UC Berkeley

UC Berkeley Previously Published Works

Title
Influence Maximization with Fairness at Scale

Permalink
https://escholarship.org/uc/item/30v888tv

Authors
Feng, Yuting
Patel, Ankitkumar
Cautis, Bogdan
et al.

Publication Date
2023-08-06

DOI
10.1145/3580305.3599847

Influence Maximization with Fairness at Scale

Yuting Feng
University of Paris-Saclay & CNRS LISN
yuting.feng@universite-paris-saclay.fr

Ankitkumar Patel
University of California - Berkeley
ankitkumar@berkeley.edu

Bogdan Cautis
University of Paris-Saclay & CNRS IPAL Singapore
bogdan.cautis@universite-paris-saclay.fr

Hossein Vahabi
University of California - Berkeley
puyavahabi@berkeley.edu

ABSTRACT

In this paper, we revisit the problem of *influence maximization with fairness*, which aims to select k influential nodes to maximise the spread of information in a network, while ensuring that selected sensitive user attributes (e.g., *gender*, *location*, *origin, race*, etc) are fairly affected, i.e., are proportionally similar between the original network and the affected users. Recent studies on this problem focused only on extremely small networks, hence the challenge remains on how to achieve a scalable solution, applicable to networks with millions or billions of nodes. We propose an approach that is based on learning node representations (embeddings) for fair spread from *diffusion cascades*, instead of the social connectivity, and in this way we can deal with very large graphs. We propose two data-driven approaches: (a) fairness-based participant sampling (FPS), and (b) fairness as context (FAC). Spread related user features, such as the probability of diffusing information to others, are derived from the historical information cascades, using a deep neural network. The extracted features are then used in selecting influencers that maximize the influence spread, while being also fair with respect to the chosen sensitive attributes. In FPS, fairness and cascade length information are considered independently in the decision-making process, while FAC considers these information facets jointly and takes into account correlations between them. The proposed algorithms are generic and represent the first policy-driven solutions that can be applied to arbitrary sets of sensitive attributes at scale. We evaluate the performance of our solutions on a real-world public dataset (Sina Weibo) and on a hybrid real-synthetic dataset (Digg), which exhibit all the facets that we exploit, namely diffusion network, diffusion traces, and user profiles. These experiments show that our methods outperform the state-the-art solutions in terms of spread, fairness, and scalability.

CCS CONCEPTS

• **Information systems** → **Social recommendation**; **Social advertising**; • **Human-centered computing** → **Social media**; *Social recommendation*; • **Networks** → *Social media networks*.

KEYWORDS

information diffusion, fairness, representation learning

1 INTRODUCTION

Social media plays a crucial role in connecting people in the Internet era, allowing everyone to exchange information easily. Importantly, social networks such as Facebook, Twitter, or Instagram can be used as a powerful medium to carry out information diffusion campaigns of various kinds, e.g., in marketing and political scenarios. Indeed, social media advertising is a booming domain, gradually replacing advertising over the more traditional channels (TV, radio, print, mail, etc). It is enabled by the highly effective word-of-mouth mechanisms that are embedded in social applications.

Social advertising campaigns are usually carried out with a limited *budget*, and so, the challenge is on how to maximize the spread of information (influence) accordingly. The family of algorithmic problems under the generic name of *influence maximization* (IM) [15] pertains to all these scenarios that aim to maximize information spread in a diffusion network under budget constraints. It aims to select the k most influential nodes from which the diffusion of a specific message should start. In many ways, IM mirrors an increasingly popular and effective form of marketing in social media, which targets a sub-population of *influential people*, instead of the entire base of users of interest, known as *influencer marketing* [8].

Furthermore, recent trends have motivated advertisers to seek more control over who may receive their messages, as opposed to simply maximizing the spread of information. Along this line, there is an increasing need for information diffusion in a fair way. Indeed, many campaigns – e.g., sharing public opinions, job or loan advertisements, news sharing, etc – may require not only a large spread, but also *fairness* in influencing users across various sensitive attributes like race, gender, or location, in order to achieve unbiased viewpoints. Therefore, a recent focus has been on how to maximize fairness along with the influence spread [6, 10, 12, 25, 28]. Yet the challenge remains on how to achieve a scalable solution for fair influence maximization, while enabling a flexible formulation that can support one or multiple sensitive attributes.

In this paper, we generalize the influence maximization problem by modeling the objective as a function of spread and fairness across a set of sensitive attributes. To quantify fairness, we propose an

analytical model supporting a set of attributes with n-ary categories. For this problem, by exploiting the historical diffusion cascades and the users' fairness information, we then propose two deep-learning based solutions: (a) fairness-based participant sampling (FPS), and (b) fairness as context (FAC). Importantly, we achieve scalability by building our predictive models from the observed diffusion cascades, instead of relying on computationally intensive simulations of diffusions. In FPS, node embeddings and diffusion probabilities are derived while considering fairness and information cascade independently. On the other hand, FAC adopts joint learning between fairness and information cascades, in order to derive the node embeddings and diffusion probabilities. Finally, in both algorithms, diffusion seeds are selected through a fair-greedy algorithm using the node embeddings and diffusion probability information.

In summary, our main contributions are as follows:

- We propose a new formulation for fair influence maximization and a model supporting a set of categorical attributes to quantify fairness. Our model is based on the notion of equity, i.e., ensuring that nodes from different groups of interest are (almost) equally likely to be influenced.
- We describe two algorithms for this problem, called FPS and FAC. These deep-learning based solutions are applicable to arbitrary sets of sensitive attributes and can scale to large, realistic social graphs. We achieve scalability by exploiting the diffusion cascades information to solve the fair influence maximization problem, as opposed to social connectivity and computationally expensive diffusion simulations.
- We propose a model-independent fairness-aware greedy algorithm to select the seed set that maximizes influence, based on the learnt graph representation, which comes with a knob to control the degree of aversion to inequality.

We evaluate these algorithms over a real-world and publicly available dataset (Weibo), and a hybrid real-synthetic dataset (Digg), in which we randomly allocate user features based on real feature distributions. By comparing with the state-of-the-art methods, we show that FPS and FAC can maximize the number of activated nodes while achieving the required fairness w.r.t. the sensitive attribute(s).

2 RELATED WORK

Influence maximization (IM) usually has as objective the *expected spread* under a stochastic diffusion model, which describes diffusions as probabilistic processes. The work of [15] introduced two such discrete-time diffusion models, Linear Threshold and Independent Cascade, which have been adopted by most of the literature that followed (see for instance the survey of [20]). Such models rely on diffusion graphs with edges weighted by a score of influence, i.e., a spread probability. Since selecting the seed nodes maximizing the expected spread is NP-hard under such diffusion models, greedy / approximation graph algorithms that rely on the monotonicity and sub-modularity of the objective have been studied extensively. However, scaling IM algorithms to large, realistic networks remains difficult. Indeed, most of the IM research focuses on improving efficiency and scalability (see the benchmarks of [3, 4]).

Influence maximization with cascades. IM studies may also differ in the underlying data assumptions. The work of [2] has argued recently that algorithms based on stochastic models that rely only on the connectivity information lead to sub-optimal results, as information related to the users' diffusion activities (information cascades) is ignored. In other studies (e.g., [9, 26]), the authors considered exploiting the historical cascades along with social connectivity information, in order to derive probabilities of influence under IC or LT assumptions. However, the simplifying assumption of diffusion independence may lead to inaccurate estimation since the correlations across participating nodes in the cascades are ignored. Recently, [7, 22] addressed IM by extracting influencer and susceptible *embeddings* from historical cascades and used such learned models to make spread predictions and to solve IM. Such deep-learning based solutions avoid the independence assumption and consider high-order correlations between users.

Graph algorithms for fair IM. As the gap between theoretical models / algorithms for IM and real-world diffusion scenarios with realistic assumptions reduces gradually, recent studies have considered the *fairness* of diffusion campaigns, by revisiting graph algorithms for IM. In [28], the authors describe two initial notions for group fairness, *maximin* (inspired by the legal notion of disparate impact) and *diversity* (every group receiving influence commensurate to what it could have generated on its own). They show that these fairness objectives – while non-submodular – can be reduced to multi-objective sub-modular optimization, and the authors describe a method for general multi-objective sub-modular optimization. In a similar vein, [12] studies the maximin criterion, presents hardness results, and analyses several adaptations of the IM greedy strategy. Similar to the two aforementioned works, [6] focuses on maximin fairness, considering probabilistic strategies for seed selection, which leads to approximating solutions for fair IM with provable guarantees, close to those for classic IM.

For a broader scope, in [10], a generic framework based on integer programming is shown to capture various flavors of group fairness in IM such as *equality*, *equity*, *maximin*, or *diversity*. As these cannot be jointly enforced, in practice one must chose based on the application scenario. [25] also provides a broad formal study on desirable fairness properties by (i) transposing existing principles (e.g., monotonicity) from social welfare theory on the "utility vectors" that describe spread across groups of interest, and (ii) proposing new principles specific to group fairness, such as the *utility gap reduction*. They propose a model of welfare functions, which allows to tune the fairness vs. spread trade-off and enables a generic algorithm, applicable for various fairness definitions. Impossibility results are also discussed, such as for jointly obeying the utility gap reduction along with some of the other principles.

Among other related graph-based algorithms on fairness and information diffusion, [1] considers a time-constrained spread formulation, along with IM algorithms for an objective function that combines spread and maximum disparity. In [6], the focus is on maximizing the diversity of the information that reaches users, instead of the diversity of the users who get a particular piece of information. As such, this is orthogonal to the goals of fair IM. [27] studies how the diversity of the seed set improves fairness of spread. [21] takes another approach for fair IM, by sampling first under fairness constraints from the initial graph. The work of [13] starts from the observation that maximizing the influence for one group may come at the cost of reducing it for other groups. They propose

a model allowing to explicitly choose the desired balance between the groups' spread objectives, where all objectives except one are turned into constraints, while the remaining objective is optimized subject to the constraints. In short, constraints stipulate that any IM solution must yield at least a given fraction of the optimal influence.

These graph algorithms for fair IM have major limitations, that we aim to address in this paper, such as (i) working exlusively with the diffusion topology, often completed by synthetic diffusion probabilities, which has limited practical relevance, (ii) often relying on fairness criteria that in practice may still lead to unbalanced spread, and (iii) failing to achieve a good trade-off between fairness and spread in a scalable manner. Indeed, scalability remains a key issue for practical purposes, as discussed for instance in [10]. The reason is that these state-of-the-art approaches work by (i) adapting existing IM algorithms for greedily selecting seeds from the diffusion graph, and (ii) by relying on computationally expensive procedures (e.g., Monte-Carlo or Reverse-Influence sampling). Moreover, such adaptations must also cope with the fact that fairness-aware IM no longer benefits from the submodularity of the influence objective.

Learning embeddings for fair IM. By exploiting cascades and node features – two data facets often available in diffusion scenarios – we propose to *learn* node embedding models allowing us to solve instances of the fair IM problem in a manner that is both efficient and flexible w.r.t. the spread effectiveness vs. fairness trade-off.

Along this research path – learning node representations for fair IM – the works that are closest to ours are [16, 17]. In [17], the authors propose an adversarial neural network-based approach. First, a set of influencers are identified using a variant of K-means clustering over node embeddings obtained by the adversarial neural network. Then, the IC model is used to identify a diverse influencer set. The follow-up work [16] revisits generic algorithms for learning node representations (e.g., DeepWalk [24], node2vec [14]) in order to "boost" fairness. This is done by biasing random-walks in the vicinity or across boundaries between the user groups for which fairness must be achieved. By assigning larger weights to edges that are at groups' peripheries, they obtain embeddings by which the selected seeds can spread the information more effectively to multiple groups. The limitations of [16, 17] are threefold: (i) as they exploit the social connectivity, they incur a high computation cost and cannot reasonably scale; in particular, training the adversarial neural network in [17] grows exponentially with the number of users in the social network, (ii) they assume that the probability of diffusion between a pair of nodes is constant and follows the Markovian assumption, and (iii) they are not generic, i.e., not applicable for fairness across any and multiple sensitive attributes.

Compared to [16, 17], we revisit IM with fairness under a more flexible formulation, by algorithmic solutions that are applicable to arbitrary sets of sensitive attributes and to large, realistic datasets. Our methods alleviate therefore the main practical limitations of the state-of-the-art solutions, including the one of scalability.

Finally, fairness-aware deep representation learning – e.g., for users in a recommender system [29] – have been proposed for unbiased representations of users, s.t. sensitive attributes cannot be inferred from the model's output. Here, the neural networks are used to learn the influencing and fairness aptitude of influencers

from the cascades, aspects which pertain to the user base (population) that may be influenced, and not to the influencers themselves.

3 PRELIMINARIES

In the fair IM literature, akin to algorithmic fairness studies, several notions of *group* fairness have been studied. We briefly discuss them next, motivating our fairness model choice in the process.

- *Equality*: conceptually straightforward, it refers to the fair allocation of seeds by the IM algorithm, i.e., each group getting the same number of seeds. Therefore, this design does not necessarily prevent unfair spread.
- *Maximin*: its goal is to minimize the gap between groups / categories w.r.t. the received influence relative to their size. Focusing on the minimum influence among the groups relative to size may drastically diminish the overall spread, e.g., when one group is less well-connected than the others.
- *Diversity*: the crux of this fairness model is to ensure that each group / category of users receives influence at least equal to their internal spread, i.e., in the graph induced by that group. Under this formulation, diversity is not "inequality averse", as it may lead to unbalanced spread.
- *Equity*: similar to maximin, equity is a model that aims to reach a fair share of influence, this time by ensuring that any node's probability of getting influenced is (almost) the same, regardless of the group it belongs to. So a group's expected number of influenced nodes should be proportional to its ratio in the graph. Equity corresponds to what is also known as *demographic parity* (DP) in algorithmic fairness studies.

Our fairness model is based on the equity notion, which strictly corresponds to the ideal fairness for a sensitive categorical attribute s (or a combination thereof) and the groups it induces in the graph by C_s, the set of categories of s. To evaluate the fairness of a set of selected influencers with respect to a set of sensitive attributes, we next derive an analytical expression to quantify fairness.

Proposed fairness model. Under the notion of equity, we consider categorical attributes that are of interest in achieving fairness. In what follows, we discuss fairness for a single sensitive attribute s, but this extends easily to a set (combination) of categorical attributes. For a fair spread, we want the proportion of users belonging to a particular category (or *group*) of s in the entire population to remain the same after that spread, i.e., in the influenced population. For example, if we have a 70% *male* population (the groups being *male* / *female*), we want, after the spread, the influenced nodes to be made of 70% *males*. This would be the ideal fairness per sensitive attribute and categories thereof. However, if we cannot achieve this, we need to assess the deviation from the ideal situation. In a spread context, for each categorical value $i \in C_s$ (e.g., *gender: female*) we will have two sets of users of interest: (1) the users $V_i \subseteq V$ belonging to category i in the entire population (e.g., *all women in V*), and (2) the influenced users (after the spread) belonging to category i, denoted by $\Omega_i \subseteq V_i$ (e.g., *all influenced women*).

For an attribute s with multiple categories C_s, the ideal fairness of the diffusion results would be

$$\frac{|\Omega_i|}{|V_i|} \approx \frac{|\Omega_j|}{|V_j|} \approx \frac{|\Omega_k|}{|V_k|} \approx \dots \forall i, j, k, \dots \in C_s. \tag{1}$$

We have perfect fairness when the proportion of influenced nodes in all the categories (groups) of s is the same. Hence the *coefficient of variation CV* can be used here to evaluate the degree of variation (unfairness) of influenced ratios for all the categories, computed as:

$$CV = \sigma/\mu, \tag{2}$$

where σ, the standard deviation of the influenced ratios, is:

$$\sigma = \sqrt{\frac{\sum_{i \in C_s} \left(\frac{|\Omega_i|}{|V_i|} - \mu\right)^2}{|C_s|}}, \tag{3}$$

and μ denotes the average of influenced ratios:

$$\mu = \frac{1}{|C_s|} \sum_{i \in C_s} \frac{|\Omega_i|}{|V_i|}. \tag{4}$$

With the relative dispersion of influenced users in the groups induced by s to capture unfairness, the fairness score can then be scaled by a *sigmoid* function and bounded between 0 and 1, by[1]

$$f_s = \frac{2}{1 + \exp(CV)}. \tag{5}$$

4 PROBLEM DEFINITION

We formulate the fair IM problem to optimize both influence spread and fairness as follows. We are given a social network $G(V, A)$, where V is a set of nodes representing social network users and A is the set of edges representing the social connectivity between users (e.g., derived from users' contacts and following information). Additionally, we have a set S of categorical, sensitive user attributes, such as *gender*, *race*, *location*, etc., which are of interest when aiming for spread fairness. Each node (user) will have a combination of categories pertaining to this set of attributes associated to it.

Besides the social structure and the users' sensitive attributes, we also represent the information from *diffusion cascades D* (logged traces of information spread in the network), with each cascade $d \in D$ being a set of pairs (v, t_v) where t_v is the timestamp when the information was diffused by user $v \in V$. In each cascade $d \in D$, the user having the smallest timestamp is the initiator of the cascade, while all the other users are participants. Let also I denote the overall set of initiator nodes (influencers) from the cascades D.

As diffusions do not necessarily follow the social topology, we aim to learn and represent diffusion probabilities between nodes based on the cascades, with both spread and fairness as objectives. Our problem is to find a set of k influencers that maximizes jointly the spread (number of influenced nodes) and its fairness.

5 ALGORITHMIC SOLUTIONS

We propose two algorithms, Fairness-based Participant Sampling (FPS) and Fairness As Context (FAC)[2], to address the fair IM problem. The crux of these proposals is to extract the influencing aptitude and fairness of users from cascades and to represent this information in a multidimensional space using a neural network. The representation of users' features in a latent space is referred to as node embedding [11]. Moreover, diffusion probabilities between pairs of cascade initiators (influencers) and cascade participants (influencees) are

derived using the neural network. Finally, the node embedding and diffusion probabilities are used to select a set of influencers by a fair-greedy approach. In the FPS approach, node embedding and diffusion probabilities are derived by considering users' spread aptitude and fairness independently. FAC jointly considers users' spread aptitude and fairness, to derive node representations and diffusion probabilities. We describe these algorithms next, w.l.o.g limiting the description to a single sensitive attribute s.

5.1 Fairness-based Participant Sampling (FPS)

Information about the influencers' aptitude in influencing others can be derived from historical cascades. FPS explicitly penalizes the spread aptitude of an influencer as a function of the influencer's fairness. Fig. 1 (left) shows the schematic representation of the FPS model. The original context is exclusively extracted from historical cascades, where the input feature is the initiator of a cascade and the output labels are the set of participating nodes in the cascades. The key idea is to first oversample the original context of a cascade $d \in D$ by $\eta\%$ and then downsample the oversampled context based on (a) the temporal information of actions taken by the participating nodes in the cascade, and (b) fairness of the cascade initiator. The rationale is that quickly responsive participating users are highly likely to be influenced. Thus, the probability of sampling a node v is inversely proportional to the elapsed time $t_v - t_u$ since d was initiated by node u. From the cascade d, the probability of sampling node v from the initiator node u is defined as follows:

$$p(v|d) = \frac{(t_v - t_u)^{-1}}{\sum_{(x, t_x) \in d} (t_x - t_u)^{-1}} \tag{6}$$

Let f_u^s be the fairness of an influencer u w.r.t. a sensitive attribute s, obtained as in Eq. (5) when restricting to the cascades initiated by u. The context is further downsampled by selecting $|d| \times f_u^s$ participating nodes, where $|d|$ is the length of cascade d and the fairness score is used as a penalty factor to downsample the influencers who started unfair information diffusions w.r.t. the attribute s.

Next, the sampled context is input to the deep neural network, to learn about the influencers' spread aptitude and the diffusion probabilities for influencer-influencee pairs. The input context for a given cascade d with sample length l is $X_d = \left\{ \left(x, y^1\right), \ldots, \left(x, y^l\right) \right\}$, for $x \in R^{|I|}$ and $y^j \in R^{|V|}$. Recall that I is the set of cascade initiators in the training dataset and V is the set of all the nodes in the social graph. A node in the context is represented using one-hot encoding. The first hidden layer contains the number of neurons equivalent to the embedding size $|E|$, and thus, the parameters at the first hidden layer are $\theta \in R^{|I| \times |E|}$. The number of neurons at the second hidden layer is equivalent to the number of nodes in the social network, and thus, the parameters in the second hidden layer are $T \in R^{|E| \times |V|}$. The number of nodes at the output layer is $|V|$ with the softmax activation function.

The network is trained to minimize the cross-entropy loss[3] function L^n. Eq. (7), (8), (9) give the analytical operations of the model. At iteration n, z_u^n denotes coefficients at the output of second hidden layer for influencer u, and o_u^n is the softmax transformation at the

---

[1]By definition, fairness is above $\sqrt{n-1}/n$, for n the number of categories of the sensitive attribute or combination thereof. E.g., *gender* fairness is always above 0.5.
[2]Code and experiments available at github.com/goldenretriever-5423/fair_at_scale.

[3]We use Noise-Contrastive Estimation (NCE) loss, which is efficient when dealing with such large softmax problems; it is fundamentally a negative sampling method to approximate the multi-class problem to a binary one.
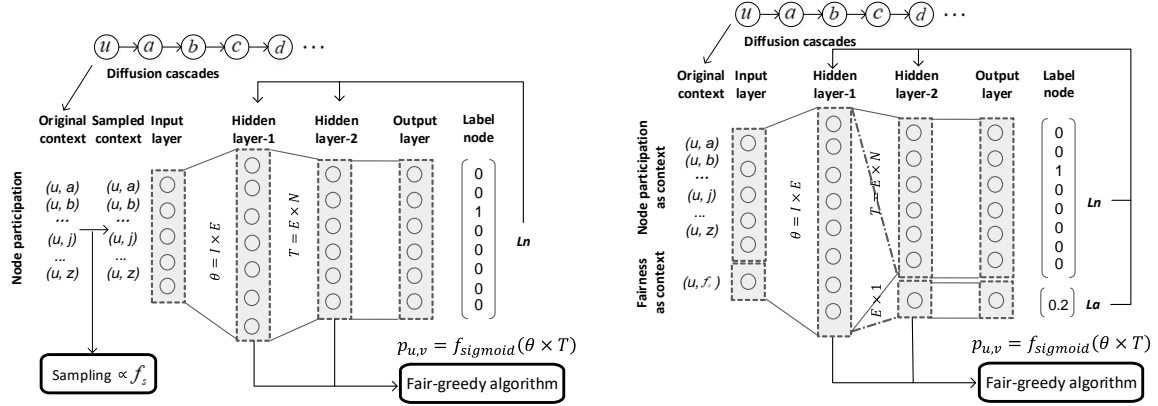
Figure 1: Illustration of our algorithmic solutions: FPS (left) and FAC (right).

output layer for user u.

$$z_u^n = \theta_u \times T + b^n \tag{7}$$

$$o_u^n = \frac{e^{z_u^n}}{\sum_{v \in V} e^{z_v^n}} \tag{8}$$

$$L^n = Y^n \log(O^n) \tag{9}$$

The parameters of the first hidden layer represent the influence embedding of cascade initiators. The second norm of an embedding captures the aptitude of a user in influencing other users. Applying the fairness-based penalty in generating contexts penalizes the spread aptitude of biased influencers. On the other hand, the parameters of the second hidden layer represent the susceptible embedding of users. Applying the sigmoid transformation to the product of both, influence and susceptible embedding results in diffusion probabilities $p_{u,v}$ that are defined as follows:

$$p_{u,v} = f_{sigmoid}(\theta \times T). \tag{10}$$

A diffusion probability $p_{u,v}$ is therefore the probability of node v appearing in cascades started by u. Such probabilities can be used to identify influenced nodes, with the assumption that there exists direct or indirect connectivity between users u and v. Therefore, unlike exploring all possible connections between users u and v [11], the implicit assumption of path existence drastically reduces the computational complexity, and the algorithm becomes scalable.

5.2 Fairness as Context (FAC)

The FAC approach jointly considers the fairness of influencers and the size of cascades initiated by them to derive embeddings and diffusion probabilities. As shown in Fig. 1 (right), the model contains two separate neural networks with a common hidden layer, where the common hidden layer captures the users' embeddings. The goal of the top neural network is to capture the influence aptitude of users, in the form of embedding and diffusion probabilities. The input to the neural network is one-hot encoded sampled contexts. Unlike the FPS approach, the original context is just sampled based on the temporal information of actions taken by users, using Eq. (6). For a given cascade d with length l, the input context is $X_d = \left\{ \left(x, y^1\right), \ldots, \left(x, y^l\right)\right\}$, for $x \in R^{|I|}$ and $y^j \in R^{|V|}$. The first hidden layer contains the number of neurons equivalent to the embedding size $|E|$, and thus, the parameters at the first hidden layer are $\theta \in$

$R^{|I| \times |E|}$. The number of neurons at the second hidden layer is equivalent to the number of nodes in the social network, and thus, the parameters in the second hidden layer are $T \in R^{|E| \times |V|}$. The number of neurons at the output layer is $|V|$ with the softmax activation function. The first neural network is trained with the objective of minimizing the cross entropy function L^n; its analytical operations are described in Eq. (7), (8), and (9).

The bottom neural network aims to capture the effects of fairness in the form of embedding and diffusion probabilities. The input to the neural network is $Z_j = \left(x_j, f_j^s\right)$, $x_j \in R^{|I|}$ and $f_j^s \in R^{|I|}$, where f_j^s is the fairness of influencer j w.r.t the attribute s, as defined before. The first hidden layer contains the number of neurons equivalent to the embedding size $|E|$, and thus, the parameters at the first hidden layer are $\theta \in R^{|I| \times |E|}$. The number of neurons at the second hidden layer is one, and thus, the parameters in the second hidden layer are $T \in R^{|E| \times 1}$. The output of the neural network contains a single neuron with sigmoid activation function. The objective of the second neural network is to minimize the mean square error, L_a. The analytical operations of the bottom neural network can be described as follows: at iteration n, z_u^n denotes coefficients at the output of second hidden layer for influencer u, o_u^n is the transformation at the output layer for influencer u, F^n is the fairness of influencers in a vector form, and L_a^n is the loss function.

$$z_u^n = \theta_u \times U + b^n \tag{11}$$

$$o_u^n = \frac{1}{1 + e^{-z_u^n}} \tag{12}$$

$$L_a^n = \left(F^n - O^n\right)^2 \tag{13}$$

Recall that, in the FAC architecture, the first hidden layer is common across both neural networks, where its parameters represent the users' embeddings. The key idea for designing a common hidden layer is to generate embeddings, while capturing the correlations between users' spread aptitude and fairness related features. After training the neural networks, the embeddings and diffusion probabilities are used by the greedy influencer selection algorithm discussed before, to identify the set of influencers.

5.3 Fair-greedy influencer selection

By the FPS and FAC models, the norm of the learned embeddings captures the aptitude of an influencer to spread information to

many nodes and in a fair way, while the inner product of the embeddings for an influencer-influencee pair gives the diffusion probability thereof. Based on these two ingredients, we describe in Algorithm 1 our fairness-aware greedy algorithm, which selects the k influencers who jointly maximize influence and fairness for the chosen sensitive attribute(s). Alg. 1 adapts [22]'s IMINFECTOR approach for influence maximization based on node embeddings learned from cascades (itself an adaptation of the CELF algorithm for IM [19]), by adding the fairness dimension to it.

In Alg. 1, $D \in \mathbb{R}^{|I| \times |V|}$ is the diffusion probability matrix, $F \in \mathbb{R}^{|I|}$ gives the average fairness score for the diffusion cascades initiated by each influencer u, and $\lambda[u]$ gives the expected influence spread of u, as the fraction of V nodes expected to be influenced by u based on the norm of embedding $\|\theta\|$:

$$\lambda[u] = \frac{\|\theta_u\|_2 \times |V|}{\sum_{u' \in I} \|\theta_{u'}\|_2}. \tag{14}$$

For the marginal gain computation at each step, the greedy algorithm works with the set of remaining, yet-to-be-influenced nodes (initially V itself). At each step, it identifies / updates for each yet-to-be-selected influencer u a number of $\lambda[u]$ yet-to-be-influenced nodes (lines 4, 20), namely those that maximize the cumulative diffusion probability from u, denoted Ω, i.e., those associated with the top $\lambda[u]$ diffusion probabilities from u (lines 5, 21).

Then, the influencer u with the largest marginal gain – which is a linear combination between Ω and the average fairness observed in u's cascades (lines 11, 23) – is selected and added to the seed set. Its top $\lambda[u]$ influenced nodes are then removed from the yet-to-be-influenced ones (line 17). In the marginal gain computation, the parameter α represents the aversion to unfairness, allowing us to tune the trade-off fairness vs. influence, after min-max scaling.

As in [22], the cumulative diffusion probability Ω, marginal gain, and set consisting of u's top $\lambda[u]$ yet-to-be-influenced nodes are updated lazily, when selecting some influencer u (lines $19 - 23$).

The complexity of Alg. 1 is $O(k|I||V| \log |V|)$. As with CELF, the algorithm benefits from (i) the fact that $|V|$ diminishes at every iteration, and (ii) much fewer influence spread evaluations than I are done (lines $19 - 23$).

6 EXPERIMENTS

We evaluated the proposed algorithms on two datasets. Weibo is the largest Chinese microblogging platform, and the Weibo dataset [30] is a publicly available one built for information diffusion studies. It includes all the facets of interest for our work, namely a diffusion graph, the information cascades, and user profiles. Despite our best effort to find a similar, real-world and publicly available dataset, we could not identify one containing these data facets, in particular ground-truth cascades and user profiles. Therefore, to reduce the risk of validation biases, we built a hybrid real-synthetic dataset, based on the publicly available Digg2009 dataset [18]. Digg is a platform where users follow others, (re)post stories, and vote on posts. This dataset contains all the necessary facets, except for the user profiles. Hence we complemented it by synthetically generated user profiles (sensitive attributes), which overall fit publicly-known distributions in Digg. The main data statistics are in Table.1. The experiments were done on a server Ubuntu 20.04.4 LTS Intel CPU E5-2620 v2 @ 2.10GHz, with 188GB RAM and Tesla k40m GPU.

Algorithm 1: Fair-greedy selection of spread seeds

Input: Probability diffusion matrix $D \in \mathbb{R}^{|I| \times |V|}$, cascade based fairness vector $F \in \mathbb{R}^{|I|}$, expected fair spread vector $\lambda \in \mathbb{R}^{|I|}$, seed set size k, parameter α

Output: seed set $\mathcal{S}$

1   $Q \leftarrow [\ ], \mathcal{S} \leftarrow \emptyset, L \leftarrow [0 : |V|]$

2   **for** $s = 0; s < |I|; s + +$ **do**

3     $R \leftarrow \text{argsort}(D[s, L])[0 : \lambda[s]]$

4     $\Omega \leftarrow \text{sum}(D[s, R])$

5     $Q.\text{append}([s, \Omega, F[s], R, 0])$

6   $\Omega_{ub} \leftarrow \max(Q, 1), \Omega_{lb} \leftarrow \min(Q, 1)$

7   $F_{ub} \leftarrow \max(F), F_{lb} \leftarrow \min(F)$

8   **for** $q \in Q$ **do**

9     $q[1] \leftarrow F_{lb} + (q[1] - \Omega_{lb})/(\Omega_{ub} - \Omega_{lb}) \times (F_{ub} - F_{lb})$

10   $Q \leftarrow \text{sort}(Q, \text{lambda } q \in Q: (1 - \alpha) \times q[1] + \alpha \times q[2])$

11   **while** $|\mathcal{S}| < k$ **do**

12     $q \leftarrow Q[0]$

13   **if** $q[4] == |\mathcal{S}|$ **then**

14       $\mathcal{S}.\text{add}(q[0])$

15       $R \leftarrow q[3]$

16       $L \leftarrow L - R$

17       $Q.\text{delete}(q)$

18   **else**

19       $R \leftarrow \text{argsort}(D[q(0), L])[0 : \lambda[q[0]]]$

20       $\Omega \leftarrow \text{sum}(D[q[0], R])$

21       $Q[0] \leftarrow [q[0], \Omega, F[q[0]], R, |\mathcal{S}|]$

22       $Q \leftarrow \text{sort}(Q, \text{lambda } q \in Q: (1-\alpha) \times q[1] + \alpha \times q[2])$

23   **return** $\mathcal{S}$

Table 1: Datasets description.

	Weibo	Digg
influencers	19,698	532
influencees	1,170,688	336,225
social links	225,877,632	2,617,993
posts	62,087	3,553
median size of cascades	50	528
maximal size of cascades	35986	24099

6.1 Baseline methods and setup

The state-of-the-art methods on fair IM based on graph algorithms [6, 10, 12, 25, 28] cannot scale to large diffusion scenarios with thousands-to-millions of users. In contrast, by using the ground-truth diffusion cascades instead of the social connectivity, and by learning node representations based on them, our models are able to deal with social graphs and information diffusion at large scale. The same scalability limits are exhibited by the two state-of-the-art methods on fair IM that do rely on node representation [16, 17], yet learned only from the social connectivity. For all these methods, their respective empirical validation results, performed on extremely small datasets, are clear indicators of their high computation cost and thus limited applicability. Therefore, for a comparison
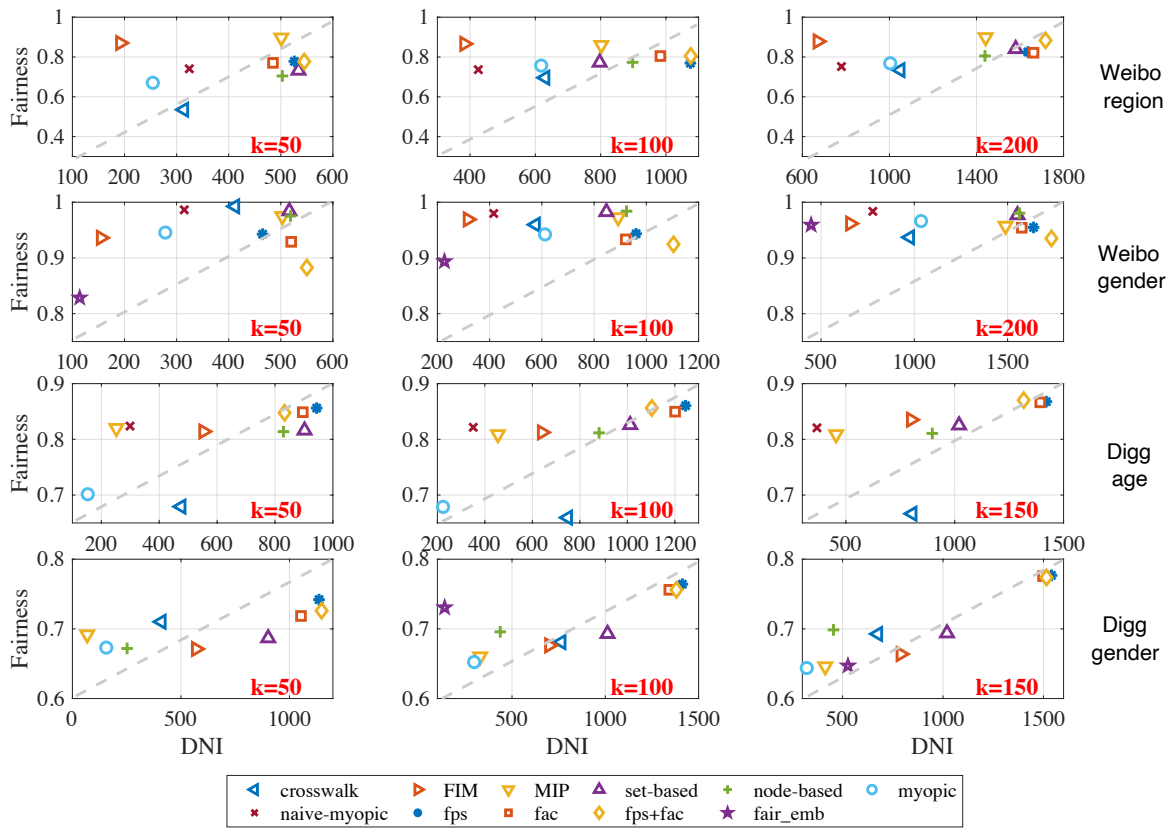
Figure 2: Validation on the sampled datasets from Weibo and Digg (the closer to the upper-right corner the better).

with these methods, we sampled smaller representative Weibo / Digg datasets, with thousands of nodes, on which their running time remained nevertheless quite high (mostly order of hours).

However, we also experimented with the complete datasets, comparing with IM models that are (i) applicable to large social graphs, (ii) exploit diffusion cascades (as we do), but (iii) are fairness agnostic. This enabled us to showcase that the fairness achieved by our models, at realistic scale, does not come at the expense of spread.

Fair IM models (on sampled datasets). **Crosswalk** [16] is a random-walk based graph representation method, which enhances fairness by re-weighting the edges between nodes from different groups. We initialize the edge weights based on the in-out degree and, with the resulting embeddings, the potential seeds are selected via clustering. **FairEmbedding** [17] uses an adversarial network for graph embedding, with a discriminator to discern the sensitive attributes. By design, it is only applicable for attributes with binary discrimination, so its experiments are only carried out on *gender*. The potential seeds are also selected by clustering based on the resulting embeddings. The graph algorithms denoted **Node-based** / **Set-based** from [6] use probabilistic strategies over nodes and communities to select seeds for fair IM under the maximin formulation. Their edge weights are also initialized based on in-out degrees. **FIM** [25] proposes a framework based on social-welfare theory to mitigate the fairness vs. spread trade-off, with a single

inequality-aversion parameter; it employs the notion of fairness under equity (our case) when the inequality-aversion parameter is set to 0. **MIP** [10] achieves fairness under equity in IM by adding linear constraints or formulating the objective function as an integer program. Finally, **Myopic** / **Naive-myopic** [12] are heuristic graph algorithms that select seeds based on maximin group fairness; the former iteratively chooses seeds based on minimum probability of reaching groups, while the later choose the seeds at once.

We also compare with **FPS+FAC**, a hybrid solution where we take the down-sampling strategy of FPS as input, while learning the fairness score as in FAC, in the output of network, in order to investigate their potential to jointly improve performance.

Fairness-agnostic IM methods (on complete datasets). Since the fair IM models cannot be evaluated on the complete datasets, we look instead at fairness-agnostic IM methods using the diffusion cascades to identify and select the spread seeds. **IMINFECTOR** [22] designs a multi-task graph neural network, which takes as input the cascades to generate node embeddings, by which both diffusion probabilities (between influencers and influencees) and cascade sizes can be predicted. A seed set is selected based on these, by a CELF-like greedy algorithm. Our work draws inspiration from [22] in both the way node representations are learned from cascades, and in the subsequent greedy algorithm for fair IM. **CELFIE** [23] is a model preceding **IMINFECTOR**, where the network is trained on

cascades to learn the node embeddings, but not the spread aptitude. Finally, we consider a data-based heuristic, **AvgCascades** [5], which ranks the seeds based on the average size of their diffusion cascades.

Setup. We split the cascades into train (60%), validation (20%), and test (20%) subsets, by occurrence time. For the deep neural networks, the embedding dimension $|E|$ is set to 50, the models are trained for 10 epochs, with a learning rate of 0.1. The original context is over-sampled by $\eta = 120\%$. In FPS, to avoid sparsity and to ensure that most influencers can be used in the training stage, we applied a threshold on the maximum penalization exerted on those having low fairness, such that at least 3 cascades are kept for training. In the fair-greedy seed selection, the parameter α indicating the aversion to unfairness is by default set to 0.2. As the sensitive attributes, in Weibo we consider *gender* and *region*, while in Digg we consider *gender* and *age*. The *gender* distribution is around 53% *female* and 47% *male* in Weibo, respectively 35% *female* and 65% *male* in Digg. Performance is evaluated on fairness and spread, based on the test cascades. Spread is given by the DNI metric (Distinct Nodes Influenced) [9], as a union over the nodes that participate in the test cascades initiated by the selected seeds.

6.2 Validation on sampled datasets

The results are presented in Fig. 2 where, for each sub-plotting, the x axis represents DNI and the y axis represents fairness. To give a clearer view on the differences in performance, we only keep the fairness region where most models lie, which means that those with significantly worse performance may be omitted in some plots.

We can see that, as the size of the seed set increases, both the number of influenced nodes and the fairness score goes up for almost all the models. Our models generally outperform the other baselines, especially on the influence dimension (DNI). As for fairness, our models remain consistently in the top-tier, except when applied to Weibo *gender*, a case which we will further explore next.

Among the baselines, **Set-based** and **Node-based** generally outperform the other baselines, with stable performance on fairness and influence, proving the effectiveness of selecting seeds with randomization. **FIM** and **MIP** perform generally well on fairness, but often fail in maximizing the spread. This may be due to the fact that their focus is on fairness under various flavors (among which the one we focus on, equity), by a generic formal and algorithmic solution, which may come at the expense of influence spread.

As to the embedding-based methods, **Crosswalk** and **FairEmbedding**, their algorithms for selecting the seeds are clustering based, leading to rather unstable performance along both dimensions of interest. In that regard, our fair-greedy selection algorithm after the embedding process proves to be much more effective.

Among the models we propose, FPS performs relatively worse in Weibo, while in Digg they are performing almost equally well. As the down-sampling strategy as a penalization on unfair influencers in FPS is independent from the fairness score learned in FAC, we do not necessarily benefit from the combination of these two methods.

6.3 Validation on the complete datasets

As the most related state-of-the-art methods are not scalable on the realistic datasets, we further validate our solutions by comparison with generic, cascade-based IM models, in order to showcase their
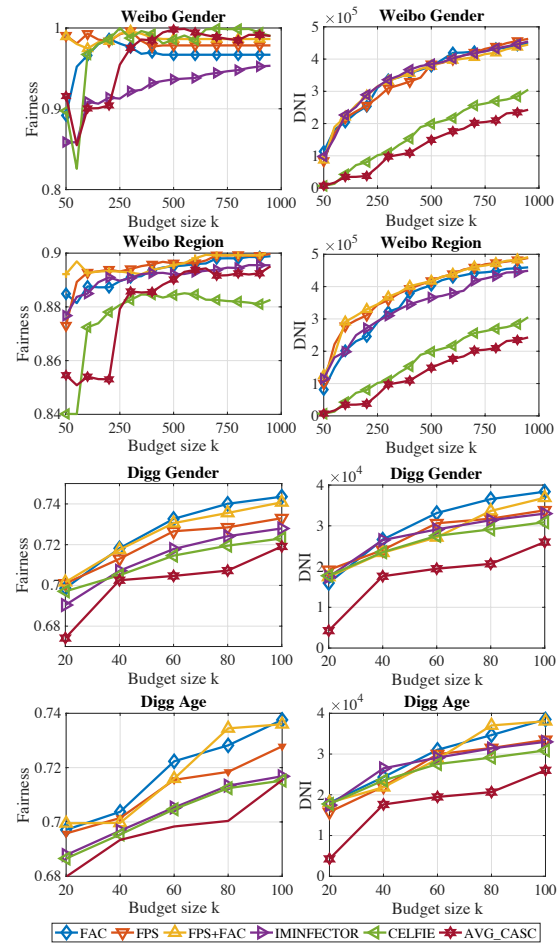

Figure 3: Validation on the complete datasets.

scalability and effectiveness (Fig. 3). As expected, our models outperform the baselines on fairness, in both Weibo and Digg. Once again, and probably even more so now, since we have large cascades, the case of Weibo *gender* is an exception, since large spread translates automatically and "effortless" to high fairness. The performance of our methods on influence remains consistently in the top-tier, proving that the fairness gains do not come at the expense of influence, and that spread maximization can remain the top objective.

It is worth noting that **IMINFECTOR** not only has good (top-tier) spread performance, but also does reasonably well on fairness. Our explanation is that, at large scale and as a feature of our datasets, the larger the audience an influencer may have, the closer that audience's distribution on sensitive attributes will be to the one of the entire population. Therefore, under the equity formulation, high fairness becomes well-correlate to high spread.

6.4 Validation on an unbalanced Weibo dataset

Recall from the results of Fig. 2 – sampled Weibo data, *gender* fairness – that although we outperform the other baselines on spread, the differences on fairness between models are limited. We can also note that the fairness scores on Weibo *gender* are higher than for other attributes. Our take on this is that, since *gender* is
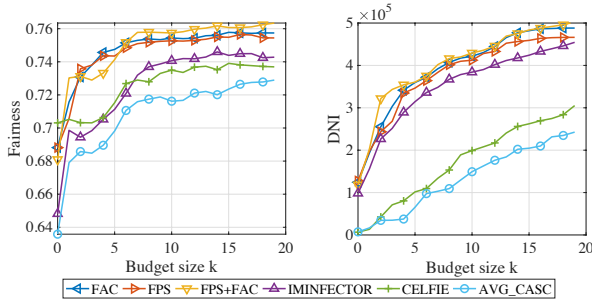
Figure 4: Unbalanced-*gender* complete Weibo dataset.

already quite balanced, the diffusions follow that same distribution and thus only minor differences in fairness performance are to be expected. In Fig. 2, the plots fall within a thin y-region close to the 1 bar, while being scattered on the x-axis, which means all models can do quite well on fairness, while exhibiting differences on spread.

To further investigate this aspect, in Weibo (complete and sampled), we did a set of experiments where we "flip" *gender* profile values in the cascades of some randomly selected influencers, for randomly selected participants thereof. The rationale was to obtain some unbalanced influencers, i.e., infuencers with unfair cascades, while reaching an overall *male / female* ratio of roughly $25\% - 75\%$.

In Fig. 4, we present the comparison with the fairness-agnostic methods. We can observe now that when we decorrelate high spread potential and high fairness, the fairness gap between our methods and the baselines (in particular **IMINFECTOR**) becomes larger.

In Fig. 5, we have the evaluation on the sampled data. We can notice the performance shifts, as our models have now a clear advantage when facing the risk of large yet unfair propagations, which also indicates they are more robust to such unbalancedness.

6.5 Combinations of sensitive attributes

Recall that our fairness model applies not only to individual sensitive attributes, but also to combinations thereof. We present in this section our evaluation results on combinations of attributes: *gender_region* in Weibo (complete), with or without *gender* flipping, and *gender_age* in Digg (complete). In Weibo, with 2 *gender* categories and 36 *region* ones, we obtain a combined sensitive attribute having 72 categories (e.g., "women in region 1"). Similarly, in Digg, we have a combined sensitive attribute with 12 categories.

The results are given in Fig.6. We can notice that all our models show once again clear advantages over the baselines on both fairness and influence. Interestingly, the results on the original Weibo
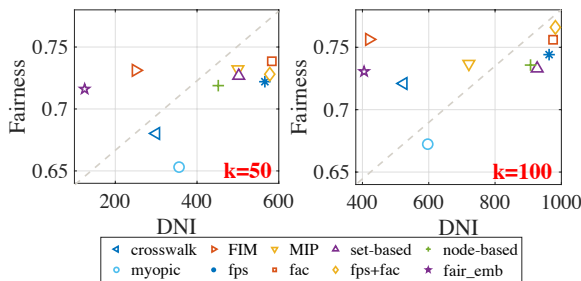

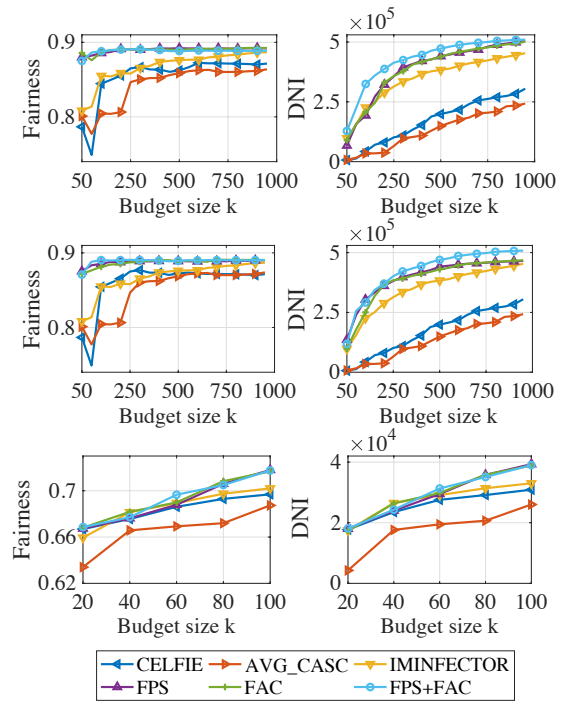Figure 5: Validation on the unbalanced-*gender* Weibo sample.

Figure 6: Evaluation for combinations of attributes (top row: Weibo, mid-row: flipped-*gender* Weibo, bottom row: Digg).

data and on the one with flipped *gender* are quite similar now, as the "effortless fairness" phenomenon caused by the balanced *gender* distribution is now attenuated by aggregation with *region*.

7 CONCLUSION

We revisit the problem of influence maximization with fairness. We propose two deep-learning based algorithms for extracting node representations from a history of diffusion cascades. Unlike the prior works, the proposed algorithms are generic – applicable to arbitrary sets of sensitive attributes – and scalable. Our embedding algorithms take into account both the spread aptitude and the fairness of influencers, and use the resulting embeddings to greedily select the optimal spread seeds. The experimental results, performed on both real-world and synthetic data, show that our methods outperform the state-of-the-art methods for fair IM (with limited scalability) or for fairness-agnostic IM based on cascades. Future work will investigate how the training could be done efficiently in a single stage for all the sensitive attributes and how learned representations pertaining to individual sensitive attributes could be "transferred" to combinations thereof.

REFERENCES

[1] J. Ali, M. Babaei, A. Chakraborty, B. Mirzasoleiman, K. Gummadi, and A. Singla. 2022. On the Fairness of Time-Critical Influence Maximization in Social Networks. *IEEE Transactions on Knowledge & Data Engineering* 01 (2022), 1–1.

[2] Sinan Aral and Paramveer S. Dhillon. 2018. Social influence maximization under empirical influence models. *Nature Human Behaviour* 2 (2018), 375–382.

[3] Akhil Arora, Sainyam Galhotra, and Sayan Ranu. 2017. Debunking the Myths of Influence Maximization: An In-Depth Benchmarking Study. In *SIGMOD*.

[4] Akhil Arora, Sainyam Galhotra, and Sayan Ranu. 2019. Influence Maximization Revisited: The State of the Art and the Gaps that Remain. In *EDBT*.

[5] Eytan Bakshy, Jake M Hofman, Winter A Mason, and Duncan J Watts. 2011. Everyone's an influencer: quantifying influence on twitter. In *WSDM*.

[6] Ruben Becker, Gianlorenzo D'Angelo, Sajjad Ghobadi, and Hugo Gilbert. 2021. Fairness in Influence Maximization through Randomization. In *AAAI*.

[7] Simon Bourigault, Sylvain Lamprier, and Patrick Gallinari. 2016. Representation Learning for Information Diffusion through Social Networks: An Embedded Cascade Model. In *WSDM*.

[8] Danny Brown and Sam Fiorella. 2013. *Influence Marketing: How to Create, Manage, and Measure Brand Influencers in Social MediaMarketing*. Que Pub.

[9] Nan Du, Le Song, Manuel Gomez Rodriguez, and Hongyuan Zha. 2013. Scalable Influence Estimation in Continuous-Time Diffusion Networks. In *NIPS*.

[10] Golnoosh Farnadi, Behrouz Babaki, and Michel Gendreau. 2020. A Unifying Framework for Fairness-Aware Influence Maximization. In *The Web Conference*.

[11] Shanshan Feng, Gao Cong, Arijit Khan, Xiucheng Li, Yong Liu, and Yeow Meng Chee. 2018. Inf2vec: Latent Representation Model for Social Influence Embedding. In *ICDE*.

[12] Benjamin Fish, Ashkan Bashardoust, Danah Boyd, Sorelle Friedler, Carlos Scheidegger, and Suresh Venkatasubramanian. 2019. Gaps in Information Access in Social Networks. In *WWW*.

[13] Shay Gershtein, Tova Milo, and Brit Youngmann. 2021. Multi-Objective Influence Maximization. In *EDBT*, Yannis Velegrakis, Demetris Zeinalipour-Yazti, Panos K. Chrysanthis, and Francesco Guerra (Eds.).

[14] Aditya Grover and Jure Leskovec. 2016. node2vec: Scalable feature learning for networks. In *SIGKDD*.

[15] David Kempe, Jon Kleinberg, and Éva Tardos. 2003. Maximizing the spread of influence through a social network. In *ACM SIGKDD*.

[16] Ahmad Khajehnejad, Moein Khajehnejad, Mahmoudreza Babaei, Krishna P Gummadi, Adrian Weller, and Baharan Mirzasoleiman. 2022. CrossWalk: fairness-enhanced node representation learning. In *AAAI*.

[17] Moein Khajehnejad, Ahmad Asgharian Rezaei, Mahmoudreza Babaei, Jessica Hoffmann, Mahdi Jalili, and Adrian Weller. 2020. Adversarial Graph Embeddings for Fair Influence Maximization over Social Networks. In *IJCAI*.

[18] Kristina Lerman, Rumi Ghosh, and Tawan Surachawala. 2012. Social Contagion: An Empirical Study of Information Spread on Digg and Twitter Follower Graphs. https://doi.org/10.48550/ARXIV.1202.3162

[19] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne M. VanBriesen, and Natalie S. Glance. 2007. Cost-effective outbreak detection in networks. In *ACM SIGKDD*.

[20] Y. Li, J. Fan, Y. Wang, and K. Tan. 2018. Influence Maximization on Social Graphs: A Survey. *IEEE Transactions on Knowledge and Data Engineering* 30, 10 (2018), 1852–1872. https://doi.org/10.1109/TKDE.2018.2807843

[21] Mingkai Lin, Wenzhong Li, and Sanglu Lu. 2020. Balanced Influence Maximization in Attributed Social Network Based on Sampling. In *WSDM '20: The Thirteenth ACM International Conference on Web Search and Data Mining, Houston, TX, USA, February 3-7, 2020*, James Caverlee, Xia (Ben) Hu, Mounia Lalmas, and Wei Wang (Eds.). 375–383.

[22] George Panagopoulos, Fragkiskos Malliaros, and M Vazirgiannis. 2020. Multi-task Learning for Influence Estimation and Maximization. *IEEE Transactions on Knowledge and Data Engineering* (2020).

[23] George Panagopoulos, Fragkiskos D Malliaros, and Michalis Vazirgianis. 2020. Influence maximization using influence and susceptibility embeddings. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 14.

[24] B. Perozzi, R. Al-Rfou, and S. Skiena. 2014. Deepwalk: Online learning of social representations. In *KDD*.

[25] Aida Rahmattalabi, Shahin Jabbari, Himabindu Lakkaraju, Phebe Vayanos, Max Izenberg, Ryan Brown, Eric Rice, and Milind Tambe. 2021. Fair Influence Maximization: a Welfare Optimization Approach. In *AAAI*.

[26] K. Saito, Masahiro Kimura, K. Ohara, and H. Motoda. 2009. Learning Continuous-Time Information Diffusion Model for Social Behavioral Data Analysis. In *ACML*.

[27] Ana-Andreea Stoica, Jessy Xinyi Han, and Augustin Chaintreau. 2020. Seeding Network Influence in Biased Networks and the Benefits of Diversity. In *WWW*.

[28] Alan Tsang, Bryan Wilder, Eric Rice, Milind Tambe, and Yair Zick. 2019. Group-Fairness in Influence Maximization. In *IJCAI*.

[29] Chuhan Wu, Fangzhao Wu, Xiting Wang, Yongfeng Huang, and Xing Xie. 2021. Fairness-aware recommendation with decomposed adversarial learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 4462–4469.

[30] Jing Zhang, Biao Liu, Jie Tang, Ting Chen, and Juanzi Li. 2013. Social Influence Locality for Modeling Retweeting Behaviors. In *IJCAI*.