A Viral Marketing-Based Model For Opinion Dynamics in Online Social Networks

Sijing Tu
KTH Royal Institute of Technology
Stockholm, Sweden
sijing@kth.se

Stefan Neumann
KTH Royal Institute of Technology
Stockholm, Sweden
neum@kth.se

ABSTRACT

Online social networks provide a medium for citizens to form opinions on different societal issues, and a forum for public discussion. They also expose users to viral content, such as breaking news articles. In this paper, we study the interplay between these two aspects: *opinion formation* and *information cascades* in online social networks. We present a new model that allows us to quantify how users change their opinion as they are exposed to viral content. Our model is a combination of the popular Friedkin–Johnsen model for opinion dynamics and the independent cascade model for information propagation. We present algorithms for simulating our model, and we provide approximation algorithms for optimizing certain network indices, such as the sum of user opinions or the disagreement–controversy index; our approach can be used to obtain insights into how much viral content can increase these indices in online social networks. Finally, we evaluate our model on real-world datasets. We show experimentally that marketing campaigns and polarizing contents have vastly different effects on the network: while the former have only limited effect on the polarization in the network, the latter can increase the polarization up to 59% even when only 0.5% of the users start sharing a polarizing content. We believe that this finding sheds some light into the growing segregation in today's online media.

CCS CONCEPTS

• **Information systems** → **Social networks**; *Data mining*; • **Theory of computation** → *Graph algorithms analysis.*

KEYWORDS

online social networks, opinion dynamics, information spread

ACM Reference Format:
Sijing Tu and Stefan Neumann. 2022. A Viral Marketing-Based Model For Opinion Dynamics in Online Social Networks. In *Proceedings of the ACM Web Conference 2022 (WWW '22), April 25–29, 2022, Virtual Event, Lyon, France.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3485447.3512203

WWW '22, April 25–29, 2022, Virtual Event, Lyon, France

ACM ISBN 978-1-4503-9096-5/22/04…$15.00
https://doi.org/10.1145/3485447.3512203

1 INTRODUCTION

Online social networks are a ubiquitous part of modern societies. In addition to connecting users with their friends, many people also use them as content aggregators, by following media outlets or reading articles shared by their peers. Clearly, engaging in social networks may impact one's opinions with respect to societal issues: users might adjust their opinions during a discussion based on arguments by their peers; or they might adapt their opinions based on new facts revealed in a news article they read.

Due to this strong connection between opinion formation and information spread, online social networks have become the target of viral disinformation campaigns. Popular examples include groups like QAnon who spread conspiracy theories and fake news about topics, such as vaccination, or state actors who try to influence election results in opposing countries. While it is well-researched how viral content spreads through social networks, such models do not consider how user opinions are impacted by the viral content. Therefore, understanding how new information influences user opinions and being able to quantify the impact of such disinformation campaigns is highly desirable.

A prominent model to quantify opinion dynamics in social networks is the Friedkin–Johnsen (FJ) model [15]. The FJ model stipulates that each user has an *expressed opinion* that the user reveals publicly and is network-dependent, and an *innate opinion* that is fixed and network-independent. However, it does not take into account how users change their opinions based on new information (e.g., viral content) that is disseminated in the network.

Furthermore, researchers have studied problems related to optimizing certain opinion-based network indices, for instance, maximizing the average opinion [17] or polarization [9, 16], or minimizing polarization and disagreement [10, 25, 27]. In this line of work, optimization occurs by nudging the expressed or innate opinions of a set of seed users towards a certain direction. However, existing works do not specify how such nudging takes place, nor do they consider the interplay of opinion nudging within a more realistic setting of information cascades.

Therefore, the current research *either* allows us to quantify user opinions and optimize opinion-based network indices without taking into account viral content *or* it allows to assess the spread of viral content without reasoning about user opinions. This limitation leads us to the following questions:

(1) *Can we quantify how viral content influences user opinions in online social networks?*
(2) *Can we study the interplay between information cascades and opinion dynamics?*
(3) *Can we optimize opinion-based network indices by taking into account the spread of viral content?*

Our contributions. We answer the above questions affirmatively by proposing a new model that combines the Friedkin–Johnsen model [15] and the influence-maximization framework of Kempe et al. [21]. To the best of our knowledge, our model is the first that allows to quantify the impact of viral content on user opinions.

Contrasted with the vanilla FJ model in which the innate user opinions are "fixed" but the expressed opinions are changing over time based on user interactions, our model considers the viral content that is shared in the network, and it assumes that for users who are exposed to this content, *there is a probability that their innate opinion changes*. This could be the case, for example, when a user reads an article that makes them reconsider their stance on a certain topic.

When a subset of users change their innate opinions, their expressed opinions will also be modified, which in turn will have an impact on the whole network via the FJ opinion dynamics. Thus, *the change of the innate opinions of few users may have an impact on the whole network*: even when a user's innate opinion does not change by the viral content (because they ignore it or the content never reaches them), they still might change their expressed opinion due to "peer-pressure."

Our model connects these two phenomena: it allows us to understand how viral content can impact individual users, while it also enables us to study how individual behavior ripples through the network and affects the overall discussion.

We consider two different types of content: *non-polarizing* and *polarizing*. For non-polarizing content, such as marketing campaigns, the innate opinions of the users can only increase. For polarizing content, we take into account the *backfire effect* [28]: interaction with opposing content may lead to a decrease in a user's innate opinion. This could be the case, e.g., in political campaigns when a party runs an ad that makes their supporters react positively but their opponents react negatively.

From an algorithmic view point, we present methods for simulating our model. Additionally, we consider the problem of optimizing certain opinion-based network indices. We present a greedy $(1 - 1/e - \epsilon)$-approximation algorithm for maximizing the sum of user opinions for non-polarizing content. We also present algorithms for maximizing the controversy and the disagreement–controversy indices [27] for non-polarizing content; our algorithms have data-dependent approximation ratios. Finally, we provide heuristics for maximizing other indices, such as polarization and disagreement, for non-polarizing and polarizing content.

To obtain our optimization algorithms, we build upon the reverse-reachable sets framework [6, 31, 32]. One challenge is that, in our setting, the arising optimization problems are based on quadratic forms and, therefore, we have to extend the reverse reachable set framework to this more general setting.

We evaluate our methods on real-world data. Our experiments reveal a striking difference between non-polarizing and polarizing content. On the one hand, non-polarizing content can significantly increase the sum of user opinions, but it has limited impact on the polarization and sometimes even *decreases* it. The situation for polarizing content is the opposite: it barely increases the sum of user opinions but it can increase the polarization significantly. We see that even when only 0.5% of the users start sharing a polarizing content, the network polarization increases by more than 20% on average and can rise up to 59%. We believe that this finding provides an explanation for the growing polarization that can be witnessed in modern day's online media.

We present the proofs of our claims in the full version [33].

2 RELATED WORK

Our aim is to quantify how viral content impacts user opinions in social networks. Naturally, our approach builds on existing models for opinion dynamics and information cascades.

Opinion dynamics have been studied in different research areas, including psychology, social sciences, and economics [7, 20]. Here we build on the popular Friedkin–Johnsen (FJ) model [15], which is an extension of a classic model by DeGroot [14]. Many extensions of the FJ model have been proposed. For example, Amelkin et al. [1] assume that the innate user opinions change over time based on the expressed opinions. We refer to the discussion in [1] for other related models. However, these works do not take into account the changes of innate opinions based on exposition to viral contents.

Recent work used these models for understanding properties of opinion dynamics and formulating natural optimization problems. Bindel et al. [5] analyze the "price of anarchy" in the FJ model by considering as cost the internal conflict of the individuals in the network and comparing the cost at equilibrium and the social optimum. Gionis et al. [17] maximize the sum of opinions of the network users. Other works study the problem of measuring and reducing polarization of opinions, or other disagreement indices, in the FJ model [12, 25, 27, 36], while adversarial settings have also been considered, aiming to quantify the power of an adversary seeking to induce discord in the model [9, 10, 16].

To model information cascades, we build on the popular independent-cascade model of Kempe et al. [21]. Many extensions and variants of this model have been proposed. For example, Sathanur et al. [30] incorporate intrinsic user activations based on external sources. Another popular extension is the topic-aware cascade model by Barbieri et al. [4], and has various applications including social advertising [2, 3]. While such models allow to model information spread, they do not allow to quantify how these change the user opinions.

The backfire effect, the tendency of individuals to hold firmly on their beliefs when faced with factual corrections, has been observed in political sciences [28], but has not been studied extensively in computational social sciences. Exceptions are the works of Chen et al. [11], who incorporate backfire in an opinion-dynamics model for biased assimilation [13], and Hirakura et al. [19], who propose a model of polarization that incorporates empathy and repulsion.

Our optimization algorithms rely on reverse reachable sets, introduced by Borgs et al. [6], and improved by subsequent techniques [31, 32]. We extend these ideas to our setting, to obtain algorithms for objectives that include quadratic terms. We note that the activity-maximization task defined by Wang et al. [34] is a special case of our setting. We apply the "sandwich" framework [24] to obtain data-dependent approximation guarantees for some of our objectives. For the efficient computation of our objective functions, we use the methods by Xu et al. [35] based on Laplacian solvers.

To our knowledge, this is the first work that studies how user opinions change due to viral information in online social networks.

Table 1: Matrices of the different indices.

Index	Notation	Matrix
Polarization	$\mathcal{P}(\mathbf{L})$	$(\mathbf{I}+\mathbf{L})^{-1}(\mathbf{I}-\frac{\mathbf{1}\mathbf{1}^\intercal}{n})(\mathbf{I}+\mathbf{L})^{-1}$
Disagreement	$\mathcal{D}(\mathbf{L})$	$(\mathbf{L}+\mathbf{I})^{-1}\mathbf{L}(\mathbf{L}+\mathbf{I})^{-1}$
Internal conflict	$\mathcal{I}(\mathbf{L})$	$(\mathbf{L}+\mathbf{I})^{-1}\mathbf{L}^2(\mathbf{L}+\mathbf{I})^{-1}$
Controversy	$\mathcal{C}(\mathbf{L})$	$(\mathbf{L}+\mathbf{I})^{-2}$
Disagreement–controversy	$\mathcal{I}^{dc}(\mathbf{L})$	$(\mathbf{L}+\mathbf{I})^{-1}$

For fixed user opinions, Monti et al. [26] studied how cascades spread through the network, based on the user opinions and the topics of the contents.

3 PRELIMINARIES

Let $G=(V,E,w)$ be an undirected weighted graph, with $n=|V|$ nodes and edge weights $w\colon E\to\mathbb{R}_{>0}$. We let $N(u)$ denote the set of neighbors of $u\in V$. The Laplacian of G is $\mathbf{L}=\mathbf{D}-\mathbf{W}$, where $\mathbf{D}$ is the $n\times n$ diagonal matrix with $\mathbf{D}_{u,u}=\sum_{v\in N(u)}w(u,v)$ for all $u\in V$ and $\mathbf{W}$ is the $n\times n$ matrix with $\mathbf{W}_{u,v}=w_{u,v}$ for all $u,v\in V$.

Friedkin–Johnsen (FJ) model. In the FJ model, we are given a weighted undirected graph $G=(V,E,w)$ with n nodes. Each node u corresponds to a user of a social network. Each user u has an *expressed* opinion $z_u\in[0,1]$, which depends on the network, and a fixed *innate* opinion $s_u\in[0,1]$. We write $\mathbf{s}\in[0,1]^n$ and $\mathbf{z}\in[0,1]^n$ to denote the vectors of innate and expressed opinions.

The expressed opinions are updated in rounds. More concretely, let $\mathbf{s}$ be the vector of innate opinions, and $\mathbf{z}^{(t)}$ be the vector of expressed opinions at time t. The update rule is given by

$$\mathbf{z}^{(t+1)}=(\mathbf{D}+\mathbf{I})^{-1}(\mathbf{W}\mathbf{z}^{(t)}+\mathbf{s}). \tag{3.1}$$

Taking the limit $t\to\infty$, the expressed opinions converge to

$$\mathbf{z}^{*}=(\mathbf{I}+\mathbf{L})^{-1}\mathbf{s}. \tag{3.2}$$

We study the following popular network indices in our model, where the matrices of the quadratic forms are as defined in Table 1:

- *sum of user opinions*, which is given by $\mathcal{S}_{\mathbf{s}}=\mathbf{1}^\intercal\mathbf{s}$, and it is well-known that $\mathcal{S}_{\mathbf{s}}=\mathbf{1}^\intercal\mathbf{z}$;
- *polarization [27]* $\mathcal{P}_{G,\mathbf{s}}=\sum_{u\in V}(z^*_u-\bar{z})^2=\mathbf{s}^\intercal\mathcal{P}(\mathbf{L})\,\mathbf{s}$, where $\bar{\mathbf{z}}=\frac{1}{n}\sum_{u\in V}z^*_u$ is the average user opinion;
- *disagreement [27]* $\mathcal{D}_{G,\mathbf{s}}=\sum_{(u,v)\in E}w_{u,v}(z^*_u-z^*_v)^2=\mathbf{s}^\intercal\mathcal{D}(\mathbf{L})\,\mathbf{s}$;
- *internal conflict [10]* $\mathcal{I}_{G,\mathbf{s}}=\sum_{u\in V}(s_u-z^*_u)^2=\mathbf{s}^\intercal\mathcal{I}(\mathbf{L})\,\mathbf{s}$;
- *controversy [10, 25]* $\mathcal{C}_{G,\mathbf{s}}=\sum_{u\in V}(z^*_u)^2=\mathbf{s}^\intercal\mathcal{C}(\mathbf{L})\,\mathbf{s}$; and
- *disagreement-controversy [27, 35]* $\mathcal{I}^{dc}_{G,\mathbf{s}}=\mathbf{s}^\intercal\mathcal{I}^{dc}(\mathbf{L})\,\mathbf{s}=\mathcal{C}_{G,\mathbf{s}}+\mathcal{D}_{G,\mathbf{s}}$.

4 MODELLING THE INFLUENCE OF VIRAL CONTENT ON USER OPINIONS

We formally introduce our model in Sec. 4.1 and show how it can be simulated in Sec. 4.2.

4.1 The spread-acknowledge model

Following the independent cascade model [21], we assume that a value $p_{u,v}\in[0,1]$ encodes the probability that user v reacts to content received from user u; we allow that $p_{u,v}\neq p_{v,u}$. Furthermore, we introduce parameters $\epsilon,\delta>0$, as explained below.

As per the FJ model, each user u has an expressed opinion z_u and an innate opinion s_u. Additionally, each user has a *state*, which is either inactive, ignore, acknowledge or spread. We order the states by "inactive < ignore < acknowledge < spread" and we follow the convention that when a user changes their state, they can only pick one that is higher with respect to this ordering. An illustration of the model with respect to state transitioning and actions performed for a single node v is provided in Figure 1.

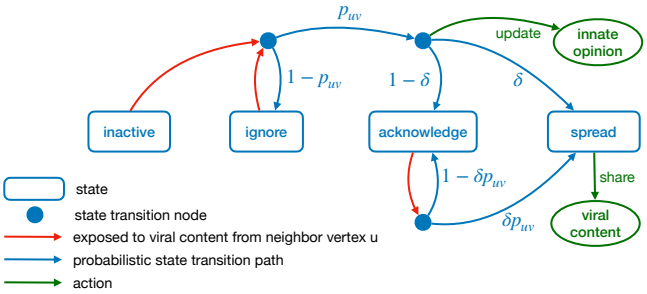


Figure 1: An illustration of the spread-acknowledge model with respect to state transitioning and actions performed for a single node v. In the initial round, k seed nodes are in state spread, while the rest of nodes are in state inactive.

Our model proceeds in rounds. Initially, in round 0, there are k users whose state is spread and all other users are inactive; in later rounds, it is possible that users change their state. We will refer to the users whose initial state is spread as *seed nodes*. Each round $t>0$ has two *phases*:[1] In the first phase, the users update their expressed opinions. In the second phase, the viral content is spread through the network and users may change their state and their innate opinion. We describe both phases below.

Phase I: Updating user opinions. The users update their expressed opinion as in the FJ model, i.e., according Eq. (3.1).

Phase II: Information spreading. Consider round $t>0$. Let U denote the set of users who have changed their state to spread in round $t-1$. If $U=\emptyset$, we consider Phase II finished. Otherwise ($U\neq\emptyset$), each user $u\in U$ shares the viral content with all of its neighbors. When a neighbor v of u is exposed to the viral content, it switches to a new state and possibly adjusts its innate opinion. If v is in state inactive or ignore, then this is done as follows:

- With probability δp_{uv}, user v switches to state *spread*; v adjusts its innate opinion (described below) and shares the content in the next round with its neighbors.
- With probability $(1-\delta)p_{uv}$, user v switches to state *acknowledge*; v adjusts its innate opinion (described below) but does *not* share the content with its neighbors in the next round.
- With probability $1-p_{uv}$, user v switches to state *ignore*; v performs no action (i.e., v does not try to share the content and v also does not adjust its innate opinion).

If v is in state acknowledge then it switches to state spread with probability δp_{uv} and remains in state acknowledge with probability $1-\delta p_{uv}$. Finally, if v is in state spread then v always stays in this state. In both of these cases, v does not adjust its opinion again.

[1] We note that for our model and our analysis it is not necessary to consider two phases, we only make this assumption for the sake of better exposition. We could as well assume that both phases are interleaved and happen simultaneously.

The above process ensures that the state ordering defined before is obeyed during state switching and that each user adjusts its innate opinion at most once. Finally, note that our model is a generalization of the independent cascade model if $\delta = 1$.

Adjusting innate opinions. Now we describe how users change their innate opinions when they are exposed to viral content.

Consider user u at state inactive or ignore whose new state becomes acknowledge or spread. Then the innate opinion $\mathbf{s}_u$ changes to a new value $\hat{\mathbf{s}}_u$. We consider two scenarios:

- *Marketing campaign:* The user's opinion becomes more positive after seeing the content, i.e., $\hat{\mathbf{s}}_u = \min\{\mathbf{s}_u + \epsilon, 1\}$ for the parameter $\epsilon > 0$ from above. Here, we use the min-operation to ensure that the new opinion $\hat{\mathbf{s}}_u$ is in the interval $[0, 1]$.
- *Polarizing campaign with backfire:* In a polarizing campaign we assume that while some users embrace the content, others will find it repelling. More concretely, we assume that there is a threshold $\tau \in [0, 1]$ such that: (1) If $\mathbf{s}_u \geq \tau$ then u embraces the content and adjusts its opinion to $\hat{\mathbf{s}}_u = \min\{\mathbf{s}_u + \epsilon, 1\}$; (2) If $\mathbf{s}_u < \tau$, then u finds the content repelling and adjusts $\hat{\mathbf{s}}_u = \max\{0, \mathbf{s}_u - \epsilon\}$.

Note that in our model with polarizing campaigns, users can still share a content they dislike. While this might seem non-intuitive at first, we believe that it is a realistic behavior: users who oppose a certain content often share it together with a counter-argument. We remark that our model can be modified to avoid this.

Finally, observe that $\hat{\mathbf{s}}$ is a random vector that depends on the outcome of the information spread. However, once we fix the randomness of the information spread, $\hat{\mathbf{s}}$ becomes deterministic. This will be a useful property in the following.

Possible model extensions. We note that our model is quite general and can be extended in various ways. First, we modelled information cascades via the independent cascade model [21]. However, our model and our result from Lemma 4.1 also hold if we used the linear threshold model [21], topic-aware versions of the independent cascade and linear threshold models [4], as well as intrinsic user activations [30]. In particular, using the linear threshold model could lead to insights on contents that spread via complex contagion [8, 18]. Second, above we considered the two relatively simple settings for adjusting the innate user opinions $\hat{\mathbf{s}}_u$. However, we note that Lemma 4.1 below generalizes to the setting when $\hat{\mathbf{s}}_u$ is any user-defined function of $\mathbf{s}_u$.

4.2 Equivalence with the two-stage model

While the spread-acknowledge model is easy to explain and motivate by real-world scenarios, it is not clear how to simulate it efficiently. If we implemented the model as described above, we would have to update the expressed opinions in each round, which can be costly. To avoid this, we now introduce a new model that can be simulated more efficiently, and we show that it produces an identical distribution over the innate and expressed opinions.

The two-stage model. Our simplified model also proceeds in rounds, but it performs the information spreading and the updating of the user opinions in two sequential *stages*. More concretely, in each round of the *first stage*, we perform the information spreading process that is described in Phase II above (and we do *not* perform the updating of the expressed opinions as per Phase I). In this process, some of the users' innate opinions and their states might change. When after a round no new users have changed their state to spread, we start the second stage. In each round of the *second stage*, we perform the same update of the expressed opinions as described in Phase I above (and we do *not* run Phase II).

Simulating the two-stage model. Next, we discuss why the two-stage model is well-suited for efficient simulations. First, observe that the first stage stops when no node changed their state to spread in the previous round, i.e., when $U = \emptyset$. Therefore, the first stage can have at most $O(n)$ rounds (since each of the n users can take at most four different states and since we assumed that users only increase their state with respect to the ordering of the states). Additionally, in each round of the first stage, we can update the states of the nodes by iterating over all nodes v that are neighbors of a node $u \in U$ and then updating the state of v with the probability described in Phase II. Since each user can become a spreader only once, the time for executing *all* rounds of the first stage is $O(m)$, where m is the number of edges in the graph.

Second, recall that the adjusted innate opinions $\hat{\mathbf{s}}_u$ only depend on the randomness from the information spreading process. Therefore, after the first stage finished, the innate opinions $\hat{\mathbf{s}}_u$ are fixed. Thus, we can assume that the vector $\hat{\mathbf{s}}_u$ is known and the expressed equilibrium opinions are given by $\hat{\mathbf{z}}^* = (I + L)^{-1}\hat{\mathbf{s}}$. The time complexity for the second stage is the time required to solve for $\hat{\mathbf{z}}^*$.

Equivalence of the opinion distributions. It remains to show that both models induce the same distribution over the innate and expressed opinions. To show this equivalence, we assume that both models are run with the same input graphs, the same seed nodes, and the same (non-adjusted) innate opinions $\mathbf{s}$. Now let us denote the adjusted innate opinions generated by the spread-acknowledge model by $\hat{\mathbf{s}}_u$ and those by the two-stage model by $\tilde{\mathbf{s}}_u$. Recall that both $\hat{\mathbf{s}}_u$ and $\tilde{\mathbf{s}}_u$ are random vectors that depend only on the outcome of the information-spreading process. The following lemma asserts the equivalence of the two models. The proof is presented in [33].

Lemma 4.1. *For all $a \in [0, 1]^n$, $\Pr[\hat{\mathbf{s}} = a] = \Pr[\tilde{\mathbf{s}} = a]$. Furthermore, let $\hat{\mathbf{z}}^* = (I + L)^{-1}\hat{\mathbf{s}}$ and $\tilde{\mathbf{z}}^* = (I + L)^{-1}\tilde{\mathbf{s}}$ be the equilibrium opinions. Then $\Pr[\hat{\mathbf{z}}^* = b] = \Pr[\tilde{\mathbf{z}}^* = b]$ for all $b \in [0, 1]^n$.*

5 ALGORITHMS

We present algorithms for maximizing the indices defined in Sec. 3. We give algorithms for approximating the indices (Sec. 5.1). Then we present our algorithms for maximizing the sum of user opinions (Sec. 5.2) and for maximizing the controversy and the disagreement–controversy indices (Sec. 5.3). We present our proofs in [33].

5.1 Estimating indices

Let $\mathcal{M}(\mathbf{L})$ be one of the matrices from Table 1, which induces the quadratic form for each of the indices that we wish to study. Recall that $\mathbf{s}$ is the non-adjusted vector of innate opinions and $\hat{\mathbf{s}}$ is the random vector of adjusted innate opinions. In the following, our goal is to compute $\mathbb{E}[\hat{\mathbf{s}}^\top \mathcal{M}(\mathbf{L})\,\hat{\mathbf{s}}]$.

Let $\Delta\hat{\mathbf{s}} = \hat{\mathbf{s}} - \mathbf{s}$ be the random vector that denotes how the users changed their opinions. Then observe that

$$\mathbb{E}[\hat{\mathbf{s}}^\top \mathcal{M}(\mathbf{L})\,\hat{\mathbf{s}}] = \mathbf{s}^\top \mathcal{M}(\mathbf{L})\,\mathbf{s} + \mathbb{E}[2\mathbf{s}^\top \mathcal{M}(\mathbf{L})\,\Delta\hat{\mathbf{s}} + \Delta\hat{\mathbf{s}}^\top \mathcal{M}(\mathbf{L})\,\Delta\hat{\mathbf{s}}].$$

Since the first term in the sum is deterministic, we drop it and focus on $\mathbb{E}[h(\Delta\hat{\mathbf{s}})]$, where $h(\Delta\hat{\mathbf{s}}) := 2\mathbf{s}^\top \mathcal{M}(\mathbf{L})\,\Delta\hat{\mathbf{s}} + \Delta\hat{\mathbf{s}}^\top \mathcal{M}(\mathbf{L})\,\Delta\hat{\mathbf{s}}$. We show that computing $\mathbb{E}[h(\Delta\hat{\mathbf{s}})]$ is #**P**-hard since our model generalizes the independent cascade model.

Lemma 5.1. *Given seed nodes S, computing $\mathbb{E}[h(\Delta\hat{\mathbf{s}})]$ is #**P**-hard.*

Monte Carlo Simulation. Since Lemma 5.1 shows that computing $\mathbb{E}[h(\Delta\hat{\mathbf{s}})]$ exactly is hard, we resort to approximations. One option is to use Monte Carlo simulations of our model. More concretely, we can simulate our model as described in Sec. 4.2 to obtain multiple samples of $\hat{\mathbf{s}}$. Now a Chernoff bound implies that we can compute an approximation of $\mathbb{E}[\hat{\mathbf{s}}]$ with high probability. Then we can compute an approximation of $\mathbb{E}[h(\Delta\hat{\mathbf{s}})]$ in near-linear time using the algorithms by Xu et al. [35], which are based on Laplacian solvers. This approach is efficient when the number of seed node sets for which we wish to compute $\mathbb{E}[h(\Delta\hat{\mathbf{s}})]$ is *small*.

Reverse reachable sets. However, in our optimization algorithms we will need to evaluate $\mathbb{E}[h(\Delta\hat{\mathbf{s}})]$ for a *large* number of different seed node sets and thus using the Monte Carlo approach is too inefficient. Therefore, we use *reverse influence sampling* [6, 31, 32], which allow us to reduce the number of simulations of our model.

Our notion of reverse reachable sets is as follows. Suppose that we want to simulate our model on a graph $G = (V, E)$. A *possible world* is a copy of G that has labels on the edges and we generate the labels as follows. For each edge $(u, v) \in E$, we pretend that u has state spread and v has state inactive. Now we sample the state of v as described in Phase II above and we *label* (u, v) with the new state of v. For example, if v changes its state to acknowledge then the label of (u, v) is acknowledge. This process is repeated for all edges $(u, v) \in E$ and we always assume that u has state spread and v has state inactive, regardless of the outcomes of previous samples.

Now consider a possible world g. We say that there exists a *live path* from u to v if there exists a path in g in which all edges have label spread except the edge incident upon v which may have label acknowledge or spread. Notice that live paths encode when users change their opinions in our model: user v adjusts its opinion if and only if there exists a live path from a seed node to v.

Next, let g be a randomly generated possible world and let u be a random node in G. A *random RR-set R for u in g* is a set of nodes in g such that there exists a live path to u.

Estimating indices. Now we turn to estimating $\mathbb{E}[h(\Delta\hat{\mathbf{s}})]$. Existing information propagation methods can be used for estimating $2\mathbb{E}[\mathbf{s}^\top \mathcal{M}(\mathbf{L})\,\Delta\hat{\mathbf{s}}]$, because $\Delta\hat{\mathbf{s}}$ is the only random quantity in this expression. However, we also need to approximate $\mathbb{E}[\Delta\hat{\mathbf{s}}^\top \mathcal{M}(\mathbf{L})\,\Delta\hat{\mathbf{s}}] = \sum_{u,v} \mathcal{M}(\mathbf{L})_{u,v}\mathbb{E}[\Delta\hat{\mathbf{s}}_u\Delta\hat{\mathbf{s}}_v]$, which involves products of random variables and which existing methods cannot do. Our main observation is that in each possible world it holds that $\Delta\hat{\mathbf{s}}_u\Delta\hat{\mathbf{s}}_v \neq 0$ if and only if there exist live paths from the seed nodes to u *and* v. Wang et al. [34] followed a similar approach but only considered pairs (u, v) for which there exist edges in the graph; here, we have to perform this operation for all pairs $(u, v) \in V^2$.

In the following, we set $\Delta s_u \in [-\epsilon, \epsilon]$ to denote how much user u adjusts its opinion once it reaches state acknowledge or spread. Note that $|\Delta s_u|$ can be smaller than ϵ because of the interval concatenation that we described in Phase II above. Next, let S be a set of seed nodes and let $\mathbf{1}_u(S)$ be an indicator with $\mathbf{1}_u(S) = 1$ if u adjusts innate opinion and otherwise $\mathbf{1}_u(S) = 0$. Let $\mathbf{1}(S)$ be a vector of $\mathbf{1}_u(S)$ consisting of each $u \in V$. Note that $\mathbf{1}(S)$ is a random vector and that $\Delta\mathbf{s}$ is deterministic. Observe that $\Delta\hat{\mathbf{s}} = \Delta\mathbf{s} \odot \mathbf{1}(S)$, where is $\odot$ the Hadamard product.

To simplify our notation, we set $w_u = (2\mathbf{s}^\top \mathcal{M}(\mathbf{L}))_u \Delta s_u$ and let $m_{u,v} = (\Delta s_u)^\top \mathcal{M}(\mathbf{L})_{u,v} \Delta s_v$. Then we obtain:

$$h(\Delta\hat{\mathbf{s}}) = \sum_{u,v \in V} \frac{1}{n} w_u \mathbf{1}_u(S) + m_{u,v}\mathbf{1}_u(S)\mathbf{1}_v(S) =: F(S).$$

Given these definitions, we let R_u and R_v be random RR-sets for u and v, respectively, and we set

$$\omega_{R_u,R_v}(S) = \mathbb{1}[(R_u \cap S) \neq \emptyset] w_u + \mathbb{1}[(R_u \cap S) \neq \emptyset, (R_v \cap S) \neq \emptyset] n\, m_{u,v},$$

and for a set $\mathcal{R}$ of random RR-sets we define

$$F_{\mathcal{R}}(S) = \frac{\sum_{(R_u,R_v)\in\mathcal{R}} \omega_{R_u,R_v}(S)}{|\mathcal{R}|}. \tag{5.1}$$

We show that $F_{\mathcal{R}}(S)$ is an unbiased estimator for $\mathbb{E}[F(S)]$.

Lemma 5.2. *Let $\mathcal{R}$ be a set of samples of pair of random RR-sets. Then $\mathbb{E}[F(S)] = \mathbb{E}_{u,g\sim\mathcal{G}}[n\, F_{\mathcal{R}}(S)]$.*

Since the previous lemma only holds in expectation, we now consider approximations that hold with high probability. Let $\ell > 0$ be an error parameter, $\theta = |\mathcal{R}|$ be the number of RR-sets and suppose we know $\mathrm{OPT} = \max_{|S|\leq k} \mathbb{E}[F(S)]$ (we show later how to obtain bounds on OPT using statistical tests). Our goal will be to pick θ large enough such that

$$\Pr[|n\, F_{\mathcal{R}}(S) - \mathbb{E}[F(S)]| \geq \frac{\epsilon}{2}\,\mathrm{OPT}] \leq \frac{1}{n^\ell \binom{n}{k}}, \tag{5.2}$$

since then a union bound implies that for any seed set S of size k, $\mathbb{E}[F(S)]$ is a good estimator for $F_{\mathcal{R}}(S)$ w.h.p. We show that if we pick θ large enough then Equation (5.2) is satisfied.

Lemma 5.3. *Let $\chi = \max_{u,v\in V} |w_u + n\, m_{u,v}|$ and $\lambda = \frac{8n\chi}{\epsilon^2}(\frac{\epsilon}{3} + 1)(\ell \ln n + \ln 2 + \ln \binom{n}{k})$. If $\theta \geq \frac{\lambda}{\mathrm{OPT}}$ then Equation (5.2) holds.*

5.2 Maximizing network indices

Now we consider the *sum of expressed opinions problem*, where we are given an undirected weighted graph $G = (V, E)$ with edge probabilities $p_{u,v}$ and a positive integer k. The goal is to find a set of seed nodes of cardinality at most k that maximizes the sum of expressed opinions $\mathbb{E}[\mathcal{S}_{\hat{\mathbf{s}}}] = \mathbb{E}[\mathbf{1}^\top \hat{\mathbf{z}}^*]$. Our main result is as follows.

Theorem 5.4. *There exists a greedy approximation algorithm that computes a $(1 - 1/e - \epsilon)$-approximation with high probability.*

Indeed, in [33] we show that our model is strictly more powerful than the FJ model in which we can increase k innate user opinions.

To obtain the theorem, we maximize the sum of the adjusted parts of the innate opinions $\mathbb{E}[\sum_u \Delta\hat{\mathbf{s}}_u]$, since it is well-known that $\sum_u \hat{\mathbf{z}}^*_u = \sum_u \hat{\mathbf{s}}_u$ and thus we can maximize $\mathbb{E}[\sum_u \Delta\hat{\mathbf{s}}_u]$. Equivalently, we can maximize $F(S) := \sum_{u\in V} \mathbf{1}_u(S)\Delta s_u$, as we show next.

Lemma 5.5. $\arg\max_S \sum_{u\in V} z^*_u(S) = \arg\max_S F(S)$

The main benefit of Lemma 5.5 is that to maximize $F(S)$, we do not have to compute the sparse matrix inverse from Equation (3.2) which is very costly. Note that if $\Delta s_u = \epsilon$ for all $u \in V$, maximizing $F(S)$ reduces to the influence maximization problem [21]. However,

Algorithm 1: RR-Greedy

input : $\mathcal{R}, k$
output : $\tilde{X}^G$

$\tilde{X}^G \leftarrow \emptyset$
while $|\tilde{X}^G| \leq k$ **do**
  $x \leftarrow \arg\max_x F_{\mathcal{R}}(\tilde{X}^G \cup \{x\}) - F_{\mathcal{R}}(\tilde{X}^G)$;
  $\tilde{X}^G \leftarrow \tilde{X}^G \cup \{x\}$;
return $\tilde{X}^G$

Algorithm 2: Sampling

input : $\tilde{\mathbf{G}}, \lambda, \beta, \epsilon_2, k, \Delta\mathbf{s}, \chi, \mathrm{LB}_0$
output : $\mathcal{R}$
$\mathcal{R} \leftarrow \emptyset, \mathrm{LB} \leftarrow \mathrm{LB}_0$;
for $i = 1, \ldots, \log_2 n - 1$ **do**
  $y \leftarrow n/2^i, \theta_i \leftarrow \frac{\beta}{y}$;
  while $|\mathcal{R}| \leq \theta_i$ **do** $\mathcal{R} \leftarrow \mathcal{R} \cup$ GenerateRR-Set;
  $\tilde{\mathbf{X}}_i \leftarrow$ RR-Greedy$(\mathcal{R}, k)$;
  if $n\, F_{\mathcal{R}}(\tilde{\mathbf{X}}_i) \geq (1 \quad \epsilon_2)\, y\chi$, **then** $\mathrm{LB} \leftarrow \frac{n\, F_{\mathcal{R}}(\tilde{\mathbf{X}}_i)}{1\ \epsilon_2}$, break;
$\theta \leftarrow \lambda/\mathrm{LB}$;
while $|\mathcal{R}| \leq \theta$ **do** $\mathcal{R} \leftarrow \mathcal{R} \cup$ GenerateRR-Set;
Return $\mathcal{R}$;

if $\Delta s_u < \epsilon$ for some u, the solutions might differ. The approximation result from the theorem stems from the following lemma.

Lemma 5.6. *The function $\mathbb{E}[F(\cdot)]$ is monotone and submodular. Thus the greedy algorithm achieves an approximation ratio of $1 - \frac{1}{e}$.*

Maximizing network indices. To estimate $F(S)$, we define $F_{\mathcal{R}}(S)$ similar to Equation (5.1). The difference is that we drop the quadratic terms $\mathbb{1}[(R_u \cap S) \neq \emptyset, (R_v \cap S) \neq \emptyset] n\, m_{u,v}$, and we set $w_u = \Delta s_u$.

Our algorithm works as follows. We sample a set $\mathcal{R}$ of RR-sets and greedily pick the nodes that maximize $F_{\mathcal{R}}(S)$. The algorithm keeps on adding RR-sets to $\mathcal{R}$ until a statistical test asserts that we have found a lower bound on OPT. More concretely, we keep on sampling if the value estimated by $nF_{\mathcal{R}}(S)$ is *not* a lower bound on OPT (see (1) in Lemma 5.7) and when we stop sampling then we obtain a good enough lower bound LB (see (2) in Lemma 5.7). Then we can apply Lemma 5.3 with $\theta \geq \lambda/\mathrm{LB}$ to obtain our approximation guarantees. We present the pseudocode including the sampling in Algorithm 2 and the greedy subroutine in Algorithm 1. We run our algorithms with parameters $\mathrm{LB}_0 = \max_u |\Delta s_u|$ and $\beta = n(\frac{4}{3}\epsilon_2 + 2)(l \ln n + \ln \log_2 2n + \ln \binom{n}{k}))/\epsilon_2^2$.

Lemma 5.7. *Let $\tilde{\mathbf{X}}$ be the output of Algorithm 1 and suppose that $|\mathcal{R}| = \theta \geq \frac{\beta}{y}$. Then with probability at least $1 - \frac{n^{-\ell}}{\log_2(n)}$: (1) If OPT $< y\chi$, then $n\, F_{\mathcal{R}}(\tilde{\mathbf{X}}) < (1 + \epsilon_2) y\chi$. (2) If $\mathrm{OPT} \geq y\chi$, then $n\, F_{\mathcal{R}}(\tilde{\mathbf{X}}) \leq (1 + \epsilon_2)\mathrm{OPT}$.*

The above approach also extends to other indices if we use $F_{\mathcal{R}}(S)$ as per Equation (5.1) and set $\mathrm{LB}_0 = \max_{u,v} |w_u + m_{u,v}|$.

Table 2: Statistics of the datasets, where n is the number of nodes and m is the number of edges.

Dataset	n	m	Dataset	n	m
Convote	219	521	NipsEgo	2888	2981
Netscience	379	914	PagesGov	7057	89429
WikiTalkHT	404	734	HepPh	11204	117619
WikiVote	889	2914	Anybeat	12645	49132
Reed98	962	18812	CondMat	21363	91286
EmailUniv	1133	5451	Gplus	23613	39182
Hamster	2000	16097	Brightkite	56739	212945
USFCA72	2672	65244	WikiTalk	92117	360767

5.3 The sandwich method

Now we present an algorithm for finding at most k seed nodes that maximize the Dis-Con Index $\mathcal{I}^{dc}_{G,\mathbf{s}}$ and the Controversy Index $C_{G,\mathbf{s}}$. Since these optimization problems are not submodular, we cannot use the greedy algorithm from above. However, the indices' matrices $\mathcal{I}^{dc}(\mathbf{L})$ and $C(\mathbf{L})$ only contain non-negative entries and this allows us to define submodular upper and lower bounds on the objective functions. Thus, we apply the sandwich method [24] to obtain data-dependent approximation guarantees.

We obtain our upper and lower bounds as follows. Let $\mathcal{M}(\mathbf{L}) \in \{\mathcal{I}^{dc}(\mathbf{L}), C(\mathbf{L})\}$. Now let $\mathcal{M}(\mathbf{L})^U$ be the diagonal matrix in which $\mathcal{M}(\mathbf{L})^U_{ii}$ is the sum of all entries in the i'th row of $\mathcal{M}(\mathbf{L})$. Let $\mu_0(S) = \mathbb{E}[2\mathbf{s}^\intercal \mathcal{M}(\mathbf{L})\, \Delta\hat{\mathbf{s}} + \Delta\hat{\mathbf{s}}^\intercal \mathcal{M}(\mathbf{L})\, \Delta\hat{\mathbf{s}}]$, $\mu_L(S) = \mathbb{E}[2\mathbf{s}^\intercal \mathcal{M}(\mathbf{L})\, \Delta\hat{\mathbf{s}}]$, $\mu_U(S) = \mathbb{E}[2\mathbf{s}^\intercal \mathcal{M}(\mathbf{L})\, \Delta\hat{\mathbf{s}} + \Delta\hat{\mathbf{s}}^\intercal \mathcal{M}(\mathbf{L})^U\, \Delta\hat{\mathbf{s}}]$. Since the entries of all of these matrices are non-negative, we obtain our desired relationship $\mu_L(S) \leq \mu_0(S) \leq \mu_U(S)$.

As both $\mu_L(S)$ and $\mu_U(S)$ are monotone and submodular, a greedy algorithm can approximate them within factor $1 - \frac{1}{e} - \epsilon$. In our *sandwich algorithm*, we greedily select nodes S_L, S_U and S_0 that maximize $\mu_L(S)$, $\mu_U(S)$ and $\mu_0(S)$, respectively. Then we evaluate each of the sets on $\mu_0(S)$ and return the one with the highest objective value, i.e., we return $\arg\max_{S \in \{S_0, S_L, S_U\}} \mu_0(S)$. We obtain the following approximation guarantees.

Theorem 5.8 (Lu et al. [24]). *Let $S^* = \arg\max_{|S| \leq k} \mu_0(S)$. Then $\mu_0(S) \geq \max\left\{\frac{\mu_0(S_U)}{\mu_U(S_U)}, \frac{\mu_L(S^*)}{\mu_0(S^*)}\right\}(1 - \frac{1}{e} - \epsilon)\, \mu_0(S^*)$.*

6 EXPERIMENTS

We proceed to the experimental evaluation. Our experiments were conducted on an Intel Xeon E5 2630 v4 at 2.20 GHz with 128GB memory. Our code is written in Julia and is available on github.[2]

Datasets. We use publicly available real-world datasets [22, 23, 29] of social networks. For each network we extracted the largest connected component. Dataset statistics are presented in Table 2.

Parameters. For each network, we set the innate opinion s_u of each user u uniformly at random in $[0, 1]$ [35]. We set the parameters $p_{u,v}$ as in the weighted cascade model [21, 31, 32], i.e., $p_{u,v} = \frac{1}{d(v)}$, where $d(v)$ is the in-degree of v. We set $w_{u,v} = 1$ for the FJ model. For polarizing campaigns with backfire, we set $\tau = 0.5$. For all of our algorithms and heuristics, we set $\epsilon = 0.1$, $\ell = 1$ and $\epsilon_2 = 0.6$.

[2] https://github.com/SijingTu/WebConf-22-Viral-Marketing-Opinion-Dynamics

Algorithms. We implemented our approximation algorithms from Sec. 5 and we denote them *Sum* for the sum index, and *DisCon* for the disagreement–controversy index. Additionally, we use heuristic versions of the greedy Algorithm 1, together with the statistical test scheme from Algorithm 2. This gives us the following algorithms: *Pol* for maximizing the polarization index, *IntCon* for maximizing the internal conflict, and *Dis* for maximizing disagreement.

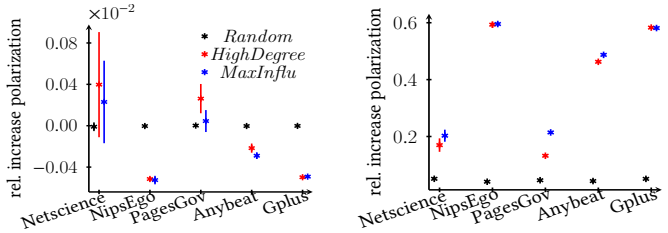
We will see below that those algorithms that include quadratic terms are very costly to run. Therefore, we introduce scalable heuristics. We make two changes in the heuristics: (1) To obtain the seed nodes, we only consider the indices' linear components, but we evaluate the final set of seed nodes on the whole function (including the quadratic part). (2) Following the sampling scheme from Algorithm 2 typically leads to large sampling sizes and sometimes caused our algorithms to run out of memory. Thus, we sample at most $200n$ RR-sets for smaller datasets, and $5n$ RR-sets when $n > 50,000$. This gives good estimates in practice (typically with $< 1\%$ error). We denote the heuristics by *LinDisCon*, *LinPol*, *LinIntCon*, and *LinDis*.

We compare our optimization algorithms against three baselines: *MaxInflu* chooses the seed nodes that maximize the influence; *HighDegree* picks the seed nodes with highest degrees; *Random* selects seed nodes uniformly randomly. Since *Random* is the only randomized baseline, we report average values over 10 runs. As these methods provide us with a fixed seed set, we use the Monte Carlo simulation from Sec. 5.1 to evaluate their results.

Additionally, we compare against a greedy heuristic *FJ* by Chen and Racz [9] that maximizes the indices from Table 1 under the vanilla FJ model. *FJ* is allowed to change k innate user opinions arbitrarily much but, unlike in our model, there is no information spread; we provide *FJ* with the same parameter k as all other algorithms. Unlike for the other methods, we do *not* take the seed set returned by *FJ* and compute its score in our model, but we report the relative increase of *FJ* in the vanilla FJ model; this will allow us to evaluate whether the information spreading makes our model more powerful. We will also include a value *FJUpp* by Gaitonde, Kleinberg and Tardos [16, Thm. 3.4] which gives an analytic upper bound on what is achievable in the setting of *FJ*; we note that this upper bound might be loose (i.e., it is possible that it is too large). Note that if our algorithms achieve values larger than *FJUpp*, our model is strictly more powerful than what is achievable in the vanilla FJ model without the information propagation step.

Evaluation. We report the relative increases of the indices from Sec. 3. That is, for $\mathcal{M}(\mathbf{L})$ being a matrix from Table 1, $\mathbf{s}$ being the non-adjusted innate opinions, and $\hat{\mathbf{s}}$ being the adjusted innate opinions, we report $(\hat{\mathbf{s}}^\intercal \mathcal{M}(\mathbf{L})\,\hat{\mathbf{s}} - \mathbf{s}^\intercal \mathcal{M}(\mathbf{L})\,\mathbf{s})/(\mathbf{s}^\intercal \mathcal{M}(\mathbf{L})\,\mathbf{s})$.

How does viral marketing change the indices? First, let us consider how our baselines influence the user opinions under the spread-acknowledge model. In Figure 2, we report how the polarization index changes when we pick 2% of the nodes as seeds. We repeat our experiments 5 times and present the mean and the variance. In Figure 2(a) we see that marketing campaigns have little effect on the polarization index in the network and increase it by less than 0.1%. However, the situation is very different when we consider polarizing campaigns with backfire (Figure 2(b)): the polarization increases up to 60% and typically increases *at least* 20%



(a) (b)

Figure 2: The relative change of the Polarization Index on different datasets with $k = \lceil 2\% \cdot n \rceil$ seed nodes. The plots show (a) marketing campaigns and (b) polarizing campaigns.

if the most influential users share the polarizing campaign. Using random seed nodes has little impact on the polarization.

Scalability and accuracy of the heuristics. In [33] we show that the heuristics scale linearly in the graph size and are up to three orders of magnitude faster than the greedy algorithms, while being of similar quality. Thus, next we focus on the heuristics that only consider the linear terms and scale to larger datasets.

Experiments for marketing campaigns. Next, we evaluate our methods for marketing campaigns with $k = \lceil 0.5\% \cdot n \rceil$ seed nodes. In Table 3 we report the results for all previously mentioned methods, excluding *HighDegree* which behaves very similarly to *MaxInflu*. We will consider the Sum Index and the Polarization Index and we will evaluate how these indices change based on solutions of algorithms with different objectives. While this might look counterintuitive at first, this approach reveals interesting connections between the different methods we consider and the indices we optimize. For *FJ*, we use two corresponding versions that maximize the Sum Index and the Polarization Index, respectively; for two large datasets, *FJ* and *FJUpp* ran out of time.

Let us consider the Sum Index. The methods *Sum*, *LinDisCon* and *MaxInflu* typically achieve the highest values and all of them are of similar quality. Not surprisingly, this suggests that for marketing campaigns maximizing the user opinions is essentially the same as maximizing influence. For nine datasets, the Sum Index increases by less than 5% but for some it increases by up to 18.75%. Quite interestingly, only on two datasets *LinPol* increases the Sum Index by more than 1%, which suggests that the solutions of *LinPol* and the other methods are quite dissimilar. Additionally, we observe that the solution by *FJ* barely increases the Sum Index.

However, the situation is quite different for the Polarization Index. Here, *LinPol* clearly achieves the biggest increases followed by *LinDis* and *FJ*. Interestingly, on several datasets the seed nodes produced by *Sum*, *LinDisCon* and *MaxInflu* even *decrease* the polarization; we explain this by the fact that if many users increase their opinions with respect to a topic, then the overall acceptance of this topic increases and the topic becomes less polarizing. Additionally, we observe that on all datasets, *LinPol* achieves slightly higher values than *FJ*, even though *FJ* can change the k innate opinions arbitrarily much, while our marketing campaign can increase each innate user opinion by at most ϵ. For both indices, *Random* and *LinIntCon* have little to no effect.

Table 3: Results for marketing campaigns with $k = \lceil 0.5\% \cdot n \rceil$ seeds. We report the relative increase of each index in percent.

Dataset	Sum Index								Polarization Index								
	Sum	*LinDisCon*	*LinPol*	*LinDis*	*LinIntCon*	*MaxInflu*	*Random*	*FJ*	*Sum*	*LinDisCon*	*LinPol*	*LinDis*	*LinIntCon*	*MaxInflu*	*Random*	*FJ*	*FJUpp*
Netscience	**2.79**	2.75	0.74	1.01	0.21	2.78	0.27	0.11	3.15	3.18	**7.54**	5.89	-0.35	3.17	-0.06	2.36	10.54
WikiVote	**4.14**	4.12	0.53	0.64	0.48	4.11	0.3	0.11	-0.64	-0.61	**3.83**	3.2	0.81	-0.58	-0.06	2.92	12.29
Reed98	3.2	**3.22**	0.28	0.3	0.3	3.2	0.27	0.1	0.14	0.09	**10.13**	8.04	0.15	0.11	-0.13	9.56	68.48
EmailUniv	3.31	**3.36**	0.37	0.41	1.18	3.35	0.29	0.11	0.51	0.42	**4.64**	4.13	0.35	0.44	-0.13	3.8	18.12
Hamster	**4.09**	4.07	0.76	0.81	0.45	4.06	0.27	0.1	0.39	0.46	**7.59**	5.82	0.43	0.72	0.01	5.18	25.70
USFCA72	3.09	3.09	0.23	0.22	0.28	**3.1**	0.3	0.11	-0.68	-0.67	**11.7**	11.01	0.47	-0.68	-0.07	11.57	82.48
NipsEgo	**18.75**	**18.75**	0.47	0.1	0.1	**18.75**	0.15	0.1	-5.3	-5.29	**1.71**	0.51	0.12	-5.29	-0.09	0.79	5.34
PagesGov	3.47	3.47	0.53	0.5	0.44	**3.48**	0.28	0.1	0.78	0.79	**7.31**	5.49	0.85	0.51	-0.06	6.96	36.87
HepPh	2.52	**2.53**	0.61	0.67	0.42	2.52	0.26	0.1	-0.3	0.01	**6.83**	4.5	0.43	-0.09	-0.05	3.26	16.03
Anybeat	**11.96**	**11.96**	0.93	1.02	0.53	**11.96**	0.25	0.1	-1.21	-1.19	**3.18**	2.58	0.52	-1.24	-0.08	1.7	7.80
CondMat	**2.79**	**2.79**	0.59	0.65	0.48	**2.79**	0.25	0.1	0.35	0.58	**6.9**	4.61	0.44	0.31	-0.07	3.42	15.69
Gplus	**18.06**	**18.06**	3.85	0.37	0.44	**18.06**	0.26	0.1	-4.98	-4.98	**6.2**	1.03	0.29	-4.98	-0.07	0.92	6.41
Brightkite	6.16	6.15	0.72	0.89	0.53	**6.17**	0.27	-	-0.17	-0.06	**4.27**	2.53	0.47	-0.24	-0.07	-	-
WikiTalk	9.27	9.27	1.73	1.59	0.71	**9.28**	0.29	-	-0.82	-0.71	**3.37**	2.63	0.62	-0.79	-0.09	-	-

Experiments with polarizing campaigns. Next, we consider polarizing campaigns with backfire and $k = \lceil 0.5\% \cdot n \rceil$ seed nodes. We report our results in Table 4.

Let us start with the Sum Index. Unlike in marketing campaigns, now *Sum* is clearly the best method overall and outperforms *MaxInflu*. However, for all methods the increase is very small, indicating that it is difficult to increase the sum of the user opinions with polarizing campaigns.

Now consider the Polarization Index where the increases compared to the marketing campaigns are startling. On 10 out of 14 datasets, *LinPol* increases the polarization by *at least* 10% and the biggest increases reach up to 59%. This is in stark contrast to marketing campaigns where on all but two datasets the polarization increased by *at most* 10%. Even *MaxInflu* always increases the polarization by more than 5% and up to 59%. Next, we observe that *LinPol* achieves much larger increases in polarization than *FJ*, typically being at least factor 2 larger and up to factor 75 (for NipsEgo). Finally, we observe that on three datasets, *LinPol* outperforms that analytic lower bound *FJUpp* and for 10 out of 12 datasets it is within factor 3. These findings suggest that for polarizing campaigns, the information spread is very powerful compared to only changing the innate opinions of a given set of users.

Further experiments are provided in [33].

7 CONCLUSIONS

We presented a novel model that allows to quantify how viral information effects user opinions in online social networks. We presented algorithms to simulate the model and to optimize different network indices. This allowed us to understand how much impact adversaries can have on the social network. Our experiments showed that marketing campaigns and polarizing contents behave very differently. While for marketing campaigns it is possible to significantly increase the user opinions, this seems very difficult for polarizing contents. However, the picture is vastly different for the polarization in the network: it barely increases for marketing campaigns but for polarizing contents the increase can be very high, even when the number of seed nodes is small. We believe that this gives an insight into the growing polarization observed in today's social media.

There are several interesting directions for future work. Obtaining approximation algorithm for polarizing contents is intriguing. Another important question is to study how the parameters of our model should be set to capture real-world behaviors as accurately as possible; beyond pure parameter estimation, this might involve replacing the independent cascade model or the FJ-model with other models from the literature.

ACKNOWLEDGMENTS

We are deeply grateful to Aristides Gionis for his mentorship and many discussions during this project. This research is supported by the Academy of Finland projects AIDA (317085) and MLDB (325117), the ERC Advanced Grant REBOUND (834862), the EC H2020 RIA project SoBigData++ (871042), and the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. Our computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX partially funded by the Swedish Research Council through grant agreement no. 2018-05973.

REFERENCES

[1] Victor Amelkin, Francesco Bullo, and Ambuj K. Singh. 2017. Polar Opinion Dynamics in Social Networks. *IEEE Trans. Autom. Control.* 62, 11 (2017), 5650–5665.
[2] Cigdem Aslay, Francesco Bonchi, Laks VS Lakshmanan, and Wei Lu. 2017. Revenue maximization in incentivized social advertising. *Proceedings of the VLDB Endowment* 10, 11 (2017), 1238–1249.
[3] Cigdem Aslay, Wei Lu3 Francesco Bonchi, Amit Goyal, and Laks VS Lakshmanan. 2015. Viral Marketing Meets Social Advertising: Ad Allocation with Minimum Regret. *Proceedings of the VLDB Endowment* 8, 7 (2015).
[4] Nicola Barbieri, Francesco Bonchi, and Giuseppe Manco. 2013. Topic-aware social influence propagation models. *Knowledge and information systems* 37, 3 (2013), 555–584.
[5] David Bindel, Jon Kleinberg, and Sigal Oren. 2015. How bad is forming your own opinion? *Games and Economic Behavior* 92 (2015), 248–265.
[6] Christian Borgs, Michael Brautbar, Jennifer Chayes, and Brendan Lucier. 2014. Maximizing social influence in nearly optimal time. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*. SIAM, 946–957.
[7] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. 2009. Statistical physics of social dynamics. *Reviews of modern physics* 81, 2 (2009), 591.
[8] Damon Centola and Michael Macy. 2007. Complex contagions and the weakness of long ties. *Amer. J. Sociology* 113, 3 (2007), 702–734.
[9] Mayee Chen and Miklos Z Racz. 2020. Network disruption: maximizing disagreement and polarization in social networks. *arXiv preprint arXiv:2003.08377* (2020).
[10] Xi Chen, Jefrey Lijffijt, and Tijl De Bie. 2018. Quantifying and minimizing risk of conflict in social networks. In *KDD*. 1197–1205.
[11] Xi Chen, Panayiotis Tsaparas, Jefrey Lijffijt, and Tijl De Bie. 2019. Opinion Dynamics with Backfire Effect and Biased Assimilation. *CoRR* abs/1903.11535 (2019).

Table 4: Results for polarizing campaigns with $k = \lceil 0.5\% \cdot n \rceil$ seeds. We report the relative increase of each index in percent.

Dataset	Sum Index								Polarization Index								
	Sum	*LinDisCon*	*LinPol*	*LinDis*	*LinIntCon*	*MaxInflu*	*Random*	*FJ*	*Sum*	*LinDisCon*	*LinPol*	*LinDis*	*LinIntCon*	*MaxInflu*	*Random*	*FJ*	*FJUpp*
Netscience	**0.48**	0.46	0.04	0.07	0.34	0.34	-0.01	0.11	2.72	5.03	**7.62**	5.54	5.65	5.66	0.62	2.36	10.54
WikiVote	**0.33**	0.25	-0.28	-0.28	-0.27	-0.33	-0.02	0.11	3.14	5.83	9.46	**9.5**	9.09	9.14	0.64	2.92	12.29
Reed98	**0.27**	**0.27**	0.02	0.18	0.16	0.16	0.01	0.1	6.6	7.68	**15.82**	11.35	8.87	8.79	0.74	9.56	68.48
EmailUniv	**0.33**	0.31	-0.23	-0.21	-0.18	-0.16	-0.02	0.11	3.39	5.23	**7.72**	7.36	6.87	6.81	0.79	3.8	18.12
Hamster	**0.3**	0.27	-0.07	-0.09	-0.11	-0.1	-0.01	0.1	4.3	5.2	**11.1**	8.98	8.55	8.56	0.66	5.18	25.70
USFCA72	**0.22**	**0.22**	0.05	-0.04	-0.02	-0.02	-0.01	0.11	6.85	10.05	**13.18**	7.39	5.68	5.74	0.74	11.57	82.48
NipsEgo	**0.46**	0.22	0.2	0.21	0.21	0.21	0.0	0.1	38.05	59.55	**59.57**	59.56	59.55	**59.57**	0.46	0.79	5.34
PagesGov	**0.29**	0.28	-0.03	0.01	0.0	-0.0	-0.0	0.1	4.78	5.89	**12.73**	10.1	6.19	7.48	0.66	6.96	36.87
HepPh	**0.33**	0.32	-0.01	-0.08	-0.1	-0.1	-0.01	0.1	3.69	5.17	**7.62**	5.91	4.61	5.16	0.66	3.26	16.03
Anybeat	**0.42**	0.3	0.11	0.12	0.14	0.12	0.0	0.1	29.3	38.14	**39.84**	39.75	39.12	39.55	0.48	1.7	7.80
CondMat	**0.36**	0.32	0.01	0.01	0.02	0.02	0.0	0.1	4.26	5.58	**8.28**	6.68	5.32	5.84	0.65	3.42	15.69
Gplus	**0.49**	0.15	0.1	0.1	0.1	0.1	0.0	0.1	29.75	57.48	**57.94**	**57.94**	57.92	57.93	0.66	0.92	6.41
Brightkite	**0.38**	0.24	-0.02	-0.0	0.0	0.01	0.0	-	5.66	13.35	**15.86**	15.65	15.17	15.58	0.7	-	-
WikiTalk	**0.49**	0.29	0.02	0.01	0.02	0.02	0.0	-	13.46	25.84	**28.79**	28.71	28.19	28.57	0.73	-	-

[12] Uthsav Chitra and Christopher Musco. 2019. Understanding filter bubbles and polarization in social networks. *arXiv preprint arXiv:1906.08772* (2019).

[13] Pranav Dandekar, Ashish Goel, and David T Lee. 2013. Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences* 110, 15 (2013), 5791–5796.

[14] Morris H DeGroot. 1974. Reaching a consensus. *J. Amer. Statist. Assoc.* 69, 345 (1974), 118–121.

[15] Noah E Friedkin and Eugene C Johnsen. 1990. Social influence and opinions. *Journal of Mathematical Sociology* 15, 3-4 (1990), 193–206.

[16] Jason Gaitonde, Jon Kleinberg, and Eva Tardos. 2020. Adversarial perturbations of opinion dynamics in networks. In *EC*. ACM, 471–472.

[17] Aristides Gionis, Evimaria Terzi, and Panayiotis Tsaparas. 2013. Opinion maximization in social networks. In *Proceedings of the 2013 SIAM International Conference on Data Mining*. SIAM, 387–395.

[18] Douglas Guilbeault, Joshua Becker, and Damon Centola. 2018. Complex contagions: A decade in review. *Complex spreading phenomena in social systems* (2018), 3–25.

[19] Naoki Hirakura, Masaki Aida, and Konosuke Kawashima. 2020. A Model of Polarization on Social Media Caused by Empathy and Repulsion. *IEICE Proceedings Series* 63, D3-4 (2020).

[20] Matthew O Jackson. 2008. *Social and Economic Networks*. Princeton University Press.

[21] David Kempe, Jon Kleinberg, and Éva Tardos. 2015. Maximizing the Spread of Influence through a Social Network. *Theory Of Computing* 11, 4 (2015), 105–147.

[22] Jérôme Kunegis. 2013. Konect: the koblenz network collection. In *Proceedings of the 22nd international conference on World Wide Web*. 1343–1350.

[23] Jure Leskovec and Andrej Krevl. 2014. SNAP Datasets: Stanford Large Network Dataset Collection. http://snap.stanford.edu/data.

[24] Wei Lu, Wei Chen, and Laks VS Lakshmanan. 2015. From competition to complementarity: comparative influence diffusion and maximization. *Proceedings of the VLDB Endowment* 9, 2 (2015), 60–71.

[25] Antonis Matakos, Evimaria Terzi, and Panayiotis Tsaparas. 2017. Measuring and moderating opinion polarization in social networks. *Data Mining and Knowledge Discovery* 31, 5 (2017), 1480–1505.

[26] Corrado Monti, Giuseppe Manco, Cigdem Aslay, and Francesco Bonchi. 2021. Learning Ideological Embeddings from Information Cascades. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 1325–1334.

[27] Cameron Musco, Christopher Musco, and Charalampos E Tsourakakis. 2018. Minimizing polarization and disagreement in social networks. In *Proceedings of the 2018 World Wide Web Conference*. 369–378.

[28] Brendan Nyhan and Jason Reifler. 2010. When corrections fail: The persistence of political misperceptions. *Political Behavior* 32, 2 (2010), 303–330.

[29] Ryan Rossi and Nesreen Ahmed. 2015. The network data repository with interactive graph analytics and visualization. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.

[30] Arun V. Sathanur, Mahantesh Halappanavar, Yi Shi, and Yalin E. Sagduyu. 2018. Exploring the Role of Intrinsic Nodal Activation on the Spread of Influence in Complex Networks. In *Social Network Based Big Data Analysis and Applications*. 123–142.

[31] Youze Tang, Yanchen Shi, and Xiaokui Xiao. 2015. Influence maximization in near-linear time: A martingale approach. In *Proceedings of the 2015 ACM SIGMOD international conference on management of data*. 1539–1554.

[32] Youze Tang, Xiaokui Xiao, and Yanchen Shi. 2014. Influence maximization: Near-optimal time complexity meets practical efficiency. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*. ACM, 75–86.

[33] Sijing Tu and Stefan Neumann. 2022. A Viral Marketing-Based Model For Opinion Dynamics in Online Social Networks. *CoRR* abs/2202.03573 (2022). arXiv:2202.03573 https://arxiv.org/abs/2202.03573

[34] Zhefeng Wang, Yu Yang, Jian Pei, Lingyang Chu, and Enhong Chen. 2017. Activity maximization by effective information diffusion in social networks. *IEEE Transactions on Knowledge and Data Engineering* 29, 11 (2017), 2374–2387.

[35] Wanyue Xu, Qi Bao, and Zhongzhi Zhang. 2021. Fast Evaluation for Relevant Quantities of Opinion Dynamics. In *WebConf*. 2037–2045.

[36] Liwang Zhu, Qi Bao, and Zhongzhi Zhang. 2021. Minimizing Polarization and Disagreement in Social Networks via Link Recommendation. *Advances in Neural Information Processing Systems* 34 (2021).