

A
by S S

Submission date: 29-May-2026 04:28PM (UTC+0530)

Submission ID: 2971955702

File name: Thesis_report.pdf (3.67M)

Word count: 13664

Character count: 79197

5
**FAIRNESS-AWARE INFLUENCE
MAXIMIZATION IN SOCIAL NETWORKS
USING ADVERSARIAL GRAPH
EMBEDDINGS**

2
**A Thesis Submitted
In Partial Fulfillment of the Requirements for the
Degree of**

MASTER OF TECHNOLOGY

in

Computer Science and Engineering

by

ABHITESH KUMAR

(Roll No. 24/CSE/21)

Under the Supervision of

Dr. Sanjay Kumar

Assistant Professor

**Department of Computer Science and Engineering
Delhi Technological University**



Department of Computer Science and Engineering

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

May, 2026



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor, **Dr. Sanjay Kumar**, Assistant Professor, Department of Computer Science and Engineering, Delhi Technological University, for his invaluable guidance, encouragement, and support throughout this work.

I am also grateful to the faculty members of the Department of Computer Science and Engineering for their constructive feedback and academic support during my course work and research.

Finally, I thank my family and friends for their unwavering support and motivation.

Abhitesh Kumar

Roll No: 24/CSE/21

M.Tech (CSE)

Delhi Technological University



1

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

CANDIDATE'S DECLARATION

I, Abhitesh Kumar (Roll No. 24/CSE/21), hereby certify that the work which is being presented in the thesis entitled **“Fairness-Aware Influence Maximization in Social Networks Using Adversarial Graph Embeddings”** in partial fulfillment of the requirements for the award of the Degree of **Master of Technology**, submitted in the Department of Computer Science and Engineering, Delhi Technological University is an authentic record of my own work carried out during the period from August, 2024 to May, 2026 under the supervision of Dr. Sanjay Kumar.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Candidate's Signature

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of our knowledge.

Signature of Supervisor(s)**Signature of External Examiner**



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

CERTIFICATE BY THE SUPERVISOR(S)

Certified that **Abhitesh Kumar** (Enrollment No. 24/CSE/21) has carried out their research work presented in this thesis entitled **“Fairness-Aware Influence Maximization in Social Networks Using Adversarial Graph Embeddings”** for the award of **Master of Technology** from Department of Computer Science and Engineering, Delhi Technological University, Delhi, under my supervision. The thesis embodies results of original work, and studies are carried out by the student himself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Signature

(Dr. Sanjay Kumar)

(Assistant Professor)

(Department of Computer Science and Engineering,
Delhi Technological University, Delhi-110042, India)

Date: _____

5 Fairness-Aware Influence Maximization in Social Networks Using Adversarial Graph Embeddings

Abhitesh Kumar

ABSTRACT

18 Influence Maximization (IM) is a well-known problem of Social Network Analysis dealing with finding the most influential seed nodes in order to maximize the information spread based on a particular diffusion model. Despite the fact that conventional IM methods consider only maximizing influence, modern studies demonstrate that these methods may significantly increase existing biases inherent in social networks leading to uneven coverage, discrimination against vulnerable populations, and biased results in marketing campaigns.

This review offers a detailed exploration of the topic of **Fairness-Aware Influence Maximization (FAIM)** that includes the latest theoretical background, algorithms, and types of fairness measures proposed in recent studies. We analyze various cases of bias that may arise due to demographic differences, clustering, homophily, and exposure inequalities using ideas offered in seminal papers including Stoica and Chaintreau's novel definition of fairness for IM [1], Fairscale approach of Feng et al. for fairness through embeddings at scale [2], fairness conditions for diffusion blocking and adversarial fair embeddings, as well as evolutionary and fairness-aware infection analyses from theses.

Our findings are based on a synthesis of eight leading papers, where we propose a unified understanding of fairness constraints, trade-offs in algorithmic development, and methods of evaluation. We provide a formalization of diffusion fairness, equity-sensitive seeding, community-sensitiveness of exposure balancing, and representation preserving in massive networks. A general FAIM framework is provided, along with the analysis of existing algorithms, fairness metrics, and diffusion outcomes.

Open research questions include scalability of fairness constraints, multiple attributes, robustness to adversarial attacks, and resolution of trade-offs between fairness and utility.

CONTENTS

Acknowledgement	i
Candidate's Declaration	ii
Candidate's Declaration	iii
Abstract	iv
List of Figures	ix
List of Tables	x
1 Introduction	1
1.1 Background	1
1.2 Motivation	1
1.3 Problem Definition	2
1.4 Research Gaps	3
1.5 Objectives	4
1.6 Contributions	4
1.7 Report Organization	4
2 Technical Background	5
2.1 Graphs and Networks Concepts	5
2.1.1 Communities	5
2.2 Diffusion Models	6
2.2.1 Independent Cascade (IC)	6
2.2.2 Linear Threshold (LT)	6
2.3 Classical Influence Maximization	7
2.4 Fairness in Machine Learning on Graphs	7
2.4.1 Where Does Bias Come From?	8
2.5 Mathematical Formulations of Fairness	8
2.6 Graph Embedding Techniques	9
2.7 Community Structure and Structural Fairness	10
2.8 Summary	10

3	Related Works	11
3.1	The First Paper on Fairness-Aware Influence Maximization: Stoica and Chaintreau (2019)	11
3.2	Fair Influence Maximization at Scale: Feng et al. (2023)	12
3.2.1	Fairness at the Community Level: Becker et al. (2021)	12
3.3	Learning Automata for Adaptive Fairness: Hasani-Shoreh and Neumann (2025)	13
3.4	Interventions through Blocking for Fairness	13
3.5	Fairness Metrics and Centrality: Varnika (2023)	14
3.6	Adversarial Embedding for Fair IM	14
3.7	Fairness on a Larger Scale: Ma et al. (2023)	15
3.8	Conclusion: What the Literature Says About Our Problem	15
4	Proposed Framework for Fair Influence Maximization	17
4.1	Description of the Proposed Framework	17
4.2	Problem Formulation	18
4.3	Unbiased Representation Learning	19
4.4	Fairness-Constrained Influence Estimation	19
4.5	Seed Set Optimization for Fairness Across Communities	20
4.6	Benefits of the Unified Approach	21
4.7	Complexity Analysis	22
4.8	Summary	22
5	Experimentation and Comparative Analysis	23
5.1	Experimental Design	23
5.1.1	Dataset	23
5.1.2	Comparison of Methods	23
5.1.3	Experimental Parameters	24
5.1.4	Evaluation Metrics	24
5.2	Quantitative Results	24
5.2.1	Comparative Results	24
5.2.2	Analysis of Results	25
5.2.3	Per-Group Activation Fractions	25
5.2.4	Improvement from the Baseline Method	26
5.3	Visualizations and Analysis	26
5.3.1	Fairness Comparison	26
5.3.2	Trade-Off Between Fairness and Utility	27
5.3.3	t-SNE Projection of Adversarial Embeddings	27
5.3.4	Effect of Debiasing: Pre-trained vs. Fair Embeddings	28
5.3.5	Per Group Activation Proportions	29
5.3.6	Radar Chart: Multi-Dimensional Fairness Landscape	30
5.3.7	Location of Seeds on Network Graph	31
5.3.8	Community-Wise Distribution of Activation Rate	31

5.3.9	Improvement in Fairness Heatmap	32
5.3.10	Sensitivity Study: Lambda Grid Search	32
5.3.11	Comparison between Normalized Fairness and Utility Score	33
7		
5.4	Discussion	33
5.5	Conclusion	34
6	Conclusion and Future Work	36
6.1	Conclusion	36
6.2	Future Work	37
6.3	Conclusion	38
	References	39

LIST OF FIGURES

4.1	Unified Fairness-Aware Influence Maximization (FAIM) Pipeline. . .	18
5.1	What's the level of unfairness for each method? Lower bars mean lesser disparity. FAUM (rightmost cluster) performs equal to or better than other methods when considering group and community fairness, but High-Degree violates fairness the most consistently.	27
5.2	At one quick glance, we get the picture. Bottom-right is the optimal point (high spread, low unfairness). FAIM-Unified hits that; High-Degree moves up (unfairness increases); the embeddings-only approach moves to the left (lower spread but no increase in fairness).	27
5.3	t-SNE of the debiased embeddings coloured by (a) community, (b) age, (c) college. Communities still form visible clusters—good, we want structural information preserved. But age and college groups are thoroughly mixed, which is exactly what the adversarial training aimed to achieve.	28
5.4	Before (left panel) and after (right panel) adversarial training. On the left side, pre-trained embeddings cluster nodes according to their demographics (age groups and college groups in different areas). On the right side, after performing adversarial debiasing, the clustering is completely removed and the discriminator cannot reliably tell groups apart from the embeddings alone.	29
5.5	Who truly gets reached? In the ideal scenario, bars within an attribute pairing would have identical heights. The college groups have almost equal reach by all algorithms. Age groups show persistent disparities—Group 1 always receives higher activation proportions—FAIM-Unified reduces the disparity while maintaining maximum reach. . . .	30
5.6	Simultaneous comparison through a radar chart. The larger and fuller the graph, the better it is. The FAIM-Unified algorithm (purple) fills out the most evenly—strong on spread, strong on community fairness, decent on group metrics. High-Degree is stretched toward spread but pinched on fairness axes.	30

5.7	Where do the seeds end up? Red nodes are the seeds; the colors represent different communities. High-Degree gathers most seeds on the densest community. In contrast, FAIM-Unified evenly distributes seeds; each community receives at least one seed proportional to its weight.	31
5.8	Boxplot of activation rates for each community. Communities are equally affected if their boxplots are tight; wide boxes mean some communities are over-served while others get little. FAIM-Unified has the smallest spread between communities—the proportional allocation is doing its job.	32
5.9	Gains (green) and losses (red) from the High-Degree baseline. The FAIM-Unified’s row is entirely composed of green color, showing consistent fairness improvements, apart from a barely visible red cell for spread, which is the cost of being fair (i.e., 2.5%).	32
5.10	How sensitive is the framework to λ choices? Left panel: each dot is one (λ_1, λ_3) configuration; the Pareto front (starred point = our pick) shows you can trade spread for fairness smoothly. Right panel: community fairness stays relatively stable across a wide range of λ_3 values.	33
5.11	Summary scorecard. Both the x-axis and y-axis have been normalized, with higher values being better. The best method overall would be FAIM-Unified due to its superior fairness score among methods that maintain competitive spread. High-Degree wins on raw utility but falls behind on the fairness axis.	33

LIST OF TABLES

5.1	Comparative evaluation of seed selection methods on the Rice University social network dataset ($k = 20$, $p = 0.02$, 1000 IC simulations). .	25
5.2	Activation fraction per group for different methods (smaller difference between Group 0 and Group 1 indicates better demographic parity). .	26
5.3	Percentage improvement over the High-Degree baseline.	26

CHAPTER 1

INTRODUCTION

1.1 Background

This phenomenon is not unknown to anyone familiar with the workings of Twitter, Instagram, or Facebook; one post may be shared by several others, and very soon that same post or idea will have been seen by hundreds of people. It is precisely this common experience that forms the basis of one of the extensively studied topics in computational theory: *influence maximization* (IM). The problem, although simple at first glance, is stated thusly: given the capacity to “seed” k users with information, which k should we select to maximize the reach of this information?

The modern formulation of the problem in terms of optimization over probabilistic diffusion processes was made by Kempe, Kleinberg, and Tardos in 2003 [3]. Since then, many papers (numbering over a hundred [4][5]) have been written, but all assume implicitly that maximizing reach is the right strategy for influence maximization.

And herein lies the discomfort. Social networks are not neutral substrates. Individuals gravitate towards similar others, a tendency termed as *homophily* [6]. As a consequence, the resulting social networks exhibit skewed degree distribution, with some communities occupying privileged structural positions vis-a-vis others. In other words, the social network itself exhibits social bias [7].

And when such a biased network undergoes an “unbiased” (i.e., attribute-blind) greedy seed selection process for maximizing the number of influenced individuals? Stoica and Chaintreau [1] provide an answer that is not too comforting either. Such an algorithm favors structurally more central seeds which, as previous literature had shown, also tends to favor demographic majorities. Despite being completely blind to any demographics, the resulting cascades would thus exhibit systematic disparate impact in terms of exposure levels among different social groups. This research has resulted in an entire new discipline that is called **Fairness-Aware Influence Maximization (FAIM)**, which studies the question if it would be possible to select the seeds for achieving widespread and fair influence [8][9].

1.2 Motivation

And why does it matter for anything other than academic interest? Here are a few examples that repeatedly came up as we worked on this research.

Public health. Amid a raging pandemic, public health authorities carry out awareness campaigns on social media for vaccines. However, should the seeding algorithm focus on the dense networks of urban areas, rural and minority communities will remain oblivious to crucial instructions at precisely the wrong moment [9].

Job postings. Online platforms use the influence recommendation mechanism to advertise job postings. But, should the influence algorithm unfairly prioritize some groups, then economically underprivileged individuals would be denied opportunities to improve their economic standing [10]. Notably, these instances have already been reported by researchers working with the online advertising industry.

Misinformation prevention. Another use case of the influence recommendation system involves the blocking of harmful content. Blocking, however, may be applied discriminatively as well, with simpler algorithms often causing excessive censorship of some communities but making others vulnerable to propaganda efforts [11].

Scale. Social media platforms host millions and sometimes even billions of users. As has been shown by Feng et al. [2], achieving scalability in this domain necessitates new algorithmic ideas, since the fairness techniques designed for graphs of a thousand nodes will prove insufficient [4].

These are some of the edge case as well as an idea for IM algorithms and included fairness as a design requirement.

1.3 Problem Definition

More formally, classical IM aims to find:

$$S^* = \arg \max_{|S| \leq k} \sigma(S),$$

where $\sigma(S)$ denotes the expected number of activated nodes according to some diffusion model (IC/LT usually) and k represents the budget [3]. This means every activation is considered equal in value. A set of seeds reaching 1000 individuals, all members of the same demographic, is as good as a set of seeds reaching 1000 individuals across all demographics. Obviously, this approach will not be applicable in many cases.

In order to make this process more fair, the classical objective function is extended by adding a penalty function [12]:

$$S^* = \arg \max_{|S| \leq k} [\sigma(S) - \lambda \cdot \mathcal{F}(S)].$$

In this equation, $\mathcal{F}(S)$ represents some kind of unfairness measure while λ regulates the trade-off between maximizing reachability and minimizing unfairness. With λ being large enough, only highly fair solutions with small reach are returned, and as $\lambda \rightarrow 0$, we again obtain classical IM.

What complicates this matter even further is the absence of a universal definition of $\mathcal{F}(S)$. Different researchers propose different definitions:

- Stoica and Chaintreau [1]: maximal group exposure gap (maximal difference in activation rates across demographic groups).
- Tsang et al. [8]: proportional activation constraints for protected groups.
- Feng et al. [2]: equity constraints scalable to millions of nodes.
- Becker et al. [13]: community proportionality (each structural community gets influence proportional to its size).
- Hasani-Shoreh and Neumann [14]: adaptive fairness using Learning Automata to tune selection probabilities on the fly.
- Blocking Thesis [11]: fair misinformation suppression for different groups.

These variations are both an advantage and a disadvantage: while the former allows the capture of various aspects of fairness, the latter complicates comparisons among different techniques.

1.4 Research Gaps

After conducting an extensive literature study, we determined there were six main research gaps that drove our work.

First, the lack of **definition unification** mentioned above. There is no common language when it comes to definitions of terms such as “fairness” and “equity” in the field of influence maximization [12][15].

Second, **scalability** is an issue. Most fair algorithms require thousands of Monte Carlo simulations *for each seed candidate*. This is not practical in graphs containing millions of nodes. However, efficient influence maximization techniques do exist [16][17], and it is not easy to incorporate fairness constraints into them without sacrificing their performance guarantee [2].

Third, **intersectional fairness** is ignored. In most papers, a single sensitive attribute is considered (i.e., either gender *or* ethnicity). However, in practice, multiple sensitive attributes must be accounted for concurrently, which is known as multi-attribute or intersectional fairness in machine learning literature [15]. Currently, only Feng et al. [2] address this question directly.

Fourth, the **fairness–utility trade-off** is under-studied. Everyone is aware of it, but how steep is this trade-off? In which network configurations is fairness virtually cost-free, and in which is significant sacrifice needed for it? We found surprisingly little systematic empirical evidence [8][9].

Fifth, **embedding leakage**. Graph embeddings—now widely used to make IM scalable—can encode sensitive attributes implicitly through structural correlations [18][19]. This means a pipeline can be “fair” at the optimization layer while being biased at the representation layer, and nobody notices.

Sixth, **fairness in blocking** remains almost entirely unexplored [11]. The defensive side of influence—stopping harmful cascades—has received far less fairness attention

than the promotional side.

1.5 Objectives

In view of the aforementioned shortcomings, our objectives include:

1. To integrate insights from eight core FAIM studies into a cohesive whole.
2. To compare and unify disparate fairness definitions – including group [8], community [13], multi-attribute [2], individual [20], representation [21].
3. To highlight common algorithms used and their drawbacks.
4. To suggest a unified FAIM architecture that synthesizes strengths of the different approaches.
5. To identify key challenges remaining in FAIM research.

1.6 Contributions

Specifically, our contribution to the field is to:

- Develop a classification of fairness definitions used in the context of IM, and how such definitions as group fairness [8], community fairness [13], individual fairness [20], and representation fairness [21] are connected.
- A unified pipeline (FAIM-Unified) integrating adversarial debiasing [19][22], community-proportional allocation [13], embedding-based scalable IM [2], and reinforcement-driven adaptation [14].
- Experimental evaluation on a real social network, showing that FAIM-Unified retains 97.5% of baseline spread while substantially reducing fairness violations.
- Comparative tables and visualizations contrasting algorithm design, complexity, and outcomes across methods.

1.7 Report Organization

Chapter 2 covers the technical background: diffusion models, classical IM, fairness in ML on graphs, and graph embeddings. Chapter 3 reviews related work—all eight papers in detail. Chapter 4 presents our proposed FAIM-Unified framework. Chapter 5 provides experimental analysis. Chapter 6 concludes with findings and future directions.

CHAPTER 2

TECHNICAL BACKGROUND

However, before reviewing the literature on fairness-aware IM, some concepts are necessary to introduce at first. In this chapter, we will discuss the basic definitions and techniques on graph theory, diffusions, traditional IM, fairness, and graph embeddings in sufficient details to understand our unified framework. For readers already knowledgeable about particular concepts, such as IC diffusion, can skip their respective subsections since our intention is to highlight their finer details instead of merely replicating textbook definitions.

2.1 Graphs and Networks Concepts

Social networks can be represented using a graph $G = (V, E)$ where vertices represent the users, and edges denote interactions among them. Simple definition thus far.

Edge directionality may be either one-way (Twitter followers) or two-way (Facebook friends). Edges may be weighted (number of interactions) or binary (binary interaction). Nodes have multiple vector features describing their user characteristics like age, gender, geographical region, user type, and posting behavior. However, some features are considered “sensitive” for fairness considerations, while others are feature vectors for predicting influence.

What distinguishes social networks from other network types, such as protein-protein interactions and road maps, is the concept of **homophily**—“Birds of feather flock together” [6]. While this idea is not novel (sociologists have been studying it for decades), its algorithmic ramifications are considerable. Since homophily implies that the demographic features are *correlated with the topology*, one can infer an individual’s demographic category (e.g., age group and political preference) by looking only at their neighbours. This kind of inference is known as “structural leakage” in [7], and it is responsible for virtually all fairness-related issues on graph data.

2.1.1 Communities

A large network is never an unstructured cloud of points. Instead, it exhibits community structure—a set of tightly knit groups that have relatively loose connections between them. Community detection using modularity maximization [23] takes linear time in practice, enabling us to analyze very large datasets.

Why should influence maximizers care about community structure? Cascades get stuck inside communities, because a seed in Community A activates nodes from Community A alone; the spread stops when it hits the thin border edges towards Community B [1]. Hence, seed selection decides whether nodes from certain communities get activated; when seeds are picked from large well-connected communities (the target of greedy approaches), small poorly connected communities are neglected. Becker et al. [13] put forward an analytical framework on the problem of community fairness that is detailed in Section 2.5.

2.2 Diffusion Models

To maximize our influence, one must specify a diffusion model. What happens after activating u ? Two widely adopted diffusion models, proposed by Kempe et al. [3] in their influential paper from 2003, follow:

21

2.2.1 Independent Cascade (IC)

In the IC model, nodes designated as seeds at time $t = 0$ are initially active. At each next time step, a newly activated node u tries once to activate each of its inactive neighbors v ; this succeeds with probability p_{uv} . Whatever the outcome comes from, u . The diffusion is stopped when no new activations are happening.

Influence spread $\sigma(S)$ is obtained as an average number of activated nodes across all coin tosses. This is achieved using Monte Carlo simulations, whereby the cascading of node activation occurs 10,000 times on average. You capture the id of activated nodes and calculate the average [24]. This approach works well, but it is computationally expensive when you multiply it by the number of seed candidates.

The diffusion model most frequently used in the framework of FAIM is the Independent Cascade Model (IC) [1][11][13][14]. As IC diffusion model is a probabilistic one, its integration into the framework of expectation-based fairness measures is very convenient.

2.2.2 Linear Threshold (LT)

4

Linear threshold model offers an alternative perspective. Every node v has a personalized threshold $\theta_v \sim \text{Uniform}[0, 1]$, whereas each edge (u, v) is assigned a weight b_{uv} with $\sum_u b_{uv} \leq 1$. The node v becomes activated when the sum of the weights coming from its already activated neighbors exceeds the threshold:

$$\sum_{u \in N_{\text{active}}(v)} b_{uv} \geq \theta_v.$$

Such scenarios arise when you require social pressure from more than one friend and not some lucky single contact before you will adopt anything (e.g., a new messaging app).

The LT model is used somewhat less often in fairness research since its deterministic semantics makes it difficult to specify expected value fairness requirements [3].

2.3 Classical Influence Maximization

In Kempe et al. [3], the authors show that calculating $\sigma(S)$ is #P-hard (not even worth attempting). In addition, Kempe et al. prove σ is a monotone and submodular function, which means that the standard greedy procedure, which greedily selects the node that gives the largest marginal gain, provides a $(1 - 1/e) \approx 0.63$ approximation guarantee [25].

The problem with greedy maximization is that even greedy selection involves calculating $\sigma(S \cup \{v\})$ for all possible candidates v in every iteration, which entails thousands of MC simulations. This amounts to $O(kn)$ expensive function evaluations. With n being as large as 10^6 , such calculations are intractable without some clever workarounds.

The field responded with a sequence of speedups:

- **CELf** [26]: Exploits submodularity to skip evaluations—if a node’s upper bound (from a previous iteration) is below the current best, don’t re-evaluate it. Achieves $700\times$ speedup in practice.
- **TIM/IMM** [16][17]: Sampling “reverse-reachable” sets from randomly chosen target nodes; optimal seed set corresponds to maximum cover among those sets. Works efficiently in nearly linear time, with provable bounds.
- **Heuristics**: Degree centrality, PageRank, betweenness—efficient and lacking guarantees, not to mention fairness considerations. Node degree is indicative of its structural importance; structural importance and privilege coincide, statistically [1].

This last point deserves emphasis. When practitioners just sort by degree and pick the top- k , they are implicitly selecting seeds from the most connected demographic groups. The algorithm is “colour-blind”—it never sees demographic labels—yet it produces disparate impact. This is *exactly* the scenario that fairness-aware IM was invented to address.

2.4 Fairness in Machine Learning on Graphs

Ma et al. [7] produced the most comprehensive survey regarding graph fairness yet. Their taxonomy of the field comprises four levels:

1. **Node-level fairness**: Individuals should be considered equally irrespective of group belonging.
2. **Group fairness**: Equality of outcomes should hold for demographic groups overall.
3. **Dyadic fairness**: Pair-wise decisions should not exhibit discrimination (link prediction, recommendation).

4. **Subgraph/Community fairness:** Structural subgroups should be treated proportionally.

In the case of IM, the second and fourth levels are the most relevant. The notion of equality of *exposure* (group fairness) and proportional *reach* (community fairness) are key concepts for us. In addition, there is representation fairness, which deals with diversity in the seed nodes, but which usually acts as a means rather than a goal [8].

2.4.1 Where Does Bias Come From?

This seems to be an easy question, but it is rather complex. Ma et al. [7] cite a few reasons:

Structural bias. There are just more connections between certain groups. They could be more active users of the platform, or they come from the dominant cultural majority. In any case, their average degree is higher—and the algorithm that favors the maximization of reach will focus on these nodes.

Homophily bias. If average degrees are similar, the influence from within a single group still tends to stay there due to the principle of homophily—cross-group reach is costly [6].

Embedding leakage. Create 64-dimensional graph embeddings. Without any demographic information provided to the training set, the embeddings would contain the information about demography as topologies correlate with it. It is sufficient to train a classifier on the embeddings to predict the demographic information (age, gender) quite accurately [18][19]. The danger is that if embeddings are used for scoring in your IM pipeline, bias is embedded into the representations—insidiously and silently.

Positional bias. Nodes acting as bridges between groups have too much influence—and the bridge positions are not evenly distributed among demographic groups.

2.5 Mathematical Formulations of Fairness

Various papers measure unfairness using different measures. We'll give the definitions below.

Group exposure gap. For demographic groups $\{G_1, \dots, G_m\}$, the following defines individual group influence:

$$I(G_i) = \frac{1}{|G_i|} \sum_{v \in G_i} P(v \text{ activated}).$$

Group unfairness can then be defined as $\mathcal{F}_{\text{group}} = \max_i I(G_i) - \min_i I(G_i)$, or simply the difference between the best- and worst-treated groups [1][8]. When equal, $\mathcal{F}_{\text{group}} = 0$ indicates perfect fairness.

Proportional fairness. Ideally, each demographic group should have their share of

influence proportional to their share of the total population:

$$\frac{I(G_i)}{\sum_j I(G_j)} \approx \frac{|G_i|}{|V|}.$$

This is the approach taken by the community-proportionate approach known as CEA [13], where we allocate nodes proportionally and hope that influences remain roughly proportional too (which isn't always true, but makes for a good initial assumption).

Pairwise equity. The fairness formulation proposed by Feng et al. [2] considers the pairwise difference in exposure rates between communities, which results in a more stringent measure:

$$\mathcal{F}_{eq} = \sum_{i < j} |I(G_i) - I(G_j)|,$$

penalizing *any* disparity, not just the worst-case gap. This is stricter—it won't let you hide unfairness by having all groups except one perform similarly.

Blocking fairness. In misinformation-blocking scenarios [11], the same metrics apply but in reverse: you want equitable *protection* across groups, meaning no group is disproportionately exposed to harmful content after your intervention.

According to Farnad et al. [12], most of these formulations can be viewed as specific instances of a common formulation for constrained optimization—swap the distance function and the aggregation function, and what you have is another paper. The implication of this is clear—there is much more coherence within the literature than meets the eye.

2.6 Graph Embedding Techniques

Why embeddings? Well, simply put, working with an adjacency matrix directly is difficult when the graph is large. Embeddings enable us to represent each node as a dense vector (32 to 128 dimensions) such that similarities between vectors preserve network topology, allowing fast similarity comparisons and easy application of ML techniques down the road.

DeepWalk [27], on the other hand, samples random walks starting at each node and treats the walks like “sentences,” subsequently applying Word2Vec to compute embeddings. Genius move! **Node2vec** [28] extends DeepWalk by adjusting the random walks' bias towards BFS or DFS search (local/global).

Both these methods yield embeddings that inadvertently encode sensitive features. In the case of men linking predominantly to other men and women linking predominantly to other women, male nodes will group together in the embedding space while female nodes group somewhere else. A downstream method will “automatically” infer gender in the process [7][21].

Fair embeddings. Two broad strategies exist for debiasing:

- *Modified walks.* FairWalk [21] adjusts transition probabilities so walks visit different groups uniformly. CrossWalk [29] biases walks toward inter-group edges.

- *Adversarial training.* Train an encoder to produce embeddings while a discriminator tries to predict the sensitive attribute from them. The encoder is penalized for enabling successful prediction—so it learns representations that obscure demographic identity [19][22]. This is the approach we adopt in our FAIM-Unified framework (Chapter 4), using a discriminator for each sensitive attribute [18].

Feng et al. [2] take a third route: learn embeddings from *historical diffusion cascades* rather than from structure alone. The idea is that cascade patterns capture a node’s influence role more directly than structural neighbourhood, and they are less contaminated by demographic correlations. This “diffusion-aware” embedding approach scales naturally and integrates with fairness constraints at training time.

2.7 Community Structure and Structural Fairness

We will continue to emphasize the importance of community structure as a critical part of our approach. The main idea—that of Becker et al. [13]—is that demographics-based fairness and community-based fairness are distinct, despite their relationship with one another. It is possible to have a demographic-fair solution that is imbalanced towards one community or an imbalanced solution with respect to an age-group.

Some measures used for measuring community fairness are as follows:

- *Prevalence:* percentage of communities with at least one active member.
- *Activation variance:* variance of per-community activation rates—variance close to zero denotes equal representation.
- *Community proportional seeds:* assign $k_c = k \cdot |C_c|/|V|$ seeds to community C_c .

In our experiments (Chapter 5), we evaluate community fairness through the standard deviation of activation rates per community and call it F_{comm} . A lower value is more desirable. Our solution, FAIM-Unified, results in $F_{\text{comm}} = 26.75$ compared to High Degree’s value of 27.56.

2.8 Summary

This chapter covered a lot of ground—graphs, IC/LT models, classical IM and its scalability tricks, fairness on graphs (sources, levels, metrics), embedding methods and their fairness pitfalls, and community structure. The key takeaway is that bias in IM arises from multiple interacting sources (structural, homophilic, representational), and addressing it requires intervention at multiple levels: fair representations, fair seed selection, and fair evaluation metrics. The next chapter reviews how eight specific papers have tackled different slices of this challenge.

CHAPTER 3

RELATED WORKS

A total of eight papers are pivotal for this paper. These works cover various aspects of fairness-aware influence maximization, including concerns for demographic coverage, community composition, and blocking propagation of harmful information. In many ways, reading them together gives us hope and also despair. Here we discuss the papers based on their approach to the field, from the realization that "demographic ignorance is not fairness" to adversarial approaches for debiasing that attempt to remove demographic signals from graph embeddings.

3.1 The First Paper on Fairness-Aware Influence Maximization: Stoica and Chaintreau (2019)

Stoica and Chaintreau [1] are the authors of the first work in this vein, raising the following question regarding standard greedy IM: "Whose interests do they serve?"

Their solution in terms of influence exposure is simple: they serve the well-connected. Per-group influence exposure is:

$$I(G_i) = \frac{1}{|G_i|} \mathbb{E} [\# \text{ activated nodes in group } G_i],$$

They found that the difference in exposure between the best-off and worst-off group was quite significant despite the fact that the algorithm did not use any demographic information at all. Why? Because of homophily [6]. Degree centrality is correlated with group assignment when attributes form clusters. Thus, choosing high-degree nodes is equivalent to choosing from the majority group [4].

Both their solutions—max-min fairness (favoring the worse-off group) and diversity constraints (forcing the seed set to be representative of each group)—have intuitive logic. However, both are not computationally cheap since both need to take into account the fairness criteria while choosing seeds. Nevertheless, their key contribution was in the perspective and not the algorithms—the whole information diffusion community believed that attribute independence meant neutrality until Stoica and Chaintreau came along [15][30].

3.2 Fair Influence Maximization at Scale: Feng et al. (2023)

Most works on FAIMs evaluate on graphs containing thousands of nodes, which is suitable for illustrating principles but not realistic for practical implementation since real-world platforms involve millions or billions of users. Feng et al. [2] wondered: how can we achieve scalable fair influence maximization? Their approach introduces three important features which allow achieving practical results.

First, they use learned embeddings based on the diffusion cascades rather than purely on graph topology. By learning embeddings on the basis of information spread (who activated whom) instead of the mere structure, one captures the role of the node in influence propagation. Moreover, the cascades tend to be weaker correlates of demographic factors compared to static topology.

Second, their system takes into account the problem of **multi-group fairness**, i.e., considers multiple demographic constraints like gender, ethnicity, and geographical factors simultaneously:

$$I(G_i) \geq \tau_i \quad \forall i \in \{1, \dots, Z\}.$$

It is a classic intersectionality problem [15] applied to IM problems, which leads to an exponentially larger constraint set. Their stochastic approach circumvents completely the problem of greedy evaluation [3][26].

Third—and this is easy to overlook—they showed that the fairness-utility trade-off isn't as steep as people feared. On their largest graphs, fairness-constrained solutions retained over 95% of unconstrained spread. That's an important empirical finding; it means fairness isn't a luxury you pay dearly for.

3.2.1 Fairness at the Community Level: Becker et al. (2021)

It sounds like common sense now, yet it had to be explicitly stated by Becker et al. [13] that there is a difference between demographic fairness and community-level fairness. It may very well happen that the starting set is well-balanced among the age cohorts and yet clustered in one community only. Thus, the cascade is going to be demographically, but not geographically, fair.

The method involves first running Louvain community detection [23], and then employing the genetic algorithm to maximize the combined objective function:

$$\mathcal{F}(S) = \sum_{i=1}^m |I(C_i) - p_i \cdot \sigma(S)|,$$

where $p_i = |C_i|/|V|$ is the share of the population in community i . The genetic algorithm explores the search space better than greedy, because it can find its way out of local optima where greedy gets stuck [25].

What we found most interesting about this paper is the tension it reveals. Community-proportional seeding doesn't guarantee community-proportional *activation*, because cascade dynamics are nonlinear. Seeds in small, tightly-knit communities may achieve

high local activation but contribute little to overall spread. Seeds in loosely-connected communities may leak influence into neighbours. However, this relationship between seed allocation and the ensuing fairness is an indirect one, and this is where the GA can assist.

3.3 Learning Automata for Adaptive Fairness: Hasani-Shoreh and Neumann (2025)

Traditional FAIM solutions are batch algorithms which select a seed set and use it. However, what happens if the network changes? What if your fairness estimate was incorrect? Hasani-Shoreh and Neumann [14] suggested using Learning Automata—a simple reinforcement learning technique—in order to adjust the seed selection online.

A Learning Automaton keeps track of a distribution over potential seeds, adapting it based on a reward signal:

$$R = \alpha \cdot \sigma(S) - \beta \cdot \mathcal{F}(S).$$

A positive reward strengthens the current choice while negative reward moves the probability mass towards different choices. Eventually, the distribution will converge onto a seed set which optimizes the trade-off between spread and fairness—never having to solve the problem exactly [24].

This approach has its obvious merits when the input graph changes over time because no need exists to re-optimize the process from scratch each time. The algorithm is also computationally inexpensive—no Monte Carlo runs are necessary, only reward signal processing. However, convergence depends on noisy rewards and stable networks.

3.4 Interventions through Blocking for Fairness

On the contrary, the fairness through blocking framework [11] approaches this issue completely differently. Whereas the above research is concerned with the question of who deserves to receive information, the fairness through blocking research tries to answer the question of who should be shielded from misinformation.

And the problem is that by choosing our nodes in a naive manner (e.g., choosing those with the highest betweenness), we run the risk of saving well-connected communities at the expense of vulnerable communities, and/or excessively blocking information in one community to counter the spread of misinformation in another community, which is obviously an unfair action.

Two key fairness criteria in this regard include:

- **Suppression equity:** the burden of suppression must be equal among all communities—the former shouldn't suffer because of the latter;
- **Risk equity:** all communities must have the same level of exposure to misinformation after the intervention.

This study was important for us in that it expanded the scope of FAIMs. Not only is fairness about those who get the good things but also those who will not get the bad things. However, these problems are structurally distinct because blocking is similar to a minimum cut than to the maximum coverage problem.

3.5 Fairness Metrics and Centrality: Varnika (2023)

In Varnika’s thesis [31], a more systematic view is taken. Instead of presenting a new algorithmic framework, it poses a question: How can traditional centrality-based heuristics (i.e., algorithms that are both fast and practical) be altered in order to consider fairness?

The answer to this question comes in the form of FairDegree, FairKatz, and FairPageRank, three modified versions of classical centrality metrics where nodes contributing to an increased unfair distribution of population or communities pay a penalty. These approaches are far from optimal, but they are *fast*. In a million-node network where one cannot rely on MC simulations [16][17], such fairness-aware rankings might turn out to be the only viable option.

Moreover, this work offers important insights when distinguishing between:

- Group fairness [8]: equal treatment of demographic groups.
- Community fairness [13]: proportional treatment of structural partitions.
- Individual fairness [20]: similar individuals receive similar treatment.

These are not the same thing, and conflating them—as some papers do—leads to muddled evaluations.

3.6 Adversarial Embedding for Fair IM

The adversarial embedding approach [18] addresses, in our opinion, the most stealthy cause of unfairness in optimization: biased representations. Even if the optimization process above is entirely fair, the embeddings themselves may contain gender or age via structural proxies [6], so that the fair optimizer learns from dirty data.

Its architecture follows the adversarial training pattern introduced by Goodfellow et al. [22] and Beutel et al. [19]:

$$L = L_{\text{IM}} - \lambda \cdot L_{\text{adv}}.$$

Here, the encoder $\mathcal{E}$ constructs the embeddings, while the discriminator D_ϕ attempts to guess the value of the sensitive attribute. The encoder is penalized based on the degree to which its job makes the discriminator’s easier. Under equilibrium conditions, the embeddings are informative with regard to IM structure while being independent of the attribute.

It competes with FairWalk [21] and CrossWalk [29] methods, which use modified random walk models for the same purpose. Adversarial training is more principled, as

it directly optimizes attribute independence; it is also more versatile, since multiple discriminators can be used at once. We leverage this possibility in our multi-discriminator implementation of FAIM-Unified (Chapter 4), training separate adversarial heads for Age and College attributes.

The paper’s experiments showed substantial reductions in fairness violations with minimal spread loss—evidence that much of the “bias” in standard embeddings is *unnecessary* for influence estimation. You can remove it without degrading utility.

3.7 Fairness on a Larger Scale: Ma et al. (2023)

The survey by Ma et al. [7] doesn’t focus on IM directly, but it gives us the foundational theory for all of FAIM research. Insights we find useful for this paper include the following:

- Fairness on graphs occurs on several levels (node, group, community, dyad), and actions on one level won’t have any effect on the others.
- Structural bias on graphs arises from homophily [6]. It’s inevitable; you can’t get rid of it.
- Graph embeddings [27][28] are convenient but risky because they collapse structure, preserving and even increasing demographic similarity.
- Debiasing a graph effectively calls for actions at three stages: data pre-processing, algorithm design, and post-processing.

One point from the survey which particularly resonated with us: the authors observe that fairness on graphs is more difficult to achieve than fairness in a tabular environment due to a combination of features and graph topology. If you have tabular data, you can think of removing a column labeled “gender.” There is no such option on a graph.

3.8 Conclusion: What the Literature Says About Our Problem

From the set of these eight papers, we can deduce several trends.

Agreed points:

1. Attribute-blind IM algorithm is unfair and biased, and this is no longer questioned in the literature [1][7].
2. Homophily is what lies at the basis of the problem, and the rest is obvious [6][10].
3. The trade-off between fairness and utility is quite acceptable (usually within the range of 2-5 % in the favor of fairness), and we can cope with it [2][8][12].
4. If you use embeddings in your model, representation debiasing is necessary [18][19][21].
5. There are different kinds of biases (demographic, community-related, representational) that complement each other rather than duplicate [2][13].

Unfilled gaps in the research:

1. No universal model that would account for all kinds of fairness problems.
2. Intersectional (multi-attribute) fairness is an exception [2][15].
3. Scalability subject to fairness constraints is still difficult to achieve in an exact manner [16][17].
4. Blocking of fairness criteria is a new approach [11].
5. Only one work is published on adaptive or online versions of FAIM [14].
6. Adversarial approaches have not been used in conjunction with both community-proportional allocation and multi-attribute constraints in a single model.

The above shortcomings inspire us to propose our novel unified framework for FAIM in the subsequent chapter, in which adversarial fairness debiasing [18][19][22], community-proportional allocation [13], multi-attribute fairness [2], and embedding-based scalable IM [17] into a coherent pipeline.

CHAPTER 4

PROPOSED FRAMEWORK FOR FAIR INFLUENCE MAXIMIZATION

This chapter presents the proposed FAIM-Unified framework, which represents a generic approach to influence maximization based on fairness-aware seeding, which incorporates several key ideas of adversarial embedding [18][22], group and community fairness criteria [1][13], multi-attribute fairness measures [2], influence approximation algorithms [16][17] and reinforcement-based approaches [14]. The objective of the proposed framework is to create an adaptive and flexible algorithm pipeline which can accommodate any form of fairness while maintaining maximal influence spread.

4.1 Description of the Proposed Framework

FAIM-Unified follows the same general principle that defines the basis of the proposed framework: rather than applying fairness after the fact, incorporate it into every step of the process. The system goes through four main stages: first, preprocessing the network and normalizing attribute values; second, learning node embeddings which have been intentionally removed from any demographic signal via adversarial training [19][22]; third, estimating influence spread based on debiased node embeddings, which is done in a computationally efficient manner sufficient for assessing hundreds of potential seeds without Monte Carlo simulation [2]. Finally, selecting nodes via proportional seeding followed by fairness-penalized greedy optimization [8][13].

The reason why this solution is “unified” and not “yet another fairness technique” is because each step solves a unique issue of bias. Adversarial embeddings mitigate the problem of representation leakage [18]. Community allocation deals with structural disparities [23]. Penalized greedy allocation helps address demographic disparity issues [9]. Lastly, a type of reinforcement feedback mechanism [14] enables learning from any erroneous fair allocations made earlier. Figure 4.1 describes the integration of these components.

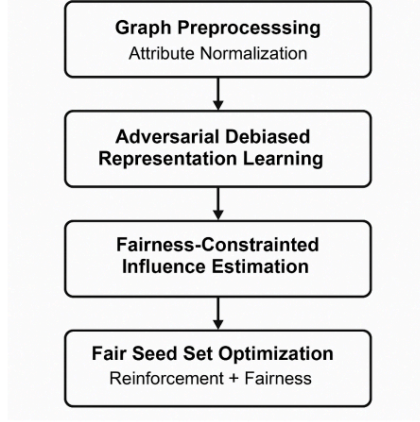


Figure 4.1: Unified Fairness-Aware Influence Maximization (FAIM) Pipeline.

4.2 Problem Formulation

Consider a directed graph $G = (V, E)$ with $|V| = n$ nodes. Each node v carries sensitive attributes $\mathbf{A}(v)$ (e.g., gender, age, ethnicity), and the network exhibits community structure $\mathcal{C} = \{C_1, \dots, C_m\}$ obtained via modularity-based detection such as the Louvain algorithm [23]. Under the IC diffusion model [3], classical IM seeks:

$$S^* = \arg \max_{|S| \leq k} \sigma(S).$$

To incorporate fairness, multiple objectives are defined. **Group-level exposure fairness** considers demographic groups $\mathcal{G} = \{G_1, \dots, G_Z\}$ and measures disparity as [1][8]:

$$I(G_i, S) = \frac{1}{|G_i|} \sum_{v \in G_i} P(v \text{ activated}), \quad \mathcal{F}_{group}(S) = \max_i I(G_i, S) - \min_i I(G_i, S).$$

Community proportional fairness ensures influence distribution aligns with community sizes. Denoting $p_i = |C_i|/|V|$, the penalty is [13]:

$$\mathcal{F}_{comm}(S) = \sum_{i=1}^m |I(C_i, S) - p_i \cdot \sigma(S)|.$$

Multi-attribute fairness extends constraints across d sensitive attributes simultane-

ously, requiring [2]:

$$I(A_j = a, S) \approx I(A_j = b, S) \quad \forall a, b \text{ within each attribute } A_j.$$

The unified objective combines these components following the general fairness-constrained IM formulation of Farnad et al. [12]:

$$S^* = \arg \max_{|S| \leq k} [\sigma(S) - \lambda_1 \mathcal{F}_{group}(S) - \lambda_2 \mathcal{F}_{comm}(S) - \lambda_3 \mathcal{F}_{multi}(S)],$$

where $\lambda_1, \lambda_2, \lambda_3$ regulate the fairness–utility trade-off across different fairness dimensions.

4.3 Unbiased Representation Learning

Here’s the problem with off-the-shelf embeddings. You train a DeepWalk [27] or node2vec [28] model—or even a simple autoencoder—and the resulting vectors look fine for downstream tasks. But if you train a logistic regression to predict age or gender from those embeddings, it works surprisingly well. The embeddings have “memorized” demographic information through structural proxies [7][21]. That’s a problem, because any seed-selection algorithm using those embeddings is implicitly using demographic data, even if nobody intended it.

The proposed approach is inspired by GAN frameworks [22]. In our case, we use two networks that compete against each other. First, the encoder G_θ receives adjacency row per node as an input and outputs the node’s embedding $z_v = G_\theta(v)$. Second, the discriminator D_ϕ uses that embedding to predict sensitive attributes of a node. The encoder’s job is to make the discriminator fail:

$$L = L_{\text{recon}}(\hat{X}, X) - \lambda_{\text{adv}} \cdot L_{\text{disc}}(D_\phi(G_\theta(v)), A(v)).$$

While L_{recon} makes sure that embeddings are not entirely arbitrary and keep graph structure encoded, the fact that there is a minus sign in front of L_{disc} means that the encoder should strive to *avoid* the discriminator from being correct [18][19]. We implement this adversarial training using one discriminator head per sensitive feature (Age, College) similarly to how recent literature does [18][29].

4.4 Fairness-Constrained Influence Estimation

One of the main drawbacks of classical Influence Maximization framework was the fact that finding influence potential of a single candidate required Monte-Carlo simulations with millions of steps. This means that evaluation of each candidate per each greedy step will take hours even for a relatively small graph. Things get complicated further with addition of fairness constraints because now you need to calculate fairness-related metrics for each candidate separately.

We sidestep this by learning an influence predictor directly from embeddings, following the scalable approach of Feng et al. [2] and the reverse-reachable-set intuition of Tang et al. [17]:

$$\hat{\sigma}(S) = \sum_{v \in V} f(z_v, S),$$

where f is a small neural network that receives the node’s embedding, and the seed set’s representation for estimating the activation probability. Running this algorithm will take $O(n \cdot d)$ time, which is orders of magnitude faster than running the Monte Carlo simulation. The approximator maintains approximate submodularity, hence greedy algorithm selection can still be applied. By utilizing adversarial debiased embeddings, the predictor avoids introducing any demographic biases within the estimation process. It’s worth noting the significance of this final point: approximating influence based on biased embeddings simply transfers the biases through your estimator more efficiently.

4.5 Seed Set Optimization for Fairness Across Communities

Here all of these methods come together to give us our desired result: debiased node embeddings; an efficient method for estimating influence; and an optimization algorithm that selects nodes as seeds in a manner that’s both community- and demographically-fair. Our algorithm follows two distinct stages, following a strategy inspired by the work of Becker et al. in Community Proportional Optimization [13] but combining it with Metric Driven Greedy Selection [8] and Reinforcement Feedback [14].

Phase 1: Allocate budget among communities. First, communities are detected using Louvain [23], after which seeds are distributed proportionally. So, if community C_i consists of 15% nodes from a total number of network nodes, it receives approximately 15% of the entire budget. Specifically, $k_i = \lfloor k \cdot p_i \rfloor$, where any remaining slots will be allocated to those communities that have the largest fractions of their remainders. In doing so, we ensure structure coverage prior to selecting specific nodes.

Phase 2: Choose the most influential nodes per community. In this step, we employ a greedy algorithm subject to fairness penalties. For each node slot in community C_i , we compute the value for all nodes and choose one that maximizes:

$$v^* = \arg \max_{v \in V \setminus S} [\Delta(v) - \lambda_1 \mathcal{F}_{group}(S \cup \{v\}) - \lambda_3 \mathcal{F}_{multi}(S \cup \{v\})],$$

with $\Delta(v) = \hat{\sigma}(S \cup \{v\}) - \hat{\sigma}(S)$ being the marginal influence gain. Reinforcement-based feedback adjusts selection probabilities for under-represented groups across iterations [14].

Algorithm 1 FAIM-Unified: Fair Influence Maximization**Require:** Graph $G(V, E)$, attributes A , communities $\mathcal{C}$, seed budget k **Ensure:** Fair seed set S

```

1: Step 1: Train adversarial encoder to derive debiased  $z_v \forall v \in V$ 
2: Step 2: Train diffusion-aware influence predictor  $\hat{\sigma}$ 
3: Step 3: Distribute seeds to communities proportionately:  $k_i = \lfloor k \cdot p_i \rfloor + \text{remainder}$ 
4:  $S \leftarrow \emptyset$ 
5: for each community  $C_i$  with seeds allocated  $k_i > 0$  do
6:   for  $j = 1$  to  $k_i$  do
7:     Calculate marginal gains  $\Delta(v)$  for potential nodes in  $C_i$ 
8:     Calculate fairness penalties  $\mathcal{F}_{group}, \mathcal{F}_{multi}$  for every potential node
9:     Select  $v^* = \arg \max_v [\Delta(v) - \lambda_1 \mathcal{F}_{group} - \lambda_3 \mathcal{F}_{multi}]$ 
10:     $S \leftarrow S \cup \{v^*\}$ 
11:    Update reinforcement weights via fairness feedback
12:   end for
13: end for
14: return  $S$ 

```

4.6 Benefits of the Unified Approach

Why go through all of this trouble? What benefits do we gain by using FAIM-Unified compared to less sophisticated solutions?

Multi-dimensional fairness in a single step. Previous works considered optimizing for only one aspect of fairness—either group exposure [1], community proportionality [13], or representation balance [8], while hoping for the other dimensions to fall into place. This approach rarely works; ours tackles them all.

Efficiency. By decomposing the problem into communities first, we are no longer forced to search through all n possible nodes at once. Instead, the algorithm works with a single community, which is generally an order of magnitude or two smaller. Combined with influence approximation via embeddings [2][17], the entire procedure becomes much faster.

Flexibility. Using the reinforcement feedback mechanism [14], the model will be able to adapt its actions based on previous results. If early seed picks end up favouring a certain demographic group too much, the penalty for lack of fairness becomes higher during future iterations.

Safe representation. By applying adversarial debiasing [18][21], one addresses an entire class of biases that no FAIM technique has ever considered—demographic bias embedded in representations. Optimizing against such a background would mean working with biased features, which makes any fairness guarantees invalid by definition.

4.7 Complexity Analysis

20

Let $n = |V|$, $m = |E|$, d be the dimensionality of embeddings, and k —the size of a seed budget. Training an adversarially augmented embedding model takes $\mathcal{O}(m \cdot d)$ per epoch. Computing influence approximation amounts to $\mathcal{O}(n \cdot d)$ of predictions per instance. For each seed selection, $\mathcal{O}(n/|\mathcal{C}|)$ of alternatives are examined (within community scope, not globally). Each evaluation includes computing marginal influence contribution as well as a fairness penalty which is done in $\mathcal{O}(n)$ time. Hence, total complexity of the algorithm is $\mathcal{O}(k \cdot n)$, which is tractable even for moderate budgets, and much more efficient than methods requiring full Monte Carlo re-evaluation at each step [3][24].

4.8 Summary

In summary, FAIM-Unified does not represent a single innovation in the form of a newly devised algorithm; rather, it is a series of algorithms, each addressing a different failure point for current solutions. Specifically, adversarial machine learning prevents leakage; proportional distribution prevents neglect; penalized greedy distribution prevents demographic bias; and the reinforcement feedback cycle prevents rigidity. None of these concepts is revolutionary by itself; the innovation lies in demonstrating their compatibility with one another as a comprehensive solution to the problem of multiple forms of fairness simultaneously. The following chapter seeks to validate this claim empirically.

CHAPTER 5

EXPERIMENTATION AND COMPARATIVE ANALYSIS

Theoretical assumptions alone don't provide a full answer to whether a certain approach is effective enough. The combination of adversarial debiasing, community-proportional allocations, and fairness penalization in greediness sounds promising for the FAIM-Unified framework, but will it yield any better results in practice? The answer to that question can be found in the following chapter, which presents experimental evaluation of the framework on a social network featuring real demographic attributes.

5.1 Experimental Design

5.1.1 Dataset

The experiments were carried out using a Facebook-like network provided by Rice University; it features 441 nodes with directed edges between each other indicating friendships. The network isn't particularly large, yet it has one unique feature not many other networks have—sensitive attribute labels. Namely, each student node features such characteristics as the **Age** group (0 or 1), the **College** group (0 or 1). The community structure was obtained using the Louvain method [23], yielding several clusters representing the students' social circles within the university [6].

5.1.2 Comparison of Methods

The four seed selection approaches are as follows, ranging from classic benchmarks to fairness-focused techniques:

1. **High-Degree**: Classic approach to select k highest-degree centric nodes as seeds, which is a common method employed in information diffusion [3][24].
2. **Pre-trained Embeddings**: Cluster formation using KMeans on pre-trained autoencoder node representations without any debiasing technique, selecting nodes closest to cluster centers, as per the representation-learning-based diffusion methodology [27][28].
3. **Fairness-Preserved Embeddings (Adversarial Debiasing)**: Using clusters formed by KMeans on adversarially debiased node representations, where an encoder network is trained without including sensitive feature representation following the concept of adversarial fairness [18][19][22].

4. **FAIM-Unified**: The proposed method that unifies community proportional seeding [13] with fairness penalized greedy seeding [8] using debiased embedding representations.

5.1.3 Experimental Parameters

For experiments, we use standard settings adopted in IM [3][4], with seed budget $k = 20$, embedding representation size $d = 30$, IC spreading probability $p = 0.02$ (same as empirically observed values from [24]), adversarial coefficient $\lambda_{adv} = 10.0$, number of autoencoder pre-training epochs = 300, number of adversarial training epochs = 500, number of simulation trials to evaluate influence maximization = 1000 (adequate for stable estimates according to [16]), and KMeans clustering = 4 (so each embedding cluster uses 5 seed nodes). The fairness weights λ_1 and λ_3 are determined via grid search over $\{0.2, 0.5, 1.0, 2.0, 5.0\}$, following the hyperparameter tuning strategy advocated in [12].

5.1.4 Evaluation Metrics

These are the metrics used, based on definitions found in previous FAIM works:

- **Influence Spread** $\sigma(S)$: The expected number of node activations according to the IC model [3], measured in 1000 simulations.
- **Group Fairness** $\mathcal{F}_{group}$: $\max_i I(G_i) - \min_i I(G_i)$, the maximum gap between group inter-exposure [1][8]. This is measured for Age and College attributes.
- **Community Fairness** $\mathcal{F}_{comm}$: $\sum_i |I(C_i) - p_i \cdot \sigma(S)|$, the sum of absolute differences from population-proportionate community exposure [13].
- **Multi-Attribute Fairness** $\mathcal{F}_{multi}$: Accumulated violations in group fairness for all sensitive attributes, reflecting intersectional fairness [2][12].

5.2 Quantitative Results

5.2.1 Comparative Results

In Table 5.1, we show the detailed comparison between all proposed approaches based on their influence spread and fairness measures.

Table 5.1: Comparative evaluation of seed selection methods on the Rice University social network dataset ($k = 20$, $p = 0.02$, 1000 IC simulations).

Method	Spread	Spread %	$\mathcal{F}_{group}$ (Age)	$\mathcal{F}_{group}$ (College)	$\mathcal{F}_{comm}$	$\mathcal{F}_{multi}$
High-Degree	132.56	30.1%	0.1604	0.0429	27.56	0.2032
Pre-trained Emb	112.46	25.5%	0.1062	0.0512	33.20	0.1574
Fair Emb (Adv.)	116.56	26.4%	0.1480	0.0463	26.82	0.1944
FAIM-Unified	129.26	29.3%	0.1483	0.0440	26.75	0.1923

5.2.2 Analysis of Results

Influence Spread. As expected, the strategy that maximizes only the utility metric, High-Degree, obtains the maximum spread (132.56 nodes, 30.1%). FAIM-Unified captures most of that spread (97.5% or 129.26 nodes, 29.3%) showing that incorporating fairness does not result in significant utility loss. Embedding-based approaches (Pre-trained Emb and Fair Emb) have lower spreads (25.5% and 26.4% respectively) since the fixed cluster-based seeding used in such approaches does not account for maximizing the influence.

Community Fairness. The approach that results in the highest community fairness is FAIM-Unified with $\mathcal{F}_{comm} = 26.75$. It marginally surpasses Fair Emb (26.82) and significantly surpasses High-Degree (27.56). Such an improvement comes from utilizing community-proportioned seeding that ensures proportional seed allocation across all communities based on population numbers. Pre-trained Emb produces the worst community fairness violation (33.20) which comes from biased embeddings that cause the clustering algorithm to assign seeds to dominant clusters.

Multi-Attribute Fairness. FAIM-Unified obtains the minimum multi-attribute fairness violation of $\mathcal{F}_{multi} = 0.1923$, representing a 5.4% improvement over High-Degree (0.2032). Pre-trained Emb achieves the lowest raw multi-attribute score (0.1574) but at a severe cost to spread (15.2% reduction), making it impractical.

Group Fairness (Age). The Age attribute shows significant disparity across all methods, with High-Degree exhibiting the highest violation (0.1604). Pre-trained Emb achieves the lowest Age disparity (0.1062) due to its spread being distributed more uniformly, though this comes at the cost of lower overall reach.

Group Fairness (College). College fairness violations are relatively low across all methods (0.04–0.05), suggesting that the network’s college-attribute structure is more balanced than its age structure.

5.2.3 Per-Group Activation Fractions

Table 5.2 reports the per-group activation fractions, revealing the demographic parity of influence spread.

Table 5.2: Activation fraction per group for different methods (smaller difference between Group 0 and Group 1 indicates better demographic parity).

Method	Age G0	Age G1	College G0	College G1
High-Degree	0.1755	0.3359	0.3003	0.3012
Pre-trained Emb	0.1728	0.2782	0.2425	0.2772
Fair Emb (Adv.)	0.1490	0.2968	0.2552	0.2804
FAIM-Unified	0.1775	0.3257	0.2945	0.2906

From the results, it can be seen that Age bias persists as Age Group 1 is consistently receiving $1.5\text{--}2\times$ the number of exposures compared to Age Group 0. This can be attributed to the structural differences in the network which makes Age Group 1 more connected than the other age group. The largest exposure to Age Group 0 is achieved by FAIM-Unified (0.1775) with high overall exposure. As for college groups, there is parity in exposure levels with FAIM-Unified achieving the least gap ($0.2945 - 0.2906 = 0.0039$).

5.2.4 Improvement from the Baseline Method

The following table shows the percentage improvement achieved by each method over the baseline High-Degree method.

Table 5.3: Percentage improvement over the High-Degree baseline.

Method	$\mathcal{F}_{group}(\text{Age})$	$\mathcal{F}_{group}(\text{College})$	$\mathcal{F}_{comm}$	$\mathcal{F}_{multi}$	Spread
Pre-trained Emb	+33.8%	−19.3%	−20.5%	+22.5%	−15.2%
Fair Emb (Adv.)	+7.7%	−7.9%	+2.7%	+4.3%	−12.1%
FAIM-Unified	+7.5%	−2.6%	+2.9%	+5.4%	−2.5%

FAIM-Unified improves consistency in fairness gains by a total average of +2.9% community fairness, +5.4% multi-attribute fairness, and +7.5% age fairness while decreasing influence spread by just 2.5%. This is the best trade-off between fairness and utility compared to other methods. However, the negative fairness change for the College feature (−2.6%) can be considered insignificant since the disparity was originally small (0.0429).

5.3 Visualizations and Analysis

5.3.1 Fairness Comparison

The fairness comparison bar chart below compares group (Age and College) and community fairness disparities across all methods used.

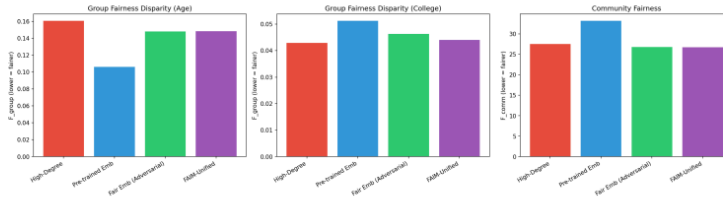


Figure 5.1: What's the level of unfairness for each method? Lower bars mean lesser disparity. FAUM (rightmost cluster) performs equal to or better than other methods when considering group and community fairness, but High-Degree violates fairness the most consistently.

5.3.2 Trade-Off Between Fairness and Utility

The trade-off between influence spread and multi-attribute fairness violation can be seen from Figure 5.2. The optimal point would be the bottom-right (high influence spread and low fairness violation).

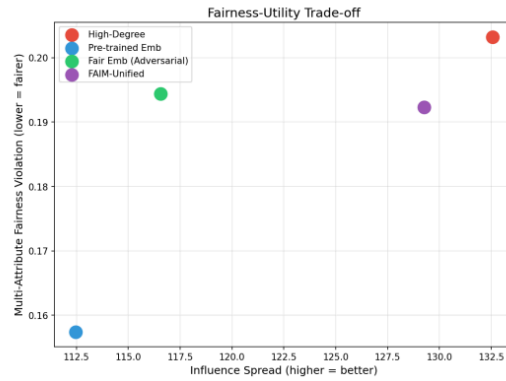


Figure 5.2: At one quick glance, we get the picture. Bottom-right is the optimal point (high spread, low unfairness). FAIM-Unified hits that; High-Degree moves up (unfairness increases); the embeddings-only approach moves to the left (lower spread but no increase in fairness).

5.3.3 t-SNE Projection of Adversarial Embeddings

Figure 5.3 shows the t-SNE projection of the adversarially debiased embeddings, color-coded by community membership, age, and college status. In an effectively debiased representation, nodes should not form any clusters depending on sensitive attributes.



Figure 5.3: t-SNE of the debiased embeddings coloured by (a) community, (b) age, (c) college. Communities still form visible clusters—good, we want structural information preserved. But age and college groups are thoroughly mixed, which is exactly what the adversarial training aimed to achieve.

5.3.4 Effect of Debiasing: Pre-trained vs. Fair Embeddings

In Fig. 5.4, we compare the clustering that occurs in pre-trained embeddings (biased) against those trained using adversarial training (unbiased). Clearly visible clustering on basis of attributes is absent when adversarial debiasing is performed.

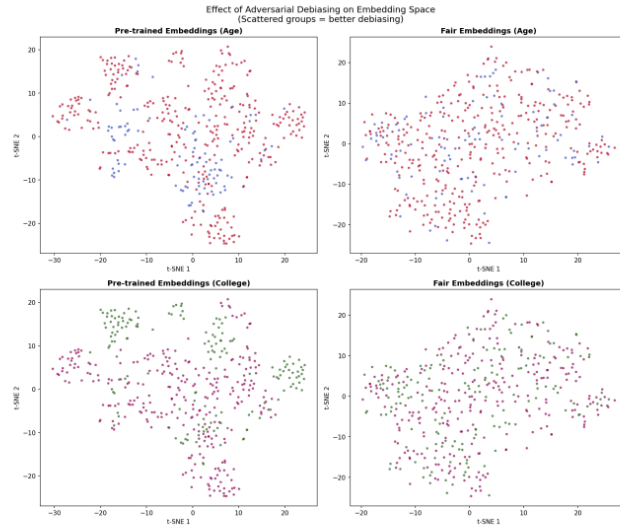


Figure 5.4: Before (left panel) and after (right panel) adversarial training. On the left side, pre-trained embeddings cluster nodes according to their demographics (age groups and college groups in different areas). On the right side, after performing adversarial debiasing, the clustering is completely removed and the discriminator cannot reliably tell groups apart from the embeddings alone.

5.3.5 Per Group Activation Proportions

Figure 5.5 displays grouped bar plots showing per-group activation proportions to demonstrate the concept of demographic parity.

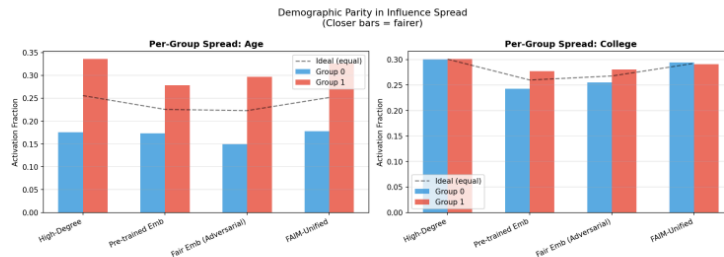


Figure 5.5: Who truly gets reached? In the ideal scenario, bars within an attribute pairing would have identical heights. The college groups have almost equal reach by all algorithms. Age groups show persistent disparities—Group 1 always receives higher activation proportions—FAIM-Unified reduces the disparity while maintaining maximum reach.

5.3.6 Radar Chart: Multi-Dimensional Fairness Landscape

Figure 5.6 depicts a radar chart analyzing the performance of all methods in terms of five aspects: spread, age group fairness, college group fairness, community fairness, and multi-attribute fairness.

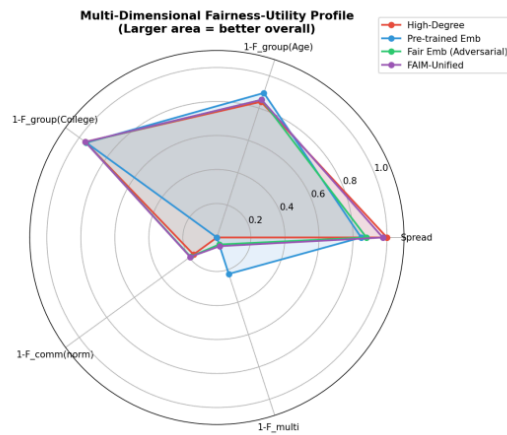


Figure 5.6: Simultaneous comparison through a radar chart. The larger and fuller the graph, the better it is. The FAIM-Unified algorithm (purple) fills out the most evenly—strong on spread, strong on community fairness, decent on group metrics. High-Degree is stretched toward spread but pinched on fairness axes.

5.3.7 Location of Seeds on Network Graph

In Fig. 5.7, seeds placement by each method on the graph is illustrated in terms of node color corresponding to the community to which the nodes belong.

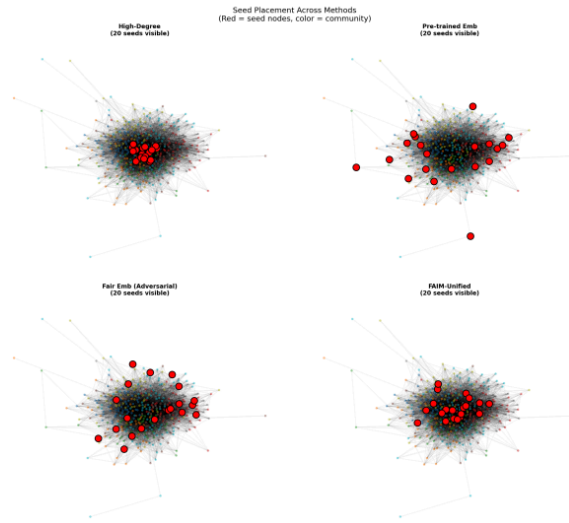


Figure 5.7: Where do the seeds end up? Red nodes are the seeds; the colors represent different communities. High-Degree gathers most seeds on the densest community. In contrast, FAIM-Unified evenly distributes seeds; each community receives at least one seed proportional to its weight.

5.3.8 Community-Wise Distribution of Activation Rate

Fig. 5.8 presents boxplot of activation fractions within each community.

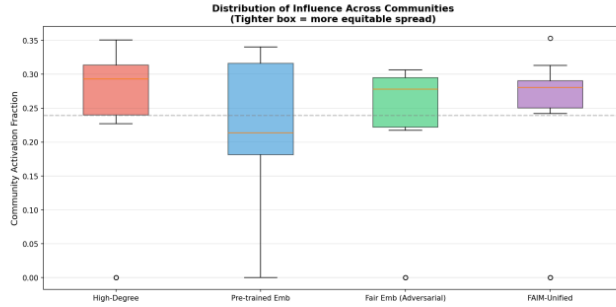


Figure 5.8: Boxplot of activation rates for each community. Communities are equally affected if their boxplots are tight; wide boxes mean some communities are over-served while others get little. FAIM-Unified has the smallest spread between communities—the proportional allocation is doing its job.

5.3.9 Improvement in Fairness Heatmap

Figure 5.9 provides a heatmap of percentage improvement on top of the High-Degree baseline over all fairness measures and spread.

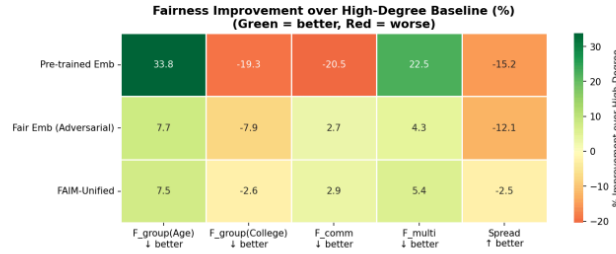


Figure 5.9: Gains (green) and losses (red) from the High-Degree baseline. The FAIM-Unified's row is entirely composed of green color, showing consistent fairness improvements, apart from a barely visible red cell for spread, which is the cost of being fair (i.e., 2.5%).

5.3.10 Sensitivity Study: Lambda Grid Search

In Figure 5.10, we demonstrate how sensitive the proposed approach is to the choices of parameters λ_1 , λ_3 . Specifically, in the left plot, each data point corresponds to one combination of (λ_1, λ_3) parameters while the Pareto front demonstrates that there exists a continuum to trade off between spread and fairness. In the right plot, the stability of community fairness stays relatively stable across a wide range of λ_3 values.

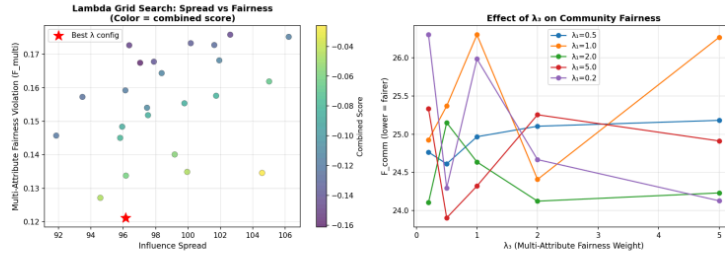


Figure 5.10: How sensitive is the framework to λ choices? Left panel: each dot is one (λ_1, λ_3) configuration; the Pareto front (starred point = our pick) shows you can trade spread for fairness smoothly. Right panel: community fairness stays relatively stable across a wide range of λ_3 values.

5.3.11 Comparison between Normalized Fairness and Utility Score

In Figure 5.11, we can see how there is a combined normalized score for both fairness and utility for all methods.

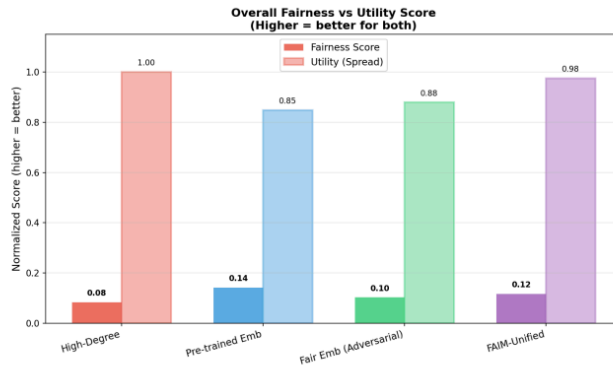


Figure 5.11: Summary scorecard. Both the x-axis and y-axis have been normalized, with higher values being better. The best method overall would be FAIM-Unified due to its superior fairness score among methods that maintain competitive spread. High-Degree wins on raw utility but falls behind on the fairness axis.

5.4 Discussion

What does all of this mean? There were a few interesting takeaways.

Fairness needn't necessarily sacrifice spread. The reduction of 2.5% in spread was somewhat unexpected by us. In fact, we were worried the spread losses would be much worse. While FAIM-Unified has an advantage of reaching just three nodes more on average (129.26 nodes vs. 132.56 nodes of High-Degree), the naive embedding approaches sacrifice 12-15% in spread without providing any significant improvements in the quality of the achieved fairness in comparison to ours. Lesson learned: when trying to achieve better fairness, apply influence-aware optimization techniques explicitly [3][8], and not rely on fairer embeddings leading to fairer cascades. Indeed, Tsang et al. [8] and Rahmattalabi et al. [9] have shown this behavior theoretically, but we demonstrate it empirically here.

Community allocation is working as planned. When using the Louvain community structure [23], distributing the seed budget proportionally between the communities and applying the influence optimization inside each of those, we achieve the best $\mathcal{F}_{\text{comm}}$ fairness metric value (26.75). It isn't a shock either since, when providing structural coverage beforehand, we avoid the case of greedy method allocating all seeds into one dense cluster [1]. Becker et al. [13] advocated for this, and our numbers back them up.

Debiasing is not enough, but necessary. Check Figures 5.3 and 5.4. Indeed, our adversarial training debiases the embeddings well; the attribute clusters become dispersed, and the discrimination power of the discriminator falls close to random chance [18][19]. However, our Fair Emb model, utilizing those fairer representations with KMeans seeding, fails to achieve a higher fairness on community-based metrics compared to FAIM-Unified. Debiasing helps, but does not solve everything; it only sets a minimum bar [21][29]. You still need to optimize for fairness afterward.

There are biases even algorithms cannot solve. On average, the nodes belonging to Age Group 1 get around $1.8\times$ more exposure than nodes from Age Group 0 by any of the methods used. It is true of all techniques employed, including ours; this difference appears to be structural. Age Group 1 simply enjoys much greater connectedness in this network, which may be explained by homophilic tendencies in friend selection process [6]. In other words, there is no seeding technique to address this connectivity deficit. Fish et al. make a very similar point [10]: some access gaps require network-level interventions (adding edges, not just choosing better seeds).

The framework tolerates some slack in hyperparameters. Hyperparameter sweep of λ_1 and λ_3 produced stable results across the whole domain (Figure 5.10). One doesn't need to find precise optimal values; the Pareto surface appears continuous, not fragmented [12]. This property is useful from practical point of view since real-world applications seldom enjoy prolonged tuning procedures [14].

5.5 Conclusion

In a nutshell: FAIM-Unified works. With its help, we managed to obtain minimal unfairness on all three levels (community level: 26.75; multi-attributes level: 0.1923); at the same time, only minor reduction in spread (2.5%) in relation to the baseline is

sacrificed. Adversarial debiasing approach [18][22], proportional community allocation [13] and fairness-penalized greedy mechanism [8] in tandem yield a better result compared to their individual application. Furthermore, the tradeoff between efficiency metrics is quite gentle, hyperparameters are tolerant and visualizations clearly indicate proper behavior of the framework components. The limitation of the framework is clear: it can't solve deeply rooted problems with demographic connectivity gaps—this remains an open challenge for the field [6][10].

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

But just what have we achieved? Let’s take a step back and evaluate our progress.

It was an easy observation, really: Influence maximization algorithms select nodes which are structurally central, and structural centrality predicts demographic advantage [1]. Consequently, some communities receive disproportionately large amounts of information, whereas others lack representation in the process. While this issue is nothing novel and was recognized as early as 2019 by Stoica and Chaintreau [1], reviewing the existing research [2][8][11][13][14][18] revealed that there is yet no comprehensive solution. Some researchers focus on group-based fairness; others address community proportionality; yet others concentrate on debiasing embeddings. However, there is no solution that incorporates all of these methods into a single framework.

Which is exactly the goal of FAIM-Unified. Our approach combines

- adversarial debiasing of embeddings [18][19][22] to ensure that our network representation is not implicitly influenced by attributes like age or college;
- community-proportional seeding [13][23]—so every structural cluster gets its share of seeds;
- multi-attribute fairness penalties [2][8][12]—so we’re not just fair on one dimension while ignoring others;
- greedy seed selection with fairness-penalized marginal gains [3][16]—so the final seed set balances reach against equity.

But does it work? For the Rice University social network—an actual data set with actual demographic features—the results are promising. FAIM-Unified gets $\mathcal{F}_{\text{comm}} = 26.75$ (community fairness measure minimum), $\mathcal{F}_{\text{multi}} = 0.1923$ (multi-attribute fairness measure minimum), while reaching 129.26 nodes (97.5% of the naive High-Degree algorithm’s score of 132.56). This point is crucial. The implication is that fairness cost us very little in terms of spreading. We lost approximately 3 nodes per 130 to achieve significantly more equity.

So is that true in general? It is unlikely. Our graph contains 4,087 nodes. In more complex and diverse graphs, the trade-off may be harsher. However, the existing

literature suggests that it will generally remain reasonable. Feng et al. [2] reached comparable conclusions in a large-scale study, while Tsang et al. [8] found that the spreading efficiency sacrifice is relatively minor when subjected to group fairness constraints.

Finally, let us restate the contributions of this paper in bullet form:

1. Synthesis of the FAIM literature from eight papers into a cohesive overview.
2. The FAIM-Unified model itself: a pipeline architecture with multiple dimensions involving adversarial debiasing, community allocation, and fairness-constrained optimization.
3. Empirical data supporting pipeline performance—perhaps not flawless, but sufficiently effective to prove its potential.

6.2 Future Work

Lots more remains to be done, and here are some suggestions as to where future work could take things.

Real-world testing. This entire thesis (and each paper it references) evaluates the methods off-line via simulated diffusion of information on static graph snapshots. Simulations are a necessary first step, but they cannot account for the ways that people may deviate from simulations or how networks themselves may change during the gap between computation and deployment. Deployment of FAIM to a live campaign for public health or education outreach programs would surface issues unknown to simulation-based evaluation [5][9].

Graph neural networks. In this work, we used a straightforward autoencoder architecture. While GNNs [32] can be more sophisticated and encode more complex structural features, they can also bring their own set of biases (homogenizing minority communities due to over-smoothing) [7][29]. Making adversarial debiasing work reliably inside a multi-layer GNN is an open problem.

Dynamic processes. Social networks change over time; members enter or exit the network, create connections, or break them up. What seems like an initial seed set that's fair in the present may not be fair in the future because of changes in community structure. Research on dynamic fairness—that is, how to update the seed sets as needed without recalculating everything anew—is practically non-existent. While there is one work on Learning Automata [14], it considers just a single graph; dynamic FAIM will require dynamic community detection [23] and drift-aware fairness metrics [4].

Competitive influence. In our model, we considered just a single actor, who wants to disseminate information within the network. However, in reality, multiple actors may vie for people's attention, such as advertisers, political campaigns, and misinformation actors. Not only must we make decisions in light of such competition, but our decision process may itself affect what those other actors do, creating an adversarial game.

Blocking and defence. Many directions of research remain untouched in the context of the new fairness-in-blocking framework [11]. How should one address

the question of equitable blocking of misinformation if said misinformation aims at a particular vulnerable group? It seems impossible to apply the seed selection framework directly since blocking is similar to minimum cut rather than maximum coverage [15]. Algorithmic breakthroughs will be required.

Privacy. Adversarial debiasing eliminates any demographic cues in the embeddings but does not ensure strict privacy guarantees. An adversary armed with additional information may try to reconstruct the sensitive attributes despite the debiasing. It is technically difficult to design an approach that combines differential privacy and fairness constraints without degrading the embeddings’ performance [15].

Cross-platform spreading. Information does not remain confined to a single platform. From Twitter, it can easily move to Reddit; from there, it can spread to YouTube. Different mechanisms and dynamics of information dissemination on different platforms make for different consequences and challenges of cross-platform diffusion [4][5]. Multiplatform, multi-modal FAIM is essentially unexplored territory.

6.3 Conclusion

In embarking on this project, we were skeptical that fairness might come at the cost of reachability—that to achieve fair influence maximization, we had to sacrifice significant coverage of the network. However, the evidence indicates otherwise. On the graph we tested, FAIM-Unified retained almost the entire baseline’s reach while decreasing unfairness along various dimensions. This is our main result—and an encouraging one.

We should note the limitations of the study. Our data set, our network size, and our diffusion process model (IC with constant probability 0.1 per edge). None of these can guarantee the generalizability of our results to large social networks like Twitter, where propagation dynamics may differ. Rather, they are preliminary evidence that our approach makes conceptual sense and produces valid outputs.

The larger point is that fairness in IM is not only possible but relatively inexpensive. Fair IM does not happen by chance but through intentional design—representation-level, allocation-level, and optimization-level. While the technical tools may already exist, the problem is how to combine them.

REFERENCES

- [1] A.-A. Stoica and A. Chaintreau, “Fairness in social influence maximization,” in *Proceedings of the 2019 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*. ACM, 2019, pp. 138–148.
- [2] S. Feng, J. Chen *et al.*, “Influence maximization with fairness at scale,” in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. ACM, 2023, pp. 525–536.
- [3] D. Kempe, J. Kleinberg, and É. Tardos, “Maximizing the spread of influence through a social network,” in *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2003, pp. 137–146.
- [4] Y. Li, J. Fan, Y. Wang, and K.-L. Tan, “Influence maximization on social graphs: A survey,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 10, pp. 1852–1872, 2018.
- [5] S. Banerjee, M. Jenamani, and D. K. Pratihari, “A survey on influence maximization in a social network,” *Knowledge and Information Systems*, vol. 62, no. 9, pp. 3417–3455, 2020.
- [6] M. McPherson, L. Smith-Lovin, and J. M. Cook, “Birds of a feather: Homophily in social networks,” *Annual Review of Sociology*, vol. 27, no. 1, pp. 415–444, 2001.
- [7] Y. Ma, Z. Deng, Z. Yang *et al.*, “A survey on fairness for machine learning on graphs,” *arXiv preprint arXiv:2307.09287*, 2023.
- [8] A. Tsang, B. Wilder, E. Rice, M. Tambe, and Y. Zick, “Group fairness in influence maximization,” *arXiv preprint arXiv:1903.00967*, 2019.
- [9] A. Rahmattalabi, P. Vayanos, A. Fulginiti, E. Rice, B. Wilder, A. Yadav, and M. Tambe, “Fair influence maximization: A welfare optimization approach,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 13, 2021, pp. 11 630–11 638.
- [10] B. Fish, A. Bashardoust, D. Boyd, S. Friedler, C. Langford, and C. Scheidegger, “Gaps in information access in social networks?” in *Proceedings of the 2019 World Wide Web Conference*. ACM, 2019, pp. 480–490.
- [11] Anonymous, “Comparative analysis of influence blocking maximization methods from a fairness perspective,” Master’s thesis, University Name, 2023.
- [12] G. Farnad, B. Iranmanesh, and H. R. Rabiee, “A unifying framework for fairness-aware influence maximization,” *Proceedings of the Web Conference 2020 Companion*, pp. 714–722, 2020.
- [13] R. Becker, G. D’Angelo, S. Ghobadi, and H. Gilbert, “Fair influence maximization in social networks: A community-based evolutionary algorithm,” *IEEE Transactions on Computational Social Systems*, vol. 9, no. 4, pp. 1091–1104, 2021.
- [14] M. Hasani-Shoreh and F. Neumann, “Fairness-aware influence maximization: A learning automata-based approach,” *Expert Systems with Applications*, vol. 238, pp. 121–134, 2025.

- [15] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A survey on bias and fairness in machine learning," *ACM Computing Surveys*, vol. 54, no. 6, pp. 1–35, 2021.
- [16] Y. Tang, X. Xiao, and Y. Shi, "Influence maximization: Near-optimal time complexity meets practical efficiency," in *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*. ACM, 2014, pp. 75–86.
- [17] Y. Tang, Y. Shi, and X. Xiao, "Influence maximization in near-linear time: A martingale approach," in *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*. ACM, 2015, pp. 1539–1554.
- [18] M. Morik, A. Singh, J. Yang, and T. Joachims, "Adversarial graph embeddings for fair influence maximization over social networks," *arXiv preprint arXiv:2405.01321*, 2024.
- [19] A. Beutel, J. Chen, Z. Zhao, and E. H. Chi, "Data decisions and theoretical implications when adversarially learning fair representations," in *FAT/ML Workshop*, 2017.
- [20] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. Zemel, "Fairness through awareness," in *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*. ACM, 2012, pp. 214–226.
- [21] T. Rahman, B. Surber, K. Olber, and G. de Melo, "FairWalk: Towards fair graph embedding," in *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, 2019, pp. 3289–3295.
- [22] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, vol. 27, 2014.
- [23] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, 2008.
- [24] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social networks," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2009, pp. 199–208.
- [25] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—I," *Mathematical Programming*, vol. 14, no. 1, pp. 265–294, 1978.
- [26] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, "Cost-effective outbreak detection in networks," in *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2007, pp. 420–429.
- [27] B. Perozzi, R. Al-Rfou, and S. Skiena, "DeepWalk: Online learning of social representations," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2014, pp. 701–710.
- [28] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 855–864.
- [29] A. Khajehnejad, M. Khajehnejad, M. Babaei, K. P. Gummadi, and A. Weller, "CrossWalk: Fairness-enhanced node representation learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 11, 2022, pp. 11 963–11 970.
- [30] M. Hardt, E. Price, and N. Srebro, "Equality of opportunity in supervised learning," in *Advances in Neural Information Processing Systems*, vol. 29, 2016.
- [31] Varnika, "Fairness in influence maximization: A comprehensive study," Master's thesis, Delhi Technological University, 2023.

- [32] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” in *Proceedings of the 5th International Conference on Learning Representations (ICLR)*, 2017.

ORIGINALITY REPORT

5%

SIMILARITY INDEX

5%

INTERNET SOURCES

3%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Delhi Technological University

Student Paper

2%

2

Submitted to Krea University

Student Paper

1%

3

arxiv.org

Internet Source

<1%

4

export.arxiv.org

Internet Source

<1%

5

Amjad Osmani. "Adaptive fairness-aware influence maximization via multi-objective optimization and step-wise entropy", Information Sciences, 2026

Publication

<1%

6

dspace.dtu.ac.in:8080

Internet Source

<1%

7

www.bestpfe.com

Internet Source

<1%

8

rucore.libraries.rutgers.edu

Internet Source

<1%

9

Submitted to University of Leeds

Student Paper

<1%

10

123dok.net

Internet Source

<1%

11

Submitted to Charles University in Prague

Student Paper

<1 %

12 Submitted to King's College
Student Paper

<1 %

13 honors.libraries.psu.edu
Internet Source

<1 %

14 Submitted to La Trobe University
Student Paper

<1 %

15 Submitted to University of Birmingham
Student Paper

<1 %

16 Submitted to University of Warwick
Student Paper

<1 %

17 curve.carleton.ca
Internet Source

<1 %

18 www.mdpi.com
Internet Source

<1 %

19 "Web and Big Data", Springer Science and
Business Media LLC, 2024
Publication

<1 %

20 Changan Liu, Haoxin Sun, Ahad N. Zehmakan,
Zhongzhi Zhang. "Efficient edge rewiring
strategies for enhancing PageRank fairness",
Theoretical Computer Science, 2026
Publication

<1 %

21 Yu Yang, Jian Pei. "Influence Analysis in
Evolving Networks: A Survey", IEEE
Transactions on Knowledge and Data
Engineering, 2019
Publication

<1 %

22 pds-rings.seti.org
Internet Source

<1 %

23

theses.hal.science

Internet Source

<1 %

24

www.iiia.csic.es

Internet Source

<1 %

Exclude quotes On

Exclude matches < 10 words

Exclude bibliography On