Rumor Initiator Detection in Infected Signed Networks

Jiawei Zhang
University of Illinois at Chicago
Chicago, IL, USA
jzhan9@uic.edu

Charu C. Aggarwal
IBM T. J. Watson Research Center
Yorktown Heights, NY, USA
charu@us.ibm.com

Philip S. Yu
University of Illinois at Chicago
Chicago, IL, USA
Institute for Data Science, Tsinghua University
Beijing, China
psyu@cs.uic.edu

***Abstract*—In many cases, the information spread in an online network may not always be truthful or correct; such information corresponds to rumors. In recent years, signed networks have become increasingly popular because of their ability to represent diverse relationships such as friends, enemies, trust, and distrust. Signed networks are ideal for information flow in a network with varying beliefs (trust or distrust) about facts. In this paper, we will study the problem of influence analysis and diffusion models in signed networks and investigate the problem of rumor initiator detection, given the state of the network at a given moment in time. Conventional information diffusion models for unsigned networks cannot be applied to signed networks directly, and we show that the rumor initiator detection problem is NP-hard. We propose a new information diffusion model, referred to as *asyMmetric Flipping Cascade* (MFC), to model the propagation of information in signed networks. Based on MFC, a novel framework, *Rumor Initiator Detector* (RID), is introduced to determine the potential number and the identity of the rumor initiators from the state of the network at a given time. Extensive experiments conducted on real-world signed networks demonstrate that MFC works very well in modeling information diffusion in signed networks and RID can significantly outperform other comparison methods in identifying rumor initiators.**

***Index Terms*—Rumor Initiator Detection, Information Diffusion, Signed Networks, Data Mining**

I. Introduction

In recent years, signed networks [29], [23] have gained increasing attention because of their ability to represent diverse and contrasting social relationships. Some examples of such contrasting relationships include friends vs enemies [25], trust vs distrust [26], positive attitudes vs negative attitudes [27], and so on. These contrasting relationships can be represented as links of different polarities, which result in signed networks. Signed social networks can provide a meaningful perspective on a wide range of social network studies, like *user sentiment analysis* [24], *social interaction pattern extraction* [15], *trustworthy friend recommendation* [14], and so on.

Rumor initiation and incorrect information dissemination are both common in social networks [22]. Due to the extensive social links among users, rumors on certain topics, e.g., politics, celebrities and product promotions, can propagate leading to a large number of nodes reporting the same (incorrect) observations rapidly in online social networks. In particular, the links in signed networks are of different polarities and can denote trust and distrust relationships among users [16], which will inevitably have an impact on information propagation. Incorrect rumors sometimes can bring about devastating effects, and an important goal in improving the credibility of the social channel is to identify rumor initiators [22], [21], [13], [19] in signed social networks.

In Figure 1, an example is provided to help illustrate the rumor initiators identification problem more clearly. In the example, users are connected to one another with signed links, depending on their trust and distrust relations. It is noteworthy that the conventions used for the direction of information diffusion in this network are slightly different from traditional influence analysis, because they represent signed links. For instance, if Alice trusts (or follows) Bob, a directed edge exists from Alice to Bob, but the information diffusion direction will be from Bob to Alice. Via the signed links, inactive users in the network can get infected by rumor information propagated from their neighbors with either a positive or negative opinion about the rumor (i.e., the green or red states in the figure). Considering the fact that it is often difficult to directly identify all the user infection states in real settings, we allow for the possibility of some user states in the network to be unknown. Activated users can propagate the rumor to other users. In general, if a user is activated with a positive or negative opinion about the rumor, she might activate one or more of her incoming neighbors to trust or distrust the rumor, depending on the sign of the incoming link. The main goal of the rumor initiators identification problem is to determine the most likely rumor initiators; those corresponding to the node in the blue circle in Figure 1.

Problem Setting: This paper studies the detection of rumor initiators in infected signed social networks, given the state of the network at a specific moment in time. The edges in the network are directed and signed, and they represent trust or distrust relationships. For example, when node i trusts or distrusts node j, we will have a corresponding positive or negative link from node i to node j. In this setting, nodes are associated with states corresponding to a prevailing opinion about the truth of a fact. These states can be drawn from $\{-1, +1, 0, ?\}$, where $+1$ indicates their agreement with a specific fact, -1 indicates their disagreement, 0 indicates the

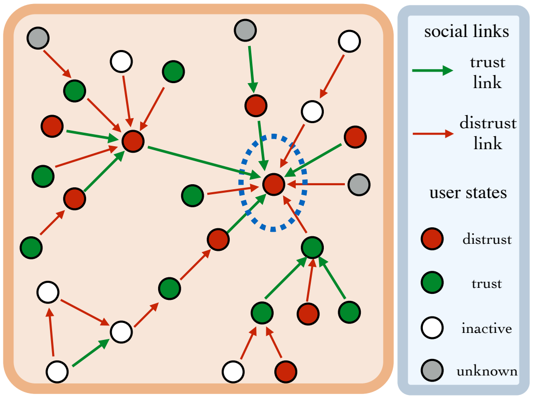


Fig. 1. Example of the ISOMIT problem.

fact that they have no opinion of the fact at hand, and ? indicates their opinion is unknown. The last of these states is necessary to model the fact that the states of many nodes in large-scale networks are often unknown. Note that the use of multiple states of nodes in the network is different from traditional influence analysis. Users are influenced with varying opinions of the fact in question, based on their observation of their neighbors (i.e., states of neighborhood nodes), and their trust or distrust of their neighbor's opinions (i.e., signs of links with them). This model is essentially a signed version of influence propagation models, because the sign of the link plays a critical role in how a specific bit of information is transmitted. Given a snapshot of the states of all nodes in the network at a given time, the main goal of this problem is to determine the most likely rumor initiators in the network. We refer to the problem as the "Infected Signed netwOrk ruMor Initiator deTection" (ISOMIT) problem.

The ISOMIT problem is very different from traditional diffusion modeling and influence analysis of unsigned networks because of its use of node states and links signs. Examples of such settings include the maximization of spread of influence through *unsigned* social networks [10], finding effectors in unsigned social networks [13], and influence maximization in signed networks [17]. Table I summarizes the differences of ISOMIT from existing works. The main challenges in addressing this problem are as follows:

- *Diffusion modeling challenges*: Most existing information diffusion models are designed for unsigned networks. In signed networks, information diffusion is also related to actor-centric trust and distrust, in which notions of node states and the signs on links play an important role.
- *Algorithmic challenges*: To infer the rumor initiators, we need to determine the number of initiators, their identities, and initial activation states. We show that the exact identification of rumor initiators in signed networks is NP-hard.

To solve the aforementioned challenges, a new method, *Rumor Initiator Detector* (RID), is introduced in this paper. We propose the *asyMmetric Flipping Cascade* (MFC) diffusion model for signed networks. Although the exact identification of the rumor initiators is NP-hard for general graphs, but it can be resolved in polynomial time for binary-tree structured networks, and it provides the insights for high-quality solutions in the general case. We leverage these insights to introduce the RID framework to identify the optimal *rumor initiators*, including their number, identities, and initial states.

II. Problem Formulation

We will first introduce the notations and definitions. Then, we will provide a mathematical formulation of the ISOMIT problem.

A. Preliminaries

Traditional social networks are unsigned in the sense that the links are assumed, by default, to be positive links. Signed social networks are a generalization of this basic concept.

Definition 1: (Weighted Signed Social Network): A *weighted signed social network* can be represented as a graph $G = (\mathcal{V}, \mathcal{E}, s, w)$, where $\mathcal{V}$ and $\mathcal{E}$ represents the nodes (users) and directed edges (social links), respectively. In signed networks, each social link has its own polarity (i.e., the sign) and is associated with a weight indicating the intimacy among users, which can be represented with the mappings $s : \mathcal{E} \to \{-1, +1\}$ and $w : \mathcal{E} \to [0, 1]$ respectively.

As discussed in Section I, we interpret the signs from a trust-centric point of view. Information propagated among users is highly associated with the intimacy scores [28] among them: information tends to propagate among close users. To represent the information diffusion process in trust-centric networks, we define the concept of *weighted signed diffusion network* as follows:

Definition 2: (Weighted Signed Diffusion Network): Given a *signed social network* G, its corresponding *weighted signed diffusion network* can be represented as $G_D = (\mathcal{V}_D, \mathcal{E}_D, s_D, w_D)$, where $\mathcal{V}_D = \mathcal{V}$ and $\mathcal{E}_D = \{(v, u)\}_{(u,v) \in \mathcal{E}}$. Diffusion links in $\mathcal{E}_D$ share the same sign and weight mappings as those in $\mathcal{E}$, which can be obtained via mappings $s_D : \mathcal{E}_D \to \{-1, +1\}$, $s_D(v, u) = s(u, v), \forall (v, u) \in \mathcal{E}_D$ and $w_D : \mathcal{E}_D \to [0, 1]$, $w_D(v, u) = w(u, v), \forall (v, u) \in \mathcal{E}_D$. For any directed diffusion link $(u, v) \in \mathcal{E}_D$, we can represent its sign and weight to be $s_D(u, v)$ and $w_D(u, v)$ respectively.

Note that we have reversed the direction of the links because of the trust-centric interpretation, in which information diffuses from A to B, when B trusts A. However, in networks with other semantic interpretations, this reversal does not need to be performed. The overall algorithm is agnostic to the specific preprocessing performed in order to fit a particular semantic interpretation of the signed network.

The social psychology literature defines a *rumor* as a story or a statement in general circulation without confirmation or certainty of facts [1]. The originators of rumors are formally defined as *rumor initiators*, which can be individuals, groups, or institutes. In this paper, we refer to *rumor initiators* as the users who initially spread the rumor to other users in online social networks. Within the *diffusion networks*, *rumors* can spread from the *initiators* to other users via diffusion links, which will lead to *infected signed diffusion networks*. Since all networks studied in this paper are all weighted and signed

TABLE I
SUMMARY OF RELATED PROBLEMS.

Property	**Rumor Initiator Detection in Signed Networks**	**Influence Maximization in Social Networks** [10]	**Influence Maximization in Signed Networks** [17]	**Finding Effectors in Social Networks** [13]
network types	signed	unsigned	signed	unsigned
problem studied	rumor initiator detection	influence maximization	influence maximization	rumor initiator detection
diffusion model	MFC model	LT and IC models	voter model model	IC model

by default, we will refer to them as *diffusion networks* for simplicity.

Definition 3: (Infected Diffusion Network): The *infected diffusion network* $G_I = (\mathcal{V}_I, \mathcal{E}_I, s_I, w_I)$ is a subgraph of the complete diffusion network G_D, where $\mathcal{V}_I \subseteq \mathcal{V}_D$ is the set of infected users, $\mathcal{E}_I \subseteq \mathcal{E}_D$ is the set of potential diffusion links among these infected users. s_I, w_I are the *sign* and *weight* mappings, whose domains are all those diffusion links in $\mathcal{E}_I$.

Definition 4: (Activation Link): Among all the links $\mathcal{E}_I$ in the infected diffusion networks, link (u, v) is called an activation link iff u activates v in the screenshot of the infected diffusion network.

Based on the MFC model to be introduced in Section III-A, each node in the infected diffusion network screenshot can be activated by exactly one node via the *activation link* and the *rumor initiators* have no incoming activation links. As a result, all the nodes in $\mathcal{V}_I$ together with the activation links among them can actually form a set of cascade trees, where nodes at higher levels are activated by nodes in the lower levels and *rumor initiators* are the roots (at level 1).

B. Problem Formulation

In this paper, we will propose diffusion models, which characterize how *rumors* can spread from the *initiators* to other users. The state of the infected diffusion network is referred to as the *infected diffusion network*. However, our main goal is to work backwards from the available state of the network given at any moment in time, and we will use the developed diffusion model to track down the rumor initiators. Let $\mathcal{I} \subseteq \mathcal{V}_I \subseteq \mathcal{V}$ be the potential set of *rumor initiators*, whose initial states towards the rumor can be represented as $\mathcal{S} = \{+1, -1\}^{|\mathcal{I}|}$, where $+1$ indicates a belief in the fact at hand, and -1 denotes belief in the opposite fact. We use binary modes of information propagation because of its relative simplicity and intuitive appeal in modeling a variety of situations. The ISOMIT problem aims at inferring the optimal *rumor initiator* set $\mathcal{I}^*$ as well as their initial states $\mathcal{S}^*$, which can maximize the likelihood that it will lead to the current state of the *infected signed network* G_I:

$$\mathcal{I}^*, \mathcal{S}^* = \arg\max_{\mathcal{I}, \mathcal{S}} \mathbf{P}(G_I | \mathcal{I}, \mathcal{S}),$$

Here, $\mathbf{P}(G_I | \mathcal{I}, \mathcal{S})$ represents the likelihood of obtaining the infected network G_I based on the influence propagated from $\mathcal{I}$ with states $\mathcal{S}$.

In summary, the input of the ISOMIT problem is the infected signed network G_I, while the objective output is the inferred rumor initiators $\mathcal{I}$ together with their initial states $\mathcal{S}$ which can maximize the likelihood $\mathbf{P}(G_I | \mathcal{I}, \mathcal{S})$.

III. PROPOSED METHOD

In this section, we will introduce the RID framework to address the ISOMIT problem. To model the information propagation process in signed networks, a new diffusion model named MFC will be introduced in Section III-A. For unsigned networks, the well-known set cover problem can be mapped to the exact rumor initiator identification problem in polynomial time, and the problem analysis and proof are available in Sections III-B-III-C. Meanwhile, for binary-tree structured signed networks, we will show in Section III-D, that the ISOMIT problem can be solved in polynomial time for a specified number of rumor initiators. Finally, in Section III-E, we will use the insights gained from the special case to introduce the RID framework to address the ISOMIT for networks beyond the binary tree structure.

A. Asymmetric Flipping Cascade Model

Many information diffusion models have been proposed for unsigned networks, including the Linear Threshold (LT) model [10], Independent Cascade (IC) model [10] and Susceptible Infectious Recovered (SIR) model [9]. An extensive survey of existing diffusion models is available in [8]. Next, we will first talk about the traditional IC model for unsigned networks, and then introduce the MFC (asy<u>M</u>metric <u>F</u>lipping <u>C</u>ascade) model proposed for the signed network setting.

1) Traditional IC Diffusion Model: In the traditional IC diffusion model, the information diffusion process usually starts with a set of seed users (e.g., rumor initiators) $\mathcal{I}$, whose influence propagates within the network in discrete steps. At step τ, any user u just activated at step $\tau - 1$ is given only one chance to activate each of its currently inactive neighbors, e.g., v, with probability $w(u, v)$ (i.e., the weight of diffusion link (u, v)). If u succeeds, v will become active in step $t + 1$; otherwise, v stays inactive and u cannot make any further attempts to activate v in subsequent rounds. All activated users will stay active and cannot be activated again. Such a process stops when no more activations are possible.

2) The MFC Diffusion model: The IC model, which assumes that social links are all of the same polarity, works for unsigned networks, but it cannot be applied to signed networks with node states to reflects beliefs of different polarities. To overcome such a shortcoming, a novel diffusion model, MFC, will be introduced in this section.

The signs associated with diffusion links denote the "positive" and "negative" relationships, e.g., trust and distrust, among users. In everyday life, people tend to believe information from people they trust and not believe the information from those they distrust. For example, if someone we trust says that "Hillary Clinton will be the new president", we believe it to be true. However, if someone we distrust says the same thing, we might not believe it. In addition, when receiving contradictory messages, information obtained from the trusted people is usually given higher weights. In other words, the effects of trust and distrust diffusion links are asymmetric in

Algorithm 1 MFC Information Diffusion Model

```
Input: input rumor initiators $\mathcal{I}$ with states $\mathcal{S}$
       diffusion network $G_D = (\mathcal{V}_D, \mathcal{E}_D, s_D, w_D)$
Output: infected diffusion network $G_I$
initialize infected user set $\mathcal{U} = \mathcal{I}$, state set $\mathcal{S}_\mathcal{U} = \mathcal{S}$
let recently infected user set $\mathcal{R} = \mathcal{I}$
while $\mathcal{R} \neq \emptyset$ do
    new recently infected user set $\mathcal{N} = \emptyset$
    for $u \in \mathcal{R}$ do
        let the set of users that $u$ can activate to be $\Gamma(u)$
        for $v \in \Gamma(u)$ do
            if $s(v) = 0$ or $\big(s_D(u,v) = +1$ and $s(u) \neq s(v)\big)$ then
                if $s_D(u,v) = +1$ then
                    $p = \min\{1.0, \alpha \cdot w_D(u,v)\}$
                else
                    $p = w_D(u,v)$
                end if
                if u activates $v$ with probability $p$ then
                    $\mathcal{U} = \mathcal{U} \cup \{v\}$, $\mathcal{S}_\mathcal{U} = \mathcal{S}_\mathcal{U} \cup \{s(v) = s(u) \cdot s_D(u,v)\}$
                    $\mathcal{N} = \mathcal{N} \cup \{v\}$
                end if
            end if
        end for
    end for
    $\mathcal{R} = \mathcal{N}$
end while
extract infected diffusion network $G_I$ consisting of infected users $\mathcal{U}$
```

activating users. For instance, when various actors assert that "Hillary Clinton will be the new president", we may tend to follow those we trust, even though the distrusted ones also say it. In addition, if someone we distrust says that "Hillary Clinton will be the new president", we may think it to be false and will not believe it. However, after being activated to distrust it, if we are exposed to contradictory information from a trusted party, we might be willing to change our minds. To model such cases, which are unique to signed and state-centric networks, we propose to follow a number of basic principles in the MFC model, (1) the effects of positive links in activating users is boosted to give them higher weights in activating users, and (2) users who are activated already will stay active in the subsequential rounds but their activation states can be flipped to follow the people they trust.

In MFC, users have 3 unique known states in the information diffusion process: $\{+1, -1, 0\}$ (i.e., trust, distrust and inactive respectively). Users with unknown states are automatically taken into account during the model construction process by assuming states as necessary. For simplicity, we use $s(\cdot)$ to represent both the sign of links as well as the states of users. If user u trusts the rumor, then user u is said to have a positive state $s(u) = +1$ towards the rumor. The initial states of all users in MFC are assigned a value of 0 (i.e., inactive to the rumor). A set of rumor initiators $\mathcal{I} \subseteq \mathcal{V}$ activated by the rumor at the very beginning will have their own attitudes towards the rumor based on their judgements, which can be represented with $\mathcal{S} = \{+1, -1\}^{|\mathcal{I}|}$. Rumor initiators in $\mathcal{I}$ spread the rumor to other users in signed networks step by step. At step τ, user u (activated at $\tau - 1$) is given only one chance to activate (1) inactive neighbor v, as well as (2) active neighbor v but v has different state from u and v trusts u, with the boosted success probability $\overline{w_D}(u, v)$, where $\overline{w_D}(v, u) \in [0, 1]$ can be represented as

$$\overline{w_D}(v,u) = \begin{cases} \min\{\alpha \cdot w_D(v,u), 1\} & \text{if } s_D(v,u) = +1, \\ w_D(v,u), & \text{otherwise.} \end{cases}$$

In the above equation, parameter $\alpha > 1$ denotes the boosting of information from u to v and is called the *asymmetric boosting coefficient*.

If u succeeds, v will become active in step $\tau + 1$, whose states can be represented as $s(v) = s(u) \cdot s(u, v)$. For example, if user u thinks the rumor to be real (i.e., $s(u) = +1$) and v trusts u (i.e., $s(u, v) = +1$), once v get activated by u successfully, the state of v will be $s(v) = +1$ (i.e., believe the rumor to be true). Otherwise, v will keep its original state (either inactive or activated) and u cannot make any further attempts to activate v in subsequent rounds. All activated users will stay active in the following rounds and the process continues until no more activations are possible.

MFC can model the information diffusion process in signed social networks much better than traditional diffusion models, such as IC. To illustrate the advantages of MFC, we also give an example in Figure 2, where two different cases: "simultaneous activation" (i.e., the left two plots) and "sequential activation" (i.e., the right two plots) are shown. In the "simultaneous activation" case, multiple users (B, C, D and E) are all just activated at step τ, who all think a rumor to be true and at step $\tau + 1$, B-E will activate their inactive neighbor A. Among these users, A trusts E and distrusts the remaining users. In traditional IC models, signs on links are ignored and B-E are given equal chance to activate A in random order with activation probabilities $w_D(\cdot, A), \cdot \in \{B, C, D, E\}$. However, in the MFC model, signs of links are utilized and the activation probability of positive diffusion (E, A) will be boosted and can be represented as $\min\{\alpha \cdot w_D(E, A), 1\}$. As a result, user A is more likely to be activated by E in MFC. Meanwhile, in the sequential activation case, once a user (e.g., F) succeeds in activating G, G will remain active and other users (e.g., H) cannot reactivate A any longer in traditional IC model. However, in the MFC model, we allow users to flip their activation state by people they trust. For example, if G has been activated by F with state $s(G) = -1$ already, the trusted user H can still have the chance to flip G's state with probability $\min\{\alpha \cdot w_D(H, G), 1\}$. The pseudo-code of the MFC diffusion model is provided in Algorithm 1.

B. The ISOMIT Problem

Given the *rumor initiators* $\mathcal{I}$ together with their initial states $\mathcal{S}$, influence can propagate from them to other users in the network via different paths. For any user u in the infected network, the influence propagation paths from *initiators* to u can be represented as the set $\{\mathcal{P}(u_i, u)\}_{u_i \in \mathcal{I}}$, where $\mathcal{P}(u_i, u)$ represents the set of paths from initiator u_i to user u specifically. Each path (e.g., $p \in \mathcal{P}(u_i, u)$) is a sequence of directed diffusion links from u_i to u. We use the notation $(x, y) \in p$ to denote the fact that the diffusion link (x, y) lies on path p. Depending on the sign of link (u, v) as well as the states

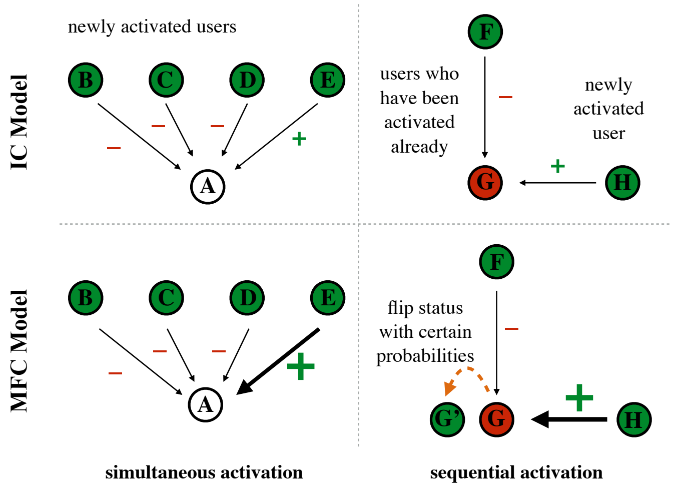


Fig. 2. Example of the binary tree transformation.

of u and v, link (u, v) can be either *sign consistent* and *sign inconsistent*.

Definition 5: (Sign Inconsistent Diffusion Link): Diffusion link (u, v) is defined to be *sign inconsistent* if $s(u) \cdot s(u, v) \neq s(v)$.

The probability that $u \in \mathcal{V}$ is infected with state $s(u)$ because of influence from the initiators $\mathcal{I}$ with state $\mathcal{S}$ can be computed as:

$$\mathbf{P}\left(u, s(u)|\mathcal{I}, \mathcal{S}\right) =$$

$$1 - \prod_{i \in \mathcal{I}} \prod_{p \in \mathcal{P}(i,u)} \left(1 - \prod_{(x,y) \in p} g\left(s(x), s_I(x,y), s(y), w_I(x,y)\right) \right),$$

where the function $g\left(s(x), s_I(x,y), s(y), w_I(x,y)\right) =$

$$\begin{cases} \min\{1, \alpha \cdot w_I(x,y)\}, & \text{if } s(x) \cdot s_I(x,y) = s(y), s_I(x,y) = +1, \\ w_I(x,y), & \text{if } s(x) \cdot s_I(x,y) = s(y), s_I(x,y) = -1, \\ 0, & \text{if } s(x) \cdot s_I(x,y) \neq s(y). \end{cases}$$

Consider a link (x, y) lying on the path from rumor initiators in $\mathcal{I}$ to u, such that states of x and y are consistent (i.e., $s(x) \cdot s_I(x, y) = s(y)$). In such a case, the probability of link (x, y) being an activation link would be $\min\{1, \alpha \cdot w_I(x, y)\}$ if (x, y) is a positive link (due to the boosting of positive links in MFC model), and it would be $w_I(x, y)$, otherwise. However, in case of inconsistency (i.e., $s(x) \cdot s_I(x, y) \neq s(y)$), link (x, y) will be either not an activation link or was an activation link originally but y's state is flipped by some other nodes. In other words, y would not be activated by x in the screenshot of the infected diffusion network, and the $g(\cdot)$ is assigned with value one in the *sign inconsistent* case.

One can model the probability of the current state of the *infected signed network* G_I, conditional on the *rumor initiators* $\mathcal{I}$ with initial states $\mathcal{S}$ as follows:

$$\mathbf{P}(G_I|\mathcal{I}, \mathcal{S}) = \prod_{u \in \mathcal{V}_I} \mathbf{P}\left(u, s(u)|\mathcal{I}, \mathcal{S}\right).$$

C. NP-hardness of Exact ISOMIT Problem

Based on the aforementioned remarks, we will show that obtaining the whole *infected networks* exactly based on $\mathcal{I}$ and $\mathcal{S}$ achieving 100% inference probability with minimum number of rumor initiators is an NP-hard problem.

Algorithm 2 Maximum Weight Spanning Graph (MWSG)

```
Input: Graph $G = (\mathcal{V}, \mathcal{E}, s, w)$
Output: Maximum weight spanning graph $G' = (\mathcal{N}, \mathcal{L}, w)$
initialize node set $\mathcal{N} = \emptyset$, link set $\mathcal{L} = \emptyset$
for $u \in \mathcal{V} \setminus \mathcal{N}$ do
    $\mathcal{N} = \mathcal{N} \cup \{u\}$
    find edge $e = \arg\max_{e \in \mathcal{E}} w(e)$
    $\mathcal{L} = \mathcal{L} \cup \{e\}$
end for
```

Lemma 3.1: Based on the MFC diffusion model, the ISOMIT problem of achieving probability $\mathbf{P}(G_I|\mathcal{I}, \mathcal{S}) = 1$ with the minimum number of initiators is NP-hard.

Proof 1: We will prove the Lemma by showing that the set-cover problem (which is known to be NP hard) can be reduced to the ISOMIT problem in polynomial time. Formally, given a set of elements $\mathcal{E} = \{e_1, e_2, \cdots, e_n\}$ and a set of m subsets of $\mathcal{E}$, $\mathcal{L} = \{\mathcal{L}_1, \mathcal{L}_2, \cdots, \mathcal{L}_m\}$, where $\mathcal{L}_i \subseteq \mathcal{E}, i \in \{1, 2, \cdots, m\}$. The set-cover problem aims at finding as few subsets as possible from $\mathcal{L}$, so that the union of the selected subsets is equal to $\mathcal{E}$, i.e., $\bigcup \mathcal{L}_i = \mathcal{E}$ [7].

For an arbitrary instance of the set-cover problem, we define an instance of the infected signed graph to be a directed graph, denoted by G_I. The graph G_I contains $n + m + 1$ nodes: (1) for each element $e_i \in \mathcal{E}$, we construct a corresponding node n_i; (2) for each set $\mathcal{L}_j \in \mathcal{L}$, we construct node n_{j+n}; and (3) a dummy node d (i.e., the $(n + m + 1)_{th}$ node) is added to the infected network. The links in G_I include: (1) for all the elements in each set, e.g., $e_i \in \mathcal{L}_j$), we add a directed link connecting their corresponding nodes in the graph from n_i to n_{j+n}; (2) all the corresponding nodes of elements in $\mathcal{E}$ are connected to d via a directed link; and (3) d connects to the corresponding nodes of sets in $\mathcal{L}$ by directed links as well. The signs of all these links are all assigned $+1$, whose weights are: (1) $w(n_i, n_{j+n}) = 1$, for $\forall e_i \in \mathcal{E}$, $\forall e_i \in \mathcal{L}_j, \mathcal{L}_j \in \mathcal{L}$; (2) $w(n_i, d) = \frac{1}{n}$, for $\forall e_i \in \mathcal{E}$; (3) $w(n_{j+n}, d) = 1$, for $\forall \mathcal{L}_j \in \mathcal{L}$.

Now, we want to activate all the nodes in G_I with state $+\mathbf{1}$ (i.e., all trust the rumor) with as few rumor initiators as possible. Based on G_I, the solution to the ISOMIT problem will be equivalent to the set-cover problem based on elements $\mathcal{E}$ and subsets $\mathcal{L}$.

D. A Special Case: k-ISOMIT-BT Problem

In the previous section, the ISOMIT problem of achieving probability 100% with the minimum number of initiators is proven to be NP-hard. In this part, we will study a special case of the ISOMIT problem, where the number of *rumor initiators* is known to be k and the network is a binary tree, i.e., the k-ISOMIT-BT (k ISOMIT on Binary Tree) problem. We will show that the k-ISOMIT-BT problem can be addressed efficiently in polynomial time. This will also provide the insight needed to solve the general case to be introduced in the next section.

Let $T_I = (\mathcal{V}_I, \mathcal{E}_I, s_I, w_I)$ be an infected signed binary tree. If the user node $u \in \mathcal{V}_I$ is regarded as the root in the tree, its left and right children can be represented as $left(u)$ and

Algorithm 3 Contract Circles (CC)

```
Input: Graph contrainig circles $G = (\mathcal{N}, \mathcal{L}, w)$
Output: Contracted graph without circles $G' = (\mathcal{N}', \mathcal{L}', w')$
$\mathcal{L}' = \emptyset$ and new link weight mapping $w'$
for each circle $O = (\mathcal{N}_O, \mathcal{L}_O)$ in $(\mathcal{N}, \mathcal{L})$ do
    contract all nodes in $O$ into a pseudo-node $u_o$
    for each link $(u_x, u_y) \in \mathcal{L}$ do
        if $u_x \notin \mathcal{N}_O$ and $u_y \in \mathcal{N}_O$ then
            $\mathcal{L}' = \mathcal{L}' \cup \{(u_x, u_o)\}$
            $w'(u_x, u_o) = w(u_x, u_y) - w(\pi(u_y), u_y)$, where $(\pi(u_y), u_y) \in \mathcal{L}$ is the link with the maximum weight linked to $u_y$
        else
            if $u_x \in \mathcal{N}_O$ and $u_y \notin \mathcal{N}_O$ then
                $\mathcal{L}' = \mathcal{L}' \cup \{(u_o, u_y)\}$
                $w'(u_o, u_y) = w(u_x, u_y)$
            else
                $\mathcal{L}' = \mathcal{L}' \cup \{(u_x, u_y)\}$
                $w'(u_x, u_y) = w(u_x, u_y)$
            end if
        end if
    end for
end for
```

$right(u)$, respectively. At the beginning, the *rumor initiator set* and the *state set* is empty, i.e., $\mathcal{I} = \emptyset$ and $\mathcal{S} = \emptyset$. The cost of the optimal solution (i.e., the inferred initiators $\mathcal{I}$ and states $\mathcal{S}$) can be recursively computed with the following dynamic programming equation:

$$\begin{aligned}
&\mathbf{OPT}(u, \mathcal{I}, \mathcal{S}, k) = \max \Bigg\{ \\
&\min_{m=0}^{k} \Big\{ \mathbf{OPT}\big(left(u), \mathcal{I}, \mathcal{S}, m\big) + \mathbf{OPT}\big(right(u), \mathcal{I}, \mathcal{S}, k-m\big) + \mathbf{P}\big(u, s(u)|\mathcal{I}, \mathcal{S}\big) \Big\}; \\
&\mathbf{P}\big(u, s(u) = +1|\mathcal{I} \cup \{u\}, \mathcal{S} \cup \{s(u) = +1\}\big) + \min_{m=0}^{k-1} \Big\{ \mathbf{OPT}\big(left(u), \mathcal{I} \cup \{u\}, \\
&\mathcal{S} \cup \{s(u) = +1\}, m\big) + \mathbf{OPT}\big(right(u), \mathcal{I} \cup \{u\}, \mathcal{S} \cup \{s(u) = +1\}, k-1-m\big) \Big\}; \\
&\mathbf{P}\big(u, s(u) = -1|\mathcal{I} \cup \{u\}, \mathcal{S} \cup \{s(u) = -1\}\big) + \min_{m=0}^{k-1} \Big\{ \mathbf{OPT}\big(left(u), \mathcal{I} \cup \{u\}, \\
&\mathcal{S} \cup \{s(u) = -1\}, m\big) + \mathbf{OPT}\big(right(u), \mathcal{I} \cup \{u\}, \mathcal{S} \cup \{s(u) = -1\}, k-1-m\big) \Big\} \Bigg\}.
\end{aligned}$$

From root u, the optimal *rumor initiator* detection can generally follow one of three cases:

- u is not the *initiator*: The root u is not added to the *rumor initiator* set, and we make recursive calls with its left and right children nodes to identify the k *rumor initiators*.
- u is the *initiator* with state $s(u) = +1$: The root u and its state are added into the *rumor initiator* set and the *state* set, respectively (i.e., $\mathcal{I} \cup \{u\}$, and $\mathcal{S} \cup \{s(u) = +1\}$). Furthermore, we make recursive calls with its left and right children nodes to identify the remaining $k-1$ *rumor initiators* based on the updated *rumor initiator* and their *state*.
- u is the *initiator* with state $s(u) = -1$: The root u and its state are added into the *rumor initiator* set and *state* set, respectively (i.e., $\mathcal{I} \cup \{u\}$, and $\mathcal{S} \cup \{s(u) = -1\}$). Furthermore, we make recursive calls with its left and right children nodes to identify the remaining $k-1$ *rumor initiators* based on the updated *rumor initiator* and their *state*.

The formal definition of $\mathbf{P}(u, s(u)|\mathcal{I}, \mathcal{S})\}$ is available in Section III-B. Meanwhile, the special case $\mathbf{P}(u, s(u)|\{u\}, \{s(u)\})$, for a *single node* u, is computed as follows:

$$\mathbf{P}(u, s(u)|\{u\}, \{s(u)\}) = \begin{cases} 1, & \text{if } s_I(u) = s(u); \\ 0, & \text{if } s_I(u) \neq s(u), \end{cases}$$

where $s_I(u)$ is the real state of u in the infected network.

The aforementioned dynamic programming objective function can be addressed in polynomial time, and we will not introduce the details involved in solving it here due to the limited space.

Algorithm 4 Infected Cascade Trees Extraction

```
Input: infected connected component set $\mathcal{C}$
Output: infected cascade tree set $\mathcal{T}$
initialize tree set $\mathcal{T} = \emptyset$
for component $C_i = (\mathcal{V}_{C_i}, \mathcal{E}_{C_i}, s_{C_i}, w_{C_i}) \in \mathcal{C}$ do
    $(\mathcal{N}, \mathcal{L}, w)$ = MWSG($C_i$)
    if $(\mathcal{N}, \mathcal{L}, w)$ contains circles $\mathcal{O}$ then
        $(\mathcal{N}', \mathcal{L}', w')$ = CC($\mathcal{N}, \mathcal{L}, w_{C_i}$)
        $(\mathcal{N}', \mathcal{L}', w')$ = MWSG(($\mathcal{N}', \mathcal{L}', w'$))
    end if
    for circle $O \in \mathcal{O}$ do
        for link $(u_x, u_o) \in \mathcal{L}'$ do
            get the corresponding link $(u_x, u_y)$, where $u_y$ is in the circle
            remove link $(\pi(u_y), u_y)$ from $\mathcal{L}$ to break the circle $O$
        end for
    end for
    $\mathcal{T} = \mathcal{T} \cup \{(\mathcal{N}, \mathcal{L})\}$
end for
```

E. RID Method for General Networks

For the ISOMIT problems in social networks of general structure and an unknown number of rumor initiators, the method introduced in the previous section cannot be directly applied. In this section, we will introduce the RID framework to address the ISOMIT problem. We propose to first detect the infected connected components from the whole network. For each detected connected component, we propose to further prune the non-existing activation links among users to extract the "*infected cascade trees*" in the signed networks. From each infected cascade tree, we introduce the objective function to detect the optimal rumor initiators (the number, identities as well as their states).

1) Infected Connected Components Detection: The infected diffusion network can contain multiple infected connected components, where users in each component can be connected to each other via potential diffusion links among them. In this part, we will introduce the method to detect the infected connected components from the network.

Definition 6: (Infected Connected Components): An *infected connected component* is a subgraph of the *infected network* and, by ignoring the directions of diffusion links, any two vertices in the component are connected to each other.

The *signed connected components* in the pruned networks can be detected with algorithms, like breadth-first search (BFS) [3] and depth-first search (DFS) [3], in linear time. For instance, based on the BFS algorithm, we will loop through all the infected vertices in the pruned infected signed network and

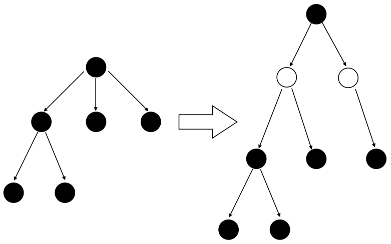


Fig. 3. Example of the binary tree transformation.

once we reach an unvisited vertex, e.g., u, we will call BFS function to find the entire connected component containing u. The time cost of BFS based connected component detection algorithm will be $O(n+m)$, where n and m are the numbers of user nodes and diffusion links in the infected diffusion network.

2) Signed Infected Cascade Forest Extraction: Let $\mathcal{C} = \{C_1, C_2, \cdots, C_l\}$ be the set of l connected components detected in the pruned infected signed network. As introduced earlier, the real information diffusion process in the infected connected component based on MFC can form a set of infected cascade trees. We show how to extract such trees later in this section.

Definition 7: (Infected Cascade Tree): The *signed infected cascade tree* summarizes the state of the information propagation and user activation process in the network. Let $T = (\mathcal{V}_T, \mathcal{E}_T, s, w)$ be an *signed infected cascade tree*. The node set $\mathcal{V}_T \subseteq \mathcal{V}_D$ consists of all the infected users in the tree and the directed activation link $(u, v) \in \mathcal{E}_T \subseteq \mathcal{E}_D$ if and only if u succeeds in activating v.

The signed infected cascade trees can be inferred from the infected network, and we propose to extract the trees capturing the most information (i.e., the most likely trees) for each connected component. Let $C_i = (\mathcal{V}_{C_i}, \mathcal{E}_{C_i}, s, w)$ be a detected connected component consisting of multiple infected cascade trees, and let $T = (\mathcal{V}_T, \mathcal{E}_T, s, w)$ be one of the tree extracted from C_i, where $\mathcal{V}_T \subseteq \mathcal{V}_{C_i}$ and $\mathcal{V}_T \subseteq \mathcal{V}_{C_i}$. The likelihood of tree T is $\mathcal{L}(T) = \prod_{(u,v)\in\mathcal{E}_T} w(u, v)$. Furthermore, the *optimal* infected cascade tree T^* in C_i can be defined as:

$$T^* = \arg\max_{T\in\mathcal{T}} \mathcal{L}(T),$$

where $\mathcal{T}$ denotes the set of all potential trees that can be detected from component C_i. The maximum likelihood *infected cascade trees* can be extracted using the Chu-Liu/Edmonds' algorithm [2], [6] from the directed connected components. The pseudo-code of the *infected cascade trees* extraction method is available in Algorithm 4, which will call the functions in Algorithms 2 and 3 to get the maximum weight spanning graphs and resolve the circles in the graph.

3) Rumor Initiator Inference: Based on the methods introduced in previous sections, we are able to detect a set of diffusion trees from the network, the roots of which without incoming edges represent the rumor initiators. Meanwhile, besides the roots, multiple *rumor initiators* can co-exist in one *infected cascade tree*. In other words, the number of extracted diffusion trees is a lower bound on the number of rumor initiators. The detected cascade tree can actually be partitioned into several isolated sub-trees instead. The roots of these sub-trees provide additional candidates for being rumor initiators. Such a partitioning process can be achieved with the algorithm introduced in Section III-D effectively. However, the extracted *infected cascade trees* from the infected signed network may not necessarily be binary trees, and this can be very complex to deal with [13]. Next, we propose to transform each cascade tree into a binary tree first and then identify the optimal rumor initiators.

TABLE II
PROPERTIES OF DIFFERENT NETWORKS

network	# nodes	# links	link type
Epinions	131,828	841,372	directed
Slashdot	77,350	516,575	directed

To transform a general tree into an binary tree without distorting information about the relative influence relationships, we propose to add extra dummy nodes to the trees, which have no effect on information diffusion, and they cannot be selected as *rumor initiators*. For example, in Figure 3, the tree in the left figure is not a binary tree, where the root node has 3 children nodes. To transform it into a binary tree, between the root and its children, $\lceil \log_2 3 \rceil$ extra nodes are added to the tree as the root's new children and the root's children nodes are assigned as the new nodes' children. These newly added nodes will not participate in the information diffusion and they cannot be selected as *rumor initiators*.

Meanwhile, to avoid the case of having too many *rumor initiators* (e.g., every user in the component is a *rumor initiators*), we will add a penalty term to constrain the number of detected rumor initiators. For each tree $T \in \mathcal{T}$ rooted at u, we can represent the optimal rumor initiators $\mathcal{I}^*$ of size k^* with initial state $\mathcal{S}^*$ as follows:

$$k^*, \mathcal{I}^*, \mathcal{S}^* = \arg\min_{k,\mathcal{I},\mathcal{S}} -\mathbf{OPT}(u, \mathcal{I}, \mathcal{S}, k) + (k-1)\cdot\beta,$$

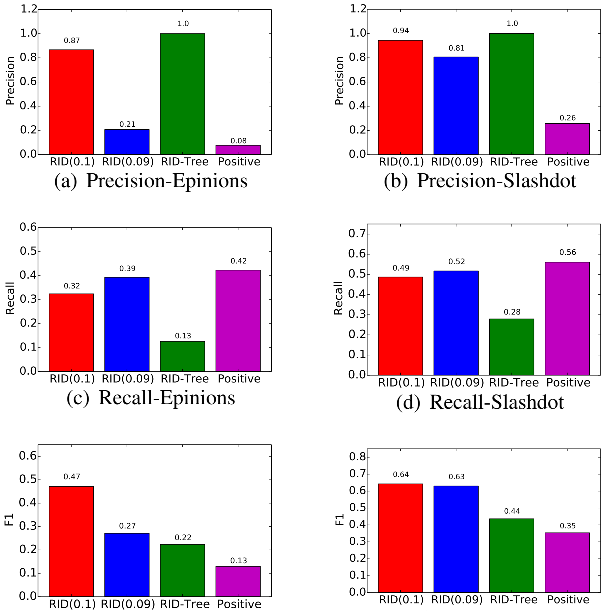
Here, parameter β denotes the penalty of each introduced rumor initiator (whose sensitivity analysis is available in Section IV-D) and term $(k-1)$ represents the extra initiators detected besides the original root of tree T. Function $\mathbf{OPT}(u, \mathcal{I}, \mathcal{S}, k)$ can be computed with the dynamic programming based method introduced in the previous section. By enumerating k from 1 to the number of nodes in T (i.e., $|\mathcal{V}_T|$), we are able to obtain the optimal solution of the above objective function. However, such a process can be very time consuming. To balance between the time cost and quality of the result, we propose to increase k from 1 to $|\mathcal{V}_T|$ and stop once the increase in k cannot lead to increase in the objective function.

IV. EXPERIMENTS

To test the effectiveness of RID in addressing the ISOMIT problem, extensive experiments were performed on real-world signed social network datasets. In this section, we will first describe the signed datasets used in this paper. We will provide the experimental settings, and give detailed analysis.

A. Dataset Description

The signed network datasets used in this paper, which include both Epinions and Slashdot. These two datasets are both

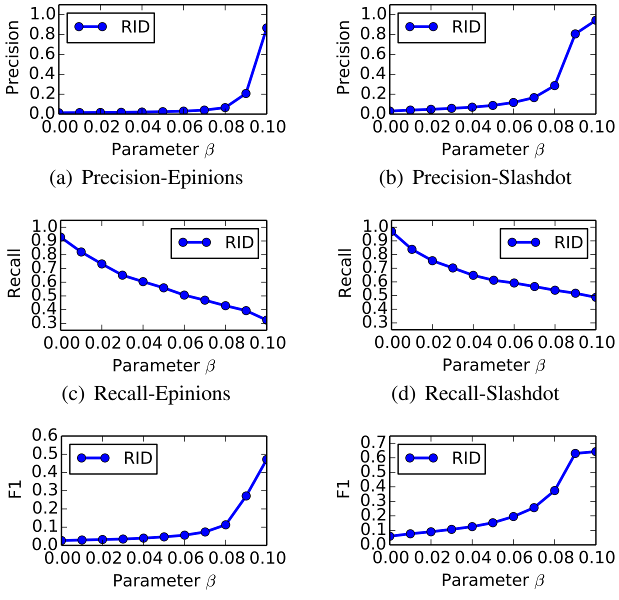


(a) Precision-Epinions (b) Precision-Slashdot

(c) Recall-Epinions (d) Recall-Slashdot

(e) F1-Epinions (f) F1-Slashdot

Fig. 4. Comparison of detected rumor initiators in each network with different methods.

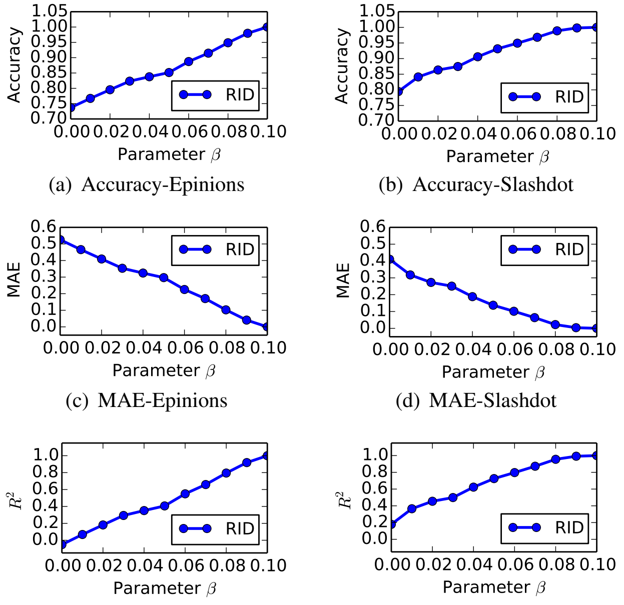


(a) Precision-Epinions (b) Precision-Slashdot

(c) Recall-Epinions (d) Recall-Slashdot

(e) F1-Epinions (f) F1-Slashdot

Fig. 5. Detected rumor initiators in each network with different β. (a-c: Epinions, d-f: Slashdot)



(a) Accuracy-Epinions (b) Accuracy-Slashdot

(c) MAE-Epinions (d) MAE-Slashdot

(e) R^2-Epinions (f) R^2-Slashdot

Fig. 6. States of detected rumor initiators in each network with different β. (a-c: Epinions, d-f: Slashdot)

public datasets and can be downloaded at site[1]. Some basic statistical information about these two datasets is available in Table II.

B. Experiment Settings

In this section, we introduce the baselines, evaluation metrics and experiment setups to examine the effectiveness of RID.

1) Comparison Methods: The comparison methods used in the experiments include:

- RID: This is the RID method proposed in our paper. By assigning parameter β with values 0.09 and 0.1, both RID($\beta = 0.09$) and RID($\beta = 0.1$) are used as the comparison methods in the experiments.
- RID-Tree: This method is obtained by simplifying the proposed RID method of our paper, and it can also be viewed as a generalization of an unsigned approach [13]. The RIDmethod contains three steps: (1) connected component extraction; (2) maximum likelihood diffusion trees extraction; and (3) rumor initiator detection from the extracted diffusion trees. The RID-Tree consists of the first two steps of RID and regards the roots of diffusion trees as the detected *rumor initiators*. RID-Tree extends and modifies the tree extraction method proposed in [13] to the signed directed networks and applies the Chu-Liu/Edmonds' algorithm [2], [6] instead.
- RID-Positive: The RID-Positive method uses the diffusion tree extraction method for regular unsigned networks in [13] and treats the roots as the *rumor initiators*. The negative links in the network are discarded and RID-Positive utilizes the positive links only.

2) Evaluation Metrics: One of our goals is to compare the *identity* of the detected rumor initiators against the ground-truth rumor initiators, which is similar to that of evaluation of other retrieval techniques. For example, one can use the precision, recall, and F1-measure. In addition, the approach also infers the *state* of the rumor initiators. This can be evaluated using metrics such as the Accuracy, MAE and R^2 (Coefficient of Determination). Considering that the methods RID-Tree and RID-Positive can only identify the *identities* of rumor initiators are but cannot determine the their initial states, the last set of metrics is designed only to measure the stability of RID, rather than against baselines.

3) Experimental Setup: Based on these social networks, we construct their corresponding *weighted signed diffusion networks* by reversing the social links among users. Meanwhile, the weight of each diffusion link, e.g., (u, v), can be

[1] http://snap.stanford.edu/data/#signnets

denoted as the Jaccard's Coefficient [18] of the corresponding social link (v, u), i.e., $JC(v, u) = \frac{|\Gamma_{out}(v) \cap \Gamma_{in}(u)|}{|\Gamma_{out}(v) \cup \Gamma_{in}(u)|}$, where $\Gamma_{out}(v)$ represents the set of users v follows and $\Gamma_{in}(u)$ denotes the followers of u. meanwhile, due to the sparsity of links in the networks, for links whose JC scores are 0, we randomly assign their weight with values randomly sampled from uniform distribution in range $[0, 0.1]$ just as what existing works do for the IC diffusion model [10]. The signs of diffusion links are identical to those of the corresponding social links.

To show how MFC works on real-world signed diffusion networks, extensive diffusion analyses have been done on these two datasets. From the signed social networks, the N rumor initiators are randomly selected from the networks. These N rumor initiators are randomly assigned to their initial state according to the positive ratio $\theta = \frac{\#positive}{N}$. For example, if $N = 10$ and $\theta = 0.5$, then 10 users will be randomly selected from the diffusion network, five of which will be assigned to the positive state and the remaining are assigned to the negative state. In the infected network simulation, the parameters N, α and θ were set to 1000, 3 and 0.5 respectively in MFC. From the rumor initiators, information can propagate to other users via the diffusion links among them in discrete steps. In MFC, the positive and negative links are asymmetric, and the weights of positive links are boosted by multiplying with the asymmetric boosting coefficient α. In addition, users can flip their states to follow their trusted neighbors. Such a process will continue until no more activations are available, and the resulting in network will be used as the input infected network for detecting the original rumor initiators.

In other words, in the experiments, the original rumor initiators who lead to the infected signed network in the simulation were used as the ground truth. The infected diffusion network introduced by the original rumor initiators is used as the input of the ISOMIT problem, based on which, RID uses the objective function in Section III-E3 to discover the identity of the rumor initiators and their states.

C. Experimental Results

The experimental results of different comparison methods achieved in networks Epinions and Slashdot are available in Figures 4.

The infected users without incoming diffusion links (i.e., the roots of extracted diffusion trees) in the network will definitely be rumor initiators. In other words, the detected rumor initiators by RID-Tree are all real rumor initiators. As a result, the precision and recall achieved by RID-Tree are 100% and 13% respectively in Figures 4(a)-4(c). The RID-Positive method can identify large number of rumor initiators in Epinions. However, only a small proportion of the identified rumor initiators are correct, which accounts for about in Epinions. Due to this reason, the precision and recall scores achieved by RID-Positive in Epinions are 8% and 42% respectively. The RID method of this paper further breaks the diffusion tree extracted by RID-Tree to locate more rumor initiators, and it will incorporate more users into the rumor initiators set, some of which are correct, and others are false. The precision achieved by RID(0.1) and RID(0.09) will be slightly lower than RID-Tree in Figure 4(a) but the recall of RID will be much larger. Taking both precision and recall into consideration, the F1 score achieved by RID(0.1) and RID(0.09) is much higher than RID-Tree and RID-Positive. Similar results can be observed in Figure 4 in network Slashdot. The sensitivity analysis of parameter β is available in the next section.

D. Parameter Sensitivity Analysis

The parameter β controls the weight of the costs introduced by the number of diffusion trees extracted from the network. Generally, when β is small, e.g., 0, the number of extracted trees will not pose any constraint on the rumor initiator detection process and RID can break each diffusion tree into a large number of smaller parts so that the cost introduced in Section III-B is minimized. On the other hand, when β is large, the constraint on the number of decomposed diffusion trees in the objective function will play a more important role. As a result, RID tends to maintain larger trees instead, even though this leads to larger structure based cost. To demonstrate these results, extensive results about the parameter β were done, and the results are available in Figures 5-6.

The choice of β has an effect on the trade-off between precision and recall. For example, as β increases, the precision increases at the expense of recall because of fewer discovered initiators. This is evident from Figure 5 for both the Epinions and Slashdot networks. Larger values of β help constrain RID in dividing the identified diffusion trees into smaller parts, and therefore the precision increases. On the other hand, with fewer diffusion trees, the number of correctly identified rumor initiators will be lower. The F1 score achieved by RID in both Epinions and Slashdot can both increase as β increases.

1) Correctness of State Identification: In addition, the RID method can not only determine the identity of the correct rumor initiators, but also also their assigned states at the very beginning. The results achieved by RID with different values of β is available in Figure 6. Among all the correctly identified rumor initiators, we calculate the accuracy, MAE and R^2 scores achieved by RID in inferring their initial states (i.e., +1 or -1). As shown in Figures 6(a) and 6(a), the accuracy achieved by RID in states inference will increase as β increases and can be very close to 100% at $\beta = 1.0$. Meanwhile, the MAE of RID drops as β increases, which will be less than 0.2 when β is greater than 0.7 in Epinions and 0.4 in Slashdot as shown in Figures 6(c) and 6(d), respectively. The inferred states of these correctly identified rumor initiators have very positive correlations with their real states. It can be demonstrated by examining the R^2 score achieved by RID in these two networks, which is available in Figures 6(e) and 6(f).

V. Related Work

Information diffusion has a rich history in research on social network analysis. Domingos and Richardson [5], [20] were the first to propose to study the influence propagation based on knowledge-sharing sites. Kempe et al. [10] were the first to study it in the specific context of social networks and propose two seminal diffusion models. These correspond to the Independent Cascade (IC) model and Linear Threshold (LT) model. These models have served as the basis of many other models.

Among these works on information diffusion, rumor propagation in online social networks is of practical importance. Kwon et al. identify characteristics of rumors by examining temporal, structural and linguistic aspects of rumors[12]. Rumors can spread very fast in online social networks, and Doerr et al. propose to study the structural and algorithmic properties of networks which accelerate such a propagation in [4]. To maximize the influence or rumors, the diffusion of competing rumors in social networks is studied in [11]

In recent years, signed networks have gained increasing attention. Li et al. [17] studied the influence diffusion dynamics and influence maximization in social networks with friend and foe relationships. Polarity related influence maximization problem in signed social networks is studied in [16], where a new diffusion model, corresponding to the Polarity Independent Cascade (P-IC) model, is proposed.

Influence source identification in regular unsigned networks has been studied in existing works. Lappas et al. [13] propose the problem of finding effectors in social networks. In LTGM10, the k-Effectors problem is formally defined and the time complexity of the problem for different types of graphs is analyzed in details. Shah et al. study similar problems in [22] to infer the sources of a rumor in a network, where a SIR-based rumor diffusion model is introduced. They propose to detect the rumor sources by identifying users with high "*rumor centrality*", which is also used in their computer virus sources discovery work [21]. Prakash et al. propose to study the culprits in epidemics in [19]. The underlying structure of cascades in online social networks is studied in [30].

VI. Conclusion

Signed networks arise in many domains, such as adversarial networks, trust/trust networks, and friend/foe networks. In many of these networks, information propagation can be affected by the signs on the links. In this paper, we present an algorithm for rumor initiator detector in infected signed networks. We propose a diffusion model for information propagation in such networks. Then we use this model to determine the initiators from a specific state of the network, with the use of the RID algorithm. We present extensive experimental results, which show the advantages of our approach over other baseline methods.

VII. Acknowledgement

This work is supported in part by NSF through grants IIS-1526499, and CNS-1626432, and NSFC 61672313.

Research of the second author was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-09-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

[1] G. Allport and L. Postman. *The psychology of rumor*. 1947.

[2] Y. Chu and T. Liu. On the shortest arborescence of a directed graph. *Science Sinica*, 1965.

[3] T. Cormen, C. Stein, R. Rivest, and C. Leiserson. *Introduction to Algorithms*. McGraw-Hill Higher Education, 2nd edition, 2001.

[4] B. Doerr, M. Fouz, and T. Friedrich. Why rumors spread so quickly in social networks. *Communications of the ACM*, 2012.

[5] P. Domingos and M. Richardson. Mining the network value of customers. In *KDD*, 2001.

[6] J. Edmonds. Optimum branchings. *Journal of Research of the National Bureau of Standards*, 1967.

[7] U. Feige. A threshold of ln n for approximating set cover. *Journal of the ACM*, 1998.

[8] A. Guille, H. Hacid, C. Favre, and D. Zighed. Information diffusion in online social networks: A survey. *SIGMOD Record*, 2013.

[9] H. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 2000.

[10] D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In *KDD*, 2003.

[11] J. Kostka, Y. Oswald, and R. Wattenhofer. Word of mouth: Rumor dissemination in social networks. In *SIROCCO*, 2008.

[12] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang. Prominent features of rumor propagation in online social media. In *ICDM*, 2013.

[13] T. Lappas, E. Terzi, D. Gunopulos, and H. Mannila. Finding effectors in social networks. In *KDD*, 2010.

[14] J. Leskovec, D. Huttenlocher, and J. Kleinberg. Predicting positive and negative links in online social networks. In *WWW*, 2010.

[15] J. Leskovec, D. Huttenlocher, and J. Kleinberg. Signed networks in social media. In *CHI*, 2010.

[16] D. Li, Z. Xu, N. Chakraborty, A. Gupta, K. Sycara, and S. Li. Polarity related influence maximization in signed social networks. *PLOS*, 2014.

[17] Y. Li, W. Chen, Y. Wang, and Z. Zhang. Influence diffusion dynamics and influence maximization in social networks with friend and foe relationships. In *WSDM*, 2013.

[18] D. Liben-Nowell and J. Kleinberg. The link-prediction problem for social networks. *Journal of the American Society for Information Science and Technology*, 2007.

[19] B. Prakash, J. Vreeken, and C. Faloutsos. Spotting culprits in epidemics: How many and which ones? In *ICDM*, 2012.

[20] M. Richardson and P. Domingos. Mining knowledge-sharing sites for viral marketing. In *KDD*, 2002.

[21] D. Shah and T. Zaman. Detecting sources of computer viruses in networks: Theory and experiment. 2010.

[22] D. Shah and T. Zaman. Rumors in a network: Who's the culprit? *IEEE Transactions on Information Theory*, 2011.

[23] J. Tang, Y. Chang, C. Aggarwal, and H. Liu. A survey of signed network mining in social media. *ACM Computing Surveys, to appear*, CoRR abs/1511.07569, 2015.

[24] R. West, H. Paskov, J. Leskovec, and C. Potts. Exploiting social network structure for person-to-person sentiment analysis. *TACL*, 2014.

[25] K. Wilcox and A. T. Stephen. Are close friends the enemy? online social networks, self-esteem, and self-control. Journal of Consumer Research, 2012.

[26] Y. Yao, H. Tong, X. Yan, F. Xu, and J. Lu. Matri: a multi-aspect and transitive trust inference model. In *WWW*, 2013.

[27] J. Ye, H. Cheng, Z. Zhu, and M. Chen. Predicting positive and negative links in signed social networks by transfer learning. In *WWW*, 2013.

[28] J. Zhang and P. S. Yu. Community detection for emerging networks. In *SDM*, 2013.

[29] J. Zhang, Q. Zhan, L. He, C. Aggarwal, and P. Yu. Trust hole identification in signed networks. In *ECMLPKDD*, 2016.

[30] B. Zong, Y. Wu, A. Singh, and X. Yan. Inferring the underlying structure of information cascades. In *ICDM*, 2012.