Enhancing Structural Minority Visibility in Link Recommendations

Shera Potka, Isla Li, Jason Kepler, and Alex Thomo

University of Victoria, British Columbia, Canada

Abstract. Traditional link recommendation algorithms in social networks often promote popular nodes, leading to societal polarization and the marginalization of minority voices. Conversely, several studies have notably emphasized fairness and diversity in recommendation algorithms. However, these studies tend not to prioritize the visibility of minority groups in their design. Instead, they focus on optimizing other objectives related to fairness and diversity, which may not necessarily align with enhancing minority visibility. Our study investigates the impact of fairness and diversity-aware algorithms on the visibility of structural minorities in social networks. We assess several of these algorithms, noting their varied effects, which range from indifference to potential backlash from the majority. Additionally, we explore the influence of these algorithms on network cohesion and popularity bias. A key contribution of our work is the introduction of MinWalk, a novel algorithm designed to enhance the visibility of minorities in a balanced manner. Our approach to algorithmic fairness specifically addresses the challenges associated with minority representation.

Keywords: Social Networks · Recommendation Systems · Minorities · Fairness

1 Introduction

Social networks, impacting everything from collaboration to well-being and societal outlooks [27], have evolved from being shaped by homophily and preferential attachment [22] to being influenced by link recommendation algorithms [2]. These algorithms tend to favor popular choices [6], leading to filter bubbles and echo chambers [16] and heightening societal polarization. Additionally, they risk marginalizing minority voices [12, 32, 11], thus perpetuating stereotypes and limiting diversity [9].

A set of recent works, [29, 30, 13, 8, 14], have concentrated their focus towards understanding the temporal effects of link recommendation algorithms in social networks. This temporal perspective is important as the dynamics of social networks evolve over time, and the long-term consequences of algorithmic interventions can be profound. In particular, the studies by Fabbri et al. and Ferrara et al. [13, 14] model the feedback loop of user interactions with link recommenders, affecting network structure and *minorities*, with Ferrara et al. offering a novel metric for minority visibility based on PageRank and tools to measure popularity bias and cohesiveness in social networks. However, these works come with limitations which we outline as follows: (1) they rely on synthetic data for their studies, (2) they do not consider natural minorities that emerge organically based on their connection density, and (3) they only examine traditional link recommender algorithms ignoring newer, fairness- and diversity-aware algorithms.

In our study, we examine the visibility of minority groups in *real-world* social networks–a shift from the *synthetic data* reliance seen in [13, 12, 8, 14]. Contrary to the attribute-based community definitions in works like [29, 12] (e.g. men vs. women), or homophilic community definitions in [13, 8, 14], our research is grounded in *structural communities* (based on link density) within networks, aiming for a more accurate reflection of network clusters and minority communities. Structural communities can reveal hidden groups that we might miss with traditional methods based on attributes like gender or age. For example, in a social network, users who interact frequently might form a structural community even if they do not share obvious attributes. These groups could be tightly-knit professional circles, niche hobby clubs, or local interest groups. By focusing on these structural communities, we make sure fairness extends to all kinds of minority groups, not just the ones that fit into conventional categories.

A key feature of our work is the use of *fairness and diversity-aware recommendation algorithms*, which is in contrast to works such as [13, 8, 14] that only consider classical recommendation algorithms. In light of recent developments in fairness and diversity-aware recommendation algorithms (cf. [26, 32, 28, 20]), our study is the first to analyse these algorithms with respect to minority visibility and popularity bias. This is important because, although these algorithms have been designed with fairness and diversity in mind, they in fact pursue different objectives. These objectives, however, do not necessarily prioritize the visibility of minorities over time as users engage with the algorithm. Therefore, in this work, we aim to examine how the various fairness and diversity objectives of previous works align with the goal of increasing minority visibility over time.

Contributions: We make the following contributions in this paper. (1) We demonstrate that algorithms designed with fairness or diversity considerations show varied impacts on the visibility of structural minorities according to algorithmic ranking. While some of these algorithms appear indifferent to structural minorities, others increase their visibility significantly, potentially to a level that could risk backlash from the majority. This extreme visibility enhancement might lead to majority users abandoning the recommendation algorithm altogether. (2) We introduce a new algorithm, MinWalk, designed to effectively enhance the visibility of minorities in a balanced manner, aiming to minimize network disruptions and reduce the potential for significant dissatisfaction among the majority user base. (3) We also analyze the popularity bias of the on-focus algorithms, examining their temporal effects over multiple iterations. This involves assessing whether these algorithms inadvertently favour more popular nodes at the expense of lesser-known ones.

2 Related Works

Impact of Recommendation Algorithms on Communities. Several studies have explored the impact of recommender systems on network communities, offering valuable insights but also presenting certain limitations.

Cinus et al. [8] investigate the effect of recommender systems on echo chambers and polarization using Monte Carlo simulations. While innovative, their study is limited to synthetic datasets, homophilic communities, and use of classical recommender algorithms. We note that homophilic communities are defined by shared attributes, whereas structural communities are characterized by dense connections among members. Espín-Noboa et al. [11] explore how preferential attachment and homophily contribute to inequality in network-based ranking algorithms, but their study is limited to the effects of a single attribute and a specific network model.

Fabbri et al. [12] look at how similarity in choices or traits (homophily) can make minority groups more visible in recommendation systems, more so than the group's size. They provide a static, single-round analysis, which does not consider the changes over time or structural communities. Their subsequent study [13] suggests that these algorithms could help small homophilic groups become more noticeable in the long run. However, this work also does not address how structural communities evolve, and it is limited to a one-iteration analysis.

Of particular relevance to our discussion is the work by Ferrara et al. [14]. They define minority visibility in networks and introduce a novel metric based on the PageRank algorithm to quantifiably measure it. Furthermore, they propose methods to capture both popularity bias and network cohesiveness using established network metrics.

Stoica et al. [29] and Su et al. [30] study biases in social media link recommendation algorithms. Stoica et al. [29] examine Instagram's "glass-ceiling" effect due to link recommendation algorithms, which may restrict the network growth of groups such as women or men. Su et al. reveal a "rich get richer" trend on Twitter, where popular users gain more from these algorithms, leading to amplified visibility for already high-profile users.

Fainess-aware link recommendation algorithms. Several studies have notably emphasized fairness and diversity in recommendation algorithms, primarily focusing on random walk-based methods [26, 28, 20, 32]. Rahman et al. [26] developed Fairwalk, enhancing the node2vec embedding approach [17] by integrating fairness metrics like statistical parity to reduce bias in friendship recommendations while maintaining utility. Saxena et al. [28] introduced NodeSim, a network embedding method that accounts for community structures to enhance both intra and inter-community link prediction in social networks, and promoting diverse link predictions in the process. In a similar vein, Khajehnejad et al. [20] proposed CrossWalk, a method that biases random walks to cross group boundaries, thus promoting fairness and diversity in these walks. Further, Tsioutsiouliklis et al. [32] modified the Pagerank algorithm to create fairness-sensitive personalized Pagerank to meet a 'universal personalized fairness' criterion.

3 Networks, Communities, and Link Recommendation

In this work, we consider both directed and undirected network graphs. A graph G is defined as a pair (V, E), where V represents the set of vertices (or nodes) and E represents the set of edges (or links). Vertices are denoted by letters such as u, v, and edges are represented as pairs (u, v). For undirected graphs, each edge is treated as two directed edges in opposite directions. Consequently, we do not make a distinction between directed and undirected graphs in our analysis.

Structural Communities. Social networks are typically sparse graphs, where the number of edges is within a constant factor of the number of vertices. However, the nodes in these graphs often cluster in regions of higher density, which are referred to as structural communities in network analysis literature (cf. [15, 19, 25]). The main characteristic of these communities is that they have more connections among their members than with nodes outside the community. Structural communities play a crucial role in understanding and optimizing social dynamics [24]. They optimize network connectivity and information flow [3, 10], aid in disease modeling [7, 18], enhance recommendation systems [1, 21], foster professional collaboration [5], and facilitate effective marketing strategies [27], while providing valuable data for sociocultural research [15].

In this work, we utilize two well-known algorithms to discover structural communities: the Louvain algorithm [4] for undirected graphs, and the Leiden algorithm [31] for directed graphs. The Louvain algorithm detects high-quality community structures by optimizing modularity in a hierarchical manner [4]. On the other hand, the Leiden algorithm, an improvement over the Louvain method, focuses on refining community detection for directed networks [31]. The output of these algorithms is a partitioning of the nodes of the graph into structural communities, i.e. each node is assigned to a community. In our approach, the minority group within each network is identified by sorting the nodes by their community size and selecting the smallest $r\%$ of communities (e.g., 15%). These are then combined to form the minority group, denoted by M. The remaining nodes are considered to be the majority group, denoted by J. We have $M \cup J = V$.

Link recommendation Link recommendation algorithms utilize the network structure to suggest top-k matches for each node u. When node u accepts these recommendations, it changes the network structure, creating a feedback loop. This altered structure is then analyzed by the algorithm for new recommendations, thus continuing the cycle.

Central to most recommendation algorithms is the concept of node similarity. Nodes are typically recommended to connect with others that are most similar to them with respect to their one- and multi-hop friends. However, the definition and computation of similarity can vary. A prominent approach, Node2Vec [17], involves performing random walks starting from each node. These walks generate "sentences," with "words" representing node IDs encountered during the walk. The sentences are then embedded into a multidimensional space, assigning a vector to each node. The similarity between nodes is inferred from the similarity of their corresponding vectors–the closer two vectors are in this space, the more similar the nodes they represent are considered.

Fairness and diversity-aware algorithms bias random walks for more equitable recommendations, evolving networks towards greater diversity and/or fairness.

4 Algorithms

In this section, we provide a more detailed description of the fairness and diversity-aware algorithms for link recommendation [26, 28, 20]. In our work, we restrict the scope of our analysis to random-walk-based algorithms due to their popularity and effectiveness in capturing the structural properties of graphs. Our aim in this paper is twofold: to present a comparative study of state-of-the-art fairness and diversity-aware algorithms for link recommendation using random walks, focused on the visibility of minorities, and to introduce our new algorithm, Min-Walk, designed to enhance minority visibility in a balanced way.

4.1 FairWalk

FairWalk, introduced by Rahman et al. in [26], offers a fairness-oriented approach to generating random walks, addressing the representation bias often seen in standard methods like node2vec. The proposed technique modifies the traditional random walk procedure by first categorizing neighboring nodes into groups according to sensitive attributes. Unlike the conventional method where any neighbor might be chosen for the next step in the walk, FairWalk guarantees that each group–regardless of its size–has an equal probability of being selected. From the chosen group, a node is then randomly picked to proceed with the walk. In our evaluation, we use two groups within our setting: the minority and the majority, as the basis for applying this method.

4.2 NodeSim

NodeSim, introduced by Saxena et al. in [28], enhances random walks by promoting diversity through steps across communities. In contrast to conventional methods like node2vec, where the transition probability from node u to node v is $\frac{1}{\deg(u)}$, NodeSim incorporates node similarity and community structure into its random walk process. The unnormalized probability p_{uv} for moving from node u to node v is defined as $\alpha \cdot (\text{Sim}(u, v) + \frac{1}{\deg(u)})$ if u and v are in the same community and $\beta \cdot (\text{Sim}(u, v) + \frac{1}{\deg(u)})$ if they are in different communities, where $\text{Sim}(u, v)$ represents the Jaccard similarity of their neighbor sets. The parameters α and β play the role of guiding the random walker: a higher value of α encourages the walker to sample similar nodes within the same community, while a higher value of β motivates the walker to explore nodes outside the community of a node.

4.3 CrossWalk

CrossWalk, introduced by Khajehnejad et. al. in [20], aims to strengthen network connectivity across diverse groups by weighting edges more heavily near or across group boundaries.

For each node v within a graph, a quantifiable metric, denoted as $m(v)$, is established to assess its proximity to nodes of different groups. $m(v)$ is defined as the expected count of encounters with nodes from other groups during r random walks of length d that start from v: $m(v) = \frac{1}{r \times d} \sum_{j \in [r]} \sum_{u \in W_j^v} I[l_v \neq l_u]$ where, W_j^v represents the set of nodes visited during the j^{th} random walk from node v, and $I[l_v \neq l_u]$ is an indicator function that assumes a value of 1 when the group label l_v of the starting node v is not equal to the group label l_u of the node u encountered during the walk, and 0 when they are the same. Consequently, this measure cumulatively evaluates the interaction of node v with nodes from other groups across all walks, providing an aggregate indicator of its connectivity to diverse groups within the graph. Nodes adjacent to group boundaries with a diverse label mix in their proximity gain a higher m value, leading to an inclination of reweighted random walks toward these boundary nodes.

Edge weights are then introduced, which are biased towards promoting diversity with different weights for same-group $(1-\alpha) \times m(u)^p$ and different-group $\alpha \times m(u)^p$ connections, controlled by α and p. While CrossWalk enhances diversity, empirical results show it can sometimes overemphasize minority communities in link recommender systems. To address this, we introduce a new algorithm, MinWalk, which seeks a more balanced approach to promoting cross-boundary walks.

4.4 MinWalk

We now introduce our algorithm, MinWalk (Minority Walk), specifically designed to weigh edges in a manner that enables random-walk-based link recommender algorithms to generate walks with an emphasis on minority communities.

MinWalk operates by adjusting the edge weights of a specific small set of majority nodes, thereby only moderately interfering with the graph's original structure. Unlike CrossWalk, which generally improves minority visibility but often goes overboard in this goal (as our experiments show), MinWalk is more targeted and less disruptive. The primary goal of MinWalk is to better align the proportion of minority nodes within the top $t\%$ of PageRank (PR) values to the graph's initial minority ratio (denoted as μ). This is achieved by selectively adjusting the weights of majority nodes that are only marginally within the top $t\%$ bracket of PR values.

Adjustment Strategy. Consider a hypothetical graph with 1000 nodes, $t = 10\%$, and where the minority ratio $\mu = 15\%$. Ideally, this ratio should be mirrored in the top 10% of PR scores. If the existing minority ratio in this bracket is only 5%, MinWalk steps in to make adjustments. It targets the lower-ranked majority nodes in this top bracket, modifying the weights of the bottom 10 (from 95 to 85) of these majority nodes, to achieve the desired 15% minority representation.

In MinWalk, edge weight adjustments are made exclusively to majority nodes, denoted as J, while minority nodes, denoted as M, remain unchanged. The algorithm retains the closeness calculation from CrossWalk, emphasizing nodes at the majority periphery. The closeness value $m(v)$ for each node v is now specifically determined by the proportion of minority nodes encountered during random walks.

Algorithm 1 MinWalk: Minority-enhanced edge weighting

```
Require: Graph $G = (V, E)$, minority set $M \subset V$, majority set $J \subset V$ ($M \cup J = V$), top ratio $t$ of interest in PR ranking.
Ensure: Weights $w_{vu}, \forall (v, u) \in E$.
$n = |V|$, $\mu = |M|/|V|$    ▷ $n$ is the number of nodes, $\mu$ is the minority ratio.
$H = \{u \in J \mid rank(u, V) \leq t \cdot n\}$    ▷ $H$ is the set of $J$ nodes in the top t% of PR ranked nodes.
$K = \{u \in J \mid rank(u, J) \leq t \cdot n \cdot (1 - \mu)\}$    ▷ $K$ is the set of $H$ nodes to stay in the top t% of rankings.
$L = H \setminus K$    ▷ $L$ is the set of $J$ nodes that will have their edges reweighted.
for $v \in V$ do
    Run $r$ random walks $W_j^v, j \in [r]$ rooted at $v$.
    $m(v) = \frac{1}{r \cdot d} \sum_{j \in [r]} \sum_{u \in W_j^v} I[u \in M]$    ▷ Closeness of $J$ node $v$ to minority $M$.
    $IS(v) \leftarrow \frac{|N_v \cap M|}{|N_v|}$    ▷ $IS$ (influence score) is the ratio of the $M$ nodes in $N_v$ (neighborhood of $v$).
for $v, u \in V$ do    ▷ Initial weights: higher weight for nodes with more $M$ neighbors.
    $w_{vu} = \frac{IS(v)+IS(u)}{2}$
for $v \in L$ do    ▷ Reweighing edges.
    $Y_v = \sum_{u \in N_v \cap J} w_{vu} \cdot m(u)$    ▷ Normalization quantity for reweights of edges incoming to a $J$ node.
    $Z_v = \sum_{u \in N_v \cap M} w_{vu} \cdot m(u)$    ▷ Normalization quantity for reweights of edges incoming to a $M$ node.
    if $v \in L$ then    ▷ Valid Majority nodes
        for $u \in N_v \cap J$ do    ▷ Edges within $J$.
            $w_{vu} = w_{vu} \cdot \frac{m(u)}{Y_v}$    ▷ Edges going to nodes closer to $M$ get higher weight.
        for $u \in N_v \cap M$ do    ▷ Edges from $J$ to $M$.
            $w_{vu} = w_{vu} \cdot \frac{m(u)}{Z_v}$    ▷ Edges going to $M$ nodes get higher weight.
```

Details. To aid in understanding MinWalk, we present annotated pseudocode in Algorithm 1. We set $n = |V|$ and the minority ratio $\mu = |M|/|V|$. The algorithm then focuses on identifying key subsets within the majority set J. Firstly, it defines H as the set of nodes in J that rank in the top $t\%$ of PageRank (PR) ranked nodes within the entire set V. Subsequently, it refines this to a subset K, comprising nodes in H that remain in the top $t\%$ after adjusting for the minority ratio. This leaves us with the set $L = H \setminus K$, which consists of majority nodes targeted for edge reweighting.

For each node $v \in V$, the algorithm conducts r random walks to calculate $m(v)$, a measure of the closeness of a majority node v to the minority set M. It also computes the Influence Score $IS(v)$, defined as the ratio of minority neighbors to the total number of neighbors of v. The initial weights for the edges (v, u) are set as $w_{vu} = \frac{IS(v)+IS(u)}{2}$, emphasizing connections to nodes with a higher proportion of minority neighbors.

The reweighting process is applied to nodes in L. For each $v \in L$, normalization quantities Y_v and Z_v are calculated, summing the weighted closeness measures $m(u)$ for majority and minority nodes, respectively. The algorithm then adjusts the weights of the edges. For edges within the majority set J, the weights are modified as $w_{vu} = w_{vu} \cdot \frac{m(u)}{Y_v}$, and for edges from the majority to the minority set M, as $w_{vu} = w_{vu} \cdot \frac{m(u)}{Z_v}$. This reweighting process selectively enhances the influence of edges that connect to the minority set, or are within the majority set. This, in turn, amplifies the representation of minority nodes in the random walks performed by a random-walk-based link recommender system.

4.5 Universally-Fair Personalized PageRank

Finally, an algorithm that does not explicitly generate traces from random walks (but still based on them) is Universal Personalized Fairness in PageRank (u-ppr), introduced by Tsioutsiouliklis et al. in [32]. It modifies the standard personalized PageRank algorithm to prioritize equitable treatment across nodes. Their adjustment results in a personalized probability vector $\mathrm{PR}_v(u)$ that emphasizes fairness by distributing probabilities evenly among groups based on their proportions.

5 Experiments

Network Evolution Methodology. In our study, we begin with a network graph G and a given recommendation algorithm A. For each node u in G, algorithm A provides a top-1 link recommendation, identifying the most suitable node v for u to connect to. Following the approach of previous studies [14, 8], we then establish an edge between u and v, and additionally, we remove a random edge incident to u. In the case of a directed graph, this involves adding a directed edge from u to v and then removing a random outgoing edge from u. This strategy of adding and removing edges is designed to prevent an excessive increase in the network's edge density. We focus on maintaining constant edge density, as our evaluation metrics are sensitive to it.

By removing one connection for each new one formed, we make sure the network's density stays stable, allowing us to attribute any observed changes solely to the recommendation algorithm A, without the influence of increased total connections. This methodology aligns with social theories that suggest individuals have a finite capacity for communication, limiting the number of active ties they can maintain [23]. The issue of whether network users will accept the top recommendations for edge additions relates to the "acceptance policy". However, studies [8, 12] suggest that this policy has minimal impact on the evolution of the network. The above process is repeated for 30 iterations to assess the recommendation algorithm's effects over time.

Structural Minorities. In order to determine structural minorities in undirected and directed graphs, we utilize the Louvain [4] and Leiden [31] algorithms, respectively, for community detection. The minority group within each network is identified by sorting nodes by community size and selecting the smallest $r\%$ of communities. In the results we show here, we used an $r\%$ value of 15%, although we also tested values of 5%, 10%, 20%, and 25%, all of which exhibited similar behavior.

Metrics. The first metric we evaluate is 'minority visibility,' as introduced by [14], which assesses the prominence and network position of minority nodes following the implementation of recommendation algorithms. This metric specifically measures the proportion of minorities within the top 10% of nodes ranked by PageRank. It is essential to determine whether the algorithms enhance or diminish the representation and influence of minority nodes.

The second metric we evaluate is the Gini Coefficient, a measure of inequality in the distribution of connections (degrees) among nodes. A higher value indicates a network where connections are concentrated among a few nodes, which could lead to the marginalization of minority nodes. This metric is useful for assessing how much recommendation algorithms either contribute to or alleviate inequalities in network connections.

Name	**Nodes**	**Edges**	**Clust. Coeff.**	**Type**
congress	475	13,289	0.2242	Directed
email-eu	1,005	25,571	0.3657	Directed
wiki-vote	7,115	103,689	0.0816	Directed
facebook	4,039	88,234	0.6055	Undirected

Table 1. Datasets used in our experiments

Datasets. We utilize datasets from the Stanford Large Network Dataset Collection (https://snap.stanford.edu), including: **Congress** (Twitter interactions of the 117th U.S. Congress), **EU-Email** (email interactions within a European research institution), **Wiki-Vote** (Wikipedia voting data until January 2008), and **Ego-Facebook** (Facebook 'circles' or 'friends lists'). Details on nodes, edges, and clustering coefficients are provided in Table 1.

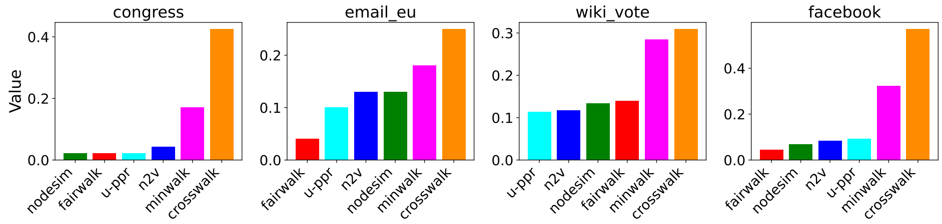

Fig. 1. Visibility after 30 iterations for minority size of 15%. In all the charts we observe that crosswalk increases the visibility well above 15%, sometimes to above 50% (see 'facebook'). A more balanced approach is our 'minwalk' algorithm which increases the visibility of minorities more moderately or keeps it close to 15%.

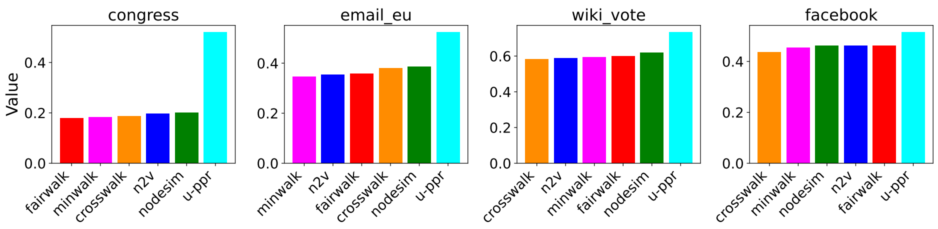

Fig. 2. Gini after 30 iterations for minority size of 15%. The greater the Gini coefficient, the more imbalanced or unequal the distribution of degrees. 'u-ppr' leads to most imbalance. 'minwalk' tends to result in Gini coefficients at the lower end, indicating its propensity to evolve the network towards a more balanced state.

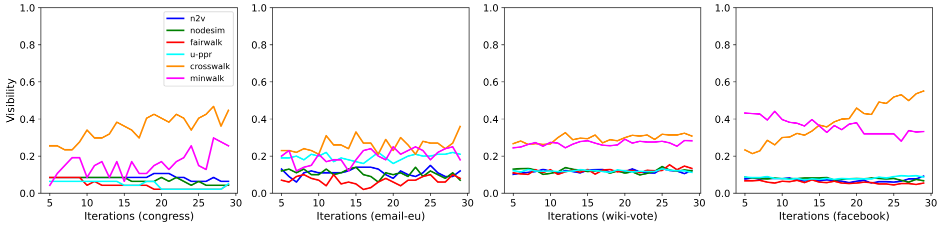

Fig. 3. Evolution of visibility over iterations for minority size of 15%. We observe that visibility remains relatively stable for all methods, with the exception of 'crosswalk', which consistently shows an increase. In some cases, this increase is concerning, such as on 'facebook', where it grows beyond 50%. Generally, our 'minwalk' algorithm generates significantly higher visibility compared to other methods, with the notable exception of 'crosswalk'. The very high and continually increasing visibility levels produced by 'crosswalk' might be considered excessive for a minority size of 15%.

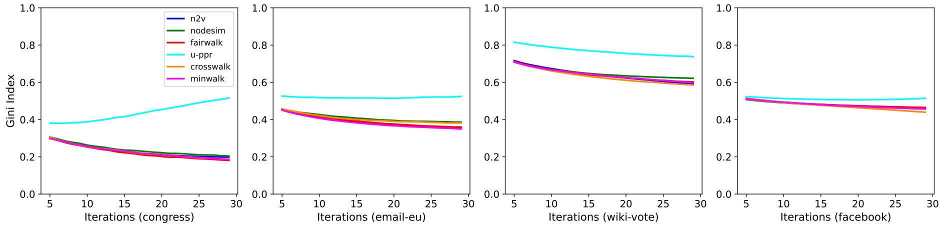

Fig. 4. Evolution of the Gini index over iterations for minority size of 15%. All algorithms, with the exception of 'u-ppr', tend to decrease the Gini index over time. A decreasing Gini indicates a more balanced degree distribution within the network.

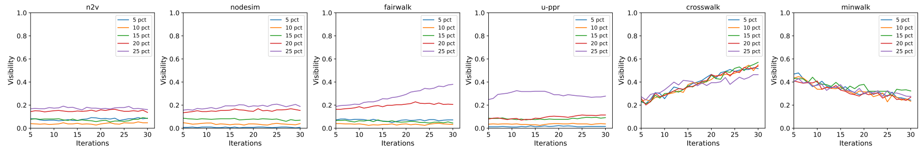

Fig. 5. Visibility evolution over iterations for different minority sizes on the 'facebook' dataset. For 'n2v', 'nodesim', and 'fairwalk', we observe a relatively flat behavior in terms of visibility change, mostly lower than the minority percentage. 'Fairwalk' exhibits an exception at the 25% minority size, where it significantly increases visibility well beyond 25%. 'u-ppr' produces strong visibility for the 25% minority size, but for other sizes, visibility is markedly reduced below their corresponding minority percentages. 'Crosswalk' seems indifferent to the minority size, concerningly increasing the visibility of minorities beyond the 50% mark. Finally, our algorithm, 'minwalk', starts somewhat high, but over time, it normalizes the visibility to more acceptable levels.

Results In our experiments, we focused on analyzing network dynamics changes after 30 iterations. Also, we set the length of random walks for the random-walk-based algorithms to 40. The results regarding minority visibility, clustering coefficient, and Gini coefficient are presented in Figures 1 and 2, specifically with a minority size of 15%.

Minority Visibility. Figure 1 helps us make the following observations related to minority visibility, particularly when the minority size is set at 15%. The first clear finding is that the 'crosswalk' algorithm significantly boosts minority visibility, often exceeding 15% and, in some cases, reaching over 50%, as observed in the 'facebook' dataset. However, we should also consider the implications of excessively amplifying minority visibility. For instance, a visibility surge beyond 50% in a 15% minority context, could be seen as counterproductive, potentially provoking a backlash from the majority. This scenario underscores the need for a balanced approach to avoid such pitfalls.

Our 'minwalk' algorithm increases the visibility of minority groups more moderately than 'crosswalk', but also significantly outperforms competing algorithms like 'n2v', 'nodesim', 'fairwalk', and 'u-ppr'. The last three algorithms, 'nodesim', 'fairwalk', and 'u-ppr', despite being designed with fairness and diversity in mind, fall short in achieving a good minority visibility. Their focus on optimizing other metrics leads to a disproportionate decrease in minority visibility, often significantly below the actual minority proportion, as evidenced in datasets such as 'congress', 'email-eu', 'wiki-vote', and 'facebook'. Moreover, these algorithms do not significantly differ in performance from the baseline 'n2v' algorithm in terms of increasing the minority visibility.

Gini Coefficient. After 30 iterations at 15% minority size, we assessed the network's degree distribution balance using the Gini coefficient, shown in Figure 2. Higher Gini values indicate greater imbalance. Our findings reveal that 'u-ppr' typically causes the greatest imbalance across all networks. Notably, 'nodesim' often emerges as the next most imbalancing algorithm, with the exception of

Facebook. In the case of Facebook, ‘nodesim’ ties for this position with ‘n2v’ and ‘fairwalk’. Our ‘minwalk’ algorithm typically results in Gini coefficients on the lower end, indicating its propensity to move the network towards a more equitable state with reduced connectivity disparities.

Evolution of Visibility and Gini Coefficient. Figures 3 and 4 illustrate the evolution of minority visibility and Gini coefficient across various datasets for a minority size of 15%.

In the evolution of visibility, we note that for most methods, visibility remains relatively constant, except for ‘crosswalk’, which shows a consistent increase over iterations. This rise is particularly notable in cases like ‘facebook’, where visibility exceeds 50%. Our ‘minwalk’ algorithm generally yields significantly higher visibility than other methods, except for ‘crosswalk’. However, the high and continually increasing visibility by ‘crosswalk’ is excessive for a minority size of 15%, potentially leading to backlash and risk of abandonment of the recommender algorithm by the majority. In datasets like Facebook, methods such as ‘n2v’, ‘nodesim’, ‘fairwalk’, and ‘u-ppr’ result in minority visibility below the 15% threshold, which is not ideal. Our goal is to achieve fair representation of minorities, approximately at their 15% proportion (or higher), within the top 10% of nodes ranked by PageRank.

Figure 5 presents the evolution of visibility across different minority sizes over iterations on the ‘facebook’ dataset. For algorithms such as ‘n2v’, ‘nodesim’, and ‘fairwalk’, we notice a relatively flat trend in visibility changes, generally staying below the corresponding minority percentage. A notable deviation is seen with ‘fairwalk’ at the 25% minority size, where it markedly boosts visibility beyond 25%. On the other hand, 'u-ppr' shows impressive visibility results for the 25% minority size, but this visibility significantly drops below the respective minority percentages for other sizes. ‘Crosswalk’ displays a concerning trend, seemingly ignoring minority size and frequently increasing minority visibility past the 50% threshold. In contrast, our ‘minwalk’ algorithm initially exhibits somewhat elevated visibility levels, yet it progressively adjusts to more appropriate levels as time progresses.

Crosswalk and Nodesim: Varying Parameters. Finally in Table 2, we study how varying the parameters of ‘crosswalk’ and ‘nodesim’ impact the visibility of the 15% minority. For both systems, altering parameters does not significantly change the results concerning minority visibility. Specifically, Crosswalk tends to increase minority visibility excessively, often exceeding the 15% mark. In contrast, NodeSim, despite parameter changes, maintains minority visibility at modest levels, typically below the 15% threshold. Similar behaviour was also observed for the clustering and gini coefficients.

6 Conclusions

Our study investigated the impact of fairness and diversity-aware algorithms on network structures, focusing on minority visibility and network dynamics. ‘Cross-

α / p	1.0	2.0	3.0
0.1	0.560794	0.526055	0.491315
0.5	0.538462	0.493797	0.531017
1.0	0.508685	0.550868	0.573201

α / β	2.0	3.0	4.0
1.0	0.121588	0.086849	0.119107
2.0	0.114144	0.089330	0.104218
3.0	0.104218	0.099256	0.099256

Table 2. [Left.] Crosswalk: Visibility of 15% minority for different values of its α and p parameters. **[Right.]** Nodesim: Visibility of 15% minority for different values of its α and β parameters.

walk' significantly enhances minority visibility, though it sometimes overemphasizes it, while 'nodesim', 'fairwalk', and 'u-ppr' often underperformed compared to 'n2v'. We developed 'minwalk' as a balanced solution, effectively increasing minority visibility proportionately and mitigating backlash risks. All algorithms, except 'u-ppr', decreased the Gini coefficient, indicating better network equity.

References

1. Gediminas Adomavicius and Alexander Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *TKDE*, 17(6):734–749, 2005.
2. Eytan Bakshy, Solomon Messing, and Lada A. Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015.
3. Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.
4. Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.
5. Stephen P Borgatti and Daniel S Halgin. Analyzing social networks. In *SAGE Publications Sage UK: London, England*, 2011.
6. Òscar Celma and Pedro Herrera. A new approach to evaluating novel recommendations. In *SIGIR*, pages 379–386. ACM, 2010.
7. Nicholas A Christakis and James H Fowler. *The spread of obesity in a large social network over 32 years.* New England J. of Medicine, 2007.
8. Federico Cinus, Marco Minici, Corrado Monti, and Francesco Bonchi. The effect of people recommenders on echo chambers and polarization. In *AAAI Conference on Web and Social Media*, pages 90–101, 2022.
9. Fernando Diaz, Michael Gamon, Jake Hofman, Emre Kıcıman, and David Rothschild. Online and social media data as a flawed continuous panel survey. In *PloS one*, volume 13, page e0190804, 2018.
10. Fatemeh Esfahani, Venkatesh Srinivasan, Alex Thomo, and Kui Wu. Nucleus decomposition in probabilistic graphs: Hardness and algorithms. In *ICDE*, pages 218–231. IEEE, 2022.
11. Lisette Espín-Noboa, Claudia Wagner, Markus Strohmaier, and Fariba Karimi. Inequality and inequity in network-based ranking and recommendation algorithms. *Scientific reports*, 12(1):2012, 2022.
12. Francesco Fabbri, Francesco Bonchi, Ludovico Boratto, and Carlos Castillo. The effect of homophily on disparate visibility of minorities in people recommender systems. In *AAAI Conference on Web and Social Media*, pages 165–175, 2020.

13. Francesco Fabbri, Maria Luisa Croci, Francesco Bonchi, and Carlos Castillo. Exposure inequality in people recommender systems: the long-term effects. In *AAAI Conference on Web and Social Media*, pages 194–204, 2022.
14. Antonio Ferrara, Lisette Espín-Noboa, Fariba Karimi, and Claudia Wagner. Link recommendations: Their impact on network structure and minorities. In *Proceedings of the 14th ACM Web Science Conference 2022*, pages 228–238, 2022.
15. Santo Fortunato. Community detection in graphs. *Physics reports*, 2010.
16. Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *WWW*, 2018.
17. Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *KDD*, pages 855–864. ACM, 2016.
18. Yang Guo, Fatemeh Esfahani, Xiaojian Shao, Venkatesh Srinivasan, Alex Thomo, Li Xing, and Xuekui Zhang. Integrative covid-19 biological network inference with probabilistic core decomposition. *Briefings in Bioinformatics*, 23(1):bbab455, 2022.
19. Joseph Howie, Venkatesh Srinivasan, and Alex Thomo. Scaling up structural clustering to large probabilistic graphs using lyapunov central limit theorem. *Proceedings of the VLDB Endowment*, 16(11):3165–3177, 2023.
20. Ahmad Khajehnejad, Moein Khajehnejad, Mahmoudreza Babaei, Krishna P Gummadi, Adrian Weller, and Baharan Mirzasoleiman. Crosswalk: Fairness-enhanced node representation learning. In *AAAI*, 2022.
21. Nikolay Korovaiko and Alex Thomo. Trust prediction from user-item ratings. *Social Network Analysis and Mining*, 3:749–759, 2013.
22. Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444, 2001.
23. Giovanna Miritello, Rubén Lara, Manuel Cebrian, and Esteban Moro. Limited communication capacity unveils strategies for human interaction. *Scientific reports*, 3(1):1950, 2013.
24. Mark EJ Newman. Modularity and community structure in networks. *Proceedings of the Nat. Academy of Sciences*, 103(23):8577–8582, 2006.
25. Shera Potka and Alex Thomo. Community structure and coherence in digital humanities works. In *2023 14th International Conference on Information, Intelligence, Systems & Applications (IISA)*, pages 1–8. IEEE, 2023.
26. Tahleen A. Rahman, Bartlomiej Surma, Michael Backes, and Yang Zhang. Fairwalk: Towards fair graph embedding. In *IJCAI*, 2019.
27. L. Rainie and B. Wellman. *Networked: The new social operating system.* MIT Press, 2012.
28. Akrati Saxena, George Fletcher, and Mykola Pechenizkiy. Nodesim: node similarity based network embedding for diverse link prediction. *EPJ Data Science*, 11(1):24, 2022.
29. Ana-Andreea Stoica, Christopher Riederer, and Augustin Chaintreau. Algorithmic glass ceiling in social networks: The effects of social recommendations on network diversity. In *WWW*, pages 923–932, 2018.
30. Jessica Su, Aneesh Sharma, and Sharad Goel. The effect of recommendations on network structure. In *WWW*, pages 1157–1167, 2016.
31. Vincent A Traag, Ludo Waltman, and Nees Jan van Eck. From louvain to leiden: guaranteeing well-connected communities. *Scientific reports*, 9(1):1–12, 2019.
32. Sotiris Tsioutsiouliklis, Evaggelia Pitoura, Panayiotis Tsaparas, Ilias Kleftakis, and Nikos Mamoulis. Fairness-aware pagerank. In *WWW*, pages 3815–3826, 2021.