Lorenzo Modotti

Decision, Risk, and Operations Division | Columbia Business School

Motivation

- Recommender systems often employ personalized rankings to display items
- Users are more likely to interact with items shown at the top
- Vast unbiased-learning-to-rank literature for search engines, less focus on two-sided platforms

Search engines

- No capacity constraints for users or web pages
- Rankings do not change often over time

Two-sided platforms

- Capacity constraints for demand and supply side
- Rankings naturally change over time due to stock-outs
- Ensuring fairness on both demand and supply side is a goal

In this work, we model a two-sided marketplace for employers and freelancers, analyzing supply-side fairness with respect to downstream metrics, in equilibrium.

Platform dynamics

- Employer $u \in \mathcal{U}$ makes requests at rate θ_u
- Available freelancers $v \in \mathcal{V}$ are ranked according to their relevance ρ_{uv}
- Starting from the topmost result v:
  - u clicks and visits the freelancer page with probability r_{uv}
  - conditional on click, u books v with probability s_{uv}; the request terminates
  - if no click or booking occurs, u keeps scrolling with probability γ; the request terminates with probability $1-\gamma$
- When v is booked, v remains unavailable for η_v units of time, with mean replenishment time $\frac{1}{\mu_v}$
- The employer-freelancer relevance is defined as $\rho_{uv} := r_{uv} s_{uv}$

Measuring unfairness

For $v \in \mathcal{V}$ define:

$$\bar{b}_v := \mathbb{E}\left[\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{I}\{v \text{ booked in } t\}\right] \qquad \textit{Booking rate}$$

$$\mathcal{B}_v := \frac{\bar{b}_v}{\mu_v} \qquad \textit{Weighted booking rate}$$

$$\mathcal{R}_v := \sum_{u\in\mathcal{U}} \theta_u \rho_{uv} \qquad \textit{Weighted relevance}$$

A fair ranking policy with respect to the booking rate should satisfy, for any $v \in \mathcal{V}$,

$$\bar{\mathcal{B}}_v := \frac{\mathcal{B}_v}{\sum_{v'\in\mathcal{V}} \mathcal{B}_{v'}} \approx \frac{\mathcal{R}_v}{\sum_{v'\in\mathcal{V}} \mathcal{R}_{v'}} =: \bar{\mathcal{R}}_v.$$

Therefore, defining

$$\psi_v := \mathcal{B}_v \sum_{v'\in\mathcal{V}} \mathcal{R}_{v'} - \mathcal{R}_v \sum_{v'\in\mathcal{V}} \mathcal{B}_{v'}$$

it follows that if $\psi_v = 0$ the policy is fair towards v, while if $\psi_v > 0$ ($\psi_v < 0$) the policy overly advantages (disadvantages) v.

Consider

$$\Psi := \sum_{v\in\mathcal{V}} \psi_v^2 \qquad \textit{Discrepancy metric}$$

Ψ penalizes large deviations of $\bar{\mathcal{B}}_v$ from $\bar{\mathcal{R}}_v$. If $\Psi = 0$, the policy is perfectly fair. As Ψ increases, the level of unfairness increases.

Equilibrium

The equilibrium booking rate of $v \in \mathcal{V}$ is

$$\bar{b}_v = \mu_v(1-q_v),$$

where $(q_v)_{v\in V}$ is the solution of the system

$$\mu_v - q_v\left[\mu_v + \sum_{u\in\mathcal{U}} \theta_u \rho_{uv} \prod_{k=1}^{k_{uv}-1} \left[\gamma(1-\rho_{uv_u^k}) q_{v_u^k} + (1-q_{v_u^k})\right]\right] = 0,\ v \in \mathcal{V}.$$

Crowding

Crowding refers to a situation where a small proportion of freelancers is highly relevant for a large proportion of employers.

For $\alpha \in [0,1]$ define $\rho := (1-\alpha)\rho' + \alpha\rho''$, where

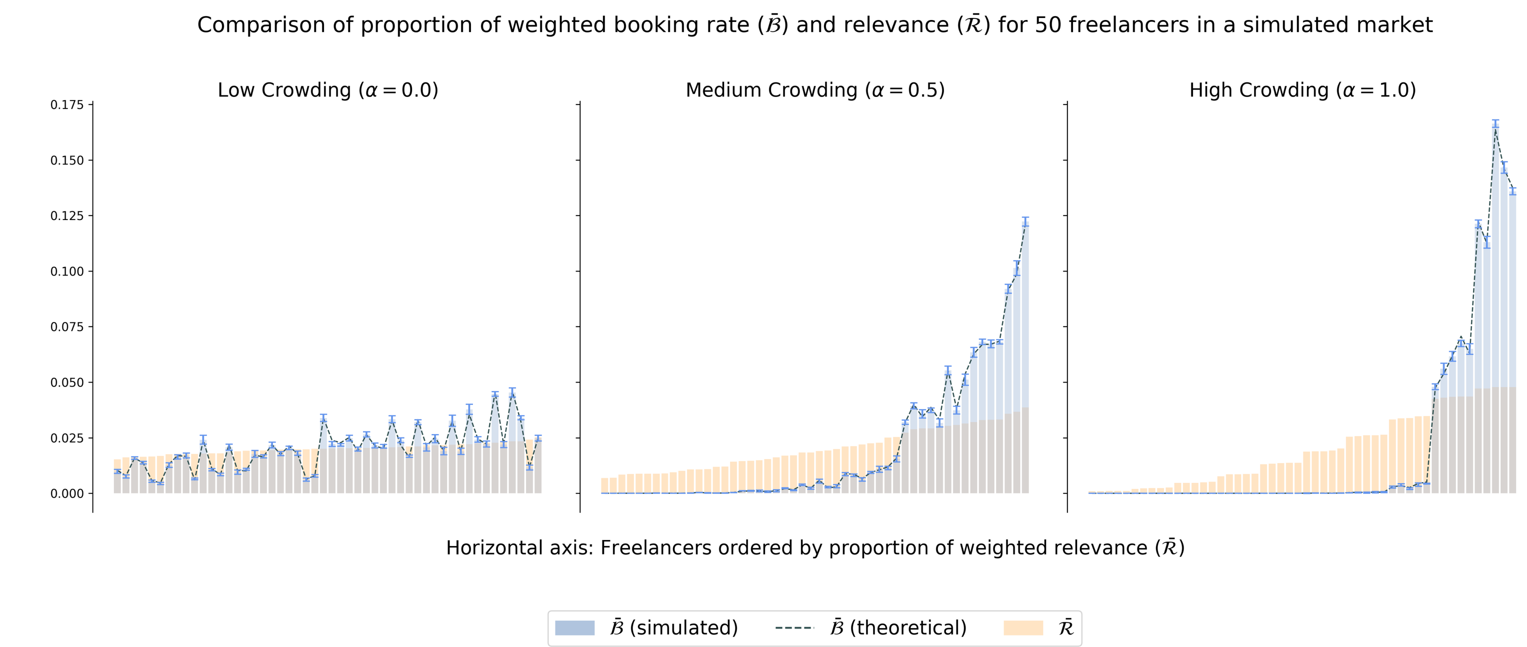
ρ' is a relevance matrix with random entries

ρ'' is a relevance matrix with the same freelancer distribution

If $\alpha = 0$, there is low crowding. As $\alpha \to 1$, crowding increases.

Numerical results

As crowding increases, highly relevant freelancers are overly advantaged by the ranking policy, at the expense of the remaining freelancers.



The level of unfairness increases with crowding and mean replenishment time.