Fairness in Social Influence Maximization

Ana-Andreea Stoica
Columbia University
astoica@cs.columbia.edu

Augustin Chaintreau
Columbia University
augustin@cs.columbia.edu

ABSTRACT

Algorithms for social influence maximization have been extensively studied for the purpose of strategically choosing an initial set of individuals in a social network from which information gets propagated. With many applications in advertisement, news spread, vaccination, and online trend-setting, this problem is a central one in understanding how information flows in a network of individuals. As human networks may encode historical biases, algorithms performing on them might capture and reproduce such biases when automating outcomes.

In this work, we study the social influence maximization problem for the purpose of designing fair algorithms for diffusion, aiming to understand the effect of communities in the creation of disparate impact among network participants based on demographic attributes (gender, race etc). We propose a set of definitions and models for assessing the fairness-utility tradeoff in designing algorithms that maximize influence through a mathematical model of diffusion and an empirical analysis of a collected dataset from Instagram. Our work shows that being feature-aware can lead to more diverse outcomes in outreach and seed selection, as well as better efficiency, than being feature-blind.

CCS CONCEPTS

• **Theory of computation** → **Network flows**; • **General and reference** → *Empirical studies.*

KEYWORDS

social influence; graphs; social networks; seeds

ACM Reference Format:
Ana-Andreea Stoica and Augustin Chaintreau. 2019. Fairness in Social Influence Maximization. In *Companion Proceedings of the 2019 World Wide Web Conference (WWW '19 Companion), May 13–17, 2019, San Francisco, CA, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3308560.3317588

1 INTRODUCTION

Social influence maximization has been a widely studied problem in online networks, having impactful applications in advertisement campaigns, viral online content, news propagation, disease spread, and many others. In spreading an idea, product, or technology through a social network, the network structure plays a crucial role in the efficient propagation of such a process: who *has* the information and who they are *connected* to determines who that information reaches. Thus, the status of being 'early-adopter', who either adopts the desired product out of their own will or receives it for free, is a privileged position for their social status and their direct impact in the diffusion process. The early-adopters promote the content to their friends, who in turn may or may not adopt it and continue the process, resulting in a cascade.

In maximizing the outreach, choosing the early-adopters strategically is crucial. Many algorithms have been proposed in finding an optimal set, from a greedy choice to leveraging the network structure through degree or distance centrality. In theory, such algorithms are blind to demographics, as they typically assess the chance for a given individual to be an early-adopter solely based on their position in a network. In practice, however, properties of social networks also encode historical biases and gender artifacts that algorithms can reinforce.

In this paper, we aim to understand the mechanisms behind information diffusion in social networks and to design algorithms that maximize social influence in a *fair* way. Fairness in machine learning and in graph algorithms is becoming an increasingly popular field in the computer science, sociology, policy, and law, as recent work shows the effect of automated algorithms in mirroring or amplifying bias in human datasets. This is an even more urgent problem in the case of relational data, where connections between individuals may be used as proxies for location, income, social status, and can thus be used by algorithms to exacerbate inequality between different demographic groups.

To our knowledge, however, fairness in social influence maximization has yet to be defined, nevertheless solved. Thus, we begin by analyzing the state-of-the-art algorithms for information diffusion and the underlying graph models, following by adapting various fairness definitions used in classification settings. We perform an empirical analysis using a collected dataset from Instagram, analyzing the effect of different seeding heuristics on the community structure. Furthermore, we develop a mathematical model to formally compare strategic and non-strategic seeding and quantify their effect on disparate outcome among different populations of a network, using as a diffusion process the independent cascade model. We conclude by formulating a way of tackling fairness in the social influence maximization problem, arguing that a strategic heuristic that is feature-aware can be fairer and more efficient.

2 RELATED WORK

Previous literature tackles this problem as an optimization problem by finding a submodular function that models the reach and optimizing it to find the best seed set. The pioneering work of social influence algorithms by [11, 12] in 2003 kicked off by proving that submodular optimization extends to a large collection of social influence processes. Amongst other more robust heuristics and efficient implementations [6], more recent papers extended such methods to

This material is based upon work supported by the National Science Foundation under Grant No.1254035 and 1514437.

WWW '19 Companion, May 13–17, 2019, San Francisco, CA, USA

ACM ISBN 978-1-4503-6675-5/19/05.
https://doi.org/10.1145/3308560.3317588

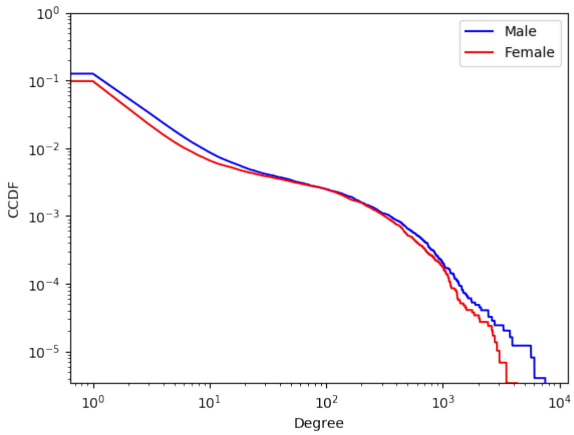


Figure 1: Complementary cumulative distribution function of the degrees in the Instagram graph, plotted for the males (blue) and the females (red).

deal with uncertain networks through online adaptive queries [18], and recently established the importance of exploiting community structure to best direct influence [4]. In this paper, we focus on the algorithmic aspect of the problem, analyzing current algorithms for maximizing influence and their implications on social bias.

A recent line of work focuses on fairness in algorithmic design for human data, with a large focus on classification tasks. From the classic example of predictive policing algorithms that reproduce bias in the data [1, 5, 22], the field of fairness in machine learning developed several ways of measuring and defining *fairness*, from defining statistical constraints on the output based on the particularities of the input data [8, 14], to pre-processing biased data [9, 10, 23], and to identifying the causal relationships between features in the data that lead to bias [13, 16]. Measurement techniques have also been developed in order to understand bias in massive data [21].

Amongst these efforts, two schools of thought emerge, one that argues for a feature-blind strategy in order to not reinforce biases from input data, and one that argues for a feature-aware strategy that mitigates bias after learning it from data. While both have applications in different contexts, an argument that [7] makes is that feature-blindness might lead to amplifying bias through confounding factors. In our work, we explore these two views and show indeed that a feature-blind strategy for social influence may cause unfair outcomes. In order to understand the effect of strategy in seed selection, we ask the question: what does it mean to be fair in information diffusion?

As mentioned in the introduction, we are interested in the effect that social influence maximization has on the different communities present in the network. As such, we may define the following:

(1) *Fairness for early-adopters:* Given a network with communities $C_1, C_2, \cdots C_k$, a social influence maximization algorithm is fair in choosing a seed set S if the proportion of all communities in the seed set is equal,

$$\frac{\mathbb{E}|\{u \in S | u \in C_i\}|}{|C_i|} = \frac{\mathbb{E}|\{u \in S | u \in C_j\}|}{|C_j|}. \quad (1)$$

(2) *Fairness in outreach:* Given a network with communities $C_1, C_2, \cdots C_k$, a social influence maximization algorithm is fair in the outreach it achieves if the cascade reaches all communities in a calibrated way,

$$\frac{\mathbb{E}(|\{u \text{ adopted } | u \in C_i\}|)}{|C_i|} = \frac{\mathbb{E}(|\{u \text{ adopted } | u \in C_j\}|)}{|C_j|} \forall i, j. \quad (2)$$

This notion can ensure that the product or the information being spread reaches the population in a calibrated manner, reaching the same percentage of each community. This is particularly important in news spread, where manipulating distribution of news can lead to misinformation and to amplifying an echo chamber effect, and maybe even more compelling for public health concerns..

Both definitions are in essence a form of calibration or statistical parity (as defined in [14]), and we argue that both are necessary: while the outreach definition ensures that the product or idea reaches all communities equally, early-adopters have a special role in the diffusion process. Having early access to products may establish their role as influencers in the network and allowing them to acquire social capital and trust, as well as to leverage that idea or product for financial gain. A classic example is being an early adopter in Bitcoin, where any user who had access to early mining had the chance of exponentially increasing their profits even when individual mining stopped being profitable. Similarly, influencers on Instagram get paid by companies to promote products. Due to the popularity of the users, the product is more likely to be seen or adopted, and thus being an early-adopter has its perks.

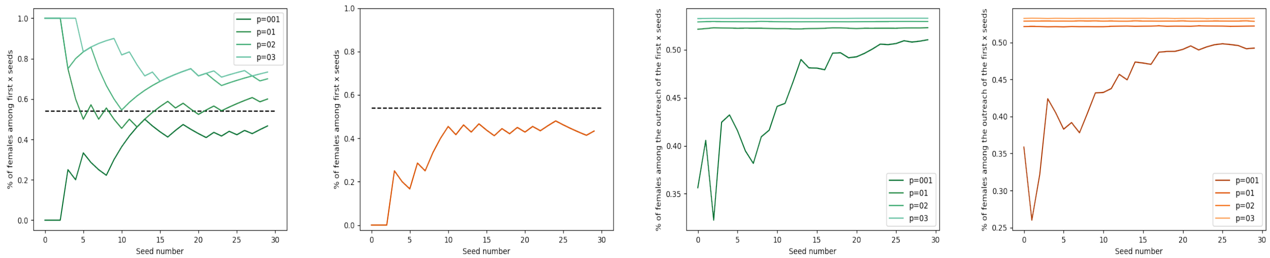
3 EMPIRICAL ANALYSIS

We performed experiments on a collected dataset from Instagram, consisting of public profiles, photos, usernames, likes and comments received on photos, and so on. As Instagram is one of the main hubs for advertisement purposes, with users continuously promoting paid content through the photos or descriptions they are posting, this network provides a comprehensive view of how information diffusion occurs.

3.1 Data Collection

In collecting the data, we used the Instagram API to crawl public profiles, starting from the founder of Instagram, Kevin Systrom. For each such public photo, we recorded the author, the number of likes and comments, and who liked and commented on the item. In total, we collected $115,796,284$ photos over multiple months of 2014 and 2015, which amounted to $539,023$ different users for whom we could infer gender from their names, as described below. While the data might not encompass the full diversity of Instagram users, it is sufficient to obtain statistically significant results most of the time.

Gender inference. We inferred the gender of the users from their first name, adapting the method from [17] that leverages first names and social security data. Filtering for names with less than 50 occurrences for either gender, we obtained a set of $32,676$ unique first



(a) Greedy (b) Degree (c) Greedy (d) Degree

Figure 2: Percentage of females among chosen seeds for the greedy algorithm (a) and the degree algorithm (b), and among the outreach for the greedy algorithm (c) and the degree algorithm (d) for different values of p .

names that we used to label about half a million users. The gender proportions are 45.57% males and 54.43% females.

The graph of likes and comments. In creating a network structure out of these aggregated users and their information, we used the likes and comments they left on eachother's posts as interactions. Whenever someone gave a like or comment on a photo, we created an edge from the originator to the recipient, ending up with **640,211 edges** between **539,023 nodes**.

We will be using this graph for the purpose of our analysis. The dataset consists of many more users and interactions for which we could not determine gender.

A first analysis reveals a disparity in the degree distribution for the two genders. As Figure 1 shows, although males are in minority, they have higher degrees (number of interactions) than females, gap that increases for highly connected nodes.

3.2 Influence maximization algorithms

In studying strategic choices of seeds versus label-blind choices, we focus on two main algorithms, the greedy algorithm as a strategic one and the high-degree algorithm as one that ignores the labels of the nodes, and their effect on a bipopulated network represented by the Instagram dataset.

Greedy algorithm. The greedy algorithm is defined as choosing the seeds sequentially, each time choosing the node that gives the best marginal increase in the resulting cascade (outreach), until we reach a seed set of a predefined size. [11] shows that this provides a $1 - 1/e$-approximation to the best choice of seeds overall. In our implementation, we choose the seedset size to be equal to $k = 30$ nodes, and assume the independent cascade model for information diffusion. Since this model is probabilistic, at each step of the process we choose the node that gives the best marginal outreach averaged over 1000 realizations, choosing conducting edges pseudo-randomly while varying conducting probability p. This algorithm is strategic since the marginal contribution of a node might be dependent on the history of chosen seeds and their labels.

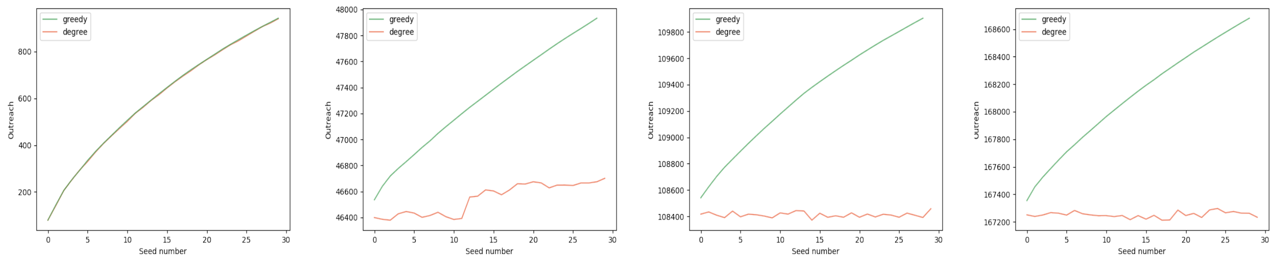
Degree algorithm. Widely used in sociology literature [20], this heuristic chooses nodes as seeds in decreasing order of their degree in the graph and is commonly known as a degree-centrality heuristic. The node selection is done independently of the labels and of the conducting probability p, yet measuring outreach is not and requires again an approximation technique due to the intractability of the problem. We again average over 1000 realizations just as above.

Fairness in seed selection. Figure 2 (a-b) shows the percentage of females in the choice of seed sets for each of the above algorithms. For the greedy algorithm, we notice that for $p = 0.01$ the fraction of females is lower than their fraction in the network, 54%, for small seedset sizes, and it slowly converges towards their true fraction as the seedset increases. However, for $p = 0.1, 0.2, 0.3$, the seedset starts by being *exclusively* female and slowly converges towards a more equilibrated gender ratio as more seeds are added.

For the degree algorithm, men are at the top with higher degrees than women, so the seed selection captures this in Figure 2 (b). We observe that even after choosing $k = 30$ seeds, women are under-represented, given their true fraction of 54%. Our results are consistent with varying the seedset size. Given that the degrees of women are not only lower at the very top but in general (Figure 1), this disparity will continue as the number of seeds increasing, deeming the degree algorithm generally unfair in representing one of the populations. It is generally difficult to assess whether this algorithm reinforces or simply perpetuates a present inequality, yet it raises the question of whether and how can we leverage the degree information in a way that is both fair and efficient.

Fairness in outreach. Figure 2 (c-d) shows the percentage of females in outreach as the seedset size increases up to $k = 30$, for the greedy and degree algorithms. Consistent with the measurements for seedset composition, the percentage of females in outreach converges to their true fraction of 54% as p increases from 0.1 to 0.3. For $p = 0.01$, however, both algorithms fail to reach a proportionate number of women in the network, the degree algorithm performing slightly worse than the greedy one.

Such results give a sense of the trade-off between fairness as diversity and efficiency. Figure 3 shows that the greedy algorithm achieves far better outreach than the degree heuristic for $p = 0.1, 0.2, 0.3$, and more or less the same for $p = 0.01$. Coincidentally, for $p = 0.01$, both algorithms prefer male over female although the male community is in minority. This is even the more meaningful as the first seed generally carries most of the marginal gain for influence for $p = 0.1, 0.2, 0.3$ (see Figure 3 (b-d)), and so



(a) (b) (c) (d)

Figure 3: Performance in terms of average outreach for the greedy algorithm (green) and the degree algorithm (orange) for $p = 0.01$ (a), $p = 0.1$ (b), $p = 0.2$ (c), $p = 0.3$ (d).

$p = 0.01$ represents a case where the influence of each subsequent seed really affects the outcome.

Thus, for such a case being strategic can also mean being more fair. As [4] argues in a similar fashion, in the case of an independent cascade, diversifying between communities gives a better outreach to nodes that otherwise would not be reached.

While the choice of the conducting probability p may differ in different scenarios and should be measured when such data is available, clearly the case in between $p = 0.01$ and $p = 0.1$ is the most notable one, since beyond that the first seed carries a disproportionate amount of the weight.

However, even in such cases being strategic doesn't necessarily mean implementing a greedy heuristic. Due to the computational difficulty of the greedy algorithm and the potential availability of other metrics in the data (such as degree, for example), it is important to understand the implications of such metrics. While an available metric might lead to a label-blind heuristic, that might also lead to disparate outcomes. As we have seen, a degree heuristic will always be unfair in the seed selection in a network whose populations follow different degree distributions (in this case, a majority population gets under-represented because of lower degrees).

In exploring this avenue of strategic seeding in a fair manner we turn to a theoretical model for understanding the cascading process.

4 THE LIMITATIONS OF LABEL-BLIND SEEDING

An established model of network growth that encapsulates clustering dynamics and a power-law degree distribution is the preferential attachment model. Prior work [2, 3, 15, 19] stresses the applicability of such a model into describing real-world network formation, having the caveat of a theoretical understanding of its dynamics. For the purpose of this study, we use the preferential attachment model with two communities and homophily, as defined below.

Preferential attachment model with homophily:

- *Minority-majority partition*: every timestep a new node appears and chooses a label, red (R) w.p. r and blue (B) w.p. $1 - r$, where $0 < r < 1$.
- Every timestep the new node connects to some other node according to *preferential attachment*, meaning that it choses a node in the network with probability proportional to that node's degree. In our model each new node forms $d > 1$ connections.
- *Homophily:* if the two nodes have the same label, connect, otherwise accept the connection w.p. ρ, where $0 < \rho < 1$.

This model results in a community structure (nodes of each label are homophilic), and when the two communities differ in sizes, so do their degree distributions. Indeed, [3] show that the degree distribution of this network follows power law for the R and B nodes with coefficients

$$\beta(B) = 1 + \frac{1}{\frac{r\rho}{2\alpha+2(1-\alpha)\rho} + \frac{1-r}{2\alpha\rho+2(1-\alpha)}}, \tag{3}$$

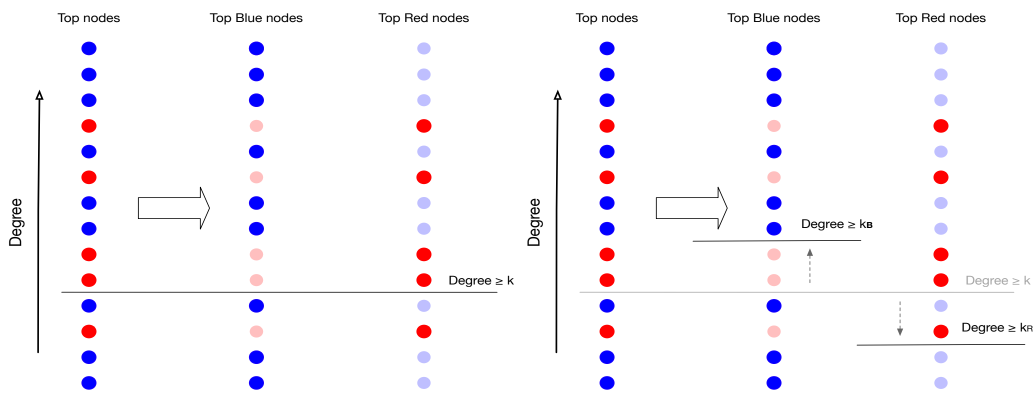
$$\beta(R) = 1 + \frac{1}{\frac{(1-r)\rho}{2\alpha\rho+2(1-\alpha)} + \frac{r}{2\alpha+2(1-\alpha)\rho}}, \tag{4}$$

where α denotes the fraction of edges that emerge from red nodes and the total number of edges.

When the proportions of R and B nodes are different ($r < 1/2$), we know that $\beta(R) > 3 > \beta(B) > 2$ and that a so-called glass ceiling effect is present (the difference in degrees for the two communities increases for more well-connected nodes). Our Instagram dataset follows a two-community structure as well (men and women), and while men are in minority, they have higher degrees than women, with a degree distribution resembling a power-law with different coefficients (Figure 1).

Independent cascade: On top of this model we simulate the independent cascade mechanism, where each node that receives the information/product/etc will adopt and promote to his friends with probability p, for $0 < p < 1$, in order to achieve an optimal seedset of size $k(n)$, a function that depends on the network size n.

Strategic seeding: In the social influence maximization problem the task is to find a set of "seeds" in the network starting from which information diffusion has maximal outreach. In the independent cascade model, the greedy algorithm is shown to approximate the optimal choice of sees [11], yet it is NP-complete and thus very hard to analyze. However, in many cases, choosing the nodes of top degree achieves a similar effect as the greedy algorithm. For the purpose of this study we assume that baseline strategic seeding



(a) Label-blind **(b) Differentiated**

Figure 4: Model of choosing seed nodes based on the degree hierarchy, in a label-blind manner (a) and in a differentiated way based on label (b).

chooses nodes from the degree ranking to maximize social influence, setting a threshold $k(n)$ above which all nodes of higher degree are chosen as seeds (Figure 4 (a)).

Hypothesis: Our claim is that color-blind seeding is not necessarily as efficient as color-aware seeding, meaning that choosing only from the highest degree nodes is equivalent to choosing only from the B community in a glass ceiling context, effectively ignoring a whole community, the R nodes. We aim to show that in some conditions choosing differentiated thresholds for the top degree nodes in each community is a better strategy that achieves higher outreach and more diversity simultaneously (Figure 4 (b)). The intuition behind is that at some point there will be more unconverted/uninfluenced nodes in the minority community.

Definition 4.1. The baseline blind strategy defines the seedset as $S_{k(n)} = \{v \in V | deg(v) \geq k(n)\}$.

Definition 4.2. The *influence* of a set S over another set T is defined as the expected number of edges with one end point in S and one in T:

$$\phi_S(T) = |\{(u, v) \in E | u \in S, v \in T\}|. \tag{5}$$

Given this influence, we know that each edge conducts with probability p, so the number of conducting edges within each community is $p \cdot \phi_{S_{k(n)}}(B/R)$.

Goal: Our goal is to find two thresholds $k^R(n)$ and $k^B(n)$ that give in expectation the same amount of seeds as a general ("blind") threshold $k(n)$ but better influence (Figure 4 (b)):

Theorem 4.3. *For the independent cascade model in a network that follows the preferential attachment model with homophily of two communities of red (R) and blue (B) nodes, and a general threshold $k(n)$ for choosing seeds according to their degree, there exists $k^R(n)$ and $k^B(n)$ with $k^R(n) < k^B(n)$, for choosing seeds according to these respective degree thresholds for which*

$$\begin{gathered}\mathbb{E}(\phi(S_{k(n)})) < \mathbb{E}(\phi(S_{k^R(n)} \cup S_{k^B(n)})), \text{ under the constraint} \\ \mathbb{E}(|S_{k(n)}|) = \mathbb{E}(|S_{k^R(n)} \cup S_{k^B(n)}|).\end{gathered} \tag{6}$$

We only give a proof sketch of the above result due to space considerations. Assume that at each timestep t, a new node comes into the network and forms exactly d edges according to the preferential model with homophily, for $d \geq 1$. Then we can write the expected influence a seed set $S_{k(n)}$ has on the network as

$$\begin{gathered}\mathbb{E}(\phi(S_{k(n)})) = \\ n \cdot \mathbb{P}(v \in B) \cdot \mathbb{P}(v \text{ influenced by one of its } d \text{ edges} | v \in B) + \\ n \cdot \mathbb{P}(v \in R) \cdot \mathbb{P}(v \text{ influenced by one of its } d \text{ edges} | v \in R).\end{gathered} \tag{7}$$

Computing these probabilities for the preferential attachment model (omitted here due to space constraints), we get that

$$\begin{aligned}\mathbb{E}(\phi(S_{k(n)})) = n \cdot (1-r) \cdot \Bigg(1 - \Bigg(1 - 2\Bigg(&\frac{\frac{1-r}{\alpha\rho+1-\alpha}}{\frac{r\rho}{\alpha+\rho(1-\alpha)} + \frac{1-r}{\alpha\rho+1-\alpha}} \cdot (1-\alpha) \\ &\cdot k(n)^{2-\beta(B)} + \frac{\frac{\rho(1-r)}{\alpha\rho+1-\alpha}}{\frac{r}{\alpha+\rho(1-\alpha)} + \frac{\rho(1-r)}{\alpha\rho+1-\alpha}} \cdot \alpha \cdot k(n)^{2-\beta(R)}\Bigg)\Bigg)^d\Bigg) + \\ n \cdot r \cdot \Bigg(1 - \Bigg(1 - 2\Bigg(&\frac{\frac{r}{\alpha+\rho(1-\alpha)}}{\frac{r\rho}{\alpha+\rho(1-\alpha)} + \frac{1-r}{\alpha\rho+1-\alpha}} \cdot (1-\alpha) \cdot k(n)^{2-\beta(B)} + \\ &\frac{\frac{r}{\alpha+\rho(1-\alpha)}}{\frac{r}{\alpha+\rho(1-\alpha)} + \frac{\rho(1-r)}{\alpha\rho+1-\alpha}} \cdot \alpha \cdot k(n)^{2-\beta(R)}\Bigg)\Bigg)^d\Bigg)\end{aligned} \tag{8}$$

We note that in computing these probabilities in closed form we are limited by the un-constrained nature of the cascading process. For the purpose of this study we make the assumption that the process of diffusion expands to two steps, friends of friends, with the hope of generalizing this work in the future. Since we want to find $k^R(n, x)$ and $k^B(n, x)$ that satisfy our goal in equation 6, we set $k^B(n,x) = k(n) \cdot x$, for a variable x, and find that solving for the constraint,

$$k^R(n,x) = \Bigg(\frac{1-\alpha}{\alpha} \cdot \frac{\beta(B)-2}{\beta(B)-1} \cdot \frac{\beta(R)-1}{\beta(R)-2} \cdot k(n)^{1-\beta(B)} \cdot (1 - x^{1-\beta(B)}) + k(n)^{1-\beta(R)} \Bigg)^{\frac{1}{1-\beta(R)}}. \tag{9}$$

Then, we can write

$$F(x) = \mathbb{E}(\phi(S_{k^B(n,x)} \cup S_{k^R(n,x)})) - \mathbb{E}(\phi(S_{k(n)})), \tag{10}$$

with the goal of finding x such that $F(x) > 0$. For such an x, it would mean that a more diverse choice of seeds can actually achieve a better outreach, compensating for the intrinsic bias of a network.

We observe that $F(1) = 0$, and if $x < 1$, we are essentially lowering the threshold for the B nodes to make it to the seed set. Since the seed set is already consisting of mostly B nodes, we would like to do the reverse in order to increase diversity, namely to choose $x > 1$ for which $F(x) > 0$. Empirical simulations of this functions show that there are ranges of parameters for which $F(x) > 0$ for $x \in [1, 1+m]$ for some $m \in \mathbb{R}_+^*$. Thus, it is sufficient to show find the ranges of parameters for which $\frac{\partial F}{\partial x}(1) > 0$, which would mean that the function F is increasing around $x = 1$, so it will be positive for a range of $x > 1$ (note: F is continuous in x). The fact that $k^R(n) < k^B(n)$ shows that there is a less stringent threshold for the red nodes, so they will get better represented in the seed selection given their deficiency in the degree distribution.

5 DISCUSSION

These results show the interplay between being fair and strategic: starting with the most basic heuristic, such as the greedy algorithm, being strategic also leads to being fairer. However, large real-world datasets may not always offer the possibility (computationally or feature-wise) to achieve this, and often times we must make use of available metrics, such as degree, distance centrality, and so on, that may be blind to sensitive features. In networks where there exists inequality based on sensitive features, we must ensure that it does not get propagated through our algorithmic design. Our work shows a first step in mitigating the bias that a feature-blind algorithm perpetuates, showing how strategic choices can achieve both diversity in information diffusion and better efficiency. We are able to design such strategic choices with a mathematical guarantee of optimality and fairness.

Future work should analyze the parameter conditions for which strategic seeding as defined above achieves better outreach and diversity in different networks and assess its quantitative benefit. Our method opens up a set of questions regarding the general nature of such trade-offs between fairness and efficiency. While the nature of the independent cascade model favors diversity in maximizing outreach, other models may do the opposite. Indeed, an interesting corollary would be to study critical mass models, such as the linear threshold model. In such cases, a strategic algorithm would focus on achieving that critical mass for information diffusion locally, which may constrain information within one community and prevent it from being equally distributed across the network.

REFERENCES

[1] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. Machine bias. *ProPublica* (May 2016).

[2] Chen Avin, Avi Cohen, Pierre Fraigniaud, Zvi Lotker, and David Peleg. 2018. Preferential attachment as a unique equilibrium. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 559–568.

[3] Chen Avin, Barbara Keller, Zvi Lotker, Claire Mathieu, David Peleg, and Yvonne-Anne Pignolet. 2015. Homophily and the glass ceiling effect in social networks. In *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science*. ACM, 41–50.

[4] Eric Balkanski, Nicole Immorlica, and Yaron Singer. 2017. The Importance of Communities for Learning to Influence. In *Advances in Neural Information Processing Systems*. 5864–5873.

[5] Solon Barocas and Moritz Hardt. 2017. Fairness in Machine Learning Tutorial. *Neural Information Processing Systems* (2017).

[6] Wei Chen, Yajun Wang, and Siyu Yang. 2009. Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 199–208.

[7] Sam Corbett-Davies and Sharad Goel. 2018. The measure and mismeasure of fairness: A critical review of fair machine learning. *arXiv preprint arXiv:1808.00023* (2018).

[8] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness Through Awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference (ITCS '12)*. ACM, New York, NY, USA, 214–226. https://doi.org/10.1145/2090236.2090255

[9] Michael Feldman, Sorelle A. Friedler, John Moeller, Carlos Scheidegger, and Suresh Venkatasubramanian. 2015. Certifying and Removing Disparate Impact. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15)*. ACM, New York, NY, USA, 259–268. https://doi.org/10.1145/2783258.2783311

[10] Moritz Hardt, Eric Price, and Nathan Srebro. 2016. Equality of Opportunity in Supervised Learning. *arXiv preprint arXiv:1610.02413* (2016).

[11] David Kempe, Jon Kleinberg, and Éva Tardos. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 137–146.

[12] David Kempe, Jon Kleinberg, and Éva Tardos. 2005. Influential nodes in a diffusion model for social networks. In *International Colloquium on Automata, Languages, and Programming*. Springer, 1127–1138.

[13] Niki Kilbertus, Mateo Rojas-Carulla, Giambattista Parascandolo, Moritz Hardt, Dominik Janzing, and Bernhard Schölkopf. 2017. Avoiding Discrimination through Causal Reasoning. *arXiv preprint arXiv:1706.02744* (2017).

[14] Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. 2016. Inherent trade-offs in the fair determination of risk scores. *arXiv preprint arXiv:1609.05807* (2016).

[15] Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, D Sivakumar, Andrew Tomkins, and Eli Upfal. 2000. Stochastic models for the web graph. In *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on*. IEEE, 57–65.

[16] Matt J Kusner, Joshua R Loftus, Chris Russell, and Ricardo Silva. 2017. Counterfactual Fairness. *arXiv preprint arXiv:1703.06856* (2017).

[17] Alan Mislove, Sune Lehmann, Yong-Yeol Ahn, Jukka-Pekka Onnela, and J Niels Rosenquist. 2011. Understanding the Demographics of Twitter Users. *ICWSM* 11, 5th (2011), 25.

[18] Lior Seeman and Yaron Singer. 2013. Adaptive seeding in social networks. In *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*. IEEE, 459–468.

[19] Ana-Andreea Stoica, Christopher Riederer, and Augustin Chaintreau. 2018. Algorithmic Glass Ceiling in Social Networks: The effects of social recommendations on network diversity. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 923–932.

[20] Stanley Wasserman and Katherine Faust. 1994. *Social network analysis: Methods and applications*. Vol. 8. Cambridge university press.

[21] Ke Yang and Julia Stoyanovich. 2017. Measuring fairness in ranked outputs. In *Proceedings of the 29th International Conference on Scientific and Statistical Database Management*. ACM, 22.

[22] Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez Rodriguez, and Krishna P Gummadi. 2017. Fairness beyond disparate treatment & disparate impact: Learning classification without disparate mistreatment. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 1171–1180.

[23] Rich Zemel, Yu Wu, Kevin Swersky, Toni Pitassi, and Cynthia Dwork. 2013. Learning Fair Representations. In *Proceedings of the 30th International Conference on Machine Learning (Proceedings of Machine Learning Research)*, Sanjoy Dasgupta and David McAllester (Eds.), Vol. 28. PMLR, Atlanta, Georgia, USA, 325–333. http://proceedings.mlr.press/v28/zemel13.html