Network bias for Job seekers: The impact of multi-hop referrals

Yiguang Zhang [1] Reetahan Mukhopadhyay [1] Augustin Chaintreau [1]

[1]Columbia University

Introduction

Through a job referral, a pair composed of a job seeker and a recruiter is assisted by a common connection to either advertise the opportunity between each other or facilitate the process of determining the right fit. The "strength of weak ties" hypothesis suggests that individuals are often referred to job opportunities informally by distant acquaintances. Having a network of professional contacts aids the referral process, and nowadays this network is tangibly realized through professional networking sites such as LinkedIn.

Job seeking is hardly a level playing field, as multiple studies proved that members of minority groups, such as women in the workplace, need to work harder to integrate within the process of job referrals. The average number of immediate connections that a woman's profile can gain on LinkedIn has been shown to be significantly less than that of a man's profile. Similar network bias among gender and race has been observed even when controlling for seniority and specialization in scientific publishing networks, as well as freelancing and content sharing platforms. Thus, a single-hop referral process, where an individual's access to job opportunities is offered primarily by their direct connections, is bound to be biased against minority groups under very general conditions. Studying fairness in networks can be difficult as there are multiple sources that contribute to bias and reduced access to opportunities that different groups are facing, and that job seekers and recruiters react to information in different manners, in that their behaviors range from passive observers waiting to be contacted, or cautious job seekers probing their network to more proactive or aggressive job-seeking strategies.

On networks such as LinkedIn, it is possible to utilize multi-hop connections, which opens up a new set of questions on whether this aids or hinders the fight against bias. In our study, we formalize a model that allows for mathematically tractable analysis of the impact of network conditions and job-seeking strategies on the expected number of referrals that job-seeker can obtain. We consider both cases of job seekers using both dissemination-based approaches as well as inquiry strategies, as well as both active and passive approaches that can consider outcomes for recruiters in addition to job-seekers. We focus on the change in bias from 1-hop to 2-hop.

Biased Preferential Attachment

We chose to use the Biased Preferential Attachment (BPA) Model to represent our network of individuals. This model is shown to exhibit key properties such as:

- *glass-ceiling effect*: the difference in the access of opportunities that is more pronounced among people with the most social capital
- *rich-get-richer*: existing members with higher degree are likely to attract new connections than the ones with lower degree
- *homophily*: members tend to connect with members from the same group.

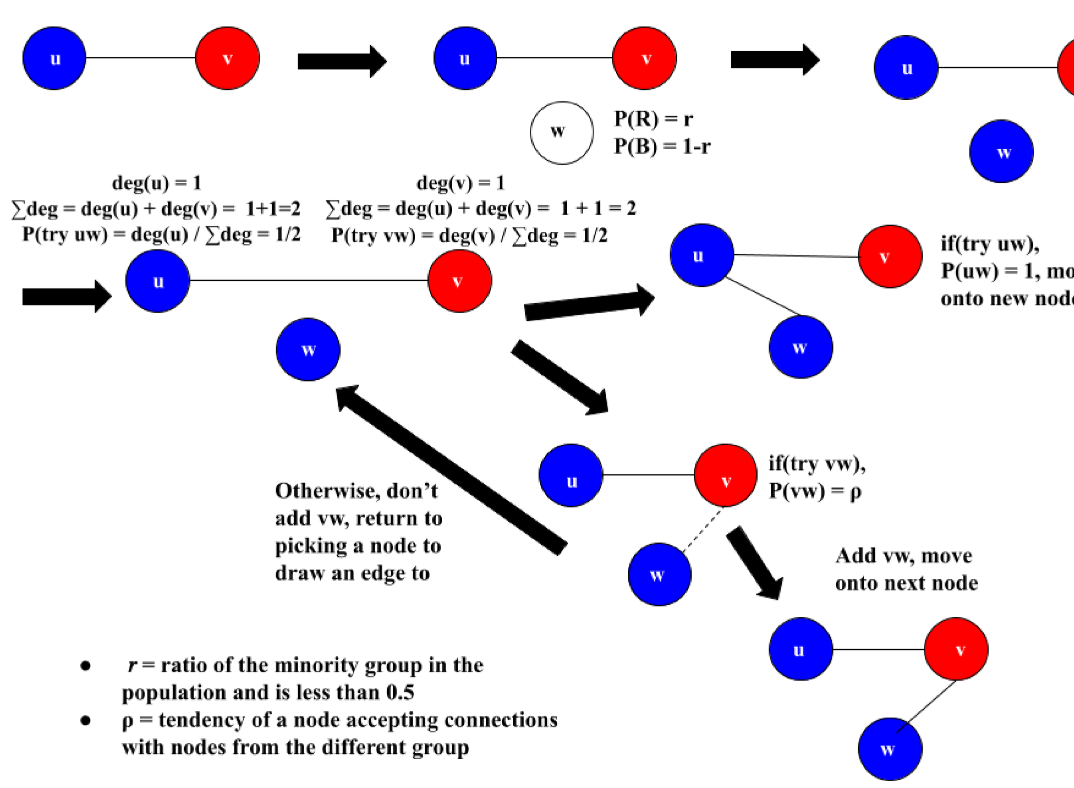

Figure 1. A flowchart showing the initialization and the progression of the first iteration of building a network under the BPA model.

Active and Passive Gain

- **k-hop successful active referral**: a job seeker initiates by sending a referral request to a recruiter, who receives it and accepts it through k hops, for a job referral strategy $\mathcal{S}$ applied by members in $\mathcal{G}(N_t, t, r, \rho)$
- **k-hop successful passive referral**: like its active analog, but for when a recruiter initiates by sending a referral offer to a job seeker u who receives it and accepts it through k hops.
- **Active Gain** ($AG^{(k)}_{\mathcal{S}(t)}(u)$): number of k-hop successful active referrals initiated by u. Let $AG^{(k)}_{\mathcal{S}(t)}(R)$ be the sum of active gains over nodes in red at time t.
- **Passive Gain** ($PG^{(k)}_{\mathcal{S}(t)}(u)$ and $PG^{(k)}_{\mathcal{S}(t)}(R)$): are like their active analogs but it's the number of k-hop successful passive referrals received by u.

Linear Model

To track the network-based unfairness in multi-hop job referral progress, we assume for simplicity that each user u in the network has an internal selective rate t_u, the internal threshold for u to send out referrals, and an accepting rate a_u, the chance that they say yes to either accept or pass a referral, where t_u and a_u are i.i.d. for all the users respectively.

Under the dissemination paradigm, initiators aim to disseminate their referrals across the network by sending out as many job referrals as they reasonably can. We assume a *linear strategy* ($\mathcal{S}_L$), where a referral sender u shares the referral to a neighbor v, if and only if $a_v > t_u$. After receiving a referral, the new sender can select their own receivers following the same pattern. **When using dissemination, the effect of multi-hop referrals on fairness is beneficial when homophily is moderate and detrimental when homophily is pronounced.**

For a sequence of networks $\{\mathcal{G}(N_t, t, r, \rho)\}$ generated by the BPA model with $r < 0.5$,

Theorem 1

$\lim_{t\to\infty} \frac{\mathbb{E}[AG^{(2)}_{\mathcal{S}_L(t)}(R)]}{\mathbb{E}[AG^{(2)}_{\mathcal{S}_L(t)}(B)]} = \frac{\beta_2}{\beta_3}$, where $\beta_2 = \frac{r\rho}{2(1-(1-\alpha^*)(1-\rho))}$ and $\beta_3 = \frac{1-r}{2(1-\alpha^*(1-\rho))}$, and α^* is the unique solution in $[0,1]$ of the equation $2\alpha^* = 1-(1-r)\frac{(1-\alpha^*)}{1-\alpha^*(1-\rho)} + r\frac{\alpha^*}{1-(1-\alpha^*)(1-\rho)}$.

As $t \to \infty$, the ratio of the expected sum for 2-hop active gains among red nodes over that among all nodes is greater than the ratio of the expected sum for 1-hop active gains for red nodes over that among all nodes if and only if $\frac{\beta_2}{\beta_2+\beta_3} > \alpha^*$.

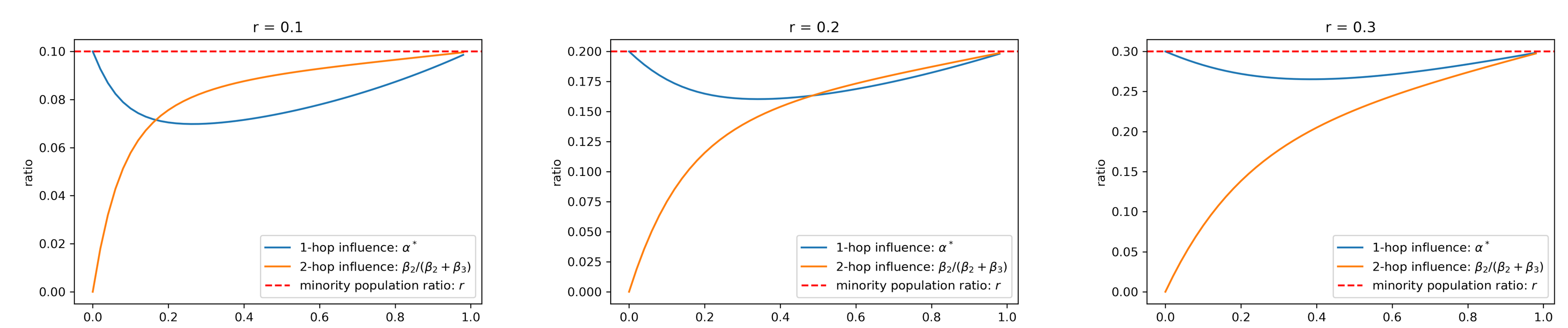

Figure 2. 2-hop connections are likely to reduce the 1-hop connections when (1) the minority ratio in the network is small, and (2) there is a moderate tendency for users to connect with people from the other group.

The passive gain case is considered from the fact that recruiters do also send out job referral offers to job seekers in reality, and this process may also generate bias from the job seeker's perspective. For all the theore tical and empirical results with $\mathcal{S}_L$, for active gains in this section also exactly apply to passive gains, which is not necessarily true for other strategies and paradigms.

By simulating $\mathcal{S}_L$ with uniformly distributed t_v and a_u values on real large network datasets with defined majority/minority groups (DBLP, men/women; Deezer, women/men; Instagram, men/women; Twitch, English/German speakers), we provide empirical evidence of this effect.

	DBLP	Deezer	Instagram	Twitch
minority population ratio (%)	20.73	44.33	54.44	7.05
cross edges (%)	26.10	47.49	41.67	5.93
homophily rarefication index	0.73	0.96	0.84	0.45
linear – 1st hop (%)	17.18	42.85	51.64	3.81
linear – 2nd hop (%)	19.89	42.91	49.17	3.37

Table 1. Under strategy $\mathcal{S}_L$, the bias found in 1-hop referral is reduced in 2-hop for Deezer and DBLP but amplified for Instagram and Twitch.

Inquiry Strategies

Alternatively to dissemination strategies, we consider the case in which an individual seeks some limited number of job referrals, or making an "inquiry" about getting a referral. We consider 3 strategies in particular:

- *Random* ($\mathcal{S}_R$): each referral sender u selects one of the neighbors to share the referral uniformly at random, and the selected receiver v accepts the referral with probability a_v. Once accepted, v repeats the same selection progress to promote the referral to the second hop
- *Popularity-based* ($\mathcal{S}_P$): each referral sender u selects the neighbor with the maximum degree among all of u's neighbors to share the referral, and the selected receiver v accepts the referral with probability a_v. Once accepted, v repeats the same selection progress to promote the referral to the second hop
- *Acceptance-based* ($\mathcal{S}_A$): each referral sender u selects the neighbor with the maximum acceptance rate among all of u's neighbors to share the referral, and the selected receiver v accepts the referral with probability a_v; if accepted, the referral stops at the first hop; otherwise, the selected receiver v passes the referral to a neighbor w selected uniformly at random, and the newly selected receiver w accepts with probability a_w.

When the number of referral requests is kept constant, the impact of multi-hop referral on fairness is determined by the inquiry strategies employed, and does not vary with network conditions.

Theorem 2

Because the chance that a red node succeeds in a 2-hop referral is the same as a blue node succeeds under both the $\mathcal{S}_R$ and the $\mathcal{S}_P$ strategies, the ratios of their expected 2-hop referral reach are just the population ratios: $\lim_{t\to\infty}\frac{\mathbb{E}[AG^{(2)}_{\mathcal{S}_R(t)}(R)]}{\mathbb{E}[AG^{(2)}_{\mathcal{S}_R(t)}(B)]} = \lim_{t\to\infty}\frac{\mathbb{E}[AG^{(2)}_{\mathcal{S}_P(t)}(R)]}{\mathbb{E}[AG^{(2)}_{\mathcal{S}_P(t)}(B)]} = \frac{r}{1-r}$.

For the $\mathcal{S}_A$ strategy, the bias found in the first hop cannot be amplified in the second hop: $\lim_{t\to\infty}\frac{\mathbb{E}[AG^{(2)}_{\mathcal{S}_A(t)}(R)]}{\mathbb{E}[AG^{(2)}_{\mathcal{S}_A(t)}(B)]} \geq \lim_{t\to\infty}\frac{\mathbb{E}[AG^{(1)}_{\mathcal{S}_A(t)}(R)]}{\mathbb{E}[AG^{(1)}_{\mathcal{S}_A(t)}(B)]}$.

We turn back to our datasets to simulate $\mathcal{S}_R$, $\mathcal{S}_P$ and $\mathcal{S}_A$ for empirical evidence to support these two theorems.

	DBLP	Deezer	Instagram	Twitch
minority population ratio (%)	20.73	44.33	54.44	7.05
cross edges (%)	26.10	47.49	41.67	5.93
homophily rarefication index	0.73	0.96	0.84	0.45
acceptance-driven – 1st hop (%)	19.59	42.20	54.0	7.00
acceptance-driven – 2nd hop (%)	21.20	42.35	55.0	7.80
The 1-hop and 2-hop ratios for the popularity-driven referral strategy and the random referral strategy are always consistent with the minority population ratios.				

Table 2. We observe that under the *acceptance-driven referral strategy* that the 2-hop active gain always alleviates the bias found in 1-hop, while the *random referral strategy* and the *popularity-driven referral strategy* remain unchanged, being the network population ratio.