Effect of Deception in Influence Maximization and Polarization on Social Networks: A Sheaf Laplacian Approach

Mehmet Emin Aktas
Georgia State University
Atlanta, Georgia, USA
maktas@gsu.edu

Esra Akbas
Georgia State University
Atlanta, Georgia, USA
eakbas1@gsu.edu

Ashley Hahn
Florida International University
Miami, Florida, USA
aaslihan@fiu.edu

Mehmet Eren Ahsen
University of Illinois Urbana-Champaign
Champaign, Illinois, USA
ahsen@illinois.edu

ABSTRACT

In the contemporary era of social media and online communication, comprehending the dynamics of information diffusion in social networks has become crucial. This research article investigates the effects of deception on information diffusion, specifically focusing on influence maximization and polarization in social networks. We propose an analytic model of deception within social networks. Building upon the sheaf Laplacian diffusion model derived from algebraic topology, we examine opinion dynamics in the presence of deception. Next, we redefine the Laplacian centrality, an influential node detection method originally designed for regular graphs, to quantify the influence of deception in influence maximization using the sheaf Laplacian. Additionally, we employ the sheaf Laplacian to model polarization in networks and investigate the impact of deception on polarization using two distinct polarization measures. Through extensive experiments conducted on synthetic and real-world networks, our findings suggest that deceptive individuals wield more influence than honest users within social networks. Furthermore, we demonstrate that deception amplifies polarization in networks, with influential individuals playing a significant role in deepening the polarization phenomenon.

CCS CONCEPTS

• **Information systems → Social networks**; • **Theory of computation → Graph algorithms analysis**; • **Human-centered computing → Social network analysis**.

KEYWORDS

Information diffusion, sheaf Laplacian, influence maximization, polarization, deception

ACM Reference Format:
Mehmet Emin Aktas, Esra Akbas, Ashley Hahn, and Mehmet Eren Ahsen. 2018. Effect of Deception in Influence Maximization and Polarization on Social Networks: A Sheaf Laplacian Approach. In *Proceedings of 19TH INTERNATIONAL WORKSHOP ON MINING AND LEARNING WITH GRAPHS (MLG 2023)*. ACM, New York, NY, USA, 9 pages. https://doi.org/XXXXXXX.XXXXXXX

MLG 2023, August 06–10, 2023, Long Beach, CA

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Social networks have become an integral part of our daily lives, revolutionizing the way we connect, communicate, and share information. With the widespread popularity of platforms such as Facebook, Tiktok, and Instagram, studying social networks has gained significant importance in understanding human behavior and the dynamics of information diffusion [5, 6, 27]. As one's beliefs and opinions emerge in discourse dialogically [17, 21], it is essential to examine the network influence by focusing on social interaction and social actors in a network.

Within a social network, not all individuals contribute equally to the network effect. Certain users possess greater influence than others when it comes to shaping the beliefs and opinions of others. Identifying these highly influential users is an essential task for various reasons, such as predicting behavior and trends [22], understanding information diffusion [7], and enhancing marketing strategies [29].

On the other hand, polarization within social networks is a significant social phenomenon that can have far-reaching implications [3]. It may occur when influential users promote polarizing views that resonate with their followers on social networks. This can lead to the formation of echo chambers, where people only consume information and interact with others who share their views, leading to further entrenchment in their beliefs [9].

There are many influence maximization and polarization models in the literature [2, 8, 23, 24]. Although these models are widely used for different graph mining problems, they suffer from an important issue: They assume that in social networks, users are honest with each other. They ignore the role of deception on social networks. However, deception appears to be surprisingly common among humans and within social networks [1, 19, 26]. Some studies report that lies are more common in online interactions than face-to-face interactions [13]. Hence, deception is inevitable in social networks and we need to take deception into consideration while modeling information diffusion and detecting influential nodes and polarization in social networks.

In this paper, we study the effect of deception in influence maximization and polarization on social networks to understand whether liars are more influential and/or cause deeper polarization in social networks. Inspiring from [16], we first model deception on social networks consisting of honest interactions and exchanging prosocial, lying to protect someone or to benefit or help others, or antisocial, lying to hurt someone else intentionally, lies between individuals [11]. We also control the honesty level of users with an honesty parameter to see the effect of different deception levels. Next, we model the opinion dynamics in social networks using the sheaf Laplacian [14]. Sheaf Laplacian provides a flexible model that allows users to express their opinion however they choose and selectively lie to their neighbors. As the next step, we redefine a node centrality measure, namely Laplacian centrality, for the sheaf Laplacian to detect influential nodes when deception is present in the network. Then, we employ this centrality measure for detecting the influence of each node to see the effect of deception on influence maximization. We further use sheaf Laplacian to model polarization when deception is present.

We conduct extensive experiments to evaluate our model on various synthetic and real-world social networks. Our experimental results show that liars, regardless of being prosocial or antisocial, are more influential than honest users. We further show that liars can cause deeper polarization in networks compared to honest people.

The paper is formatted as follows. In Section 2, we discuss the preliminary concepts for sheaf Laplacian and the related work. In Section 3, we present our methodology on modeling deception in social networks and constructing the sheaf Laplacian. We also explain how we extend Laplacian centrality to the sheaf Laplacian and model polarization. In Section 4, we explain our evaluation method and present our results on various synthetic and real-world datasets. Our final remarks with future work directions are found in Section 5.

2 BACKGROUND

In this section, we present the preliminary concepts for sheaf data structure and the sheaf Laplacian that we use to model the opinion dynamics on social networks when deception is present. We further present the related work on the effects of deception.

2.1 Preliminaries

A sheaf is a data structure associating data spaces to vertices and edges of a graph, with further telling how the data over different parts of the graph should be related. More formally, we can define a sheaf as follows.

Definition 1. *For a given graph, $G = (V, E)$, a sheaf $\mathcal{F}$ on G consists of a vector space $\mathcal{F}_v$ for each vertex $v \in V$, a vector space $\mathcal{F}_e$ for each edge $e \in E$, and a linear transformation $\mathcal{F}_{v\to e} : \mathcal{F}_v \to \mathcal{F}_e$ for each incident vertex-edge pair.*

In the notion of modeling opinion dynamics on social networks, the vector space $\mathcal{F}_v$ over each vertex $v \in V$ is the opinion space of the vertex. Mathematically, this is a real vector space with a basis of the collection of topics where the basis consists of users' social, demographic, and cultural dynamics and the moral assemblages/bases behind their opinions. The scalar values on each basis

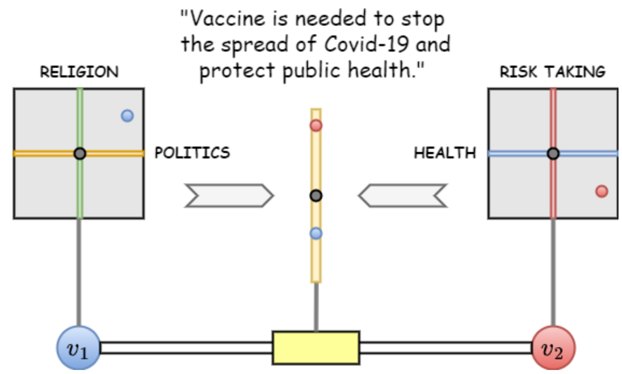


Figure 1: A sheaf structure on an edge

element correspond to negative, neutral, or positive opinions about the topic. We would like to note here that these base elements are not necessarily the same nor the same number for each user. For example, let the topic of discussion be vaccination during a pandemic. While for one user, the opinion basis could be politics and religion, for another user, it could be health and risk-taking. Then, they share their private opinion publicly based on their opinion bases. See Figure 1 for an illustration of this example. The vertex on the left (v_1) thinks both politics and religion are important in their opinion about the vaccination, i.e., its opinion base is {religion, politics}, and the user does not support it. On the other hand, the vertex on the right (v_2) takes health very important but the user also does not want to take risks, i.e., its opinion base is {health, risk taking}. As a result, that user supports the vaccination.

Furthermore, in this opinion dynamics model, the vector space $\mathcal{F}_e$ over each edge $e \in E$ is the discourse space where each user represents their opinions on the topics of discussion by formulating stances as a linear combination of existing opinions on personal opinion basis. This expression of opinions is modeled using the linear transformations $\mathcal{F}_{v\to e} : \mathcal{F}_v \to \mathcal{F}_e$. For example, let u and v be two users that are connected with an edge e in a social network. Let $x_u \in \mathcal{F}_u$ and $x_v \in \mathcal{F}_v$ be their opinions. If $\mathcal{F}_{u\to e}(x_u) = \mathcal{F}_{v\to e}(x_v)$, then there is a local consensus between u and v. For example, in Figure 1, the users do not have a consensus initially since their public discourse on vaccination does not coincide.

The sheaf Laplacian is defined similarly to the graph Laplacian. We first bundle all the data over vertices and over edges into a grouped vector spaces as follows

$$C^0(G; \mathcal{F}) = \bigoplus_{v\in V(G)} \mathcal{F}_v, \quad C^1(G; \mathcal{F}) = \bigoplus_{e\in E(G)} \mathcal{F}_e.$$

C^0 is called 0-cochains and it consists of a choice of data, $x_v \in \mathcal{F}_v$, for every vertex $v \in V$. Similarly, C^1 is called 1-cochains and it consists of a choice of data over each edge $e \in E$. Then, we tie the data over vertices (0-cochains) and edges (1-cochains) together with a linear transformation, called the ***coboundary map***, $\delta : C^0(G; \mathcal{F}) \to C^1(G; \mathcal{F})$. For an (arbitrarily) oriented edge $e = u \to v$, we define δ explicitly as follows:

$$(\delta x)_e = \mathcal{F}_{v\to e}(x_v) - \mathcal{F}_{v\to e}(x_u).$$

Then, the **sheaf Laplacian** is given by

$$L_{\mathcal{F}} = \delta^T \delta : C^0(G; \mathcal{F}) \to C^0(G; \mathcal{F}).$$

The sheaf Laplacian does not depend on the choice of orientations while constructing the coboundary map.

EXAMPLE 2. *The sheaf in Figure 2 has the following coboundary map*

$$\delta = \left[\begin{array}{c|cc|c|cc} -1 & -2 & 1 & 0 & 0 & 0 \\ \hline 0 & -2 & 3 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 3 & -1 & 1 \\ \hline 2 & 0 & 0 & 0 & -1 & 0 \end{array}\right]$$

and the sheaf Laplacian

$$L_{\mathcal{F}} = \left[\begin{array}{c|cc|c|cc} 5 & 2 & -1 & 0 & -2 & 0 \\ \hline 2 & 8 & -8 & 2 & 0 & 0 \\ -1 & -8 & 10 & -3 & 0 & 0 \\ \hline 0 & 2 & -3 & 10 & -3 & 3 \\ \hline -2 & 0 & 0 & -3 & 2 & -1 \\ 0 & 0 & 0 & 3 & -1 & 1 \end{array}\right].$$

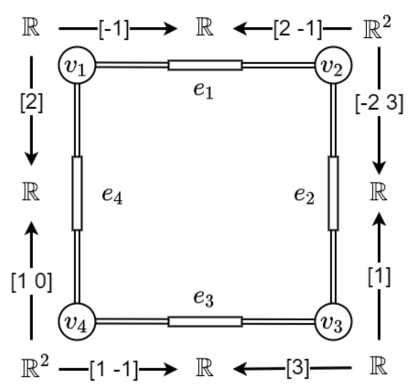


Figure 2: A sheaf structure on a 4-cycle graph. The dimension of vector spaces over v_1 and v_3 are 1 where it is 2 for v_2 and v_4. The dimension of vector spaces over edges are 1 as well.

After getting the sheaf Laplacian $L_{\mathcal{F}}$, we can modify the heat equation for graph Laplacian as

$$\frac{dx}{dt} = -\alpha L_{\mathcal{F}} x, \alpha > 0. \tag{1}$$

Here, x represents an opinion distribution over vertices and $x_v \in \mathcal{F}(v)$ represents the opinion of individual v. This heat equation has the following solution

$$x(t) = \exp(-t\alpha L_{\mathcal{F}})x(0)$$

where $x(0)$ is the initial opinion distribution.

2.2 Related Work

Trust plays an important role in social relations since it affects individuals' willingness to engage in exchange interactions [18]. Hence, deception becomes also important since it may destroy the stability of trust-based relationships [30]. Deception has also effects on social networks. For example, in [16], the authors show how lying can cause social networks to become fragmented. They also study the effects of prosocial and antisocial lies separately. Furthermore, in [4], they find that lies shape the topology of social networks and cause the formation of tightly linked, small communities. They also find that liars are the ones that connect communities of different opinions, hence they have substantial centrality in the network. [28] uses the positivity bias for predicting the use of prosocial lies on Facebook. The authors in [10] study the retweeting activity on Twitter to detect deception. We refer the readers to the survey paper [12] for more studies on deception on social networks.

3 METHODOLOGY

In this section, we first explain how we model deception in social networks. Next, we discuss how we construct the sheaf Laplacian using our deception model. Then, we redefine the centrality method originally defined for the graph Laplacian for the sheaf Laplacian to detect influential nodes in the network. We further model diffusion where there is a user that does not change his opinion (i.e., stubborn) and use this model on modeling polarization.

3.1 Modelling deception in social networks

For a vertex $v_i \in G$, let $x_i(t)$ represent the opinion of v_i about a topic at time t. We can take $x_i \in [-1, 1]$ with -1 meaning total disagreement and 1 meaning total agreement. These are the private opinion of users. To find the sheaf Laplacian of the graph, we need to know how each user discloses his opinion publicly. For the disclosing process, in this paper, we assume users are categorized into three groups: honest, prosocial liar and antisocial liar following [16]. As explained in the introduction, prosocial lies are said to benefit someone whereas antisocial lies are intended to hurt.

Inspiring by Iniguez et. al. [16], we model these three different opinion disclosure models based on both users' and their friends' opinions. The amount of the information, w_{ji}, flowing from i to j can be defined as

$$w_{ji} = \begin{cases} x_i & \text{if } i \text{ is honest} \\ \tau x_i + (1-\tau)x_j & \text{if } i \text{ is prosocial liar} \\ \tau x_i - (1-\tau)x_j & \text{if } i \text{ is antisocial liar} \end{cases} \tag{2}$$

where $\tau \in [0, 1]$ is the honesty parameter. When $\tau = 1$, liars are also honest and $\tau = 0$, they are completely dishonest. In this model, while what others think does not change how they disclose their opinion for honest users, liars (both prosocial and antisocial) express their opinion based on the opinions of their neighbors, i.e., x_j in w_{ji}. This is an issue since users cannot know the private opinions of their neighbors, instead, they can only know how they disclose their opinion publicly, i.e., their public opinion. To tackle this issue, instead of using private opinion, we define the public opinion of the user i, y_i, by taking the average amount of the information flowing from this user to his neighbors as follows

$$y_i = \frac{1}{k_i} \sum_{j \in N_i} w_{ji} \tag{3}$$

where k_i is the degree of vertex v_i and N_i is the set of neighbors of v_i in G. In other words, while users within a social network may be unaware of their friends' personal viewpoints on a particular subject, such as the importance of a flu vaccine, they can still gather insights into their friends' broader perspectives, such as their political inclinations, whether they lean towards conservatism or liberalism. Hence, y_i keeps the overall perspective of the ith user in the network. As we will explain in the next section, our deception model assumes that users communicate with their neighbors based on their neighbors' overall perspective, i.e., y_i's.

3.2 Sheaf Laplacian construction

The key information we need to construct the sheaf Laplacian $L_{\mathcal{F}}$ is the linear transformations $\mathcal{F}_{v\to e} : \mathcal{F}_v \to \mathcal{F}_e$ between a vertex and its neighbors. In other words, we need to know how each user discloses his opinion publicly with his neighbors using his opinion basis. As we discuss in the previous section, opinion disclosures (i.e., linear transformations) depend on whether the user is honest, a prosocial liar, or an antisocial liar. Based on the model in the previous section, we combine Equations 2 and 3 and obtain the linear transformation from v_i to v_j through the edge e as follows

$$\mathcal{F}_{v\to e}(x_i) = \begin{cases} x_i & \text{if user } i \text{ is honest} \\ \tau x_i + (1-\tau)y_j & \text{if user } i \text{ is prosocial liar} \\ \tau x_i - (1-\tau)y_j & \text{if user } i \text{ is antisocial liar.} \end{cases}$$

An illustrative example is available in Figure 3. In the matrix notation, the linear transformation is given as

$$\mathcal{F}_{v\to e} : \begin{cases} [1] & \text{if user } i \text{ is honest} \\ [\tau + (1-\tau)\frac{y_j}{x_i}] & \text{if user } i \text{ is prosocial liar} \\ [\tau - (1-\tau)\frac{y_j}{x_i}] & \text{if user } i \text{ is antisocial liar.} \end{cases}$$

In this model, we take the opinion space over each vertex and the discourse space over each edge as 1-dimensional for simplicity. The opinion space over each vertex simply takes the private opinion x_i for each user i, and the discourse space over each edge $e = v_i \to v_j$ takes the public disclosure of ith user's private opinion, x_i, based on the private jth user's opinion, y_j. On the other hand, this model can be generalized to any dimension of opinion and disclosure spaces.

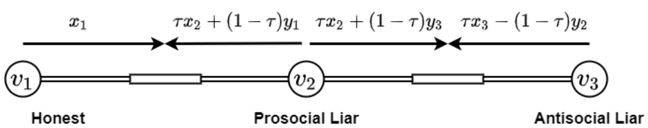


Figure 3: Linear transformations between vertices in the presence of a honest, a prosocial liar and an antisocial liar.

Next, to construct the sheaf Laplacian, we need to define the coboundary map δ for each edge using linear transformations. Let $e = v_i \to v_j$ be an oriented edge. Then, the coboundary map on e is defined as

$$(\delta x)_e = \mathcal{F}_{v\to e}(x_i) - \mathcal{F}_{v\to e}(x_j).$$

An example of how to construct the coboundary map and sheaf Laplacian matrices is available in Example 2.

3.3 Influential node detection

Sheaf Laplacian allows us to model information diffusion when deception is present. As the next step of this research, we need to detect the influential nodes in the network with deception. Most of the influence maximization methods use graph adjacency information as input. In other words, they use graph topology to detect influential users. On the other hand, it is not possible to reverse engineer sheaf Laplacian and generate a regular graph that keeps the sheaf data structure. Hence, we need to develop an influence maximization model that can take sheaf Laplacian as an input directly. In the literature, there is an influential node detection method that uses the graph Laplacian as the input, namely Laplacian centrality [24]. However, we still need to prove that Laplacian centrality can be extended to sheaf Laplacian as well. In this section, we first define what Laplacian centrality is and prove it is well-defined and can be used when we input sheaf Laplacian.

Laplacian centrality is based on the Laplacian energy of the network. The Laplacian energy is defined as follows.

Definition 3. *Let G be a weighted network on n vertices and L be the graph Laplacian of G with the eigenvalues $\lambda_1, ..., \lambda_n$. Then, the Laplacian Energy of G is given by*

$$E_L(G) = \sum_{i=1}^{n} \lambda_i^2.$$

Next, the Laplacian centrality of a node is measured as the drop of Laplacian energy in the network when that node and its adjacent edges are removed.

Definition 4. *Let G be a weighted graph and G_i be the network obtained by deleting the vertex v_i and its adjacent edges from G. Then, the Laplacian centrality $C_L(v_i, G)$ of v_i is given by*

$$C_L(v_i, G) = \frac{E_L(G) - E_L(G_i)}{E_L(G)}.$$

We give an example of the Laplacian centrality on a randomly generated scale-free graph in Figure 4. Next, we prove that we can redefine this centrality to the sheaf Laplacian as well.

Theorem 5. *Let $G = (V, E)$ be a graph and $\mathcal{F}$ be a sheaf defined on G. Then, for a vertex $v_i \in V$, the Laplacian centrality based on the sheaf Laplacian, $C_{L_{\mathcal{F}}}(v_i, G)$ is well-defined.*

Proof. Let L and $L_{\mathcal{F}}$ be the graph Laplacian and the sheaf Laplacian of G, respectively. There are two basic differences between these two matrices. First, the off-diagonal entries of L are all non-positive where $L_{\mathcal{F}}$ may have positive off-diagonal entry. Second, the sum of the off-diagonal entry in a row in L equals to the negative of the diagonal entry on that row, but this is not necessarily true for $L_{\mathcal{F}}$. Here, we prove that these two properties of $L_{\mathcal{F}}$ do not have an effect on defining the Laplacian energy.

In Theorem 1 of [24], they show that

$$E_L(G) = \sum_{i=1}^{n} d_i^2 + 2\sum_{i<j} w_{ij}^2$$

where d_i and w_{ij} are on and off diagonal of L, respectively. As we see in this definition, we take the square of the off-diagonal entries, i.e., sign of these entries has no importance. This addresses the first difference. Moreover, in the proof of Theorem 1, they do not use the fact that $d_i = -\sum_{j,j\neq i}^{n} w_{ij}$, i.e., this difference is again no importance. This addresses the second difference. As a result, we can redefine Laplacian centrality to the sheaf Laplacian. □

This theorem proves that we can use Laplacian centrality via sheaf Laplacian to detect the influential nodes when deception is present in social networks.

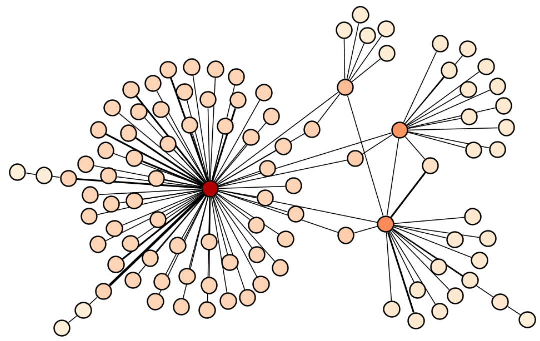


Figure 4: A randomly generated scale-free graph. Node colors represent Laplacian centrality scores. The darker color shows a larger score.

3.4 Polarization detection

The sheaf Laplacian allows us to model information diffusion when there are users that do not change their opinions in response to communication with their neighbors. We can further use this model to study the effect of deception on polarization. Our methodology here is as follows. We first select the most influential user based on Laplacian centrality as the seed node and set its opinion to 1 and the rest of the users' opinion to 0. Then, we assign the deception type of this most influential user as honest, prosocial, and antisocial liars, and in each case, we model diffusion using sheaf Laplacian where it allows to keep the seed node's opinion as 1 and let others' opinion change. When diffusion reaches a steady state level, we check the distribution of all opinions and measure polarization within the network. Below, we explain the details of the information diffusion model when there are "stubborn" users that do not change their opinions in the network and how we measure the polarization after we model the diffusion process.

3.4.1 Stubbornness and Sheaf Laplacian. To consider stubborn users in diffusion, the authors in [14] slightly change Equation 1. Let $U \in V$ be a subset of vertices in G that corresponds to the stubborn users. Then, we have

$$\left.\frac{dx}{dt}\right|_v = \begin{cases} \alpha(L_{\mathcal{F}}x)_v & \text{if } v \notin U \\ 0 & \text{if } v \in U. \end{cases} \quad (4)$$

This new heat equation can model diffusion so that the vertices in U do not change their opinion, whereas other vertices follow the regular diffusion model. Since stubborn vertices do not change opinion, we need to solve this equation for other vertices $Y = V \setminus U$. In the same paper, the solution to this equation is given as

$$y(t) = e^{-t\alpha L_{\mathcal{F}}[Y,Y]} y_0^{\|} - L_{\mathcal{F}}[Y,Y]^{\dagger}(I - e^{-t\alpha L_{\mathcal{F}}[Y,Y]t}) L_{\mathcal{F}}[Y,U]u$$

where $L_{\mathcal{F}}[\cdot,\cdot]$ denotes the block submatrix restricted to the indicated vertex set, $L_{\mathcal{F}}[Y,Y]^{\dagger}$ is the Moore-Penrose pseudoinverse matrix of $L_{\mathcal{F}}[Y,Y]$, $y_0^{\|}$ is the orthogonal projection of the initial opinion vector y_0 to $\mathrm{im} L_{\mathcal{F}}[Y,Y]$ (the proof of this solution is available in [14]). Thanks to this solution, we can model diffusion after setting the seed vertex opinion to 1 and the rest of the users' opinion to 0. Here, while the seed vertex's opinion does not change, the other vertices' opinion follows the solution in Equation 4.

3.4.2 Measuring Polarization. After obtaining the opinion distribution over vertices, we measure polarization in the network with two different methods. As the first method, we check the standard deviation of the opinion distribution. As the second method, we use the GE polarization index. This index is based on the generalized Euclidean distance, which estimates how much effort it would take to travel from one opinion to another in the network [15]. In other words, it estimates the distance between two vectors on a network, e.g., representing people's opinions. To define this index, we first split the opinion distribution vector x into two vectors: x^+ and x^-. x^+ contains all positive opinions and zero otherwise; x^- contains the absolute value of all negative opinions and zero otherwise. Then, the GE index $\delta_{G,x}$ is defined as

$$\delta_{G,x} = \sqrt{(x^+ - x^-)^T L_{\mathcal{F}}^{\dagger}(x^+ - x^-)} \quad (5)$$

where $L_{\mathcal{F}}^{\dagger}$ is the Moore-Penrose pseudoinverse matrix of $L_{\mathcal{F}}$. This index is considering how extreme the opinions of the people are, how much they organize into echo chambers, and how these echo chambers organize in the network.

4 EXPERIMENTS

In this section, we first introduce the datasets and evaluation settings in our experiments. We then present the results for the Laplacian with different honesty parameter values (τ) and determine the most effective relation type (honest, prosocial liar, and antisocial liar) on synthetic and real-world networks. We further show how deception affects polarization.

4.1 Experimental Setup

4.1.1 Datasets. We consider nine real-world undirected social networks [20, 25], which have been widely adopted in the studies of influential node detection. The general statistics of the datasets used for experiments are reported in Table 1.

Table 1: Basic properties of the real-world datasets we use are provided here. $\langle k \rangle$ is the average degree, and k_{max} is the maximum degree.

Dataset	Vertices	Edges	$\langle k \rangle$	k_{max}
Train	64	243	7.59	29
Highschool	70	366	10.46	23
Lesmis	77	254	6.60	36
Copper	112	425	7.59	49
Jazz	198	2742	27.69	100
Oz	217	1839	16.94	56
Congress	219	764	6.97	50
Innovation	244	925	7.58	29
Netscience	379	914	4.00	34

4.1.2 Settings. In this section, we will explain our experimental settings and designs for influence maximization and polarization problems.

Influence maximization: For each user i in a network, we randomly select an opinion about a topic, x_i, within the interval $[-1, 1]$. As the next step, we randomly divide the vertices into three equal

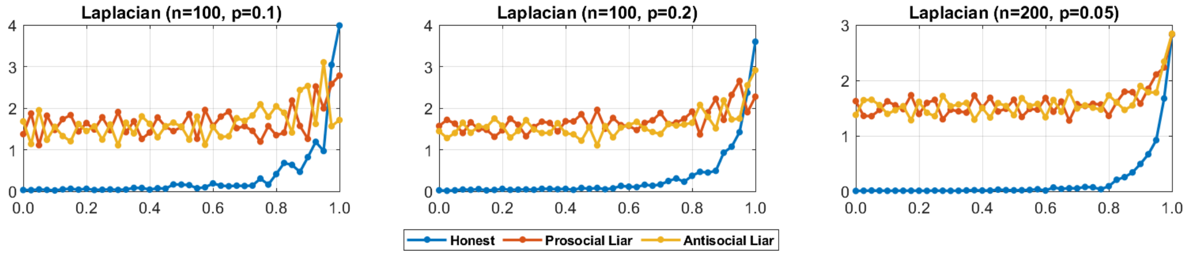


Figure 5: The average centrality scores (y-axis) with respect to the different honesty parameters τ (x-axis) for Laplacian centrality on three different Erdos-Renyi random graphs. Here, n is the number of vertices in the random graph and p is the probability of edge creation.

parts and label each part as honest, prosocial liar, and antisocial liar. Then, we analyze centrality scores of each label. There is a possible issue here that when we randomly divide the network into three parts, influential nodes when the deception is not present may accumulate in one of the parts. To avoid this, we first rank vertices from the most influential to the least and divide them into 10 equal parts, i.e., the first part includes the top 10% influential interactions and the last part includes the bottom 10% influential vertices. We then randomly divide each part into three and label them as honest, prosocial liar and antisocial liar.

After assigning opinions and relation types (honest, prosocial liar, and antisocial liar) to each user, we define the linear transformation and the coboundary map and use the boundary map to generate the sheaf Laplacian. The sheaf Laplacian also depends on the honesty parameter $\tau \in [0, 1]$. To see the effect of the different honesty levels, we partition the interval $[0, 1]$ into 40 equal parts and input each value in the sheaf Laplacian. The experiment is run 100 times for each dataset, and the average of the 100 trials is taken to obtain more reliable results.

As the final evaluation step, to get the influentiality score S_R of a relation type R, we obtain the centrality scores in each simulation, take the average of the scores for each relation type, and normalize it with the number of vertices in the network. Mathematically, the influentiality score is obtained by

$$S_R = \frac{1}{|V_R|} \sum_{i=1}^{N} \sum_{j}^{|V_R|} c_{ij}$$

where $|V_R|$ is the number of vertices of a given relation type, N is the number of runs, and c_{ij} is the centrality score of the jth vertex in V_R in the ith run. Hence, the larger S_R means a more influential relation in Laplacian centrality.

Polarization: We create the sheaf Laplacian as explained in the previous section. To study polarization, we first find the most influential user in the network based on Laplacian centrality. We set this user's opinion to 1 and the rest to 0. Our goal here is to keep the influential user's opinion fixed at 1 and observe how other users' opinions are changing, and explore whether polarization occurs. To make this happen, we use the diffusion model with stubbornness explained in Section 3.4.1. After getting the opinion distribution, we use the polarization measures explained in Section 3.4.2 to check the polarization. Again, this experiment is run 100 times for each dataset, and the average polarization measures of the 100 trials is taken to obtain more reliable results.

4.2 Results

4.2.1 Influence Maximization. In this section, we present our results on synthetic networks that we create using Erdos-Renyi random graphs and nine real-world networks that we present in Table 1. We start discussing the results for synthetic networks.

Erdos-Renyi random graphs take two parameters: the number of nodes (n) and the probability for edge creation p. To address different cases, we study three random graphs: (1) $n = 100, p = 0.1$, (2) $n = 100, p = 0.2$ and (3) $n = 200, p = 0.05$. If we take the first random graph as the test case, the second random graph has more density (i.e., the average degree) with keeping the size (n) the same and the third random graph is larger with keeping the density the same. Our goal here is to see how influentiality changes based on the density and the size of a network. Then, we evaluate the performance of each relation type (honest, prosocial liar, antisocial liar) on each network by following the outline in Section 4.1. The results are available in Figure 5.

As we see in Figure 5, liars, regardless of being prosocial and antisocial, have bigger centrality scores than honest users on average for all random graphs. As we explain in the previous section, the bigger Laplacian centrality score implies being more influential, so liars are more influential than honest users. Furthermore, density and size do not change this conclusion.

Secondly, we present the simulation results on real social networks in Figure 6. As we clearly see in the figure, liars, regardless of being prosocial and antisocial, again have bigger Laplacian centrality scores than honest users on average for all datasets. The difference gets smaller with a higher honesty level τ for each dataset. Furthermore, the total centrality score gets bigger for all deception types as the honesty parameter τ gets closer to 1.

Besides the figures, we also present the total centrality score for each real social network in Table 2. As we see in the table, based on the Laplacian centrality, prosocial and antisocial liars are the most influential interchangeably, and on average, prosocial liars are slightly more influential than antisocial liars.

4.2.2 Polarization. In this part, we present our results on a synthetic network and nine real-world networks presented in Table 1. The reason we choose a scale-free graph here is that scale-free graphs have a few highly connected nodes, called "hubs," which are connected to many other nodes with fewer connections. These hub nodes can cause polarization deeper than other nodes, so, observing polarization becomes easier.

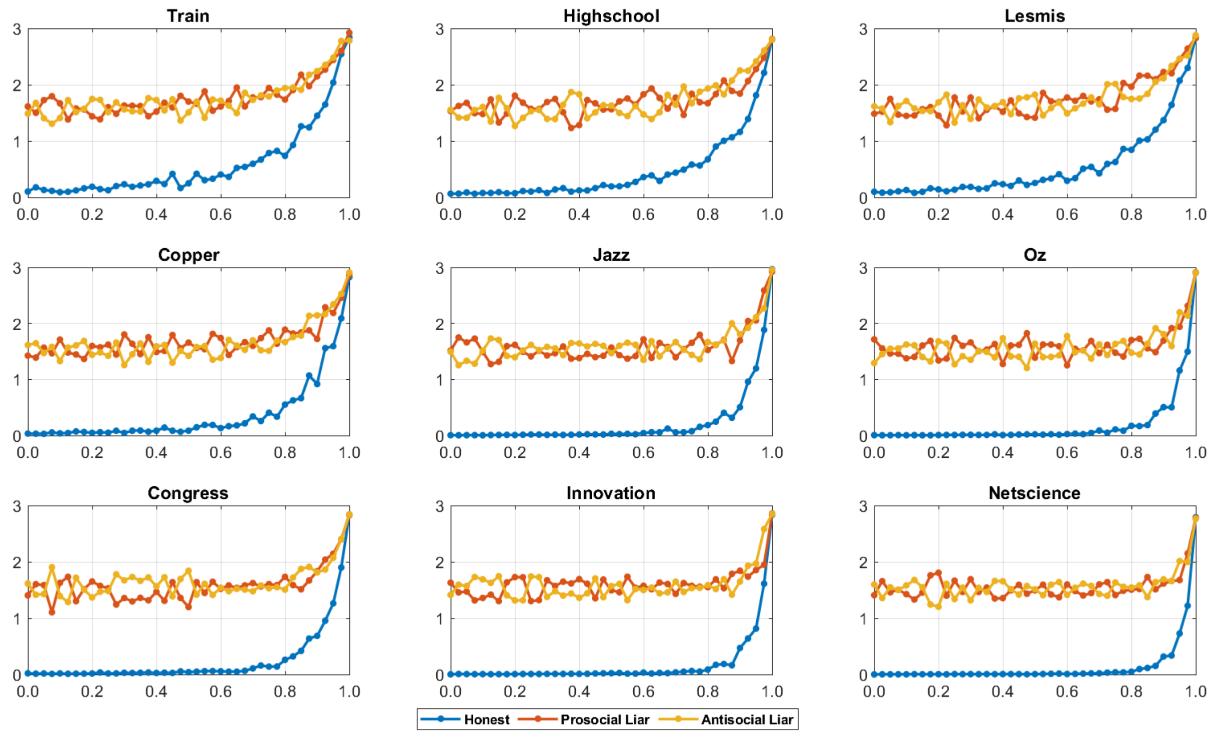


Figure 6: The average centrality scores (y-axis) with respect to the different honesty parameters τ (x-axis) for Laplacian centrality on real social networks. The bigger Laplacian centrality scores imply being more influential.

Table 2: The total centrality scores based on Laplacian centrality in Figure 6. The darker cells correspond to the more influential relation type.

	Honest	Prosocial	Antisocial
Train	24.33	**72.81**	72.41
Highschool	19.56	**70.88**	70.35
Lesmis	22.91	72.11	**72.22**
Copper	15.82	**69.86**	68.19
Jazz	9.64	65.86	**66.86**
Oz	8.21	**66.41**	64.82
Congress	10.72	65.24	**67.98**
Innovation	7.53	**66.05**	65.77
Netscience	6.11	64.33	**64.45**
Average	13.87	**68.17**	68.12

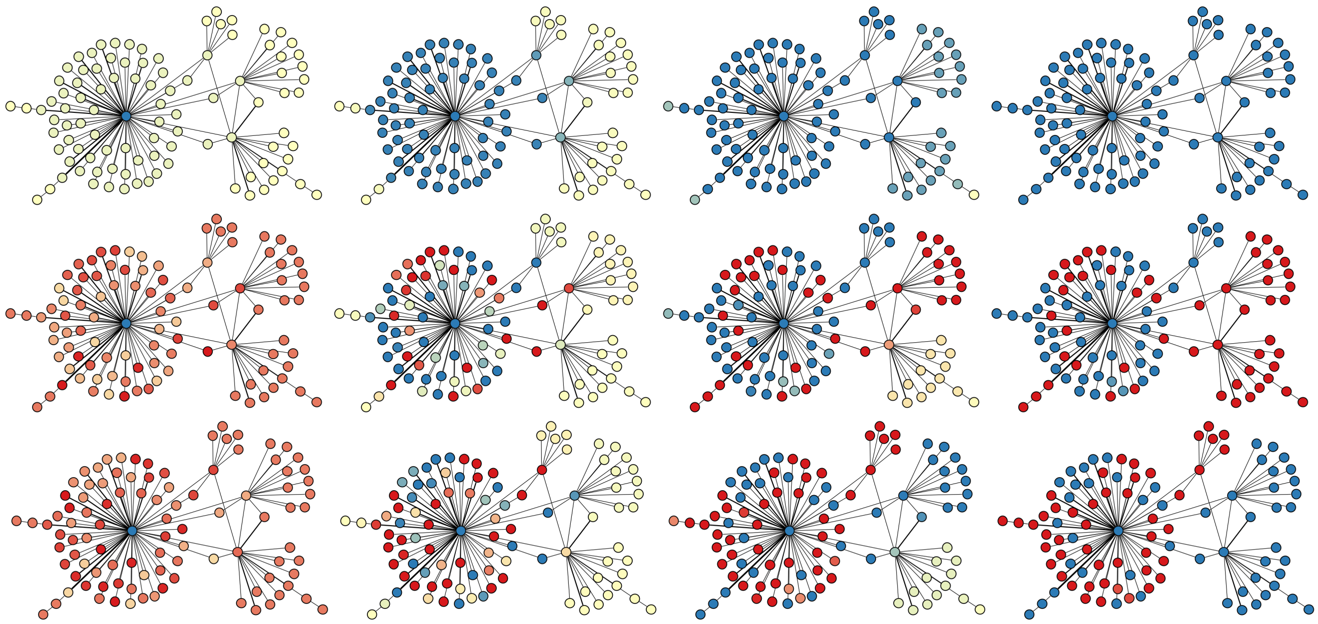
We first synthetically create a scale-free graph. The resulting graph is available in Figure 7. Next, we find the most influential user in the graph which is the darker-colored node at the center of the cluster on the left. We set this user's opinion to 1 and the rest to 0. Then, we set this user as honest, prosocial, and antisocial liar and the rest of the network as all honest, and check the opinion distribution in the graph for different time values $t = \{10^{-2}, 10^{-1}, 1, 10, 100\}$. We should also note here that whenever a user's opinion gets bigger than 1 (or smaller than -1), we set his opinion to 1 (or -1). The results are available in Figure 7.

As we see in the figure, when the influential user is honest, the rest of the users' opinions get closer to 1 as time passes. Hence, there is a consensus in the opinion dynamics without any polarization. This case actually represents the classical opinion dynamics when everyone in the network is an honest user.

Secondly, when the influential user is a prosocial or antisocial liar, we observe a clear polarization in the network. Opinions are getting accumulated at 1 and -1, which means we observe a set of users that totally agree and another set of users that totally disagree with the seed user. Furthermore, we also observe that for the antisocial lying case, other users' opinions are leaning more toward -1 than the prosocial lying case. This is correlated with the nature of deception types since prosocial lies are intended to help other users, whereas antisocial lies are intended to hurt.

In more detail, the network has four clusters, the biggest one has the seed node at the center, and there are three other relatively smaller clusters on the right. As we see in the honest case, all the nodes and clusters reach a consensus when time $t = 10$. For the prosocial liar case, the bigger cluster mostly agrees with the seed node and while one of the clusters (the smallest one) totally agrees with the seed node, the other ones totally disagree, and as a result, we see a clear polarization. On the other hand, for the antisocial liar case, we observe the opposite situation with the prosocial liar case in opinion distribution but again observe a clear polarization.

Moreover, we repeat a similar experiment on nine real-world networks. This time we choose $t = 1000$ to reach the steady state level in diffusion. Then, we calculate the standard deviation and the



(a) $t = 10^{-2}$ (b) $t = 10^{-1}$ (c) $t = 1$ (d) $t = 10$

Figure 7: Diffusion for various time scales (a) $t = 10^{-2}$, (b) $t = 10^{-1}$, (c) $t = 1$ and (d) $t = 10$ on different deception types honest (top row), prosocial liar (middle row) and antisocial liar (bottom tow) when initially the seed node has opinion 1 and the rest of the network have opinion 0. The opinion distribution for each time scale is available in Figure ??.

Table 3: Standard deviation and GE polarization index for the datasets. The darker cells correspond to more polarization within networks.

	Standard Deviation			GE Polarization Index		
	Honest	Prosocial	Antisocial	Honest	Prosocial	Antisocial
Train	0	0.259	**0.261**	0	**1.572**	1.571
Highschool	0	0.193	**0.197**	0	0.848	**0.866**
Lesmis	0	**0.443**	0.441	0	**4.866**	4.865
Copper	0	**0.258**	0.257	0	**1.655**	**1.655**
Jazz	0	**0.138**	0.136	0	**0.805**	0.803
Oz	0	0.104	**0.110**	0	0.402	**0.415**
Congress	0	0.273	**0.274**	0	**4.243**	4.241
Innovation	0	0.248	**0.253**	0	0.733	**0.738**
Netscience	0.0005	**0.153**	**0.153**	0.068	**12.471**	12.462
Average	0.00005	0.2307	**0.2313**	0.008	**3.170**	3.159

GE polarization index for each dataset, run the experiment 100 times and get the average polarization scores. As a side note, Innovation network is not connected, hence, while calculating polarization measures, we just use the connected component that includes the seed node. The results are available in Table 3.

As we see in the table, there is a total consensus for the first seven networks when the influential user is honest based on both standard deviation and the GE polarization index. Also, there is a clear polarization when the influential user is a prosocial or antisocial liar as we observe in the synthetic graph.

5 CONCLUSION

In this paper, we study the effect of deception on the influence maximization and polarization problems on social networks. We develop a method to model deception in social networks, employ the sheaf Laplacian to model the opinion dynamics when deception is present, redefine Laplacian centrality for the sheaf Laplacian to detect influential nodes, and model polarization via sheaf Laplacian. Our results show that liars are more influential than honest people in social networks. We also show that liars make networks more polarized. As future tasks, we plan to apply our method to understand the construction and circulation of truth in social media when deception is present.

REFERENCES

[1] Johannes Abeler, Anke Becker, and Armin Falk. 2012. Truth-telling: A representative assessment. (2012).

[2] Mehmet E Aktas, Sidra Jawaid, Ebony Harrington, and Esra Akbas. 2021. Influential nodes detection in complex networks via diffusion Fréchet function. In *Proceedings of the 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE. https://doi.org/10.1109/ICMLA52953.2021.00269

[3] Swapan Deep Arora, Guninder Pal Singh, Anirban Chakraborty, and Moutusy Maity. 2022. Polarization and social media: A systematic review and research agenda. *Technological Forecasting and Social Change* 183 (2022), 121942.

[4] Rafael A Barrio, Tzipe Govezensky, Robin Dunbar, Gerardo Iniguez, and Kimmo Kaski. 2015. Dynamics of deceptive interactions in social networks. *Journal of the Royal Society Interface* 12, 112 (2015), 20150798.

[5] Manuel Castells. 2000. Toward a sociology of the network society. *Contemporary sociology* 29, 5 (2000), 693–699.

[6] Damon Centola. 2010. The spread of behavior in an online social network experiment. *science* 329, 5996 (2010), 1194–1197.

[7] Christian Decker and Roger Wattenhofer. 2013. Information propagation in the bitcoin network. In *IEEE P2P 2013 Proceedings*. IEEE, 1–10.

[8] Linton C Freeman. 1978. Centrality in social networks conceptual clarification. *Social networks* 1, 3 (1978), 215–239.

[9] Ivan Garibay, Alexander V Mantzaris, Amirarsalan Rajabi, and Cameron E Taylor. 2019. Polarization in social media assists influencers to become more influential: analysis and two inoculation strategies. *Scientific reports* 9, 1 (2019), 1–9.

[10] Maria Giatsoglou, Despoina Chatzakou, Neil Shah, Christos Faloutsos, and Athena Vakali. 2015. Retweeting activity on twitter: Signs of deception. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 122–134.

[11] Uri Gneezy. 2005. Deception: The role of consequences. *American Economic Review* 95, 1 (2005), 384–394.

[12] Chungu Guo, Liangwei Yang, Xiao Chen, Duanbing Chen, Hui Gao, and Jing Ma. 2020. Influential Nodes Identification in Complex Networks via Information Entropy. *Entropy* 22, 2 (2020), 242.

[13] Jeffrey T Hancock, Jennifer Thom-Santelli, and Thompson Ritchie. 2004. Deception and design: The impact of communication technology on lying behavior. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 129–134.

[14] Jakob Hansen and Robert Ghrist. 2021. Opinion dynamics on discourse sheaves. *SIAM J. Appl. Math.* 81, 5 (2021), 2033–2060.

[15] Marilena Hohmann, Karel Devriendt, and Michele Coscia. 2023. Quantifying ideological polarization on a network using generalized Euclidean distance. *Science Advances* 9, 9 (2023), eabq2044.

[16] Gerardo Iñiguez, Tzipe Govezensky, Robin Dunbar, Kimmo Kaski, and Rafael A Barrio. 2014. Effects of deception in social networks. *Proceedings of the Royal Society B: Biological Sciences* 281, 1790 (2014), 20141195.

[17] Webb Keane. 2008. The Evidence of the Senses and the Materiality of Religion. *Journal of the Royal Anthropological Institute* 14 (2008), S110–S127.

[18] Whan-Seon Kim. 2009. Effects of a trust mechanism on complex adaptive supply networks: An agent-based social simulation study. *Journal of artificial societies and social simulation* 12, 3 (2009), 4.

[19] Catherine King, Daniele Bellutta, and Kathleen M Carley. 2020. Lying about lying on social media: a case study of the 2019 Canadian Elections. In *Social, Cultural, and Behavioral Modeling: 13th International Conference, SBP-BRiMS 2020, Washington, DC, USA, October 18–21, 2020, Proceedings 13*. Springer, 75–85.

[20] KONECT. 2007. Chicago Network Dataset. http://konect.cc/networks/.

[21] Michael Lambek. 2010. *Ordinary ethics: Anthropology, language, and action*. Fordham Univ Press.

[22] Mei Li, Xiang Wang, Kai Gao, and Shanshan Zhang. 2017. A survey on information diffusion in online social networks: Models and methods. *Information* 8, 4 (2017), 118.

[23] Linyuan Lü, Tao Zhou, Qian-Ming Zhang, and H Eugene Stanley. 2016. The H-index of a network node and its relation to degree and coreness. *Nature communications* 7, 1 (2016), 1–7.

[24] Xingqin Qi, Eddie Fuller, Qin Wu, Yezhou Wu, and Cun-Quan Zhang. 2012. Laplacian centrality: A new centrality measure for weighted networks. *Information Sciences* 194 (2012), 240–253.

[25] Ryan A. Rossi and Nesreen K. Ahmed. 2015. The Network Data Repository with Interactive Graph Analytics and Visualization. In *AAAI*. http://networkrepository.com

[26] Kim B Serota, Timothy R Levine, and Franklin J Boster. 2010. The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research* 36, 1 (2010), 2–25.

[27] Kirsten P Smith and Nicholas A Christakis. 2008. Social networks and health. *Annu. Rev. Sociol* 34 (2008), 405–429.

[28] Erin L Spottswood and Jeffrey T Hancock. 2016. The positivity bias and prosocial deception on Facebook. *Computers in Human Behavior* 65 (2016), 252–259.

[29] Wenjun Wang and W Nick Street. 2018. Modeling and maximizing influence diffusion in social networks for viral marketing. *Applied network science* 3, 1 (2018), 1–26.

[30] Timothy F Wright. 2006. The Evolution of Animal Communication: Reliability and Deception in Signaling Systems.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009