

# Fair Influence Maximization: A Welfare Optimization Approach

Aida Rahmattalabi, Shahin Jabbari, Himabindu Lakkaraju  
Phebe Vayanos, Eric Rice, Milind Tambe

June 16, 2020

## Abstract

Several social interventions (e.g., suicide and HIV prevention) leverage social network information to maximize outreach. Algorithmic influence maximization techniques have been proposed to aid with the choice of “influencers” (often referred to as “peer leaders”) in such interventions. Traditional algorithms for influence maximization have not been designed with social interventions in mind. As a result, they may disproportionately exclude minority communities from the benefits of the intervention. This has motivated research on *fair influence maximization*. Existing techniques require committing to a single domain-specific fairness measure. This makes it hard for a decision maker to meaningfully compare these notions and their resulting trade-offs across different applications.

We address these shortcomings by extending the principles of cardinal welfare to the influence maximization setting, which is underlain by complex connections between members of different communities. We generalize the theory regarding these principles and show under what circumstances these principles can be satisfied by a welfare function. We then propose a family of welfare functions that are governed by a single *inequity aversion* parameter which allows a decision maker to study task-dependent trade-offs between fairness and total influence and effectively trade off quantities like influence gap by varying this parameter. We use these welfare functions as a fairness notion to rule out undesirable allocations. We show that the resulting optimization problem is monotone and submodular and can be solved with optimality guarantees. Finally, we carry out a detailed experimental analysis on synthetic and real social networks and show that high welfare can be achieved without sacrificing the total influence significantly. Interestingly we can show there exists welfare functions that empirically satisfy all of the principles.

## 1 Introduction

The success of social interventions relies heavily on effectively leveraging social networks [23, 39, 40]. For instance, health interventions such as prevention of suicide, HIV, and substance abuse involve finding a small set of well-connected individuals who can effectively act as peer-leaders to either detect warning signs beforehand and respond appropriately (suicide prevention) or disseminate relevant information (HIV or substance abuse prevention). Such interventions are prone to fairness and ethical concerns as individuals from racial minorities or LGBTQ communities may be disproportionately excluded from the benefits of the intervention [36, 39]. The celebrated influence maximization framework has been employed to find such a set of individuals [24]. However, this framework primarily optimizes for maximizing the influence spread and does not account for fairness.

Recent work has incorporated fairness directly into the influence maximization framework by proposing various notions of fairness such as maximin fairness [36], and diversity constraints [39]. Maximin fairness aims at improving the minimum amount of influence that any community receives. On the other hand, diversity constraints are inspired by the game theory literature and ensure that each community is at least as well-off had they received their share of resources proportional to their size and allocated them internally. Each of these notions comes with drawbacks for example maximin fairness typically will result in significant degradation in total influence due to its stringent requirement while diversity constraints further exacerbate the influence gap between the communities (see Sections 1.1 and 4 for more details). Moreover, while each

notion offers a unique perspective on fairness, there is no universal agreement on what fairness means as fairness is known to be domain specific (see e.g., [31]). Therefore, it is crucial for the policymakers to be able to compare these notions on different applications. Indeed different applications have different fairness requirements e.g., the risk of excluding vulnerable communities from suicide prevention which has life and death impacts might have a higher weight compared to interventions for a healthier life style.

**Proposed Approach and Contributions** Instead of committing to a single definition of fairness for influence maximization, we build on cardinal welfare theory from the economics literature and provide a family of measures to evaluate fairness in this context. At a high level, the aim of cardinal welfare theory is to design a welfare function to measure the *goodness* or *desirability* of an allocation by proposing a set of desiderata that the welfare function should satisfy. However, due to the structural dependencies induced by the underlying social network, these principles are not directly applicable to our problem. We make the following contributions: (i) We extend the cardinal welfare principles including the *transfer principle* to the influence maximization framework, which is otherwise not applicable. We also propose a new principle which we refer to as *utility gap reduction*. This principle is particularly relevant in the graph-based problems and aims to avoid situations where high aversion to inequity leads to even more influence gap; (ii) We generalize the theory regarding these principles and describe the circumstances under which there exists a welfare function satisfying the principles. We show that if all community members are disconnected from one another (no underlying social network), our results reduce to the existing theory of social welfare; (iii) We propose a natural family of welfare functions that are governed by a single *inequity aversion* parameter which allows a decision maker to study task-dependent trade-offs between fairness and total influence and effectively trade off quantities like influence gap or influence degradation by varying this parameter in the welfare function. We then incorporate these welfare functions as objective into an optimization problem to rule out undesirable allocations. We show that the resulting optimization problem is monotone and submodular and, hence, can be conveniently solved using a greedy algorithm with optimality guarantees; (iv) Finally, we carry out a detailed experimental analysis on synthetic graphs and real social networks of homeless youth participating in substance use interventions to study the trade-off between total influence and welfare. We show that this trade-off is rather mild, as high welfare can be achieved without sacrificing the total influence significantly. Interestingly we can trade off influence gap for total influence by varying our inequity aversion parameter and show there exists welfare functions that empirically satisfy all of the principles.

## 1.1 Related Work

Artificial Intelligence and machine learning algorithms hold great promise in addressing many pressing societal problems. These problems often pose complex ethical and fairness issues which need to be addressed before the algorithms can be deployed in the real world. The nascent field of algorithmic fairness has emerged to address these fairness concerns. In particular, there is a growing body of work focusing on the well-studied machine learning settings such as classification and regression. To this end, different notions of fairness are defined based on one or more *sensitive attributes* such as race or gender. These notions mainly aim at equalizing a statistical quantity across different communities or populations [10, 20, 42]. While surveying the entirety of this field is out of our scope (see e.g., [6] for a recent survey), we point out that there is a wide range of fairness notions defined for these settings and it has been shown that the right notion is problem dependent [5, 31] and also different notions of fairness can be incompatible with each other [26]. Thus, care must be taken when we employ these notions of fairness across different applications including resource allocation problems and in particular, influence maximization.

Motivated by the importance of fairness when conducting interventions in social initiatives [27], fair influence maximization has received a lot of attention recently [1, 17, 36, 39]. These works have incorporated fairness directly into the influence maximization framework by (1) relying on either Rawlsian theory of justice [36, 37], (2) game theoretic principles [39] or (3) equality based notions [1, 38]. We will discuss the first two approaches in more details in Sections 2.1 and 3, as well as in our experimental section. Equality based approaches do not leverage social network information and imposing strict equality leads to wastage of

resources. We will discuss this in more details in Appendix A.

Fish et al. [17] investigate the notion of information access gap, where they propose maximizing the minimum probability that an individual is being influenced/informed to constrain this gap. As a result they study fairness at an individual level while we study fairness at the group level. Also, their notion of access gap is limited to the gap in a bipartition of the network which is in principle different from utility gap that we study which accommodates arbitrary number of protected groups.

Similar to our work, Ali et al. [1] also study utility gap. They propose an optimization model that directly penalizes utility gap which they solve via a surrogate objective function. Their surrogate functions are in the form of a sum of concave functions of the group utilities which are aggregated with arbitrary weights. Unlike their work, our approach takes an axiomatic approach with strong theoretical justifications and it does not allow for arbitrary concave functions and weights as they violate the welfare principles.

There has also been a long line of work considering fairness in resource allocation problems (see e.g., [7–9, 25]). More recently, group fairness has been studied in the context of resource allocation problems [4, 12, 15] and specifically in graph covering problems [36]. In resource allocation setting, *maximin fairness* and *proportional fairness* are widely adopted fairness notions. Proportional fairness is a notion introduced for bandwidth allocation [9]. An allocation is proportionally fair if the sum of percentage-wise changes in the utilities of all groups cannot be improved with another allocation. In classical resource allocation problems, each individual or group has a utility function that is independent of the utilities of others individuals or groups. However, this is not the case in influence maximization due to the underlying social network structure. We note that, while in the bandwidth allocation setting there is also a network structure, the utility of each vertex is still independent of the other vertices and is only a function of the amount of resources that the vertex receives.

Finally, Heidari et al. [22] have recently proposed to study inequity aversion and welfare through cardinal welfare theory in the context of regression problems. Their main contribution is to use this theory to draw attention to other fairness considerations beyond equality. However, we use the welfare theory to propose a tool for policymakers to study the task-specific trade-offs between various quantities like influence gap and total influence in different domains. Moreover, the classical social welfare theory does not readily extend to our setting due to dependencies induced by the social network. Indeed, extending those principles is a contribution of our work.

## 2 Problem Formulation

We use  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  to denote a graph (or network) in which  $\mathcal{V}$  is the set of  $N$  vertices and  $\mathcal{E}$  is the set of all edges. In the *influence maximization problem*, a decision maker chooses a set of at most  $K$  vertices (or seeds) to influence (or activate). The selected vertices then spread the influence in rounds according to the *Independent Cascade Model* [24]. According to this model, each newly activated vertex will attempt to spread the influence to its neighbors independently and with a fixed probability  $p \in (0, 1]$ . The process continues until no new vertex is influenced. We use  $\mathcal{A}$  to denote the initial set of seeds. The goal of the decision maker is to select a seed set  $\mathcal{A}$  to maximize the expected number of vertices that are influenced at the end of this process.

To address the ethical concerns mentioned in Section 1 regarding excluding minority communities, we take a group fairness approach. More concretely, we assume each vertex of the graph belongs to one of the disjoint communities  $c \in \mathcal{C} := \{1, \dots, C\}$  such that  $\mathcal{V}_1 \cup \dots \cup \mathcal{V}_C = \mathcal{V}$  where  $\mathcal{V}_c$  denotes the set of vertices that belong to community  $c$ . This partitioning can be induced by e.g., the intersection of a set of (protected) attributes such as race or gender for which fair treatment is important. Indeed, our model allows for the case that a vertex belongs to multiple communities by considering the intersections of communities. We use  $N_c$  to denote the size of community  $c$  i.e.,  $N_c = |\mathcal{V}_c|$ . Furthermore, we define  $\mathcal{A}^* := \{\mathcal{A} \subseteq \mathcal{V} \mid |\mathcal{A}| \leq K\}$  as the set of feasible allocations. Finally, for any allocation  $\mathcal{A} \in \mathcal{A}^*$ , we let  $u_c(\mathcal{A})$  denote the utility, i.e., the expected fraction of the influenced vertices of community  $c$ , where the expectation is taken over the spread of influence.

The standard influence maximization problem solves the following optimization problem.

$$\underset{\mathcal{A} \in \mathcal{A}^*}{\text{maximize}} \quad \sum_{c \in \mathcal{C}} N_c u_c(\mathcal{A}).$$

When it is clear from the context we drop the dependencies of utilities on  $\mathcal{A}$ .

## 2.1 Existing Notions of Fairness in Influence Maximization

The optimization problem for the influence maximization solely attempts to maximize the total influence. Many fairness-enforcing interventions can be interpreted as an added constraint to this optimization problem. We discuss a couple of these notions next (see Section 1.1 for more discussion).

**Maximin Fairness** *Maximin fairness* [36, 39] is a generalization of Rawlsian theory of justice [37] aiming to maximize the minimum utility that any of the communities obtains. For any allocation  $\mathcal{A} \in \mathcal{A}^*$ , we can compute the minimum utility that any community will receive under  $\mathcal{A}$ . Let  $\gamma$  denote the maximum of this minimum utility where the maximization is taken over all allocations in  $\mathcal{A}^*$ . Then the maximin fairness only allows allocations  $\mathcal{A} \in \mathcal{A}^*$  that satisfy the following constraint.

$$\min_{c \in \mathcal{C}} u_c(\mathcal{A}) \geq \gamma, \quad \mathcal{A} \in \mathcal{A}^*.$$

**Diversity Constraints** Inspired by the game theoretic notion of core, *diversity constraints* require that no community should be able to obtain a higher utility with a share of resources proportional to their size and allocating them internally [39]. This is illustrated by the following constraint where  $U_c$  denotes the expected utility of the best allocation of  $\lfloor KN_c/N \rfloor$  seeds to vertices from community  $c$ .

$$u_c(\mathcal{A}) \geq U_c, \quad \forall \mathcal{A} \in \mathcal{A}^* \text{ and } c \in \mathcal{C}.$$

Diversity constraints set utility lower bounds for the communities only based on their relative sizes. The main drawback of this approach is that it can further exacerbate the influence gap between the communities. We empirically compare our approach with both of these notions in Section 4.

## 3 Cardinal Welfare and Fairness

Following the cardinal welfare theory, our aim is to design welfare functions to measure the goodness of allocations for the population as a whole [30]. To do so, cardinal welfare theory propose a set of principles and welfare functions are expected to satisfy these principles.

The interpretation of welfare function as a notion of fairness is justified by the concept of *veil of ignorance* from moral philosophy [37] which is best described with the following thought experiment. Suppose the decision maker knows nothing about the position (e.g., community in our case) that they will be born in. Crucially the decision maker has the full knowledge of the distribution of positions (i.e., the utility of communities and the chance that he will be born into any of these communities). According to the utilitarian doctrine in this hypothetical position a rational decision maker would consider inequity aversion (or *risk* as they call it) and depending on their level of inequity aversion can aim to guard against the undesirable outcomes (e.g., that they will be born to low utility communities). Note that although this is a purely hypothetical exercise, it can help the decisions maker in making fair decisions by considering the possibility of taking positions of low utility community.

To formally incorporate inequity aversion we follow the constant relative risk aversion utility models and model inequity aversion through a single parameter  $\alpha$  where  $\alpha < 1$  models inequity aversion,  $\alpha = 1$  models inequity-neutrality and  $\alpha > 1$  models inequity seeking behavior [30]. The welfare function then is simply a mapping from the utilities and the inequity aversion parameter to a single number. More formally, let  $W : \mathbb{R}^N \times \mathbb{R} \rightarrow \mathbb{R}$  denote the welfare function. The welfare function is defined over the entire population ( $N$ )

instead of over the communities. This ensures that the welfare function takes into account the size of the communities in addition to their average utilities. Intuitively, given the utility values for each community and an equity aversion parameter, the welfare function assigns a score to that combination as a measure of goodness in terms of fairness or distributive justice and a higher number indicates a more desirable combination.

To define the utility vector  $\mathbf{u}$  for the entire population from the utility values for the communities we have two sensible choices. (1) We can replicate the average utility  $u_c$  of community  $c$  by the number of members of that community  $N_c$ . This will model “group fairness” where the goal is to improve the utility of the community as a whole. (2) We can let  $\mathbf{u}$  to be the utility of individuals across the population where the utility of an individual is defined as the probability that the individual is influenced by the intervention. This corresponds to “individual fairness”. While the cardinal welfare theory can account for both of these situations, we opt for the former. As pointed out by Dwork et al. [14], individual fairness is notoriously hard to achieve in practice. This problem is further exacerbated in influence maximization problems due to scarcity of resources and the network structure. Hence we leave the study of individual fairness to future work. For convenience we write  $W(\mathbf{u}, \alpha)$  as  $W_\alpha(\mathbf{u})$  and omit the dependency on  $\alpha$  when it is clear from the context.

### 3.1 Principles of Welfare

The welfare economics literature proposes a set of principles (or desiderata) that any welfare function should satisfy [21]. These principles aim to rank different social states, i.e., given any two utility vectors, these principles determine if they are indifferent or one of them is preferred. We revisit these notions and incorporate necessary changes to make them applicable to our influence maximization problem. Throughout, without loss of generality, we assume all utility vectors belong to  $[0, 1]^N$ .

**(1) Monotonicity:** If  $\mathbf{u} < \mathbf{u}'$ , then  $W(\mathbf{u}') > W(\mathbf{u})$ . In other words, if  $\mathbf{u}'$  Pareto dominates  $\mathbf{u}$ , then  $W$  should strictly prefer  $\mathbf{u}'$  to  $\mathbf{u}$ . This principle also appears as *levelling down* objection in political philosophy [34]. Equality enforcing notions has been recently proposed for influence maximization problems [1, 38]. Unlike many of the machine learning settings, these notions might not be desirable for influence maximization as they violate monotonicity (see Section 1.1 and Appendix A).

**(2) Symmetry:**  $W(\mathbf{u}) = W(P(\mathbf{u}))$ , where  $P(\mathbf{u})$  is any element-wise permutation of  $\mathbf{u}$ . According to this desideratum,  $W$  does not depend on the naming or labels of the communities (or individuals), but only on their utility levels.

**(3) Independence of Unconcerned Communities:** Let  $(\mathbf{u}|^c b)$  be a utility vector that is identical to  $\mathbf{u}$ , except for the utility of community  $c$  which is replaced by a new value  $b$ . The property requires that for all  $c, b, b', \mathbf{u}$  and  $\mathbf{u}'$ ,  $W(\mathbf{u}|^c b) < W(\mathbf{u}'|^c b) \Leftrightarrow W(\mathbf{u}|^c b') < W(\mathbf{u}'|^c b')$ . Informally, this desideratum states that  $W$  should be independent of communities whose utilities remain the same.

**(4) Affine Invariance:** For any  $\alpha > 0$  and  $\beta$ ,  $W(\mathbf{u}) < W(\mathbf{u}') \Leftrightarrow W(\alpha\mathbf{u} + \beta) < W(\alpha\mathbf{u}' + \beta)$  i.e., the relative ordering of two utility vectors is invariant to a choice of numeraire.

It is well-known that any welfare function  $W$  that satisfies the above desiderata is additive and can be written as  $W_\alpha(\mathbf{u}) = \sum_{i=1}^N u_i^\alpha$  for  $\alpha > 0$ ,  $W_\alpha(\mathbf{u}) = \sum_{i=1}^N \ln(u_i)$  for  $\alpha = 0$  and  $W_\alpha(\mathbf{u}) = -\sum_{i=1}^N u_i^\alpha$  for  $\alpha < 0$  [13], where with slight abuse of notation we used  $u_i$  to denote the  $i$ th element of the utility vector  $\mathbf{u}$ . To highlight this notational difference we use the subscript  $i$  to denote the utility of the  $i$ th element (or individual) in  $\mathbf{u}$  and subscript  $c$  to denote the (average) utility for community  $c$ . With this in mind, the welfare function can also be written based on the utility of communities e.g.,  $W_\alpha(\mathbf{u}) = \sum_{i=1}^N u_i^\alpha = \sum_{c \in C} N_c u_c^\alpha$  for the case of  $\alpha > 0$ .

These desiderata are general in the sense that they do not depend on the influence maximization problem and can be applied to any resource allocation problem (graph-based or not). We next study another two

desiderata which focus on the transfer of resources between two communities and unlike the previous desiderata crucially depend on the network structure.

**(\*5) Transfer Principle [19, 35]** : We start by describing the original transfer principle and then describe how we can modify this principle to account for network structure and communities. Consider two individuals  $i$  and  $j$  in utility vector  $\mathbf{u}$  such that  $u_i < u_j$ . Let  $\mathbf{u}'$  be another utility vector that is identical to  $\mathbf{u}$  in all elements except entries  $i$  and  $j$  where  $u'_i = u_i + \delta$  and  $u'_j = u_j - \delta$  for some  $\delta \in (0, (u_j - u_i)/2)$ . Then by the transfer principle  $W(\mathbf{u}') > W(\mathbf{u})$ . Informally, this states that transferring utility from a high-utility to a low-utility individual should increase social welfare.

We are concerned with communities rather than individuals so instead we focus on transfer of utilities from one community to another. In the presence of network structure it might not always be feasible to directly transfer utilities from one community to another without modifying the utilities of other communities. Furthermore, even if the direct transfer is possible, it can be the case that there is no symmetry in the amount of utility gained by the low-utility community and the amount of utility lost by the high-utility community. The generalized transfer principle addresses both of these shortcomings. In the spirit of the original transfer principle we only consider allocations in which some units of resources are transferred from individuals in one community to individuals in another community. We refer to such allocations as neighboring allocations.

**(5) Generalized Transfer Principle:** Suppose  $\mathcal{A}$  and  $\mathcal{A}' \in \mathcal{A}^*$  are two neighboring allocations. Let  $\mathbf{u} = \mathbf{u}(\mathcal{A})$  and  $\mathbf{u}' = \mathbf{u}(\mathcal{A}')$  denote the utility vectors from these allocations. Suppose  $\mathbf{u}'$  is sorted in ascending order and  $\mathbf{u}$  is permuted so that, for all  $c \in \mathcal{C}$ , index  $c$  in both vectors corresponds to the same community. If  $\sum_{\kappa \in \mathcal{C}: \kappa \leq c} N_\kappa (u_\kappa - u'_\kappa) \geq 0, \forall c \in \mathcal{C}$  and  $u_c > u'_c$  for some  $c \in \mathcal{C}$ , then  $W(\mathbf{u}) > W(\mathbf{u}')$ .

Informally, the generalized transfer principle states that in a desirable transfer of utilities the magnitude of the improvement in lower-utility communities should be at least as high as the magnitude of decay in higher-utility communities while enforcing that at least one low-utility community is receiving a higher utility after the transfer. In the absence of the network structure, the generalized transfer principle becomes equivalent to the original transfer principle.

We next study whether any of the welfare functions that satisfy the first four desiderata can also satisfy the generalized transfer principle. In particular in Proposition 1 we show that any additive and strictly concave welfare function satisfies the generalized transfer principle. Since the functions that satisfy the first three desiderata are strictly concave for  $\alpha < 1$  (the inequity neutral and inequity averse behaviors), the transfer principle can also be simultaneously satisfied in this regime.

**Proposition 1.** *Any strictly concave and additive function satisfies the generalized transfer principle.*

Proof are deferred to Appendix A. A key issue in fairness considerations in resource allocation problems is the gap in the utility of communities and in particular the gap between the communities with the highest and lowest utilities (which we simply refer to as utility gap henceforth). For a utility vector  $\mathbf{u}$ , we define  $\Delta(\mathbf{u}) = \max_{c \in \mathcal{C}} u_c - \min_{c \in \mathcal{C}} u_c$  to denote the utility gap. Imposing fairness interventions in allocation problems are usually motivated by the large utility gap before the intervention [33]. In the context of influence maximization, Fish et al. [17] has shown that in social networks the utility gap can further increase after an influence maximizing intervention. Thus, we introduce another desideratum to address this issue and again focus on neighboring allocations.

**(6) Utility Gap Reduction:** Suppose  $\mathcal{A}$  and  $\mathcal{A}' \in \mathcal{A}^*$  are two neighboring allocations. Let  $\mathbf{u} = \mathbf{u}(\mathcal{A})$  and  $\mathbf{u}' = \mathbf{u}(\mathcal{A}')$  denote the utility vectors from these allocations such that  $\sum_{c \in \mathcal{C}} N_c u_c \geq \sum_{c \in \mathcal{C}} N_c u'_c$ . If  $\Delta \mathbf{u} < \Delta \mathbf{u}'$  then  $W(\mathbf{u}) > W(\mathbf{u}')$ . The utility gap reduction simply states that the welfare function should prefer the utility vector with the smaller utility gap if the total utility of the vector with the smaller gap is at least as large as the total utility of the other vector.

We next show that, in general, there is no hope to design a welfare function that obeys the utility gap reduction along with the other desiderata. However, it is possible to satisfy all the desiderata if the communities are disjoint (i.e., there is no edge between individuals from different communities).

**Proposition 2.** *Let  $W$  be a welfare function that obeys desiderata 1-5. If the communities are disjoint then  $W$  also obeys the utility gap reduction.*

However, this result does not hold for general networks as we show next.

**Proposition 3.** *Let  $W$  be a welfare function that satisfies desiderata 1-5. Then there exists an instance of the influence maximization problem and two neighboring allocations  $\mathcal{A}$  and  $\mathcal{A}' \in \mathcal{A}^*$  with corresponding utility vectors  $\mathbf{u} = \mathbf{u}(\mathcal{A})$  and  $\mathbf{u}' = \mathbf{u}(\mathcal{A}')$  such that  $\sum_{c \in \mathcal{C}} N_c u_c \geq \sum_{c \in \mathcal{C}} N_c u'_c$  and  $\Delta(\mathbf{u}) < \Delta(\mathbf{u}')$  but  $W(\mathbf{u}') > W(\mathbf{u})$ .*

We showed that the network structure make it impossible to design a welfare function that satisfies all of the principles in the general case. This is perhaps consistent with the emerging theory demonstrating the fundamental inconsistency of different fairness desiderata for classification [11, 18, 26]. In Section 4 we empirically show that there exists welfare functions that obey all of these principles.

### 3.2 Fair Influence Maximization Through Welfare Maximization

A welfare function measures the desirability of an allocation. Hence as a natural notion of fairness we can require an allocation with the highest welfare as stated in the following optimization problem.

$$\underset{\mathcal{A} \in \mathcal{A}^*}{\text{maximize}} \quad W_\alpha(\mathbf{u}(\mathcal{A})).$$

Lin and Bilmes [28] show that the composition of a non-decreasing concave function (in our case  $\ln(x)$ ,  $x^\alpha$  for  $\alpha \in (0, 1)$  or  $-x^\alpha$  for  $\alpha < 0$ ) and a non-decreasing submodular function (in our case  $u_c(\mathcal{A})$ ) is submodular. Hence our proposed class of welfare functions is submodular. Our welfare functions also satisfy monotonicity. It is well-known that for maximizing any monotone submodular function, there exists a greedy algorithm with a  $(1 - 1/e)$  approximation factor [32] which we can also use to solve the welfare maximization problem.

Each choice of the inequity aversion parameter  $\alpha$  will result in a different welfare function and hence a fairness notion. For any such notion, the decision maker can easily impose a minimum threshold on the welfare function and study the trade-off between welfare (fairness) and total utility by varying this minimum threshold. As we show in our experimental evaluations in Section 4, high welfare is usually achievable without sacrificing the total influence significantly in our real-world networks.

### 3.3 Connection to Existing Notions of Fairness

The welfare maximization framework allows for a spectrum of fairness notions as a function of  $\alpha$ . While maximin fairness cannot be cast in this framework, the specific case of *leximin fairness*<sup>1</sup> corresponds to the welfare function for  $\alpha \rightarrow -\infty$  [22]. Proportional fairness [9], a notion for resource allocation problems, is also closely connected to welfare function for  $\alpha = 0$  (see Section 1.1).

## 4 Computational Results

We investigate the effect of the inequity aversion parameter  $\alpha$  on both the utility gap and total influence and show that by varying  $\alpha$  we can effectively trade off utility gap with total influence. Furthermore, we compare our approach against proposed notions of fairness for influence maximization. Finally, we explore the effectiveness of the greedy algorithm in solving the welfare maximization problem.

**Baselines** As baselines, we compare against diversity constraints (DC) and Maximin fairness. We follow the approach of Tsang et al. [39] for both problems and view these problem as a multi-objective submodular optimization with utility of each each community being a separate objective. They propose an algorithm

---

<sup>1</sup>Among two utility vectors, leximin prefers the one where the worst utility is higher. If the worst utilities are equal, leximin repeats this process by comparing the second worst utilities and so on.

and implementation with asymptotic  $(1 - 1/e)$  approximation guarantee which we also utilize here.<sup>2</sup> We also compare our approach against the solution of the influence maximization problem with no fairness considerations (IM). Particularly, we compare solutions in terms of utility gap, total influence spread and *Price of Fairness* (PoF) defined as the percentage loss of the total influence due to fairness. Precisely,  $\text{PoF} := 1 - \text{OPT}^{\text{fair}} / \text{OPT}^{\text{IM}}$  in which  $\text{OPT}^{\text{fair}}$  and  $\text{OPT}^{\text{IM}}$  are the total influence in influence maximization problems with and without fairness. Hence  $\text{PoF} \in [0, 1]$ , with smaller values being more desirable. As we can not solve the influence maximization problem optimally, in the PoF calculations we utilize the standard greedy algorithm [24]. We use the same greedy algorithm for welfare maximization. We set  $p = 0.2$  across all experiments.

**Synthetic Experiments** We start by demonstrating our setup. The synthetic networks are random instances of the Stochastic Block Model (SBM) network, a widely used model for networks with community structure [16]. In SBM networks, vertices are partitioned into disjoint communities. Within each community  $c$ , an edge between two vertices is present independently with probability  $q_c$ . Between any two vertices in communities  $c$  and  $c'$ , an edge exists independently with probability  $q_{cc'}$  and typically  $q_c > q_{cc'}$  to capture homophily. SBM is suitable as it captures the community structure.

The first result is on SBM networks where we consider 3 communities each of size 60 where the communities differ in their degree of connectedness. In particular, we set  $q_1 = 0.1, q_2 = 0.08, q_3 = 0.02$ . We choose these values to reflect asymmetry in the structure of different communities with at least one community having sparse connections. This reflects real world scenarios since not every community is equally connected. Later, we explore the effect of different connectivity parameters. We also keep the between-community edge probabilities the same and equal to 0.01. All synthetic results are averaged over 25 different realizations of the SBM networks.

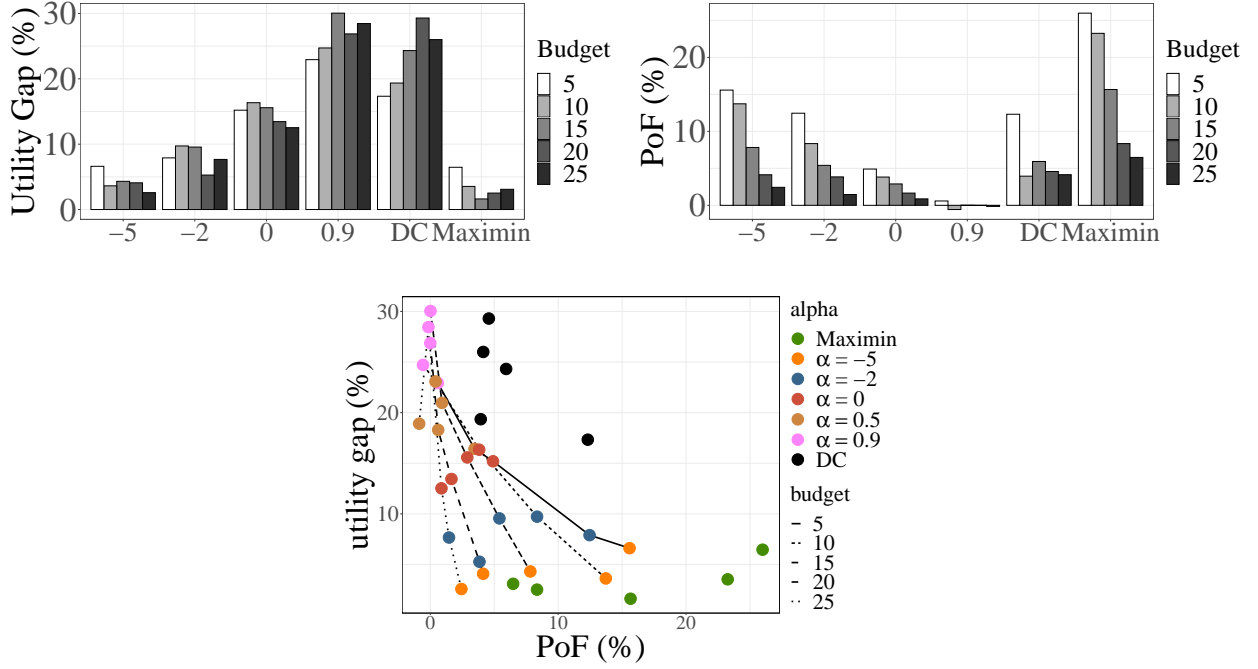


Figure 1: Top left and right panels: utility gap and PoF for different  $K$ ,  $\alpha$  values and baselines. Bottom panel: the trade-off between PoF and utility gap. Each line corresponds to a different budget.

In Figure 1, we investigate the effects of the inequity aversion parameter  $\alpha$  on influence and utility gap.

<sup>2</sup>[https://github.com/bwilder0/fair\\_influmax\\_code\\_release](https://github.com/bwilder0/fair_influmax_code_release)



More specifically in the top left and right panels of Figure 1, for five budget values  $K$  of 5, 10, 15, 20 and 25, we depict the utility gap and PoF for different levels of  $\alpha$  ranging from -5 to 0.9 and our baselines. According to the top left panel of Figure 1,  $\alpha = -5$  has a relatively small utility gap and as we increase  $\alpha$  (lower levels of inequity aversion) the utility gap increases. At  $\alpha = 0.9$ , which is close to being inequity-neutral ( $\alpha = 1$ ), the utility gap is almost as high as DC. Finally, Maximin has the lowest utility gap which is explained by an extreme level of inequity aversion. Indeed, decreasing the utility gap comes at a cost. As shown in the top right panel, Maximin exhibits the highest PoF values. Also increasing  $\alpha$  generally results in a decrease in PoF. For example for  $\alpha = 0.9$ , PoF is very close to zero. We note that the negative values are due to the fact that we used the solution to the greedy as a proxy for  $\text{OPT}^{\text{TM}}$ . This inherent trade-off between utility gap and PoF is depicted in the top right panel in which each color corresponds to a different fairness approach plotted for budget values of 5 to 25 (lines connecting the points). According to this figure, for any given budget and a PoF tolerance level, one can choose an approach with minimum utility gap. In particular, if lower PoF is expected, the decision maker can choose, among different values of  $\alpha$ , one that minimizes the gap. Moreover, the bottom panel shows that the PoF decreases as  $K$  grows. This captures the intuition that fairness becomes a less costly constraint when resources are in greater supply. We also observe that DC is dominated by the other methods both in terms of utility gap and PoF. When deciding the lower-bound utility that every community should receive, DC does not take the between-community edges into account and this will lead to lower quality solutions. Also, by definition DC provides more resources to communities that “will do better with the resources” so it does not show an aversion to inequity. Hence, it does not satisfy the utility gap reduction.

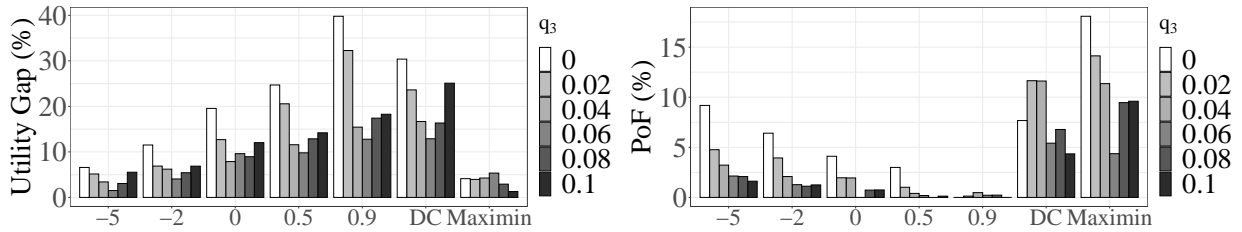


Figure 2: Left and right panels: utility gap and PoF for various connectivity levels  $q_3$ ,  $\alpha$  and baselines.

Next, we study how the relative degree of connectedness of communities affects these two measures. We again consider 3 communities of size 60 and fix  $q_1 = 0.1, q_2 = 0.05, q_3 = 0.0$  to obtain three communities with low, mid and high relative connectivity. Between-community probabilities are set to 0.005 and  $K$  to 25. We gradually increase  $q_3$  from 0.0 to 0.1. Results are summarized in Figure 2 where each group of bars correspond to a different approach for different values of  $q_3$ . The left panel shows that, for all approaches expect Maximin, by increasing  $q_3$  the utility gap gradually decreases until it reaches its minimum level at around  $q_3 = 0.06$ , after which it shows an increasing trend. This is due to structural changes in the network. More precisely, for  $q_3 < 0.06$  we are in the low-mid-high connectivity regime, where the third community receives the minimum utility. As a result, as  $q_3$  increases it becomes more favorable to allocate more resources to the this community which in turn reduces the utility gap. For  $q_3 > 0.06$ , the second community will be become the new worse-off community due its lowest connectivity. Hence, further increase in  $q_3$  causes more separation in connectedness and we see the previous behavior in reverse. While the utility gap for Maximin remain unaffected due its extreme inequity aversion, it exhibits a similar  $U$ -shaped behavior in PoF (right panel). Unlike Maximin, across different  $\alpha$  values PoF shows a decreasing trend with  $q_3$ .

**Real World Experiments** The data set consists of six different social networks of homeless youth from a major city in US as described in detail in Barman-Adhikari et al. [3]. Each social network consists of 7 racial groups, namely, *Black or African American*, *Latino or Hispanic*, *White*, *American Indian or Alaska Native*, *Asian*, *Native Hawaiian or Other Pacific Islander* and *Mixed race*. Each individual belongs to a single racial

group. We use these partitioning by race to define our communities. However, to avoid misinterpretation of the results, we combine racial groups with a population  $< 10\%$  of the network size  $N$  under the “Other” category. After this pre-processing step, each dataset will contain 3 to 5 communities. The size of these communities (in percentage), as well as the sizes of the networks and their edge densities are summarized in Table 1. We remark that the absent of a racial category in a given network is due to their small sizes and hence being merged into the “Other” category after pre-processing (e.g., Hispanic in network W2SPY.)

Network Name	# of Vertices	# of Edges	White	Black	Hispanic	Mixed Race	Other
W1MFP	219	217	16.4	41.5	20.5	16.4	5.0
W2MFP	243	214	16.8	36.6	21.8	22.2	2.4
W3MFP	296	326	22.6	34.4	15.2	22.9	4.7
W2SPY	133	225	55.6	10.5	–	22.5	11.3
W3SPY	144	227	63.0	–	–	16.0	20.0
W4SPY	124	111	54.0	16.1	–	14.5	15.3

Table 1: Racial composition (%) after pre-processing as well as the number of vertices and edges of the social networks [3].

Table 2 provides a summary of the results averaged over all network instances and the numbers in bold highlight the best values (minimum utility gap and PoF) for each budget value and across different fairness approaches. IM typically has a large utility gap (up to 17.2% for  $K = 20$  which is significant because the total influence is only 28.40%). By imposing fairness we can reduce this gap. In fact, we observe that across different values of  $\alpha$  ranging from -5 to 0.5, there is a decreasing trend in utility gap, where for  $K = 20$  and with  $\alpha = -5$ , we are able to decrease the utility gap by 13.6%. Consistent with previous results on SBM networks, both Maximin and  $\alpha = -5$  exhibit very low utility gaps, however, Maximin results in higher PoF. Furthermore, across the range of  $\alpha$  we observe a mild trade-off between fairness and utility.

Measure (%)	$K$	Fairness Approach							
		$\alpha = -5$	$\alpha = -2$	$\alpha = 0$	$\alpha = 0.5$	$\alpha = 0.9$	DC	Maximin	IM
utility gap	5	4.8	7.5	8.5	9.7	11.4	8.2	<b>3.5</b>	12.5
	10	4.6	6.6	7.3	9.5	12.9	6.9	<b>2.0</b>	11.7
	15	3.6	5.2	5.9	8.9	13.5	7.6	<b>2.4</b>	15.3
	20	3.6	4.4	5.9	7.3	14.0	5.8	<b>2.3</b>	17.2
	25	2.6	3.5	4.6	5.8	13.2	7.0	<b>2.0</b>	16.6
	30	2.4	3.2	4.3	6.4	8.6	8.2	<b>2.0</b>	15.7
PoF	5	6.9	5.1	3.4	1.5	<b>0.7</b>	14.4	16.6	0.0
	10	6.6	3.8	2.8	0.7	<b>0.1</b>	14.3	13.1	0.0
	15	3.8	2.5	1.6	1.1	<b>0.1</b>	14.6	10.5	0.0
	20	4.6	3.8	2.9	2.0	<b>1.0</b>	13.9	10.5	0.0
	25	4.0	3.2	2.5	1.9	<b>1.0</b>	13.4	9.9	0.0
	30	3.9	3.4	2.9	2.3	<b>1.8</b>	12.7	10.8	0.0

Table 2: Summary of the utility gap and PoF results averaged over 6 different real world social networks for various budget, fairness approaches and baselines. Numbers in bold highlight the best values in each setting (row) across different approaches.

The results for each network separately (X axis) is shown in Figure 3 for a fixed budget  $K = 30$ . Figure 3 shows that that the trade-offs can also be very network-dependent (compare e.g. W2SPY and W3MFP). This highlights the crucial need for a flexible framework that can be easily adjusted to meaningfully compare these trade-offs.

Finally, while theoretical results indicate that in the network setting, no welfare function can satisfy all principles including utility gap reduction, we observe empirically, both in synthetic and real world settings,

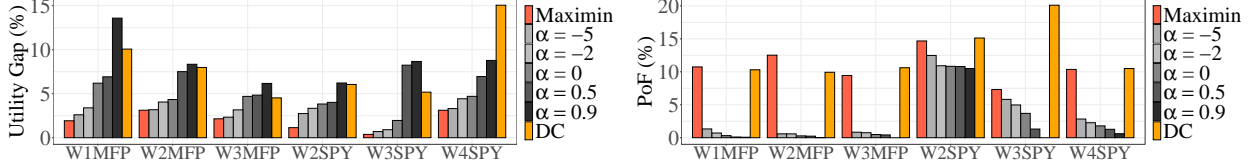


Figure 3: Left and right panels: utility gap and PoF for each real world network instances ( $K = 30$ ).

that we can systematically control the gap by choosing  $\alpha$  appropriately.

## 5 Discussion and Future Work

As the empirical evidence highlighting ethical side effects of algorithmic decision-making is growing in many different domains [2, 29, 41], the nascent field of algorithmic fairness has also seen witnessed a significant growth. It is well-established by this point that there is no universally agreed-upon notion of fairness, as fairness concerns can vary from one domain to another [5, 31]. This need for different fairness notions can also be explained by theoretical studies that show that different fairness definitions are often in conflict with each other [11, 18, 26]. To this end, most of the literature on algorithmic fairness proposes different fairness notions motivated by different ethical concerns. A major drawback of this approach is the difficulty of comparing these methods against each other in a systematic manner. Instead of following this trend, we propose a unifying framework controlled by a single parameter that can be used by a decision maker to systematically compare different fairness measures which typically result in different (and possibly also problem-dependent) trade-offs (e.g., between utility and fairness violation). Our framework also accounts for the social network structure while designing fairness notions – a consideration that is mainly overlooked in the past. Given these two contributions, it is perceivable that our approach can be used in many of the public health interventions such as HIV or Tuberculosis prevention that rely on social networks. This way, the decisions makers, such as social workers, can compare a menu of fairness-utility trade-offs proposed by our approach and decide which one of these trade-offs are more desirable for their domain without a need to understand the underlying mathematical details that are used in deriving these trade-offs.

There are crucial considerations when deploying our system in practice. First, cardinal welfare is one particular way of formalizing fairness considerations. This by no means implies that other approaches for fairness e.g. equality enforcing interventions should be completely ignored specially in domains that pure equality is more desirable. Second, we have assumed that the decision maker has the full knowledge of the network structure as well as possibly protected attributes of the individuals which can be used to define communities. Third, while our experimental evaluation is based on utilizing a greedy algorithm to solve our optimization problem, it is conceivable that this greedy approximation can create complications by imposing undesirable biases that we have not accounted for. Intuitively (and as we have seen in our experiments) the extreme of inequity aversion ( $\alpha \rightarrow -\infty$ ) can be used as a proxy for pure equality. However, the second and third concerns require more care and we leave the study of such questions as interesting future work directions.

**Acknowledgement** We would like to thank David Gray Grant, Matthew Joseph, Andrew Perrault and Bryan Wilder for helpful discussions about this work. This work was supported in part by the Smart & Connected Communities program of the National Science Foundation under NSF award No. 1831770 and by the US Army Research Office under grant number W911NF1710445.

## References

- [1] Junaid Ali, Mahmoudreza Babaei, Abhijnan Chakraborty, Baharan Mirzasoleiman, Krishna Gummadi, and Adish Singla. On the fairness of time-critical influence maximization in social networks. *arXiv*, abs/1905.06618, 2019.
- [2] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. Machine bias. *Propublica*, 2016.
- [3] Anamika Barman-Adhikari, Stephanie Begun, Eric Rice, Amanda Yoshioka-Maxwell, and Andrea Perez-Portillo. Sociometric network structure and its association with methamphetamine use norms among homeless youth. *Social Science Research*, 58:292–308, 2016.
- [4] Nawal Benabbou, Mithun Chakraborty, Xuan-Vinh Ho, Jakub Sliwinski, and Yair Zick. Diversity constraints in public housing allocation. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 973–981, 2018.
- [5] Richard Berk, Hoda Heidari, Shahin Jabbari, Matthew Joseph, Michael Kearns, Jamie Morgenstern, Seth Neel, and Aaron Roth. A convex framework for fair regression. In *4th Workshop on Fairness, Accountability, and Transparency in Machine Learning*, 2017.
- [6] Richard Berk, Hoda Heidari, Shahin Jabbari, Michael Kearns, and Aaron Roth. Fairness in criminal justice risk assessments: The state of the art. *Sociological Methods & Research*, 2018.
- [7] Dimitris Bertsimas, Vivek Farias, and Nikolaos Trichakis. The price of fairness. *Operations Research*, 59(1):17–31, 2011.
- [8] Arpita Biswas and Siddharth Barman. Fair division under cardinality constraints. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 91–97, 2018.
- [9] Thomas Bonald and Laurent Massoulié. Impact of fairness on internet performance. In *Proceedings of the Joint International Conference on Measurements and Modeling of Computer Systems*, pages 82–91, 2001.
- [10] Toon Calders, Faisal Kamiran, and Mykola Pechenizkiy. Building classifiers with independency constraints. In *Proceeding of the 9th International Conference on Data Mining Workshops*, pages 13–18, 2009.
- [11] Alexandra Chouldechova. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. In *3rd Workshop on Fairness, Accountability, and Transparency in Machine Learning*, 2017.
- [12] Vincent Conitzer, Rupert Freeman, Nisarg Shah, and Jennifer W. Vaughan. Group fairness for the allocation of indivisible goods. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, 2019.
- [13] Hugh Dalton. The Measurement of the Inequality of Incomes. *The Economic Journal*, 30(119):348–361, 1920.
- [14] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science*, pages 214–226, 2012.
- [15] Hadi Elzayn, Shahin Jabbari, Christopher Jung, Michael Kearns, Seth Neel, Aaron Roth, and Zachary Schutzman. Fair algorithms for learning in allocation problems. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, pages 170–179, 2019.
- [16] Stephen Fienberg and Stanley Wasserman. Categorical data analysis of single sociometric relations. *Sociological Methodology*, 12:156–192, 1981.

- [17] Benjamin Fish, Ashkan Bashardoust, Danah Boyd, Sorelle Friedler, Carlos Scheidegger, and Suresh Venkatasubramanian. Gaps in information access in social networks? In *Proceedings of the 28th World Wide Web Conference*, pages 480–490, 2019.
- [18] Sorelle Friedler, Carlos Scheidegger, and Suresh Venkatasubramanian. On the (im)possibility of fairness. *ArXiv*, abs/1609.07236, 2016.
- [19] William Gorman. The structure of utility functions. *The Review of Economic Studies*, 35(4):367–390, 1968.
- [20] Moritz Hardt, Eric Price, and Nati Srebro. Equality of opportunity in supervised learning. In *Advances in neural information processing systems*, pages 3315–3323, 2016.
- [21] John Harsanyi. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, 63(4):309–321, 1955.
- [22] Hoda Heidari, Claudio Ferrari, Krishna Gummadi, and Andreas Krause. Fairness behind a veil of ignorance: A welfare analysis for automated decision making. In *Advances in Neural Information Processing Systems 31*, pages 1265–1276, 2018.
- [23] Michael Isaac, Brenda Elias, Laurence Katz, Shay-Lee Belik, Frank Deane, Murray Enns, Jitender Sareen, and Swampy Cree Suicide Prevention Team. Gatekeeper training as a preventative intervention for suicide: a systematic review. *The Canadian Journal of Psychiatry*, 54(4):260–268, 2009.
- [24] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 137–146, 2003.
- [25] Jon Kleinberg, Yuval Rabani, and Éva Tardos. Fairness in routing and load balancing. In *40th Annual Symposium on Foundations of Computer Science*, pages 568–578, 1999.
- [26] Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. Inherent trade-offs in the fair determination of risk scores. In *Proceedings of the 8th Conference on Innovations in Theoretical Computer Science*, pages 43:1–43:23, 2017.
- [27] Amanda Kube, Sanmay Das, and Patrick Fowler. Allocating interventions based on predicted outcomes: A case study on homelessness services. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019.
- [28] Hui Lin and Jeff Bilmes. A class of submodular functions for document summarization. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 510–520, 2011.
- [29] Clair Miller. Can an algorithm hire better than a human? *The New York Times*, June 25 2015. URL <http://www.nytimes.com/2015/06/26/upshot/can-an-algorithm-hire-better-than-a-human.html/>. Retrieved 4/28/2016.
- [30] Herve Moulin. *Fair Division and Collective Welfare*. The MIT Press, 2004.
- [31] Arvind Narayanan. 21 fairness definitions and their politics. 2018.
- [32] George Nemhauser and Laurence Wolsey. Maximizing submodular set functions: Formulations and analysis of algorithms. 59(11):279 – 301, 1981.
- [33] Commission on Social Determinants of Health. Closing the gap in a generation: Health equity through action on the social determinants of health, 2008.
- [34] Derek Parfit. Equality and priority. *Ratio*, 10(3):202–221, 1997.

- [35] Arthur Pigou. *Wealth and welfare*. Macmillan and Company, 1912.
- [36] Aida Rahmattalabi, Phebe Vayanos, Anthony Fulginiti, Eric Rice, Bryan Wilder, Amulya Yadav, and Milind Tambe. Exploring algorithmic fairness in robust graph covering problems. In *Proceedings of the Advances in Neural Information Processing Systems 32*, pages 15750–15761, 2019.
- [37] John Rawls. *A theory of justice*. Harvard University Press, 2009.
- [38] Ana-Andreea Stoica, Jessy Xinyi Han, and Augustin Chaintreau. Seeding network influence in biased networks and the benefits of diversity. In *Proceedings of The Web Conference 2020*, pages 2089–2098, 2020.
- [39] Alan Tsang, Bryan Wilder, Eric Rice, Milind Tambe, and Yair Zick. Group-fairness in influence maximization. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 5997–6005, 2019.
- [40] Thomas Valente, Anamara Ritt-Olson, Alan Stacy, Jennifer Unger, Janet Okamoto, and Steve Sussman. Peer acceleration: effects of a social network tailored substance abuse prevention program among high-risk adolescents. *Addiction*, 102(11):1804–1815, 2007.
- [41] Benjamin Wilson, Judy Hoffman, and Jamie Morgenstern. Predictive inequity in object detection. *ArXiv*, abs/1902.11097, 2019.
- [42] Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez-Rodriguez, and Krishna Gummadi. Fairness constraints: Mechanisms for fair classification. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, pages 962–970, 2017.

## A Omitted Details from Section 3

We start by providing the proofs for Propositions 1, 2 and 3.

*Proof of Proposition 1.* Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a strictly concave function. Let  $\mathbf{u} = \mathbf{u}(\mathcal{A})$  and  $\mathbf{u}' = \mathbf{u}(\mathcal{A}')$  denote the utility vectors from allocations  $\mathcal{A}$  and  $\mathcal{A}'$ . Suppose we sort  $\mathbf{u}'$  and permute entries of  $\mathbf{u}$  so that, for all  $c \in \mathcal{C}$ , index  $c$  in both vectors corresponds to the same community. Furthermore, assume  $\sum_{\kappa \in \mathcal{C}: \kappa \leq c} N_{\kappa}(u_{\kappa} - u'_{\kappa}) \geq 0$ ,  $\forall c \in \mathcal{C}$  and  $u_c > u'_c$  for some  $c \in \mathcal{C}$ . Clearly  $\mathbf{u}$  and  $\mathbf{u}'$  satisfy the assumptions of the transfer principle. We need to show that  $\sum_{c \in \mathcal{C}} N_c f(u_c) > \sum_{c \in \mathcal{C}} N_c f(u'_c)$  or  $\sum_{c \in \mathcal{C}} N_c (f(u_c) - f(u'_c)) > 0$ .

The proof is by induction. We iteratively sweep the vectors  $\mathbf{u}$  and  $\mathbf{u}'$  from the smallest index to the largest and show that for any value  $\kappa \in \mathcal{C}$ ,  $\sum_{c \leq \kappa} N_c (f(u_c) - f(u'_c)) \geq 0$  with inequality becoming strict for at least one  $\kappa$ . To do so we repeatedly use a property of strictly concave functions known as decreasing marginal returns. According to this property  $f(x + \delta_x) - f(x) > f(y + \delta_y) - f(y)$  for  $x < y$  and  $\delta_x \geq \delta_y > 0$ .

More specifically, in our inductive step, we keep track of a “decrement budget” which we denote by  $\Delta$ . Intuitively if we can show that  $\sum_{c \leq \kappa} N_c (f(u_c) - f(u'_c)) > 0$  with budget  $\Delta$  for some  $\kappa$ , we can then use the decreasing marginal return property along with the assumption that  $\mathbf{u}'$  is sorted to show that as long as  $N_{\kappa+1}(u'_{\kappa+1} - u_{\kappa+1}) \leq \Delta$  it is the case that  $\sum_{c \leq \kappa+1} N_c (f(u_c) - f(u'_c)) > 0$ . After each round we update the  $\Delta$  and move on to the next element in the utility vectors.

Formally, let  $\Delta = 0$  to start at the beginning of this inductive process. After visiting the  $c$ th community, we simply update  $\Delta$  by  $\Delta \leftarrow \Delta + N_c(u_c - u'_c)$ . By the assumption of the transfer principle  $\Delta$  is non-negative at all points of this iterative process and is strictly positive at some point during the process. Observe that  $f(u_1) \geq f(u'_1)$  since  $u_1 \geq u'_1$  by the assumption of the transfer principle. We can use this as the base case. Since  $\mathbf{u}'$  is sorted, given that  $\Delta$  is non-negative, the fact that  $f$  is strictly concave (so that the decreasing marginal return property can be used) immediately implies that  $\sum_{c \leq \kappa} N_c (f(u_c) - f(u'_c)) \geq 0$  at any iteration  $\kappa$  of the process. The inequality becomes strict for some  $\kappa$  given the assumption of the transfer principle. This proves the claim.  $\square$

*Proof of Proposition 2.* Let  $\mathcal{A}$  and  $\mathcal{A}'$  denote two neighboring allocations with corresponding utility vectors  $\mathbf{u} = \mathbf{u}(\mathcal{A})$  and  $\mathbf{u}' = \mathbf{u}(\mathcal{A}')$ . Let  $\mathbf{u}$  denote any of the two utility vectors such that  $\sum_{c \in \mathcal{C}} N_c u_c \geq \sum_{c \in \mathcal{C}} N_c u'_c$ . Without loss of generality, we assume  $\mathbf{u}'$  is sorted in ascending order of the utilities and  $\mathbf{u}$  is permuted so that index  $c \in \mathcal{C}$  in both  $\mathbf{u}$  and  $\mathbf{u}'$  corresponds to the same group. This is because we assume that  $W$  satisfies the symmetry principle due to which by permuting a utility vector the value of the welfare function does not change. Let  $\nu$  and  $\kappa \in \mathcal{C}$  denote the communities whose utilities are changed between  $\mathbf{u}$  and  $\mathbf{u}'$ , i.e., we assume  $\nu$  and  $\kappa$  are the two communities where taking resources from  $\nu$  and giving them to  $\kappa$  will transfer  $\mathbf{u}'$  into  $\mathbf{u}$ .

To satisfy the condition of the utility gap reduction principle, it should be the case that  $u'_\nu \geq u'_\kappa$  (i.e., we transfer resources from the group with higher utility to a group with lower utility), otherwise after the transfer from  $\mathbf{u}'$  to  $\mathbf{u}$  the utility gap could not get smaller (i.e.,  $\Delta(\mathbf{u}) \geq \Delta(\mathbf{u}')$  in which case the utility gap reduction is not applicable).

Assuming  $u'_\nu \geq u'_\kappa$ , if  $\Delta(\mathbf{u}) \geq \Delta(\mathbf{u}')$ , again the assumption of the utility gap reduction principle is not satisfied, hence the principle is not applicable and there is no need to study this case. Therefore, we further assume  $\Delta(\mathbf{u}) < \Delta(\mathbf{u}')$ . We would like to show in this case a welfare function  $W$  that satisfies all the 5 other desiderata witnesses  $W(\mathbf{u}) > W(\mathbf{u}')$ .

By assumption  $\sum_{c \in \mathcal{C}} N_c u_c \geq \sum_{c \in \mathcal{C}} N_c u'_c$ . From this, it follows that:

$$\sum_{c \in \mathcal{C}} N_c (u_c - u'_c) \geq 0 \Leftrightarrow N_\nu (u_\nu - u'_\nu) + N_\kappa (u_\kappa - u'_\kappa) \geq 0 \quad (1)$$

$$\Leftrightarrow \sum_{y \in \mathcal{C}: y \leq x} N_y (u_y - u'_y) \geq 0, \forall x \in \mathcal{C}, \quad (2)$$

where both inequalities (1) and (2) follow directly from the fact that the utilities of all the other communities are the same in both  $\mathbf{u}$  and  $\mathbf{u}'$ . Finally, since  $u_\kappa > u'_\kappa$  (we are transferring resources to the community  $\kappa$ ), we can apply the generalized transfer principle to show that  $W(\mathbf{u}) > W(\mathbf{u}')$  as claimed.  $\square$

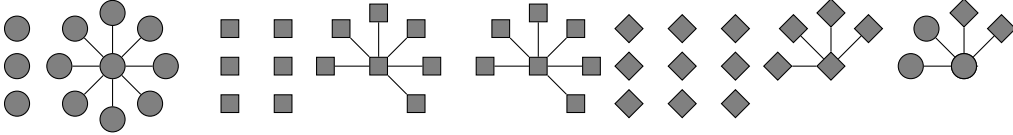


Figure 4: An illustration for the graph used in the proof of Proposition 3 without the correct scaling. There are three communities (circle, square and diamond) and they all have size 100. The circle community consists of an “all circle” star structure with 80 vertices, 14 isolated vertices and a mixed star structure (shared with the diamond community) with 6 circle vertices. The square community consists of two “all square” star structures with sizes 60 and 10 plus a set of 30 isolated vertices. The diamond community consists of an “all diamond” star structure with 30 vertices, 66 isolated vertices and a mixed star structure (shared with the circle community) with 4 diamond vertices.

*Proof of Proposition 3.* Figure 4 is an illustration of the graph that is used in the proof to witness the statement. We set  $p = 1$  (deterministic spread) and number of initial seeds  $K = 4$ . Consider two allocations  $\mathcal{A}$  and  $\mathcal{A}'$ . Let  $\mathcal{A}$  denote the allocation that consists of the center of all star structures that consist of a single community. Let  $\mathcal{A}'$  denote the allocation that is identical to  $\mathcal{A}$  with the sole difference that only the center of one of the all square structures is chosen and the last seed is selected to be the center of the star structure that is the mix of circle and diamond communities. Clearly these two allocations are neighboring. The average utilities for these allocations are (diamond = 0.3, square = 0.7, circle = 0.8) in  $\mathbf{u}$  and (diamond = 0.34, square = 0.6, circle = 0.86) in  $\mathbf{u}'$ , respectively. Both allocations correspond to a total utility of 180 but the utility gap is  $\Delta(\mathbf{u}) = 0.5$  for  $\mathbf{u}$  as opposed to the utility gap of  $\Delta(\mathbf{u}') = 0.52$  for  $\mathbf{u}'$ . So a welfare function that obeys the utility gap reduction should prefer  $\mathbf{u}$  over  $\mathbf{u}'$ .

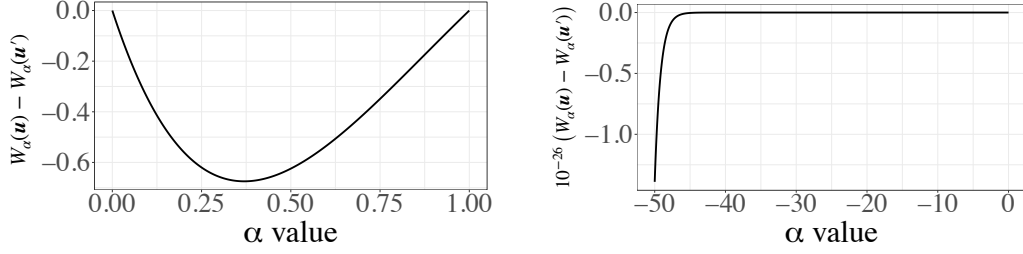


Figure 5: The difference of  $W_\alpha(\mathbf{u}) - W_\alpha(\mathbf{u}')$  on the vertical axis versus  $\alpha$  on the horizontal axis for different welfare functions (this difference is scaled by a factor of  $10^{-26}$  on the right panel). Left panel:  $W_\alpha(\mathbf{u}) = \sum_{c \in \mathcal{C}} N_c u_c^\alpha$  for  $\alpha \in (0, 1)$ ; right panel:  $W_\alpha(\mathbf{u}) = -\sum_{c \in \mathcal{C}} N_c u_c^\alpha$  for  $\alpha < 0$ .

We now show that no welfare function that satisfies the first 5 desiderata will prefer  $\mathbf{u}$  over  $\mathbf{u}'$ . Recall that such welfare functions include  $W_\alpha(\mathbf{u}) = \sum_{c \in \mathcal{C}} N_c u_c^\alpha$  for  $\alpha \in (0, 1)$ ,  $W_\alpha(\mathbf{u}) = \sum_{c \in \mathcal{C}} N_c \ln(u_c)$  for  $\alpha = 0$  and  $W_\alpha(\mathbf{u}) = -\sum_{c \in \mathcal{C}} N_c u_c^\alpha$  for  $\alpha < 0$ . We verify this claim numerically. In particular Figure 5 plots the difference of  $W_\alpha(\mathbf{u}) - W_\alpha(\mathbf{u}')$  for  $W_\alpha(\mathbf{u}) = \sum_{c \in \mathcal{C}} N_c u_c^\alpha$  when  $\alpha \in (0, 1)$  (left panel) and  $W_\alpha(\mathbf{u}) = -\sum_{c \in \mathcal{C}} N_c u_c^\alpha$  when  $\alpha < 0$  (right panel). This difference is always negative so  $\mathbf{u}'$  is preferred by these welfare functions. For  $W_\alpha(\mathbf{u}) = \sum_{c \in \mathcal{C}} N_c \ln(u_c)$ ,  $W_\alpha(\mathbf{u}) - W_\alpha(\mathbf{u}') \approx -4.3$ .

We point out that the instance used in the proof (graph structure, probability of spread and the number of seeds) used in this proof is designed with the sole purpose of simplifying the calculations of the utilities. It is possible to modify this instance to more complicated and realistic instances.  $\square$

Next, we study equality-based notions of fairness for influence maximization [1, 38]. In particular, we show through an example that the fair allocation violates the monotonicity principle (a condition known as levelling down in political philosophy [34]).

**Equality Enforcing Influence Maximization and Levelling Down** We start by introducing the most commonly used equality-based notion of fairness known as demographic parity or statistical parity.

**Demographic Parity** Demographic parity requires that the utility of each community to be roughly the same. This notion formalizes the legal doctrine of disparate impact [42]. Let  $\delta \in [0, 1)$ . Then the demographic parity implies the following constraint on the influence maximization problem.

$$|u_\nu(\mathcal{A}) - u_\kappa(\mathcal{A})| \leq \delta, \quad \forall \mathcal{A} \in \mathcal{A}^* \quad \nu, \kappa \in \mathcal{C}.$$

We use  $\mathcal{A}^{\text{DP}}$  to denote the above set of constraints.

The parameter  $\delta$  in the definition captures the amount of relaxation that is permitted by the demographic parity from strict equality and larger  $\delta$  values correspond to more relaxation (equivalently, less strictness). We next show that demographic parity does not satisfy monotonicity.

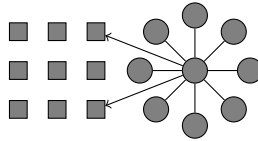


Figure 6: Companion figure to Proposition 4 where we choose  $\max(3p/(p - \delta), (p - 4p^2)/(\delta - p^2)) < N$ ,  $\delta < p < \sqrt{\delta}$ ,  $K = 2$ . The network consists of two communities circle and square, each of size  $N$ . All edges except the two shown by arrows are undirected meaning that influence can spread both ways.



**Proposition 4.** *For all  $\delta \in [0, 1)$ , there exists an instance of the influence maximization problem such that for any fair allocation  $\mathcal{A}$  satisfying demographic parity, there exists an allocation  $\mathcal{A}'$  that Pareto dominates  $\mathcal{A}$ . Mathematically, this can be written as:  $\forall \mathcal{A} \in \mathcal{A}^{\text{DP}}, \exists \mathcal{A}' \in \mathcal{A}^* : \mathbf{u}(\mathcal{A}') > \mathbf{u}(\mathcal{A})$ .*

*Proof.* Consider a graph  $\mathcal{G}$  as shown in Figure 6 consisting of two communities, square and circle, each of size  $N$ . We choose an arbitrary  $\delta \in [0, 1)$ , to reflect the arbitrary strictness of a decision maker. Let  $\delta < p < \sqrt{\delta}$ ,  $K = 2$  and  $N$  be large enough, i.e.,  $N > \max(3p/(p - \delta), (p - 4p^2)/(\delta - p^2))$ . The optimal solution of the influence maximization problem chooses the center of the star and any disconnected square vertex. In the optimal solution, the utility of circle and square communities are  $(1 + (N - 1)p)/N$  and  $(1 + 2p)/N$ , respectively and the utility gap exceeds  $\delta$  (so this solution does not satisfy demographic parity constraints). By imposing demographic parity, any fair solution is to choose one vertex from the periphery of the circle community and one from the isolated square vertices. The utilities of circle and square are  $(1 + p + p^2(N - 2))/N$  and  $(1 + 2p^2)/N$ , respectively. Given the range of  $N$ , the utility gap is less than  $\delta$ . It is easy to verify the utility of both communities have degraded so the fair solution does not satisfy the monotonicity.

We point out that the graph used in the proof is directed. This is for ease of exposition. It is possible to create a more complex example with an undirected graph.  $\square$