Fairness in a real social network

Leonidas Zafeiriou
Supervisors: Evaggelia Pitoura, Panagiotis Tsaparas
Course: Online social networks and media
Department of Computer Science and Engineering, University of Ioannina, 2022

Abstract

Ranking algorithms provide solutions to various real-life problems, such as friend suggestions in a social network or a hiring recommendation in a corporation. Given that these algorithms make people-related decisions, it is necessary that these decisions are fair. In this work, we analyzed the fairness in the rankings produced by widely used algorithms for a real social network. More specifically, we used ego-networks dataset from Google+ which include the genders of the users, among other features. To rank the users, we used *PageRank* algorithm and *eccentricity centrality* score. The two concepts of fairness under research are *inequality* and *inequity.* The former refers to the degree that the results of the algorithms are equal among all users, while the latter refers to the proportion of the population in the top ranks based on a feature (in particular *gender*) with respect to the whole population. Finally, the results of our experiments implies that there is global *inequality* and local *inequity* for both ranking strategies.

Introduction

Nowadays, the use of ranking algorithms has a great diversity of applications. These applications could be anything from recommending for buying a product in an e-shop to an interview appointment from an HR department in a corporation. While both use cases mentioned are good examples, it is clear that people-related situations as the latter are more important. To elaborate on this, ranking algorithms are used to assist people on making decisions. When this decision making will result to choosing one person over another, we want this choice to be as *fair* as possible. Therefore, if a ranking algorithm is *unfair,* then decisions that will be made will be unfair and people will be mistakenly excluded from a process. This is a situation that can possibly lead or strengthen issues like the *glass ceiling effect* and the *invisibility syndrome*, as mentioned in [1].

To understand if an algorithm is unfair in any way, we should define quantifiable concepts for this lack of fairness. We use two well-known concepts for injustice, *inequality* and *inequity.* From a social aspect, we could define inequality as a product of inequity. Inequity refers to unfair opportunities to different groups among a society which can lead in social inequality for the group of interest which means unbalance conditions. In the context of ranking algorithms, to study these concepts we follow the approach of [1]. *Inequality* is the unfair distribution of ranking scores among subjects, while *inequity* is about unfair representation of minority in top-k% in respect to the overall representation of minority. *Minority* is defined as a group of individuals, created based on a feature, with smaller population than the other groups with different values for this feature. In our case, the groups are based on the *gender* of the individuals and the minority is the less populated gender among all people.

Based on the aforementioned concepts, we studied the fairness of the results of two ranking strategies, *PageRank* and *eccentricity centrality.* Both are widely used algorithms for various applications. To test the results of these algorithms against the two fairness concepts, we used a dataset from a real social network, i.e. Google+. The dataset consists of ego-networks and has a lot of features information, including gender. The results that we got from the experiments show that indeed the algorithms produced unfair results to a certain degree. There is moderate global inequality and slight inequity. The rest of this work, consists of sections about *Methodology* that we followed, *Results* of the experiments and *Conclusions.*

Methodology

i. Dataset pre-processing

The dataset that we used for our experiments consists of ego-networks from Google+ social network [2]. The networks are *directed* since Google+ is built with the following-follower structure. For every user, there are several features in the form of binary vectors for presence or absence of the feature. Among these features are the gender, location, university, etc. For every feature, there is a great number of possible values. For our experiments, we use only **gender** feature. The possible values for gender are: [1, 2, 3]. We ignored from the analysis the value 3 of gender, supposing that it is the “don’t reveal” option so it doesn’t provide additional information.

Some of the networks contained in the dataset have an extremely small number of nodes or little to none features information. For this reason, they were manually eliminated from the overall dataset. With this process, from the 132 networks initially contained in the dataset, we worked with the 125 of them. In the remaining networks, a significant number of users was lacking gender information. These nodes were removed from the ego network they belong.

ii. Inequality

To quantify *inequality* in our graphs, we followed the approach from [1]. The authors define inequality as the inequal assignment of scores among entities. In our case, the entities are the nodes of our graphs that represent people, and the scores are the ranking results from the algorithms. Therefore, we utilized a simple version of **Gini coefficient**, which is a very popular measure of statistical dispersion in a population. The coefficient’s values are in [0, 1] interval, where 0 denotes *total equality* and 1 *total inequality.* The way we calculate the Gini coefficient for a population of ranking scores is described by the following formula:

$$Gini(s) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|s_i - s_j|}{2n^2\overline{s}}$$

where s is a vector with the scores produced by a ranking algorithm, n is the number of scores (i.e. the number of nodes of the graph) and $\overline{s}$ is the average score.

To analyze the behavior of ranks distribution among the nodes of the networks, instead of calculating the Gini coefficient for the overall population, we calculated for the top-k%, for a range of k values. Specifically, we calculated for k: [10, 20, 30, …, 100]. This way, we could make richer observations about both the global inequality and local inequality in specific subset of the nodes.

iii. Inequity

As far as the second evaluation concept, *inequity,* we followed again the definition given in [1]. Specifically, inequity refers to the fair representation of the groups given a ranking algorithm. Using an algorithm to produce scores for, in our case, nodes of a graph implies that decisions will be made according to the importance of the nodes. Equity expresses that in the top-k% positions the representation of the groups, and especially the minority, is equal to the representation in the whole population.

The inequity is measured in our experiments is described by the following formula:

$$Inequity_{score} = \frac{top\_k\_minority - overall\_minority}{overall\_minority}$$

Basically, what $Inequity_{score}$ measures is the percentage of how much more or less the minorities are represented in the top-k%. If the score is negative, then it means the minorities are *under-represented* in the top-k% positions, otherwise it means that they are *over-represented.* For example, a score of 0.5 means that there are 50% more minority presence in top-k%.

Finally, as we did with inequality, we calculated inequity score for several values of top-k%. This time we tested inequality scores for k in: [5, 10, 15, …, 100].

iv. PageRank

The first algorithm that we used to evaluate the *importance* of the nodes is the widely used *PageRank* algorithm. It is one of the most popular ranking algorithms, as it has been used from web page ranking from in Google search engine to follow recommendation in social media. As a high-level idea, PageRank evaluates both the *quantity* and the *quality* of a node in a graph. As a result, a node that it is pointed by *many high-quality* nodes should be an important node. Thus, for every ego-network of the dataset the PageRank score was calculated, and the results were exported in .txt files for later processing.

v. **Eccentricity centrality**

The second ranking technique that we used is *eccentricity centrality,* which is a measure of how central a node is in a graph. The eccentricity centrality score for a node v of a graph is the *maximum shortest path* between v and every other node u. Essentially, the smaller eccentricity value a node has means that its more distanced node is relatively closer to it which subsequently means that the node is more central in respect to the other nodes. Eccentricity centrality is also used as a metric for ranking of users in social media. To express produced scores as percentage values, the reciprocal of eccentricity values was used. This way, the smaller the eccentricity value the bigger the reciprocal and the importance of the node. Lastly, the results for each ego-network were again exported as a .txt file for later processing.

Results

As far as inequality is concerned, the results for PageRank and eccentricity centrality are depicted in Figure 1 and Figure 2, respectively. The first observation that can be made is that in the case of k equal to 100%, i.e. inequality in the whole population, the Gini coefficient is moderate. Inequality for the whole population is approximately 0.34 for PageRank and 0.39 for eccentricity centrality. This means that there never global equality, neither for PageRank nor for eccentricity. This observation agrees with the results in [1].

The second observation is that, in the case of PageRank results in Figure 1, the inequality is constantly increasing with respect to top-k% with lowest inequality to be achieved at top-10%. This situation also agrees with results in [1] and is explained by the fact that a lot of the importance of the PageRank algorithm is gathered in the very top positions. This is of course problematic since only a few nodes dominate in the whole network. The behavior of the distribution is slightly difference in the case of eccentricity, though, as it is shown in Figure 2. Because only a few nodes have extremely high scores and approximately 35% of the nodes has equal eccentricity centrality scores, there is higher inequality in top-10%. This is stabilized around 55% where most of the scores are more or less equal.

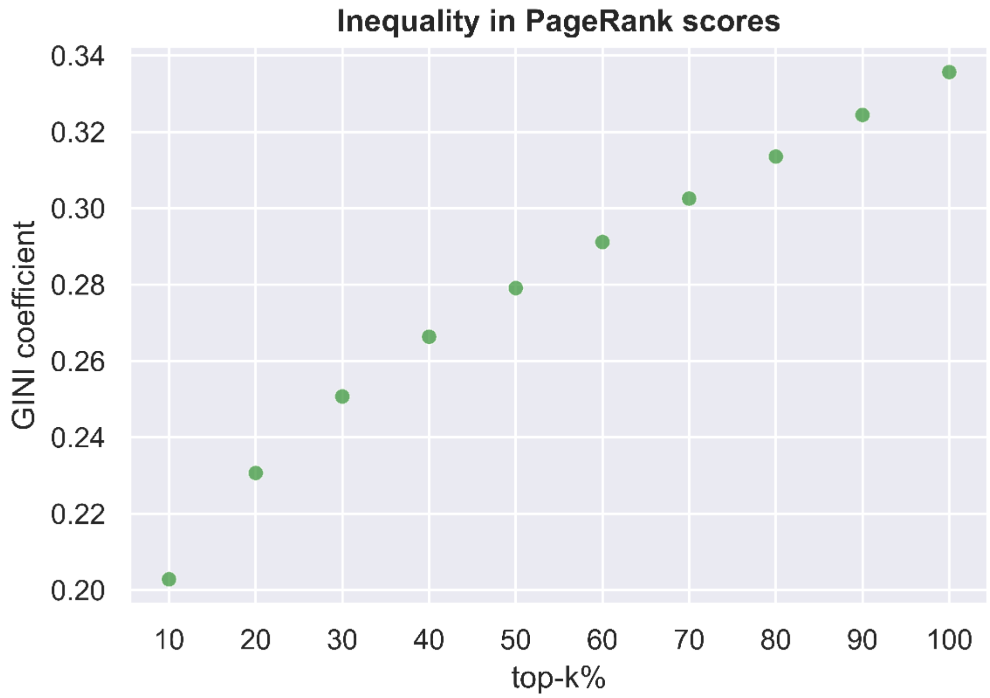

Figure 1: *Inequality in the top-k% positions for the scores produced by PageRank algorithm in Google+ dataset.*

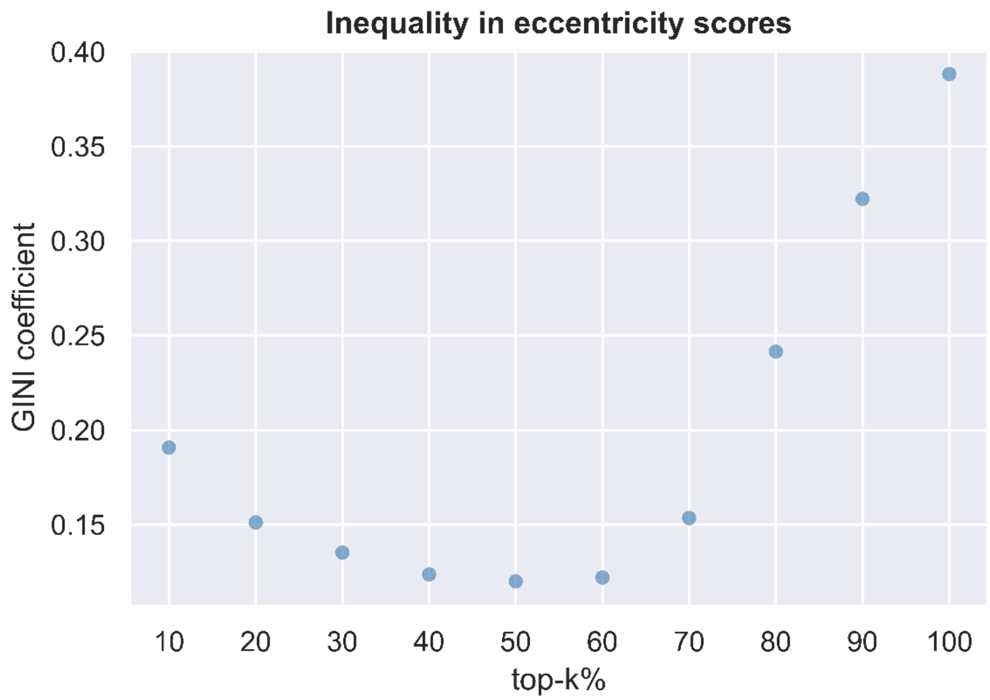

Figure 2: *Inequality in the top-k% positions for the scores produced by the reciprocal of eccentricity centrality in Google+ dataset.*

As far as inequity evaluation is concerned, the results on Google+ dataset are shown in Figure 3 and Figure 4. Generally, the results again follow the conclusions made in the reference work [1]. In [1], the authors set an arbitrary margin of *beta = |0.05|* so that inequity considered not normal. Indeed, inequity scores greater than beta are observed in both ranking strategies that we tested. Here, it should be noted that, majorities and minorities are decided based on the value of gender for each node, regardless of which gender is the minority.

In Figure 3 we can clearly see that there is an *under-representation* of nearly 0.1 inequity score in top-5%. This means, that there are 10% less minority population in the very top positions than in the whole population. Except for that, in the rest of k values *over-representation* of minorities is observed, with inequity score greater than beta in top-20% to top-40%.

In eccentricity centrality scores, shown in Figure 4, only *under-representation* of minorities is reached. The highest value of inequity is observed at top-10% of the results, which is approximately -0.09. In eccentricity centrality evaluation, there is a bigger interval than PageRank with values greater than beta. Specifically, we have worth-noticing inequity scores all the way to top-55%.

Finally, we can conclude that in both cases, at top-5% in PageRank and top-10% in eccentricity centrality scores, there is significant under-representation of minorities. This result leads to the fact that in the highest positions, which may be the ones that decisions will be made upon them, unfair representation of minorities exists.

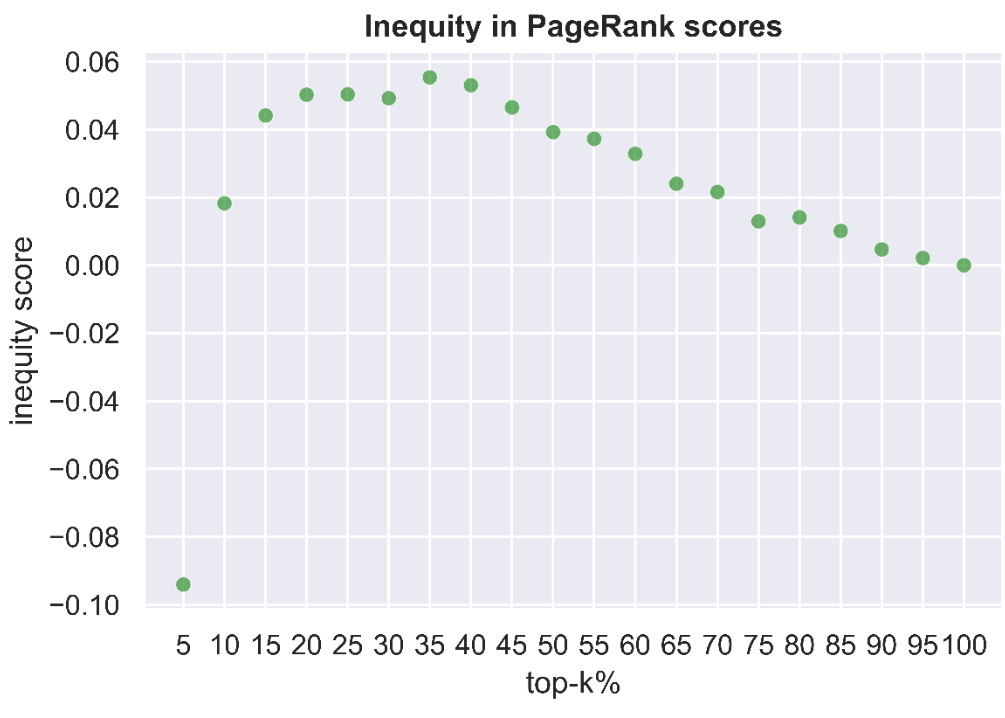

Figure 3: *Inequity in the top-k% positions for the scores produced by PageRank algorithm in Google+ dataset.*

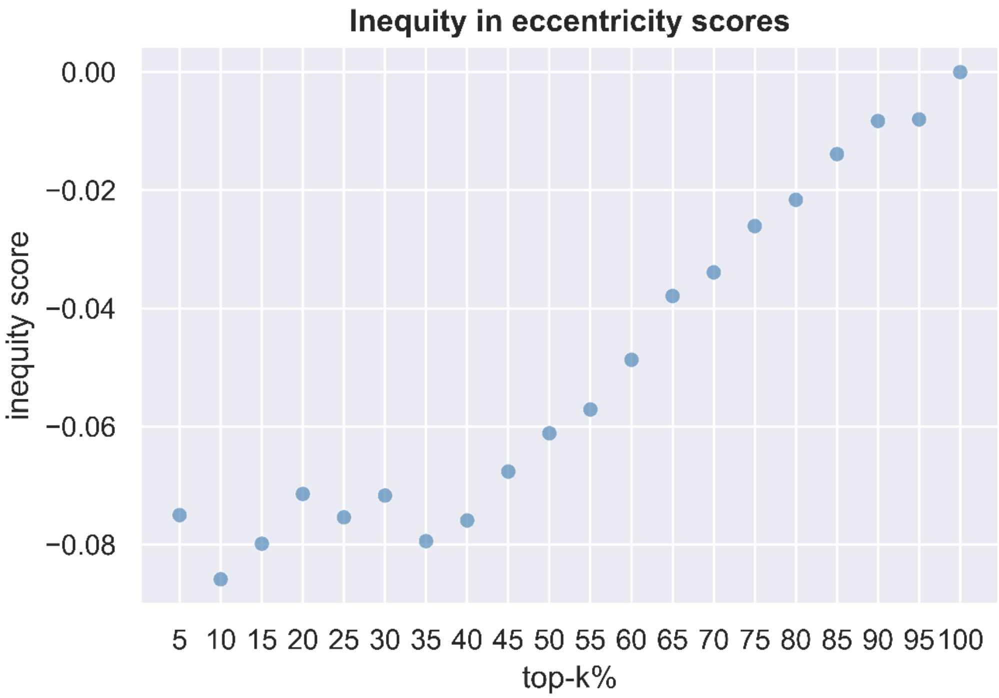

Figure 4: *Inequity in the top-k% positions for the scores produced by the reciprocal of eccentricity centrality in Google+ dataset.*

Conclusions

Altogether, the results we obtained from our experiments show the presence of unfair ranking scores from well-known ranking algorithms. It is worth noticing that our conclusions of inequality and inequity agree with the ones in [1], while the experiments were made in a different type of dataset. *Global inequality* is depicted by the moderate Gini coefficient values for the whole graph as input. On the other hand, slight *local inequity* is clear in the top-5% and top-10% of the two ranking strategies. Given that there is a level of unfairness in the results of the algorithms, more research should be made in this field and strategies to make algorithms more fair should be suggested.

References

1. Tóth, G., Wachs, J., Di Clemente, R. et al. Inequality is rising where social network segregation interacts with urban topology. Nat Commun 12, 1143 (2021).

2. J. McAuley and J. Leskovec. Learning to Discover Social Circles in Ego Networks. NIPS, 2012.