On the Information Unfairness of Social Networks

Zeinab S. Jalali* Weixiang Wang* Myunghwan Kim† Hema Raghavan‡
Sucheta Soundarajan*

Abstract

Social networks play a vital role in the spread of information through a population, and individuals in networks make important life decisions on the basis of the information to which they have access. In many cases, it is important to evaluate whether information is spreading *fairly* to all groups in a network. For instance, are male and female students equally likely to hear about a new scholarship? In this paper, we present the *information unfairness* criterion, which measures whether information spreads fairly to all groups in a network. We perform a thorough case study on the **DBLP** computer science co-authorship network with respect to gender. We then propose `MaxFair`, an algorithm to add edges to a network to decrease information unfairness, and evaluate on several real-world network datasets.

1 Introduction

Social networks play a critical role in society, and individuals in networks often make decisions on the basis of interactions and communications with other individuals in the network [4]. Access to information flowing through a network plays a role in many important aspects of a person's life: for example, job-seekers and students may hear of employment or scholarship opportunities by 'networking', and scientists learn about research ideas through their professional connections.

However, in many social networks, the topology of the network may be such that various groups (as defined by demographic features or other sensitive properties) have unequal access to information. This is of particular concern when the information of interest has the potential to affect the life trajectory of an individual, such as knowledge of employment opportunities. For example, in professional networks, central positions in a company (e.g., executive positions) are often occupied by white men, while women and minorities are on the fringes [10]. If news of promotion opportunities tends to flow disproportionately to white men, then this further consolidates that group's power and disadvantages other groups. This has been observed in real settings: for example, students from poor backgrounds may be unaware of options for attending selective colleges [3].

It is thus valuable to understand whether information flows equally to all groups. We are primarily interested in groups that are defined by *sensitive* attributes, such as race or gender. If information does not flow fairly, then one can take proactive measures to distribute that information (e.g., by making special efforts to reach out to the neglected groups).

In this paper, we introduce *information unfairness*, which quantifies the extent to which groups in a network have equal access to information. We perform a case study on the **DBLP** co-authorship network, and discuss the properties that lead to unfairness. Next, we consider the problem of adding edges to a network to reduce its information unfairness, and propose `MaxFair`, a novel algorithm, for this task. Experimental results show that `MaxFair` can obtain large decreases in information unfairness when adding only a small number of edges.

The major contributions of this paper are as follows:

1. We introduce the novel information unfairness measure, which describes whether information flows equally to all groups in a network.
2. We perform a case study on the **DBLP** coauthorship network with respect to gender, in which we compare various subfields to each other.
3. We introduce `MaxFair`, an algorithm for decreasing the information unfairness of a network by adding a small number of edges, and show that it achieves significant reductions in information unfairness.

2 Related Work

Recently, the topic of algorithmic 'fairness' has attracted a great deal of attention [3]. Researchers have studied problems relating to fairness in algorithms used in the criminal justice system, hiring, credit scoring, and many other domains [3]. At a high level, the guiding intuition behind many of these methods is that individuals should not be treated differently due to their membership in a protected group. (A protected group is a group defined on the basis of a protected attribute like gender: e.g., both the group of men as well as the group of women are considered 'protected'. In contrast to the

*Syracuse University, Syracuse, NY, USA. Email: (*zsaghati, wwang69, susounda@syr.edu*)
†Mesh Korea, Korea. Email: (*mykim@cs.stanford.edu*)
‡LinkedIn, USA. Email: (*hraghavan@linkedin.com*)

Table 1: Notation

Symbol	Definition
$G(V, E)$	Unweighted, undirected attributed graph
u, v	Nodes in V
n, m	Number of nodes and edges in G
$\mathbf{M}$	Adjacency matrix of G
p	Propagation probability along each edge
k	Maximum length cascade considered
$\mathbf{S}$	$n \times n$ normalization matrix
$\mathbf{A}, \mathbf{A}'$	Accessibility & normalized accessibility matrix
$c(v)$	Categorical attribute value of node v
C_f	Attribute group
D_{fg}	Joint attribute accessibility distribution

way the term 'protected group' is sometimes used colloquially, the term does not refer only to *underprivileged* groups.) However, these works generally do not view problems from a network perspective. One exception is the work by Fish *et al.* in [8], which studies whether individuals have fair access to information in social networks. This work differs from ours in that they study access by *individuals*, not *groups*.

Our goal with the proposed *information unfairness* measure is to allow for evaluation of fairness in network structure, based on the rates of information flow between different protected groups.

The concepts of homophily and echo chambers are related to, but substantially different from, information unfairness. Homophily is a measure of the extent to which individuals tend to associate with others that are like them. In networks, it is often measured using the assortativity coefficient [14]. Homophily has been studied extensively, and is closely tied to segregation. For example, both gender and race-based homophily can be present in professional networks [9]. The negative side effects of segregation, such as reduced health outcomes [12] and increased educational inequality [16], are well-known. Homophily is related to information access: e.g., Simpson, *et al.* found that the racial homophily of African-American communities made it difficult for advertisers to reach that group [19].

Also tied closely to information unfairness is the notion of echo chambers, which occur when opinions 'echo' within a community, amplifying those beliefs to members of that community [2]. However, as we will see in Section 3.7, neither of these concepts fully capture information unfairness.

In the latter part of this paper, we present a method for adding edges to a network to decrease its information unfairness. This is related to the topic of characterizing and improving flow through a network [20]. D'Angelo *et al.* considered the problem of adding b connections to a network to maximize flow [7]. They showed that this problem is NP-hard.

3 Information Unfairness: Definitions

Information unfairness measures the extent to which different groups have different levels of exposure to a piece of information spreading through the network. While it is unlikely that *every* individual in a network will participate in an information cascade, unfairness may occur if certain groups are consistently and disproportionately excluded from access to information. To compute information unfairness, a user provides:

1. $G = (V, E)$, an undirected graph with n nodes and m edges, with specified categorical attributes and adjacency matrix $\mathbf{M}$. We are interested primarily in attributes representing sensitive properties of individuals, such as race.
2. $k \in \mathbb{N}$, the maximum length cascade considered.
3. Function $Dist(D_1, D_2)$ that computes the distance between two distributions (see Section 3.2).
4. *Optional*: $p \in [0, 1]$, describing the propagation probability along an edge.[1]
5. *Optional*: normalization matrix $\mathbf{S}$ of size $n \times n$ (see Section 3.3).

Table 1 lists the notations used in this paper.

The intuition behind the proposed formulation of information fairness is this: we wish to know whether different groups benefit differently from the topology of the graph, in terms of access to information. If, e.g., a random node in group C_1 receives more information (in expectation) from a random node in group C_3 than a random node in C_2 receives from a random node in C_3, then the graph exhibits information unfairness.[2]

To compute this, we first compute the *accessibility matrix*, which uses the topology of the graph to characterize the flow between each pair of *nodes*. Then, using attributes, we compute distributions characterizing the flow between nodes from each pair of *groups*. Finally, we calculate the distances between these distributions in order to identify whether there are differences in flow levels between nodes from different groups.

In the proposed formulation, high values of information unfairness indicate *more* unfairness. A perfectly fair network will have an information unfairness of 0.

Step 1: Constructing the Accessibility Matrix. We first construct the $n \times n$ *accessibility matrix* $\mathbf{A}$. $\mathbf{A}_{uv}$ is the amount of flow that node u is expected to receive from node v. Because G is undirected, $\mathbf{A}$ is symmetric.

To construct $\mathbf{A}$, first note that for adjacency matrix $\mathbf{M}$, $\mathbf{M}^i$ is the number of length-i walks between a pair

[1] Here, we assume the same propagation probability along each edge. In Section 3.5, we discuss how one could generalize this.

[2] Recall that we are primarily interested in groups defined by protected attributes, regardless of whether they are 'underprivileged' (e.g., both the group of men and the group of women are of interest).

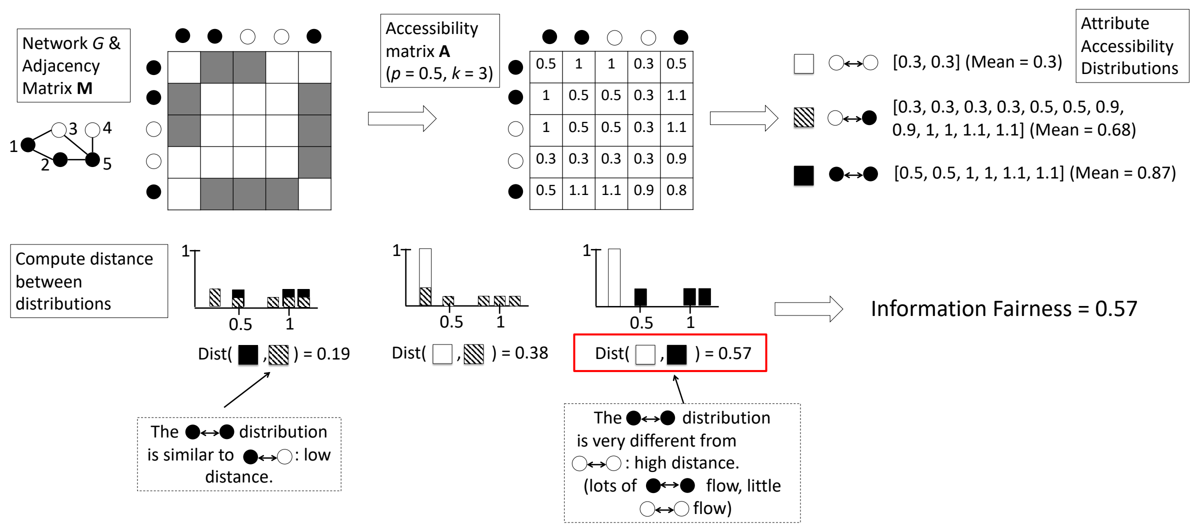


Figure 1: Overview of information unfairness computation process.

of nodes. By multiplying $\mathbf{M}^i$ by p^i, we measure the probability that information is transmitted along a walk of length i from node u to v, where p represents the probability of propagation along an edge.[3]

Thus, if p is specified by the user, we define $\mathbf{A} = p\mathbf{M} + p^2\mathbf{M}^2 + ... + p^k\mathbf{M}^k$. $\mathbf{A}_{uv}$ measures the expected number of times that node v will hear about a cascade originating at node u, using walks of length up to k [1]. If p is not specified by the user, then $\mathbf{A}$ can be found by integrating the above expression over p ranging from 0 to 1, so $\mathbf{A} = \frac{1}{2}\mathbf{M} + \frac{1}{3}\mathbf{M}^2 + ... + \frac{1}{k+1}\mathbf{M}^k$.

This particular cascade model makes finding $\mathbf{A}$ tractable. There are many other cascade models for cascades, but in general, finding the nodes that are expected to be influenced by a particular set of seeds is difficult [21]. However, if a different cascade model is more appropriate, one can define matrix $\mathbf{A}$ accordingly.

If desired, one can normalize $\mathbf{A}$ to account for various properties of the graph. Normalization is discussed in more depth in Section 3.3, but the basic idea is that we wish to compare the observed flow between each pair of nodes to what we would have expected in a random graph with some of the same properties as G. Given a normalization matrix $\mathbf{S}$, the normalized accessibility matrix $\mathbf{A}'$ is defined as the element-wise quotient of $\mathbf{A}$ with $\mathbf{S}$; that is, $\mathbf{A}'_{uv} = \mathbf{A}_{uv}/\mathbf{S}_{uv}$. If no normalization matrix $\mathbf{S}$ is provided, then $\mathbf{A}' = \mathbf{A}$.

Step 2: Characterizing Flow Between Groups. $\mathbf{A}$ allows us to characterize flow between pairs of *nodes*; but we are interested in understanding how different *groups* are affected. Suppose that each $v \in V$ has categorical attribute value $c(v) \in C = \{c_1, ..., c_t\}$. Then we define **attribute group** $C_f = \{v : c(v) = c_f : f \in \{1, .., t\}\}$, and the **joint attribute accessibility distribution** D_{fg} for attribute groups C_f and C_g is given by $\{\mathbf{A}'_{uv} : c(u) = c_f, c(v) = c_g, u \neq v\}$ [4] The joint attribute accessibility distributions characterize how well information flows between two attribute groups: each value in D_{fg} is the flow between a pair of nodes $\{(u, v) : u \in C_f, v \in C_g\}$.

Step 3: Computing the Information Unfairness of G. Given the preceding definitions and notation, the information unfairness $IU_{G,p,k}$ of a graph is given by:

$$IU_{G,p,k} = max(\{Dist(D_{f_1 g_1}, D_{f_2 g_2}) : f_1, f_2, g_1, g_2 \in \{1, .., t\}\})$$

Here, *Dist* is some function to compute the distance between two distributions. Informally, the information unfairness of a network measures the extent to which the joint accessibility distributions differ from one another. For example, does information originating at a member of a minority group have the same 'reach' as information originating at a member of the majority group? An overview of this process is given in Figure 1. For the sake of simplicity, no normalization is performed.

3.1 Example Figure 2 depicts two graphs with very different values of information fairness. For $p = 0.1, k = 3$, for each graph, we compute $\mathbf{A} = 0.1\mathbf{M} + 0.01\mathbf{M}^2 + 0.001\mathbf{M}^3$. We then identify three distributions characterizing flow between the various group pairs. We take the values from $\mathbf{A}$ to generate distributions characterizing flow between all red-red node pairs, all blue-blue node pairs, and all red-blue node pairs. Both networks have the same number of nodes and edges, so can be

[3] $\mathbf{M}^i$ and p^i refer to the i^{th} powers of $\mathbf{M}$ and p.

[4] (We exclude the elements on the diagonal because we are not interested in whether information flows from a node to itself.).

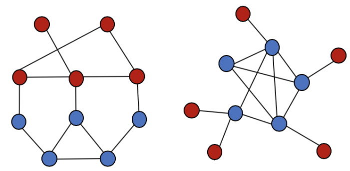


Figure 2: The left network has information unfairness of 0.04, while the right network has information unfairness of 0.1. It is clear that for the right network, very little information flows between red nodes.

compared directly without normalizing for density.

For the graph on the left, the flow between blue node pairs, flow between red node pairs, and flow between red/blue node pairs are all high: the red-red distribution has a mean of 0.06, the blue-blue distribution has a mean of 0.06, and the red-blue distribution has a mean of 0.02. These numbers represent the average number of cascades that a member of one group will receive from a member of the other group. The maximum difference between these means, and thus the information unfairness of the network, is 0.06 - 0.02 = 0.04. In contrast, on the graph on the right, there is good flow between blue nodes (mean of 0.11), good flow between red and blue nodes (mean of 0.03), and minimal flow between red nodes (mean of 0.00). This leads to an information unfairness of 0.11.[5]

It is clear that the right network should have much higher information unfairness than the left network. While information flows freely between all groups in the left network, red nodes are on the fringes of the right network, and are isolated and unable to effectively communicate with one another.

3.2 Choice of Distance Function In this work, we define the distance between two distributions as the distance between their means. Naturally, summarizing a distribution using its mean can hide important information; thus, we also considered the Earth Mover's Distance (EMD), but on the networks considered in this paper, we saw similar results. The distance between means is much faster to calculate, but we recommend that before selecting a distance function, one should examine the distributions under question and see whether a more sophisticated distance calculation (such as EMD) is necessary. (Note that the popular K-L Divergence does not take the distance between values into account- for example, the divergence between [2, 2, 2, ...] and [3, 3, 3, ...] is equal to the divergence between [2, 2, 2, ...] and [10, 10, 10, ...]- and is thus not a good choice.)

[5]These raw numbers are only meaningful in relation to one another; in Sections 3.3 and 3.4, we discuss how to normalize so that the information unfairness values are interpretable.

3.3 Normalization When computing information unfairness, the user may choose to provide a normalization matrix $\mathbf{S}$. The goal of such normalization will generally be to compare the flow between two nodes in a real network to the flow between those two nodes that would be expected in a random graph that shares desired properties with the network being studied. In such a way, it becomes possible to tease out the specific topological properties that lead to information unfairness. First, we discuss how to normalize with respect to random graphs that share the same density as G. Second, we show how to normalize with respect to random graphs that share the same degree distribution as G, allowing us assess the extent to which differences in flow are due to difference in node degrees.

Density-Based Normalization A drawback in using the values from accessibility matrix $\mathbf{A}$ directly is that in a dense graph, nodes will naturally receive more cascades than in a sparse graph, and so the unnormalized information unfairness values will be higher in the dense graph.

Here, we describe how to normalize so that it is possible to compare graphs of different densities. Define matrix $\mathbf{M}^{rand}$ so that each element of $\mathbf{M}^{rand}$ is $2m/n^2$, where m represents the number of edges in the graph and n represents the number of nodes (so $2m/n^2$ is simply the density of the graph). $\mathbf{M}^{rand}$ can be thought of as the average of all adjacency matrices of random graphs with the same number of nodes and edges as G, but is produced without actually generating any of those random graphs. From matrix $\mathbf{M}^{rand}$, one can generate random accessibility matrix $\mathbf{A}^{rand}$ in the same way that $\mathbf{A}$ was generated from $\mathbf{M}$. Then by defining $\mathbf{S} = \mathbf{A}^{rand}$, each value $\mathbf{A}_{uv}/\mathbf{A}^{rand}_{uv}$ tells us the number of cascades received by node u from node v, as compared to what one would expect in a random graph of the same density.

Degree-Based Normalization. In some cases, it may be useful to understand whether differences in accessibility are due solely to the number of connections (degrees) of members of the attribute groups, as opposed to the positioning of those members within the network. For example, in a network describing interactions within the computer science research community, it is possible that the highest degree individuals are disproportionately men (because high degree individuals are more likely to be senior researchers, and until fairly recently, there were many fewer women in computer science).[6]

We define matrix $\mathbf{M}^{deg}$ so that each element of $\mathbf{M}^{deg}$ is $d_u d_v/2m$, where d_u and d_v represent the degrees

[6]We are not making a *normative* claim about whether or not information unfairness due to degree differences should be considered acceptable; we are only seeking to explain the topological causes of information unfairness.

of nodes u and v and m represents the number of edges in the graph.[7] $\mathbf{M}^{deg}$ can be thought of as the average of the adjacency matrices of random graphs with the same degree distribution as G (this also normalizes with respect to density). Then as before, we can define $\mathbf{S}$ as the accessibility matrix obtained from $\mathbf{M}^{deg}$.

This normalization tells us the extent to which the number of cascades transmitted between two nodes depends on the degrees of those two nodes, as opposed to their position in the network.

3.4 Interpretation Higher values of information unfairness indicate that the network is less fair. The interpretation of an information unfairness value is easiest when some normalization with respect to a null model is performed (e.g., the density- or degree-based normalizations above). Such normalization allows comparison of graphs of different sizes, and information unfairness is not affected by differences in group sizes.

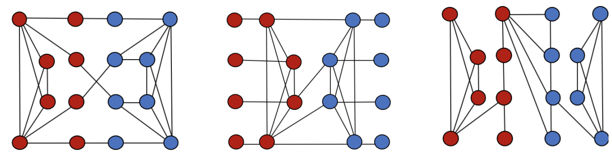
A joint attribute accessibility distribution describes how well information flows from members of one attribute group to members of the other. If flow between the groups is high, the values in the distribution are high; and vice versa. Information unfairness deals with the distance between the most different pairs of attribute group pairs. For instance, a very segregated network might see good flow within group C_f (i.e., group C_f to group C_f), moderate flow within group C_g (i.e., group C_g to group C_g), and weak flow between groups C_f and C_g. The greatest distance occurs when comparing the C_f-C_f flow to the C_f-C_g flow.

When we normalize the accessibility matrix with respect to some null model (e.g., $\mathbf{M}^{rand}$ and $\mathbf{M}^{deg}$), we divide each value of $\mathbf{A}_{uv}$ by the value that we would have expected in a random graph.[8] Thus, value $\mathbf{A'}_{uv}$ in the normalized accessibility matrix tells us the ratio of the actual amount of flow between u and v, versus the amount we would have expected in a random graph with the same high-level properties.

In a fair network, the flow between nodes u and v need not necessarily be what is expected in a random graph: what matters is that nodes in each group have been benefited or harmed by the topology equally. For instance, if flow between group C_f and group C_g is 50% lower than what one would expect in a random graph, then in order for the network to be fair, flow between groups C_g and C_h should *also* be 50% lower than expected. The distance between two distributions can thus be viewed as being relative to the null model.

[7] This is the same normalization that is famously used by the modularity metric for community quality [15].

[8] This description is a slight oversimplification. The processes described construct a single normalization matrix from the average of random adjacency matrices, as opposed to averaging the normalization matrices corresponding to all possible random adjacency matrices. The latter method would be preferable, but is not tractable for large networks.



(a) Graph(a) (b) Graph(b) (c) Graph(c)

Figure 3: Graphs with same echo chamber or assortativity values, but different information unfairness scores.

3.5 Variations on Information Unfairness It is easy to modify information unfairness for other settings.

Directed Graphs: By considering two versions of information unfairness- *transmitter unfairness* and *receiver unfairness*- we can generalize information unfairness to directed graphs. Depending on the setting, one or the other of these might be preferable: for example, a female employee in a professional network might want to know that her ideas will not be unheard (transmitter unfairness), while a minority student in a college social network may want to be confident that he will not miss out on scholarship opportunities (receiver unfairness).

Thresholded Accessibility Matrix: Elements in $\mathbf{A}$ can be greater than 1, indicating that a node is expected to receive multiple cascades originating from another node. In various real-world settings, the number of cascades from one node to another may be irrelevant: what matters is the existence of such a cascade. To account for this, one can truncate the accessibility matrix by replacing all values greater than 1 with 1.

Varying Edge Propagation Probabilities: In many social networks, edges may have different values of p, corresponding, e.g., to different levels of communication. While it is hard to know these values in practice, if they *are* known, it is easy to modify the information unfairness computation for this case by setting the elements of adjacency matrix $\mathbf{M}$ to these probabilities.

3.6 Measuring Strength of Echo Chambers Although information unfairness is not intended to measure the strength of echo chambers in a network, it is possible to use the accessibility matrix to do so. Note that when computing the accessibility distributions, we excluded elements from the diagonal of the accessibility matrix, because we didn't care about whether a node transmitted information to itself. However, by averaging these diagonal values, we estimate how information from a node comes back to that node, thus measuring the strength of echo chambers in the network.

3.7 Information Unfairness vs. Assortativity and Echo Chambers The information unfairness of

Table 2: Dataset statistics

Name	Description	#nodes	#edges	Assort	Attributes
Parallel-Full	Par. Proc. network	14316	44929	0.05	Men (82%), Women (18%)
Graphics-Full	Graphics network	12846	38320	0.07	Men (76%), Women (24%)
Security-Full	Security network	5135	15089	0.01	Men (79%), Women (21%)
Databases-Full	Databases network	17559	69731	0.09	Men (71%), Women (29%)
Data Mining-Full	Data mining network	13358	40128	0.08	Men (74%), Women (26%)
Parallel-2015	Par. Proc. network	1251	4356	0.06	Men (82%), Women (18%)
Graphics-2015	Graphics network	3525	10399	0.06	Men (71%), Women (29%)
Security-2015	Security network	1962	5976	0.07	Men (80%), Women (20%)
Databases-2015	Databases network	3185	9386	0.10	Men (68%), Women (32%)
Data Mining-2015	Data mining network	2272	7643	0.08	Men (66%), Women (34%)
Enron	Enron e-mail	144	1344	0.03	Men (76%), Women (24%)
Norway Directorate	BoD co-serving	1421	3855	-0.15	Men (63%), Women (37%)

a network is related to two other important network properties: assortativity (a measure of homophily) and echo chambers. In Figure 3, Graphs (b) and (c) have the same assortativity (0.66), and Graphs (a) and (b) have the same echo chamber properties (0.60). However, Graph (b) has significantly lower information unfairness (0.19) than the others (0.45 and 0.32 for (a) and (c), respectively, for $p = 0.3$ and $k = 4$).In all graphs, information flows easily between nodes in the same group, but it is easier for information to flow from a blue node to a red node Graph (b).

4 Datasets

We use three categories of datasets, representing different professional contexts where access to information is especially important. First, we examine subfield networks from the **DBLP** computer science co-authorship network. To identify these subfields, we extract the papers published in the top three conferences from each subfield[9] in each subfield, using ArnetMiner. We generate two versions of each subfield: using data from 2000-2019 ('Full'), and using data from 2015-2019 ('2015'). Second, we study the **Enron** e-mail network [18]. For **DBLP** and **Enron**, we infer gender from names using the Genderize.io library. (Previous work has found that Genderize.io has a fairly high accuracy of around 0.82, though results vary depending on a name's country of origin [11].) Third, we use the Norwegian Interlocking Directorate datasets (**Norway**), which includes gender [17]. Statistics are shown in Table 2.

5 Case Study: Information Unfairness of Co-Authorship Networks

In recent years, there has been a huge push to increase the representation of women in Computer Science (CS) and tech fields [6]. Our goal with this case study is to further investigate gender inequality in the CS collaboration network. We analyze the 2015-2019 **DBLP** subfield co-authorship networks with respect to gender.[10] Although these subfields are of different sizes, normalization allows us to compare across datasets. Figure 4 depicts results for $k \in \{2, 6\}$ and $p \in \{0.1, 0.3, 0.5, 0.9\}$ (results for $k \in \{4, 10\}$ were similar).[11]

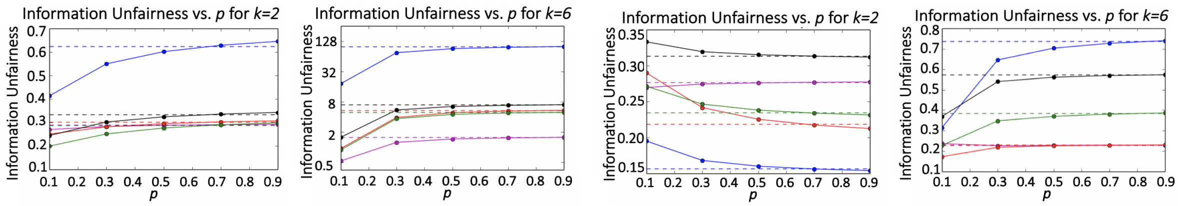
Observation 1: None of the subfields are perfectly fair. For both types of normalization, all networks exhibit non-zero information unfairness. However, the reasons for this unfairness vary. We explored the results on each subfield in more detail to determine whether the male-male, female-female, or male-female flow was lowest. On the **Parallel Processing** network, for $k = 4, p = 0.5$ with density-based normalization, the male-male, female-male, and female-female distributions have means, respectively, of 27.0, 10.8, and 7.5. Women in general do not receive very much information on this network, whether from men or from other women. On the **Graphics** network, with the same parameters, female-female flow is the *highest*, with a mean of 4.9. Both male-male and male-female distributions have a mean close to 3.5. The same general pattern is observed on **Data Mining**. On the **Database** and **Security** networks, the male-male and female-female means are close to each other, but the male-female flow is the lowest, indicating segregation between the two groups.

Observation 2: As k and p increase, information unfairness tends to increase. This is because information unfairness uses the *total number* of cascades between two nodes. If a network is homophilic, then for

[9] `https://webdocs.cs.ualberta.ca/~zaiane/htmldocs/ConfRanking.html`

[10] We restrict to 2015-2019 because we wish to investigate the current state of the subfields.

[11] We only consider k up to 10 because long cascades are uncommon in practice [13].



(a) Density-normalized results for varying k and p. (b) Degree-normalized results for varying k and p.

Figure 4: Information unfairness results for **DBLP** subfields. (Dashed lines show information unfairness value for unspecified p (i.e., when information unfairness is computed by taking the integral over all $p \in [0, 1]$).

very small k, most cascades will be to other members of the same group; however, as k increases, while cascades are now able to reach members of other groups, the number of walks within the immediate neighborhood grows *exponentially*, and so the difference between two distributions increases.[12] For small p, even for large k, cascades will not become long (e.g., if $p = 0.1$, the probability of a length-3 cascade is only 0.001).

Observation 3: In the degree-normalized case, for sufficiently low k, information unfairness *decreases* as p increases. For large k, cascades are able to travel farther from the originating node, and the local effects of homophily are diminished. As discussed above, this can only happen for large p. However, even for large p, this effect is countered by the combinatorial explosion of cascades (walks) that stay in the local neighborhood of the originating node. For each network, we observe some 'balance point' between k and p where cascades can grow long enough to overcome the local effects of homophily, but are not so long as to encounter such combinatorial explosion. With such cascades, information unfairness decreases.

Observation 4: Degree accounts for some, but rarely all, of the networks' information unfairness. The information unfairness values in Figure 4b are much lower than those in Figure 4a. For example, in Figure 4a, with density-based normalization the **Parallel Processing** subfield is extremely unfair. In this network, men and women have mean degrees, respectively, of 7.0 and 6.3. This difference is by far the largest of the subfields. With degree-based normalization, we see that for low k, this subfield has the *lowest* unfairness. This indicates that differences in distributions are due almost entirely to degree: men receive more information because they have more connections.

As k increases, **Parallel Processing** again has high information unfairness. Further exploration reveals that this subfield has very high degree assortativity (0.81 vs. -0.012 to 0.029 for the other subfields). Normalizing by degree removes the effect of a node's degree, but not its neighbors' degrees. In high-degree regions of the networks, the number of cascades is huge, and so for larger k, this network has high unfairness.

[12] Information unfairness is not equivalent to homophily; but homophily largely explains this particular behavior.

6 MaxFair: An Algorithm for Reducing Information Unfairness

In some application domains, it may be possible to add edges to a network to reduce its information unfairness. For example, a company may compute the information unfairness of its e-mail network, and then add key employees to a mailing list or ensure that they are invited to meetings with specific individuals to reduce unfairness. Here, we present `MaxFair`, an algorithm for adding b edges to a network to minimize its information unfairness.

6.1 Problem Statement The user provides an undirected network G with adjacency matrix $\mathbf{M}$, optional propagation value p, cascade length k, and budget b. The goal is to output set B of b edges such that the information unfairness of $G_b = (V, E \cup B)$ is minimized.

`MaxFair` uses the distance function that computes the difference between the means of the two distributions. It uses unnormalized information unfairness, because as long as b is small relative to the total number of edges, the normalization matrix will not change much.

6.2 Challenges A first challenge is that the problem of reducing information unfairness is not submodular. There are cases where no single edge will decrease information unfairness, but adding multiple edges will decrease unfairness. Second, estimating the change in flow when adding b edges to G is difficult in general. Prior work has characterized the flow through a network using its spectral properties [5] For example, Tong *et al.* consider the related problem of adding edges to a network to maximize *overall* flow through the network, and score each candidate edge (u, v) by the product of the eigenvector centralities of the endpoints [20]. This method cannot be directly used for our problem,

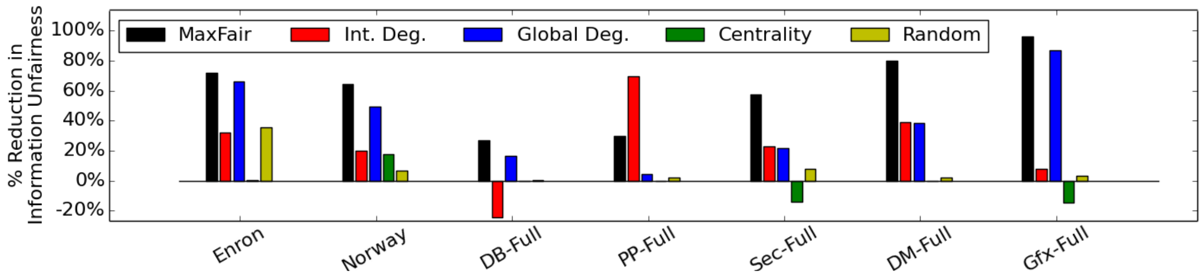


Figure 5: Information unfairness results for `MaxFair` and baseline algorithms.

because we do not seek to increase flow *generally*, but only between specific attribute groups.

6.3 Attribute-Based Centrality When estimating the effect of adding edge (u, v) on flow between attribute groups C_f and C_g, we must quantify how well nodes u and v spread information to groups C_f and C_g. If u has good flow to C_f, and v has good flow to C_g, then adding edge (u, v) will facilitate flow between the two groups.

To capture this concept, `MaxFair` performs a power iteration-type method. Suppose that we wish to create a vector vec_f containing each node's centrality with respect to group C_f. First, we initialize $vec_{f,0}$ so that the j^{th} element is 1 if $j \in C_f$ and 0 otherwise. We perform $k-1$ iterations, where in each iteration j, we set $vec_{f,j}(u) = sum([vec_{f,j-1}(v) : (u,v) \in G])$. Then we define $vec_f = p \times vec_{f,0} + p^2 \times vec_{f,1} + ... + p^k \times vec_{f,k}$. If p is not provided, one can integrate over $p \in [0, 1]$, as before. Next, we divide by $|C_f|$ to compute the mean. Finally, for each node $u \in C_f$, we add 1 to $vec_f(u)$ (because if $u \in C_f$, adding an edge to v increases flow to C_f, even if u does not subsequently spread to other members of C_f). This parallels the computation of the accessibility matrix, by finding the number of length-i cascades from u to nodes in C_f, weighted by powers of p, and summing over i from 1 to k.

6.4 `MaxFair` Let $IU_{mean}(G)$ represent the information unfairness of G using a distance function that computes the means of each joint attribute accessibility distribution. (p and k remain fixed throughout.)

`MaxFair` consists of b iterations. Let G_j denote the graph at the beginning of iteration j ($G_1 = G$). In iteration j, `MaxFair` performs the following, using G_j:

1. Compute the attribute-based centrality vector vec_f for each attribute group C_f.
2. Compute the joint attribute accessibility distributions D_{fg} for all group pairs C_f and C_g.
3. Compute the mean of each D_{fg} distribution, and the mean *all_mean* of the distribution means. Let $s_{fg} = all\_mean - mean(D_{fg})$.
4. Iterate over all pairs of nodes (u, v) that are not already connected in G_j. Define $score(u, v) = \sum_{f,g} s_{fg} * (vec_f(u) * vec_g(v) + vec_g(u) * vec_f(v))$.
5. Select the highest scoring edge to add to G_j.

The fourth step is the heart of `MaxFair`. Here, edges receive a reward for increasing flow between group pairs that have below-average flow, and a penalty for increasing flow between above-average group pairs.

We make two efficiency improvements to `MaxFair`. First, instead of recomputing information unfairness in every step, we recompute every j iterations. Second, instead of scoring all candidate edges, we prune the set by finding the top-q% highest scoring nodes with respect to attribute centrality for each attribute group, and only consider candidate edges between nodes in those sets. We discuss the effects of these improvements in Section 6.6. In our experiments, we set j to be one-tenth the number of desired edges and $q = 25\%$ (except for the small **Enron** and **Norway** networks, where $j = 1, q = 100\%$).

6.5 Results We evaluate `MaxFair` on the **DBLP-**Full, **Enron**, and **Norway** datasets. We compare `MaxFair` against four baseline algorithms: `InternalDegree` is the same as `MaxFair`, except that instead of using attribute centrality with respect to each group C_f, it uses the degree of each node to C_f. `GlobalDegree` identifies the D_{fg} joint attribute accessibility distribution with the lowest mean, and then connects the two nodes with the highest product of degrees from groups C_f and C_g. `Random` also identifies the D_{fg} distribution with the lowest mean, and then connects a random node from C_f to a random node from C_g. `Centrality` connects the node pairs with the highest eigenvector centrality product. For each network, we set the budget b to be 1% of the number of edges in the network. Figure 5 shows results for $p = 0.1, k = 5$, but we observed similar results for other combinations. In all but one case, `MaxFair` results in the largest decrease in information unfairness.

6.6 Running Time The major contributors to `MaxFair`'s running time are the information unfairness

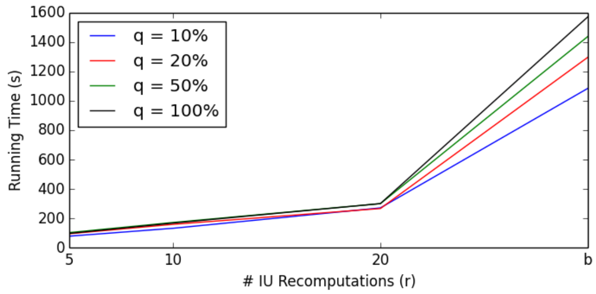


Figure 6: `MaxFair` running time for varying r, q on **Graphics-2015** network.

recalculation and the attribute centrality computation. We perform experiments in which we recompute information unfairness only r times over the course of adding b edges (as opposed to after each edge addition), for $r \in \{5, 10, 20, b\}$ (the latter value of r corresponds to recomputing in each iteration), and we prune to only consider the top-q% of nodes as scored by attribute centrality, for $q \in \{10\%, 25\%, 50\%, 100\%\}$. Figure 6 shows results on the **Graphics-2015** network. Similar results were observed in other networks. (Experiments run on a 2015 MacBook Pro with a 2.8 GHz i7 processor.)

For these values of q, pruning has no effect on solution quality, and r has a minimal effect on solution quality (e.g., for $r = 5$, `MaxFair` reduces information unfairness by 61%, and at $r = b$, it reduces it by 62%).

7 Conclusion

In this work, we introduced *information unfairness*, which evaluates whether information flows equally between all groups in a network. We perform a case study on the **DBLP** co-authorship network, in which we compare several subfields of computer science. Next, we introduced `MaxFair`, an algorithm to reduce the information unfairness of a network by adding a specified number of edges. We see that `MaxFair` results in up to a large reduction in information unfairness.

References

[1] A. Banerjee, A. G. Chandrasekhar, E. Duflo, and M. O. Jackson, *Gossip: Identifying central individuals in a social network*, tech. report, National Bureau of Economic Research, 2014.

[2] P. Barberá, J. T. Jost, J. Nagler, J. A. Tucker, and R. Bonneau, *Tweeting from left to right: Is online political communication more than an echo chamber?*, Psychological Science, 26 (2015), pp. 1531–1542.

[3] S. Barocas, M. Hardt, and A. Narayanan, *Fairness in machine learning*, NIPS Tutorial, (2017).

[4] D. J. Brass, K. D. Butterfield, and B. C. Skaggs, *Relationships and unethical behavior: A social network perspective*, Academy of Management Review, 23 (1998), pp. 14–31.

[5] G. Cencetti and F. Battiston, *Diffusive behavior of multiplex networks*, New Journal of Physics, 21 (2019).

[6] S. Cheryan, A. Master, and A. N. Meltzoff, *Cultural stereotypes as gatekeepers: Increasing girls interest in computer science and engineering by diversifying stereotypes*, Frontiers in Psychology, 6 (2015), p. 49.

[7] G. D'Angelo, L. Severini, and Y. Velaj, *Recommending links through influence maximization*, Theoretical Computer Science, 764 (2019), pp. 30–41.

[8] B. Fish, A. Bashardoust, D. boyd, S. A. Friedler, C. Scheidegger, and S. Venkatasubramanian, *Gaps in information access in social networks*, in WWW, 2019.

[9] H. Ibarra, *Paving an alternative route: Gender differences in managerial networks*, Social Psychology Quarterly, (1997), pp. 91–102.

[10] S. Jones, *White men account for 72% of corporate leadership at 16 of the Fortune 500 companies*, Fortune Magazine, (2017).

[11] F. Karimi, C. Wagner, F. Lemmerich, M. Jadidi, and M. Strohmaier, *Inferring gender from names on the web: A comparative evaluation of gender detection methods*, in WWW, 2016, pp. 53–54.

[12] M.-A. Lee and K. F. Ferraro, *Neighborhood residential segregation and physical health among hispanic Americans: Good, bad, or benign?*, Journal of Health and Social Behavior, 48 (2007), pp. 131–148.

[13] J. Leskovec, L. A. Adamic, and B. A. Huberman, *The dynamics of viral marketing*, ACM Transactions on the Web (TWEB), 1 (2007), p. 5.

[14] M. E. Newman, *Mixing patterns in networks*, Physical Review E, 67 (2003), p. 026126.

[15] M. E. Newman, *Modularity and community structure in networks*, PNAS, 103 (2006), pp. 8577–8582.

[16] G. Orfield and C. Lee, *Why segregation matters: Poverty and educational inequality.*, Civil Rights Project at Harvard University, (2005).

[17] C. Seierstad and T. Opsahl, *For the few not the many? The effects of affirmative action on presence, prominence, and social capital of women directors in norway*, Scandinavian Journal of Management, 27 (2011), pp. 44–54.

[18] J. Shetty and J. Adibi, *The enron email dataset database schema and brief statistical report*, Information Sciences Institute Technical Report, University of Southern California, 4 (2004), pp. 120–128.

[19] E. M. Simpson, T. Snuggs, T. Christiansen, and K. E. Simples, *Race, homophily, and purchase intentions and the black consumer*, Psychology & Marketing, 17 (2000), pp. 877–889.

[20] H. Tong, B. A. Prakash, T. Eliassi-Rad, M. Faloutsos, and C. Faloutsos, *Gelling, and melting, large graphs by edge manipulation*, in CIKM, 2012.

[21] F. Xu, B. Desmarais, and D. Peuquet, *Stand: A spatio-temporal algorithm for network diffusion simulation*, ArXiv: 1904.05998, (2019).