

# Rumor Blocking through Online Link Deletion on Social Networks

RUIDONG YAN, Renmin University of China

YI LI and WEILI WU, University of Texas at Dallas

DEYING LI and YONGCAI WANG, Renmin University of China

In recent years, social networks have become important platforms for people to disseminate information. However, we need to take effective measures such as blocking a set of links to control the negative rumors spreading over the network. In this article, we propose a *Rumor Spread Minimization* (RSM) problem, i.e., we remove an edge set from network such that the rumor spread is minimized. We first prove the objective function of RSM problem is not submodular. Then, we propose both submodular lower-bound and upper-bound of the objective function. Next, we develop a heuristic algorithm to approximate the objective function. Furthermore, we reformulate our objective function as the DS function (the Difference of Submodular functions). Finally, we conduct experiments on real-world datasets to evaluate our proposed method. The experiment results show that the upper and lower bounds are very close, which indicates the good quality of them. And, the proposed method outperforms the comparison methods.

CCS Concepts: • **Networks** → **Network algorithms**; • **Theory of computation** → **Design and analysis of algorithm**;

Additional Key Words and Phrases: Social network, rumor blocking, approximation algorithm, non-submodularity

## ACM Reference format:

Ruidong Yan, Yi Li, Weili Wu, Deying Li, and Yongcai Wang. 2019. Rumor Blocking through Online Link Deletion on Social Networks. *ACM Trans. Knowl. Discov. Data* 13, 2, Article 16 (March 2019), 26 pages. <https://doi.org/10.1145/3301302>

## 1 INTRODUCTION

In the past decade, social networks such as Twitter and Facebook have become an important part of people's daily lives. According to the statistics, there have been 2.13 billion monthly active users

This work is partly supported by National Natural Science Foundation of China under Grant 11671400 and 61672524. The Fundamental Research Funds for the Central University, and the Research Funds of Renmin University of China, 2015030273 and National Science Foundation under Grant 1747818.

Authors' addresses: R. Yan, D. Li (Corresponding author), and Y. Wang, Renmin University of China, NO.59 Zhongguancun Street, Beijing, 100872, China; emails: {yanruidong, deyingli, ycw}@ruc.edu.cn; Y. Li and W. Wu, University of Texas at Dallas, 800 W. Campbell Road, Richardson, TX 75080; emails: {yi.li, weiliwu}@utdallas.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2019 Association for Computing Machinery.

1556-4681/2019/03-ART16 \$15.00

<https://doi.org/10.1145/3301302>

and 1.4 billion daily active users on Facebook.<sup>1</sup> On one hand, these social networks not only help people stay in touch with family and friends, but also keep abreast of breaking news and emerging contents. And, they have become significant platforms for people to generate, search, and share multiple social contents. On the other hand, online social networks are used by some people or organizations to disseminate negative influence such as malicious rumors we want to block or contain [7, 14, 22, 27].

With widespread of the malicious rumors, social network would lose its reliability and even cause public panic. There are some characteristics of social networks would speed up and expand the diffusion of rumors. As we all know, one of the most valuable characteristics of social networks is its capability for users generating contents to circulate rapidly through the whole network. But when it comes to rumors this valuable characteristic will make things even worse. For example, when the devastating wildfires happens in California in October 2017, the officers were evacuating residents and searching through the burned ruins of homes for missing persons they still had to deal with the fake news.<sup>2</sup> Although the rumor was shot down by the officers and was debunked by some government websites afterwards, the original story was shared 60,000 times and similar stories were shared 75,000 times on Facebook.

Another characteristic of social networks is the openness. Whatever government organizations or ordinary citizens, they can disseminate their influence such as news, events and experience [27]. Due to the openness, anyone could share any influence without validating. For example, a tweet “Two Explosions in the White House and Barack Obama is injured” appears on the Twitter which turns out to be a fake since the Twitter account was hacked. The news leads social media and the stock market went wild. The temporary loss of market cap in S&P 500 alone was totally \$136.5 billion.<sup>3</sup> Therefore, it is very important to adopt effective strategies to block rumors and to minimize negative influence on social networks.

Existing researches have discussed the rumor blocking problem from different perspectives. Some studies show that removing nodes with high out-degree can often be an effective strategy [18, 24]. However, removing high degree nodes may cause dramatically change in network structure and may not work once the rumor has already spread out by those high degree nodes. Other studies either consider to find a minimal set of protectors to limit the negative influence of rumor [9] or introduce a positive cascade competing against the rumor [21]. In addition, there are extensive studies about the influence diffusion phenomenon in social networks [4, 12].

We use a vivid example in Figure 1 to illustrate the drawbacks of the existing methods. First, removing finite neighbors of the rumor seed node seems to be a feasible approach. However, the social network has the property of power-law distribution. Once the seed node has high degree, this approach may not work. For example, removing any two neighbors is no better than removing two protectors. The reason is that removing these two protectors will “protect” more inactive nodes. Second, removing the protectors may cause dramatically change in network structure. In reality, this method is expensive such as deleting the users’ Twitter or Facebook accounts. Third, using protectors to spread the positive influence fights against the rumor. This method usually has a strong assumption: if the rumor and positive influence reach a node at the same time, the node will adopt the positive influence. However, this assumption is not realistic. The researches show a node’s adoption depends on its neighbors. For example, a node has seven neighbors in which five neighbors spread rumor to the node and two neighbors spread positive influence to the node. As a result, this node has a higher probability to adopt the rumor instead of the positive influence.

<sup>1</sup><https://expandedramblings.com/index.php/by-the-numbers-17-amazing-facebook-stats/>.

<sup>2</sup><https://www.sfgate.com/bayarea/article/The-worst-rumors-about-the-Wine-Country-fires-12270530.php>.

<sup>3</sup><https://www.cnn.com/id/100646197>.

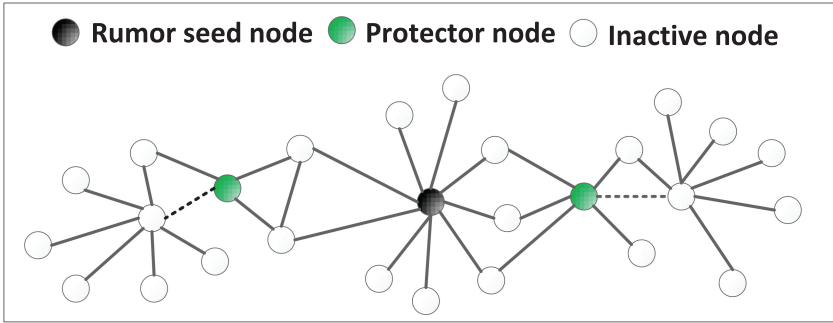


Fig. 1. An example to illustrate drawbacks of the existing methods.

Different from the above methods, our method is to remove some edges (dashed lines) such that the rumor spread as little as possible. On one hand, our method does not cause dramatically change in the network structure. On the other hand, our method is a direct control of rumor rather than an indirect control like protectors spreading the positive influence.

In this article, we address a novel problem: How can we effectively control or block the rumor spreading over the network? More specifically, given a social network  $G$ , a rumor seed set  $S$ , a candidate edge set  $E'$ , a positive integer parameter  $K$  and an influence diffusion model  $\mathcal{M}$ , we aim at identifying  $K$  edges from the candidate set and removing them from network such that the rumor spread value is minimized. In other words, we want to minimize the total probability that the nodes are activated by the seed set at the end of the rumor propagation. We call this problem *Rumor Spread Minimization* (RSM).

We consider RSM problem from a marginal decrement perspective to compute the rumor propagation as accurately as possible. As opposed to other edge blocking methods [13, 14, 23, 26], the process of edge blocking in this article could be considered as an online process. More specifically, we obtain the marginal decrement of rumor spread at each node by removing one edge from the candidate edge set at one time. In each iteration, we remove an edge that minimizes the rumor spread value. In addition, if there are multiple rumors spreading over the network, our method is still applicable. We summarize main contributions in this article as follows:

- We propose a novel RSM problem and formalize it from marginal decrement perspective for the first time.
- We analyze properties of the objective function and prove objective function is not submodular. Without the submodularity, the basic greedy algorithm could not provide performance guarantee.
- Considering the objective function without submodularity, we give a submodular lower-bound and an upper-bound. Furthermore, we also propose a heuristic algorithm to approximate the original objective function.
- We explore the relationships between the original objective function and the lower-bound as well as the upper-bound. And we prove the following: (1) The original function minus the lower-bound is a supermodular function; (2) The upper-bound minus the original function is a submodular function. Based on this, we discuss RSM problem with DS function (the Difference of Submodular functions).
- We evaluate the proposed method and compare with other popular approaches on four real-world social networks in experiments. Experiment results show that our method is outstrip to the comparison approaches.

The rest of this article is organized as follows. In Section 2, we begin by recalling some existing related work. In Section 3, we introduce influence diffusion model and problem description. In Section 4, we construct a submodular upper-bound and a submodular lower-bound, respectively. Algorithms are designed for solving RSM problem in Section 5. The experiment results are shown in Section 6. We draw conclusions in Section 7.

## 2 RELATED WORK

Domingos and Richardson [8] first study the influence between users in social networks. Kempe et al. [12] model viral marketing as a discrete optimization problem named *Influence Maximization* (IM). They propose two influence propagation models: the *Independent Cascade* (IC) model and the *Linear Threshold* (LT) model. Based on Kempe's results, other variant models have been proposed such as [2, 22, 25]. In our study, it is in the same fashion that the rumor spreads from users to users via influence propagation. We summarize the related work of the rumor blocking in the following three categories.

### 2.1 Removing Node Based Method

The first category is removing the global influential nodes such that the rumor spread is minimized [6, 15, 24]. Wang et al. in [24] try to discover and block  $k$  uninfected users to minimize the size of ultimate contaminated users. Comin et al. in [6] analyze three spreading schemes. And they propose and validate an effective methodology for the identification of the source nodes. Their method is based on the calculation of the centrality such as degree, betweenness, closeness, and eigenvector of the nodes on the sampled network. Kitsak et al. in [15] show that the most efficient spreaders are those located within the core of the network as identified by the  $k$ -shell decomposition analysis [1, 3]. The way to remove the global influential nodes is expensive, e.g., deleting users' Facebook or Twitter accounts may not be desirable for both companies.

### 2.2 Removing Edge Based Method

The second category is removing some edges that play a key role in rumor propagation [13, 14, 23, 26]. Kimura et al. in [14] propose a link blocking method to minimize the expected contamination area of the network. Tong et al. in [23] address the *NetMelt* problem, i.e., delete  $k$  edges from the original graph so that the rumor spread as less as possible. They show which edges being removed depends on the eigenvalue of the adjacency matrix of the network. Khalil et al. in [13] propose a greedy algorithm to address *Edge-deletion* problem, that is, find a set of  $k$  edges to remove such that rumor spread is minimized under the LT model. Yao et al. in [26] develop a greedy and two heuristic algorithms for the problem of minimizing the rumor spreading by blocking a limited number of links in a network. However, the above methods are all coarse-grained. In other words, they do not consider the network structure changing after each edge is removed in each iteration. Different from the above edges deletion strategy, our method is from a fine-grained level.

### 2.3 Spreading Positive Influence Based Method

The third category is spreading positive influence against the rumor, such that the positive influence is adopted by as many nodes as possible [2, 9, 19]. In [9], Fan et al. identify a minimal subset of individuals as initial protectors to minimize the number of people infected in neighbor communities at the end of rumor diffusion processes. Authors propose algorithms under two influence diffusion models and show the theoretical analysis in detail. Budak et al. in [2] propose a prediction algorithm that is based on generating random spanning trees and evaluate the performance of this approach. Nguyen et al. in [19] study a set of problems named node protectors, which aims to find the smallest set of highly influential nodes whose decontamination with good information helps

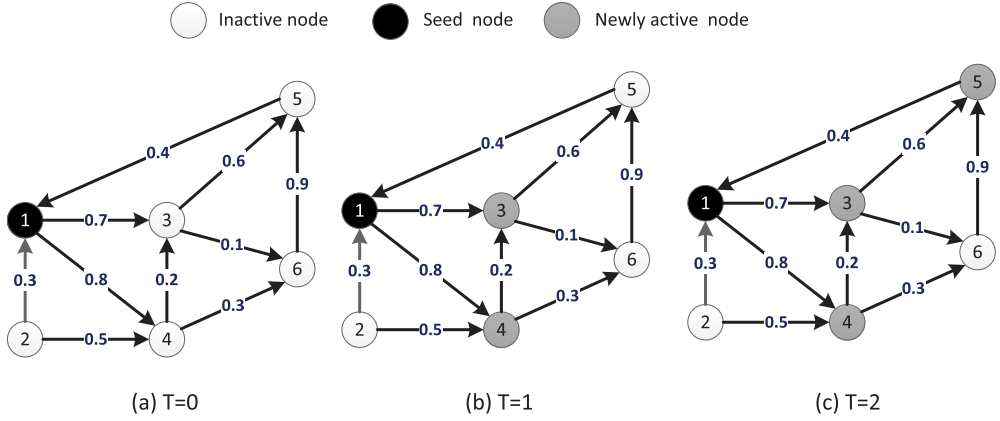


Fig. 2. The propagation process of the IC model.

to contain the viral spread of rumor. Compared with the first two of categories rumor blocking methods, this category is an indirect control of rumors.

### 3 INFLUENCE DIFFUSION MODEL AND PROBLEM DESCRIPTION

#### 3.1 Influence Diffusion Model

We briefly introduce the widely used influence diffusion model: IC model. Given a directed social network  $G = (V, E, p)$ , where  $V$  is the node set (users) and  $E \subseteq V \times V$  is the edge set (the relationships between users),  $e_{vu} = (v, u) \in E$  denotes an arbitrary edge, and  $p_{vu} \in p$  of the edge  $e_{vu}$  denotes the probability that node  $v$  activates node  $u$ . We call a node *active* if it accepts information (influence) from other nodes, *inactive* otherwise. Influence propagation process unfolds in discrete time steps. The initial seed set is  $S_0$ . Let  $S_T$  denotes the *active* nodes in time step  $T$ , and each node  $v$  in  $S_T$  has single chance to activate each *inactive* neighbor  $u$  through its out-edge  $(v, u)$  with probability  $p_{vu}$  at time step  $T + 1$ . But whether or not  $v$  succeeds, it cannot make any further attempts in subsequent rounds. Repeat this process until no more new nodes can be activated. Note that a node can only switch from *inactive* to *active*, but not in reverse direction. Figure 2 illustrates an example of the propagation process of the IC model.

In Figure 2, nodes are divided into three categories: seed node, newly active node, and inactive node. And the number embedded on each edge indicates the propagation probability, e.g.,  $p_{13}$  of the edge  $(1, 3)$  denotes the probability that the node 1 activates the node 3. According to the previous description, we know the influence propagation process is a random process. In Figure 2(a), without loss of generality, we assume node 1 is the seed node, i.e.,  $S = \{1\}$  at time step  $T = 0$ . In this time, the node 1 is ready to activate its inactive neighbors node 3 and node 4 through out-edge  $(1, 3)$  and  $(1, 4)$ , respectively. At time step  $T = 1$ , the node 3 and node 4 become newly active.<sup>4</sup> In Figure 2(b), the newly active node 3 and node 4 are ready to activate their respective inactive neighbor nodes. More specifically, on one hand, the node 3 respectively attempts to activate inactive neighbor node 5 and node 6 through out-edge  $(3, 5)$  and  $(3, 6)$ . On the other hand, the node 4 only attempts to activate inactive neighbor node 6 (since the node 3 is already active). Consequently, the node 5 becomes newly active but the node 6 is still inactive at time step  $T = 2$  in Figure 2(c). When the

<sup>4</sup>Here, we assume that node 3 and node 4 become newly active. In fact, node 3 may become newly active or inactive at time step  $T = 1$ . The state of node 4 is similar to the node 3.

Table 1. Frequently Used Notations

Notation	Description
$G = (V, E)$	social network $G$ with node set $ V  = n$ and edge set $ E  = m$
$p_{st}$	the propagation probability from node $s$ to node $t$
$e_{st}$	the directed edge from node $s$ to node $t$
$S$	the rumor seed node set
$\mathcal{M}$	information diffusion model
$E'$	the candidate edge set
$K$	a positive integer parameter
$\theta_E(v, S, \mathcal{M})$	the probability $v$ is activated by $S$ on topology $E$ under $\mathcal{M}$
$\theta_{E \setminus \{e_{st}\}}(v, S, \mathcal{M})$	the probability $v$ is activated by $S$ on topology $E \setminus \{e_{st}\}$ under $\mathcal{M}$
$\theta_E(v)$	the abbreviation for $\theta_E(v, S, \mathcal{M})$
$\mathcal{E}$	the deleted edge set from candidate edge set $E'$ and $ \mathcal{E}  = K$
$N_E^{out}(v)$	the out-neighbors of $v$ on topology $E$
$N_{E \setminus \{e_{st}\}}^{out}(v)$	the out-neighbors of $v$ on topology $E \setminus \{e_{st}\}$
$\Delta_{\mathcal{E}} \theta_E(v)$	the marginal decrement of node $v$ after removing edge set $\mathcal{E}$ from $E$
$\Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(v)$	the marginal decrement of node $v$ after removing edge $e_{st}$ from $E$
$V_s$	the ancestor node set of the node $s$
$d$	the network diameter
$\sigma_E(v)$	the rumor spread ability when $v$ is a seed node on topology $E$

time step  $T \geq 3$ , the influence propagation process stops because no node becomes a newly active node.

The readers may ask how to specifically compute the probability that a node  $v$  can be activated when given a seed set  $S$ . To address this issue, we first list frequently used notations in Table 1, then we show how to compute based on these notations. In [12], authors provide the following formula to compute the probability that a node  $v$  is activated successfully if the seed set is  $S$  under the IC model

$$\theta_E(v, S, \mathcal{M}) = \begin{cases} 1, & \text{if } v \in S \\ 0, & \text{if } N^{in}(v) = \emptyset \\ 1 - \prod_{u \in N^{in}(v)} (1 - \theta_E(u, S, \mathcal{M})p_{uv}), & \text{otherwise.} \end{cases}$$

Where  $N^{in}(v)$  is the in-neighbor set of  $v$  and  $\mathcal{M}$  denotes the IC model. It can be seen that the probability of a node  $v$  being activated depends on its all in-neighbors. Let's take the node 1 activating the node 3 as an example, i.e.,  $\theta_E(3, 1, \mathcal{M})$ .

According to the above formula, we have

$$\theta_E(3, 1, \mathcal{M}) = 1 - \prod_{u \in N^{in}(3)} (1 - \theta_E(u, 1, \mathcal{M})p_{u3}),$$

where  $N^{in}(3) = \{1, 4\}$ . Thus, we rewrite as follows:

$$\theta_E(3, 1, \mathcal{M}) = 1 - (1 - \theta_E(1, 1, \mathcal{M})p_{13}) \cdot (1 - \theta_E(4, 1, \mathcal{M})p_{43}),$$

where  $p_{13}$  and  $p_{43}$  are given. In addition,  $\theta_E(1, 1, \mathcal{M}) = 1$  since  $\theta_E(v, S, \mathcal{M}) = 1$  if  $v \in S$ . Therefore, we only need to compute  $\theta_E(4, 1, \mathcal{M})$ . Analogously,

$$\theta_E(4, 1, \mathcal{M}) = 1 - \prod_{u' \in N^{in}(4)} (1 - \theta_E(u', 1, \mathcal{M})p_{u'4}),$$

where  $N^{in}(4) = \{1, 2\}$ . Thus, we have

$$\theta_E(4, 1, \mathcal{M}) = 1 - (1 - \theta_E(1, 1, \mathcal{M})p_{14}) \cdot (1 - \theta_E(2, 1, \mathcal{M})p_{24}),$$

where  $p_{14}$  and  $p_{24}$  are given. Furthermore,  $\theta_E(1, 1, \mathcal{M}) = 1$  and  $\theta_E(2, 1, \mathcal{M}) = 0$  since  $N^{in}(2) = \emptyset$ . Therefore,

$$\theta_E(4, 1, \mathcal{M}) = 1 - (1 - 1 \cdot p_{14}) \cdot (1 - 0 \cdot p_{24}) = p_{14} = 0.8.$$

Adding  $\theta_E(4, 1, \mathcal{M}) = 0.8$  in  $\theta_E(3, 1, \mathcal{M})$ , we have

$$\begin{aligned} \theta_E(3, 1, \mathcal{M}) &= 1 - (1 - \theta_E(1, 1, \mathcal{M})p_{13}) \cdot (1 - \theta_E(4, 1, \mathcal{M})p_{43}) \\ &= 1 - (1 - 1 \cdot 0.7) \cdot (1 - 0.8 \cdot 0.2) \\ &= 0.748. \end{aligned}$$

### 3.2 Problem Description

Given a directed acyclic social network  $G = (V, E, p)$ , where  $V$  represents users and  $E$  represents relationships between users, an information diffusion model  $\mathcal{M}$  (IC model), a candidate edge set  $E' \subseteq E$  and a predetermined seed set  $S$  for rumor. Further, each node  $v \in V$  has the following parameters: (1) Let  $p_{vu} \in p$  denotes the probability that  $v$  independently disseminate rumor or influence to the neighbor  $u$ ; (2) Let  $\theta_E(v, S, \mathcal{M})$  denotes the probability that the node  $v$  is activated by the seed set  $S$  on topology  $E$  under information diffusion model  $\mathcal{M}$ . We omit parameters  $S$  and  $\mathcal{M}$  if the context is clear, i.e.,  $\theta_E(v)$ .

The goal of this article is to identify  $K$  edges denoted by  $\mathcal{E}$  from  $E'$  and to remove from the original graph  $G$  such that the rumor spread is minimized. We define RSM problem as follow.

*Definition 3.1 (RSM).* Given a directed acyclic social network  $G = (V, E, p)$ , an influence diffusion model  $\mathcal{M}$ , a predetermined rumor seed set  $S$ , a candidate edge set  $E'$  and a positive integer parameter  $K$ , RSM finds an edge set  $\mathcal{E}$  with  $|\mathcal{E}| = K$  from the candidate edge set  $E'$  such that  $f(\mathcal{E}) = \sum_{v \in V} \theta_{E \setminus \mathcal{E}}(v)$  is minimized, namely, it is equivalent to seeking

$$\mathcal{E}^* = \arg \min_{\mathcal{E} \subseteq E', |\mathcal{E}|=K} \sum_{v \in V} \theta_{E \setminus \mathcal{E}}(v), \quad (1)$$

where  $\theta_{E \setminus \mathcal{E}}(v)$  denotes the probability that the node  $v$  is activated by the seed set  $S$  on topology  $E \setminus \mathcal{E}$  under the influence diffusion model  $\mathcal{M}$ .

To effectively address this problem, we consider it from the perspective of marginal decrement. More specifically, it's satisfactory if the marginal decrement of each node in the network can be accurately calculated when an edge set is removed from the current network. Let  $\theta_{E \setminus \{e_{st}\}}(v)$  denote the probability that the node  $v$  is activated by the seed set  $S$  when the directed edge  $e_{st} = (s, t)$  is removed from current topology  $E$ . And, we have following lemma to calculate the marginal decrement of each node iteratively.

**LEMMA 3.2.** *If a directed edge  $e_{st} = (s, t)$  is removed from current topology  $E$ , we have the following two steps to calculate the marginal decrement of each node:*

—Step 1: For the node  $t$ , let  $\Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(t)$  be the marginal decrement of  $t$ , then

$$\Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(t) = \frac{p_{st} \cdot \theta_E(s) \cdot (1 - \theta_E(t))}{1 - p_{st} \cdot \theta_E(s)}, \quad (2)$$

where  $\theta_E(s)$  and  $\theta_E(t)$  denote the probabilities that node  $s$  and node  $t$  are activated by the seed set  $S$  on topology  $E$ , respectively.

—Step 2: For any node  $v \in N_{E \setminus \{e_{st}\}}^{out}(t)$ , let  $\Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(v)$  be the marginal decrement of  $v$ , then

$$\Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(v) = \frac{1 - \theta_E(v)}{1 - p_{tv} \cdot \theta_E(t)} \cdot p_{tv} \cdot \Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(t), \quad (3)$$

where  $N_{E \setminus \{e_{st}\}}^{out}(t)$  denotes the out neighbor set of  $t$  on topology  $E \setminus \{e_{st}\}$  and  $\theta_E(v)$  denotes the probability that node  $v$  is activated by the seed set  $S$  on topology  $E$ .

PROOF. We first prove Step 1. Note that the probability of node  $t$  being activated by seed set  $S$  after removing the edge  $e_{st}$  plus the marginal decrement of  $t$ , which is equal to the probability of  $t$  being activated by seed set  $S$  before removing the edge  $e_{st}$ , i.e.,

$$\theta_{E \setminus \{e_{st}\}}(t) + \Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(t) = \theta_E(t). \quad (4)$$

We assume the maximum effect of node  $s$  on node  $t$  is 1 if we can remove the edge  $e_{st}$ . Thus, the marginal decrement of node  $t$  satisfies  $\Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(t) = (1 - \theta_{E \setminus \{e_{st}\}}(t)) \cdot p_{st} \cdot \theta_E(s)$ . Adding  $\Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(t)$  to (4), then we have following equation:

$$\begin{aligned} \theta_{E \setminus \{e_{st}\}}(t) + (1 - \theta_{E \setminus \{e_{st}\}}(t)) \cdot p_{st} \cdot \theta_E(s) &= \theta_E(t). \\ \Rightarrow \theta_{E \setminus \{e_{st}\}}(t) + p_{st} \cdot \theta_E(s) &= \theta_E(t) + \theta_{E \setminus \{e_{st}\}}(t) \cdot p_{st} \cdot \theta_E(s). \\ \Rightarrow \theta_{E \setminus \{e_{st}\}}(t) &= \frac{\theta_E(t) - p_{st} \cdot \theta_E(s)}{1 - p_{st} \cdot \theta_E(s)}. \end{aligned} \quad (5)$$

Combining (4) and (5), it holds following equation:

$$\begin{aligned} \Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(t) &= \theta_E(t) - \theta_{E \setminus \{e_{st}\}}(t) \\ &= \theta_E(t) - \frac{\theta_E(t) - p_{st} \cdot \theta_E(s)}{1 - p_{st} \cdot \theta_E(s)} \\ &= \frac{p_{st} \cdot \theta_E(s) \cdot (1 - \theta_E(t))}{1 - p_{st} \cdot \theta_E(s)}. \end{aligned} \quad (6)$$

Now, we prove Step 2. On one hand, removing a directed edge  $e_{st}$  and removing two sequential directed edges  $e_{st}$  as well as  $e_{tv}$  satisfies following equation:

$$\begin{aligned} \theta_{E \setminus \{e_{st}\}}(v) &= \theta_{E \setminus \{e_{st}, e_{tv}\}}(v) + (1 - \theta_{E \setminus \{e_{st}, e_{tv}\}}(v)) \cdot p_{tv} \cdot \theta_{E \setminus \{e_{st}\}}(t), \\ \Rightarrow \theta_{E \setminus \{e_{st}, e_{tv}\}}(v) &= \frac{\theta_{E \setminus \{e_{st}\}}(v) - p_{tv} \cdot \theta_{E \setminus \{e_{st}\}}(t)}{1 - p_{tv} \cdot \theta_{E \setminus \{e_{st}\}}(t)}, \end{aligned} \quad (7)$$

where  $v \in N_{E \setminus \{e_{st}\}}^{out}(t)$ .

On the other hand, we know the following equation by (4):

$$\begin{aligned} \theta_E(v) &= \theta_{E \setminus \{e_{st}\}}(v) + \Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(v) \\ &= \theta_{E \setminus \{e_{st}, e_{tv}\}}(v) + (1 - \theta_{E \setminus \{e_{st}, e_{tv}\}}(v)) \cdot p_{tv} \cdot (\theta_{E \setminus \{e_{st}\}}(t) + \Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(t)). \end{aligned} \quad (8)$$

Adding (7) into (8), we have

$$\begin{aligned}
 \Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(v) &= \theta_E(v) - \theta_{E \setminus \{e_{st}\}}(v) \\
 &= \frac{1 - \theta_{E \setminus \{e_{st}\}}(v)}{1 - p_{tv} \cdot \theta_{E \setminus \{e_{st}\}}(t)} \cdot p_{tv} \cdot \Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(t) \\
 &= \frac{1 - \theta_E(v)}{1 - p_{tv} \cdot \theta_E(t)} \cdot p_{tv} \cdot \Delta_{e_{st}} \theta_{E \setminus \{e_{st}\}}(t), \tag{9}
 \end{aligned}$$

where  $v \in N_{E \setminus \{e_{st}\}}^{out}(t)$ . Proof is complete.  $\square$

From Lemma 3.2, we recursively use these two steps to compute the probability of each node being activated by seed set in social network when the edge  $e_{st}$  is removed from current topology until no more nodes can be updated. In particular, if there exists multiple paths from node  $t$  to node  $w$  (e.g.,  $path_1 : t \rightarrow v_1 \rightarrow w$  and  $path_2 : t \rightarrow v_2 \rightarrow w$ ), computing the marginal decrement of  $w$  is independent of the updating order. More specifically, computing the marginal decrement of  $w$  first passing  $path_1$  and then passing  $path_2$  is equal to calculating the marginal decrement of  $w$  first passing  $path_2$  and then passing  $path_1$ , where  $w \in N_{E \setminus \{e_{st}\}}^{out}(v_1) \cap N_{E \setminus \{e_{st}\}}^{out}(v_2)$ .

Based on above discussion, the objective function  $f(\mathcal{E})$  can be rewritten with marginal decrement of each node in the network, that is,

$$f(\mathcal{E}) = \sum_{v \in V} (\theta_E(v) - \Delta_{\mathcal{E}} \theta_{E \setminus \mathcal{E}}(v)). \tag{10}$$

It can be interpreted as the objective function value is equal to the probability that nodes in the network are activated by the seed set on the original topology minus the corresponding marginal decrement after removing an edge set from original topology. In (10),  $\theta_E(v)$  denotes the probability of node  $v$  being activated by the seed set  $S$  on original network,  $\mathcal{E}$  denotes the edge set removed from the original network and  $|\mathcal{E}| = K$ , and  $\Delta_{\mathcal{E}} \theta_{E \setminus \mathcal{E}}(v)$  denotes the marginal decrement of node  $v$  after removing the edge set  $\mathcal{E}$ . The item  $\sum_{v \in V} \theta_E(v)$  is fixed when given the original network  $G$ . Note that minimizing the function (10) is equivalent to maximizing total marginal decrement. Therefore, we focus on total marginal decrement caused by removing edge set  $\mathcal{E}$ , i.e.,

$$g(\mathcal{E}) = \sum_{v \in V} \Delta_{\mathcal{E}} \theta_{E \setminus \mathcal{E}}(v). \tag{11}$$

#### 4 SUBMODULAR LOWER-BOUND AND UPPER-BOUND

In this section, we first show that the objective function  $f(\mathcal{E})$  is not submodular. Based on non-submodularity, then we propose a submodular lower-bound and a submodular upper-bound. The main results are as follows.

*Example 4.1.* The object function  $f(\mathcal{E})$  is not submodular under the IC model.

**PROOF.** Submodularity indicates a natural diminishing returns property. Specifically, a submodular function is a set function  $F : 2^E \rightarrow \mathbb{R}$ , where  $2^E$  denotes the power set of  $E$ , which satisfies for every  $A \subseteq B \subseteq E$  and  $e \in E \setminus B$ ,  $F(A \cup \{e\}) - F(A) \geq F(B \cup \{e\}) - F(B)$ . We prove objective function  $f(\mathcal{E})$  is not submodular by the counterexample in Figure 3.

In Figure 3, the nodes are divided into three categories: seed node, newly active node and inactive node. The original social network is Figure 3(a). And, the number embedded on each edge indicates the propagation probability. Here, we set the propagation probability of all edges to be 1, so the propagation process is deterministic once the seed node set is selected. Without loss of generality, let  $S = \{1\}$  be the rumor seed set and other nodes are inactive at time step  $T = 0$ .

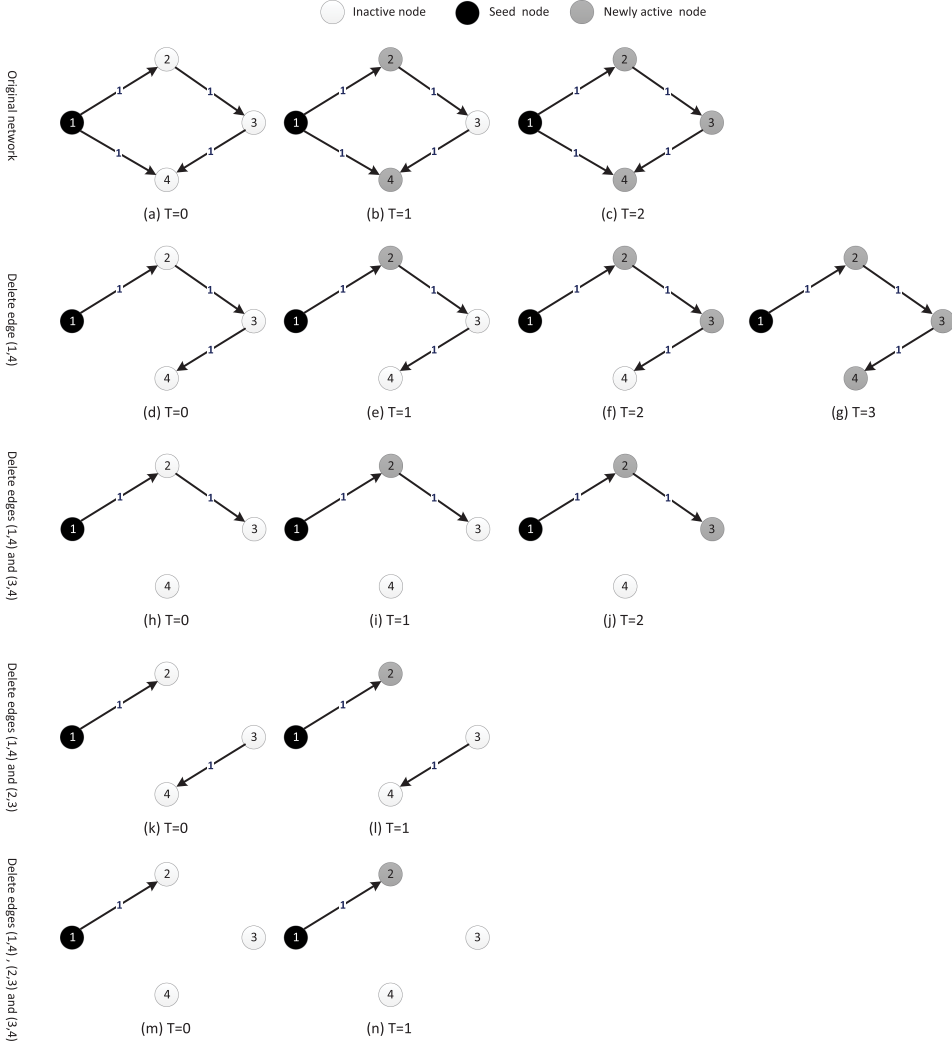


Fig. 3. A counterexample to show non-submodularity.

- *Case 1:* Without deleting any edges, the influence propagation process is shown in Figure 3(a)–(c). More specifically, at time  $T = 0$ , the node 1 is ready to activate its inactive neighbor node 2 and node 4, respectively. At time  $T = 1$ , the node 2 and node 4 become newly active since the propagation probabilities are 1. At this time, the node 2 attempts to activate node 3. At time  $T = 2$ , the node 3 becomes newly active and the influence propagation process stops.
- *Case 2:* Let  $A = \{(1, 4)\}$  and delete  $A$  from the original network (see Figure 3(d)). At time  $T = 0$ , the node 1 attempts to activate node 2. At time  $T = 1$ , the node 2 becomes active and it attempts to activate the node 3. At time  $T = 2$ , the node 3 become active and it attempts to activate the node 4. At time  $T = 3$ , the node 4 becomes active and the influence propagation process stops (see Figure 3(d)–(g)). Let function  $f(A)$  denote the sum of the probabilities that each node is activated by the seed node after removing the edge set  $A$ , that is,

- $f(A) = \sum_{v \in V} \theta_{E \setminus A}(v)$ . Apparently,  $f(A) = 4$  since all nodes in network can be activated by the seed node.
- *Case 3:* Let  $B = \{(1, 4), (3, 4)\}$  and delete  $B$  from the original network (see Figure 3(h)). At time  $T = 0$ , the node 1 attempts to activate node 2. At time  $T = 1$ , the node 2 becomes active and it attempts to activate the node 3. At time  $T = 2$ , the node 3 become active and the influence propagation process stops (see Figure 3(h)–(j)). Similarly,  $f(B) = \sum_{v \in V} \theta_{E \setminus B}(v) = 3$  since there are three nodes in network activated by the seed node.
  - *Case 4:* Let  $e = (2, 3)$  and delete  $A \cup \{e\}$  from the original network (see Figure 3(k)). At time  $T = 0$ , the node 1 attempts to activate node 2. At time  $T = 1$ , the node 2 becomes active and the influence propagation process stops (see Figure 3(k) and (l)). Similarly,  $f(A \cup \{e\}) = \sum_{v \in V} \theta_{E \setminus \{A \cup \{e\}\}}(v) = 2$  since there are two nodes in network activated by the seed node.
  - *Case 5:* Let  $e = (2, 3)$  and delete  $B \cup \{e\}$  from the original network (see Figure 3(m)). At time  $T = 0$ , the node 1 attempts to activate node 2. At time  $T = 1$ , the node 2 becomes active and the influence propagation process stops (see Figure 3(m) and (n)). Similarly,  $f(B \cup \{e\}) = \sum_{v \in V} \theta_{E \setminus \{B \cup \{e\}\}}(v) = 2$  since there are two nodes in network activated by the seed node.

In summary, we observe that  $A \subseteq B \subseteq E$  and  $e \in E \setminus B$ .  $f(A) = 4$ ,  $f(A \cup \{e\}) = 2$ ,  $f(B) = 3$  and  $f(B \cup \{e\}) = 2$ . Thus,  $f(A \cup \{e\}) - f(A) < f(B \cup \{e\}) - f(B)$  indicates function  $f(\cdot)$  is not submodular.  $\square$

From Example 4.1, the objective function is not submodular. We analyze the reasons that lead to its without submodularity. Let  $\mathcal{E} = \{e_{s_1 t_1}, \dots, e_{s_K t_K}\}$  be the set of  $K$  removed edges and  $\mathcal{E}^k = \{e_{s_1 t_1}, \dots, e_{s_k t_k}\}$  be the set of first  $k$  removed edges where  $k = 1, \dots, K$ . In particular,  $\mathcal{E}^0 = \emptyset$ . Thus, we rewrite the function (11) as follow

$$g(\mathcal{E}) = \sum_{k=1}^K \Delta_{e_{s_k t_k}} g(\mathcal{E}^k), \quad (12)$$

where  $\Delta_{e_{s_k t_k}} g(\mathcal{E}^k)$  denotes the marginal decrement when we remove the  $k$ th edge  $e_{s_k t_k} = (s_k, t_k)$  on topology  $E \setminus \mathcal{E}^k$ , that is,  $\Delta_{e_{s_k t_k}} g(\mathcal{E}^k) = \sum_{v \in V} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v)$ . Let  $N_E^{out}(v)$  be the out-neighbor set of  $v$  on topology  $E$ . Apparently, as the number of removed edges increases, they satisfy  $N_E^{out}(v) \supseteq N_{E \setminus \mathcal{E}^1}^{out}(v) \supseteq \dots \supseteq N_{E \setminus \mathcal{E}^K}^{out}(v)$ . According to influence propagation process unfolding in discrete time steps,  $\Delta_{e_{s_k t_k}} g(\mathcal{E}^k)$  can be rewritten as follows:

$$\begin{aligned} \Delta_{e_{s_k t_k}} g(\mathcal{E}^k) &= \sum_{v \in V} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v) \\ &= (\Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(t_k)) \\ &\quad + \sum_{v_1 \in N_{E \setminus \mathcal{E}^k}^{out}(t_k)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_1) \\ &\quad + \sum_{v_2 \in N_{E \setminus \mathcal{E}^k}^{out}(v_1)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_2) \\ &\quad \dots + \sum_{v_d \in N_{E \setminus \mathcal{E}^k}^{out}(v_{d-1})} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_d), \end{aligned} \quad (13)$$

where  $d$  is the network diameter and  $k = 1, \dots, K$ . From Equation (13), we can observe the following:

- The out-neighbors of  $v \in V$  become fewer when the more edges are removed, i.e., the number of items in the summation symbol is reduced. As a result, it causes  $g(\mathcal{E})$  to decrease.
- $\Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v)$  is monotone increasing with respect to  $\mathcal{E}^k$ . This is because the more edges that are deleted, the greater the total marginal decrement. It causes  $g(\mathcal{E})$  to increase.

Based on these two observations, whether the function  $g(\mathcal{E})$  increasing or decreasing is uncertain with respect to  $\mathcal{E}^k$ . Exactly, this reason makes  $g(\mathcal{E})$  non-submodular. Since function (12) is not submodular, there is no approximate guarantee if we adopt the basic greedy algorithm. Therefore, we propose a submodular upper-bound and a submodular lower-bound such that the *Sandwich Algorithm* (SA) [16] can be applied.

#### 4.1 Submodular Upper-Bound

Based on analysis in previous section, we construct a submodular upper-bound of the function (12) as  $\bar{g}(\mathcal{E}) = \sum_{k=1}^K \Delta_{e_{s_k t_k}} \bar{g}(\mathcal{E}^k)$ , where

$$\begin{aligned}
 \Delta_{e_{s_k t_k}} \bar{g}(\mathcal{E}^k) &= (\Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(t_k) \\
 &\quad + \sum_{v_1 \in N_E^{out}(t_k)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_1) \\
 &\quad + \sum_{v_2 \in N_E^{out}(v_1)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_2) \\
 &\quad \dots + \sum_{v_d \in N_E^{out}(v_{d-1})} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_d)) \\
 &\geq \sum_{v \in V} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v) \\
 &= \Delta_{e_{s_k t_k}} g(\mathcal{E}^k).
 \end{aligned} \tag{14}$$

The key idea is if we fix the out-neighbors of  $v \in V$  on original topology  $E$ , then Equation (14) is obviously the upper-bound of function (12). Furthermore, the upper-bound is submodular as long as  $\Delta_{e_{s_k t_k}} \bar{g}(\mathcal{E}^k)$  is monotone decreasing with respect to  $E \setminus \mathcal{E}^k$ . We prove this result by following theorem.

**THEOREM 4.2.** *The upper-bound function  $\bar{g}(\mathcal{E}) = \sum_{k=1}^K \Delta_{e_{s_k t_k}} \bar{g}(\mathcal{E}^k)$  is submodular with respect to  $E \setminus \mathcal{E}^k$ .*

**PROOF.** For any edge set  $A \subseteq B \subseteq E \setminus \mathcal{E}^k$  and any edge  $e_{st} = (s, t) \in \{E \setminus \mathcal{E}^k\} \setminus B$ , we have  $\Delta_{e_{st}} \theta_{\{E \setminus \mathcal{E}^k\} \setminus A}(v) \geq \Delta_{e_{st}} \theta_{\{E \setminus \mathcal{E}^k\} \setminus B}(v)$  since  $\Delta_{e_{st}} \theta_{E \setminus \mathcal{E}^k}(v)$  is monotone decreasing with respect to  $E \setminus \mathcal{E}^k$ , where  $\Delta_{e_{st}} \theta_{\{E \setminus \mathcal{E}^k\} \setminus A}(v) = \bar{g}(A \cup \{e_{st}\}) - \bar{g}(A)$  and  $\Delta_{e_{st}} \theta_{\{E \setminus \mathcal{E}^k\} \setminus B}(v) = \bar{g}(B \cup \{e_{st}\}) - \bar{g}(B)$ . Furthermore, the number of marginal decrement terms remain unchanged during the removing edges process. Therefore, we have

$$\begin{aligned}
 \bar{g}(A \cup \{e_{st}\}) - \bar{g}(A) &= (\Delta_{e_{st}} \theta_{\{E \setminus \mathcal{E}^k\} \setminus A}(t) \\
 &\quad + \sum_{v_1 \in N_E^{out}(t)} \Delta_{e_{st}} \theta_{\{E \setminus \mathcal{E}^k\} \setminus A}(v_1) \\
 &\quad \dots + \sum_{v_d \in N_E^{out}(v_{d-1})} \Delta_{e_{st}} \theta_{\{E \setminus \mathcal{E}^k\} \setminus A}(v_d))
 \end{aligned}$$

$$\begin{aligned}
&\geq (\Delta_{e_{st}} \theta_{\{E \setminus \mathcal{E}^k\} \setminus B}(t)) \\
&+ \sum_{v_1 \in N_E^{out}(t)} \Delta_{e_{st}} \theta_{\{E \setminus \mathcal{E}^k\} \setminus B}(v_1) \\
&\dots + \sum_{v_d \in N_E^{out}(v_{d-1})} \Delta_{e_{st}} \theta_{\{E \setminus \mathcal{E}^k\} \setminus A}(v_d) \\
&= \bar{g}(B \cup \{e_{st}\}) - \bar{g}(B), \tag{15}
\end{aligned}$$

where  $d$  is the network diameter. It indicates upper bound function  $\bar{g}(\mathcal{E})$  is submodular.  $\square$

#### 4.2 Submodular Lower-Bound

Analogously, we can also construct a submodular lower-bound of function (12) as  $\underline{g}(\mathcal{E}) = \sum_{k=1}^K \Delta_{e_{s_k t_k}} \underline{g}(\mathcal{E}^k)$ , where

$$\begin{aligned}
\Delta_{e_{s_k t_k}} \underline{g}(\mathcal{E}^k) &= (\Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(t_k)) \\
&+ \sum_{v_1 \in N_{E \setminus E'}^{out}(t_k)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_1) \\
&+ \sum_{v_2 \in N_{E \setminus E'}^{out}(v_1)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_2) \\
&\dots + \sum_{v_d \in N_{E \setminus E'}^{out}(v_{d-1})} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_d) \\
&\leq \sum_{v \in V} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v) \\
&= \Delta_{e_{s_k t_k}} \underline{g}(\mathcal{E}^k). \tag{16}
\end{aligned}$$

When we fix the out-neighbors of  $v \in V$  on the topology  $E \setminus E'$ , then function (16) is obviously the lower-bound of function (12). Furthermore, the lower-bound is submodular as long as  $\Delta_{e_{s_k t_k}} \underline{g}(\mathcal{E}^k)$  is monotone decreasing with respect to  $E \setminus \mathcal{E}^k$ . We use the following theory to prove this result.

**THEOREM 4.3.** *The lower-bound function  $\underline{g}(\mathcal{E}) = \sum_{k=1}^K \Delta_{e_{s_k t_k}} \underline{g}(\mathcal{E}^k)$  is submodular with respect to  $E \setminus \mathcal{E}^k$ .*

The idea of proof is similar to Theorem 4.2, therefore we omit the proof. From above two theorems, we can obtain the removed edge set  $\mathcal{E}^k$  if we take complementary set of  $E \setminus \mathcal{E}^k$ . In addition, it's worth noting that

- The function  $\bar{g}(\mathcal{E})$  corresponds to the lower-bound of original objective function, i.e.,  $\bar{f}(\mathcal{E}) = \sum_{v \in V} \theta_E(v) - \bar{g}(\mathcal{E})$ ;
- The function  $\underline{g}(\mathcal{E})$  corresponds to the upper-bound of original objective function, i.e.,  $\bar{f}(\mathcal{E}) = \sum_{v \in V} \theta_E(v) - \underline{g}(\mathcal{E})$ .

## 5 ALGORITHM

The importance of submodularity in machine learning and data mining applications has been demonstrated in many literatures [5, 13, 16, 25]. They apply submodularity to tasks such as finding a set of nodes in a social network that maximize influence while satisfying a cardinality constraint. Narasimhan and Bilmes in [17] propose an algorithm for minimizing the difference between two

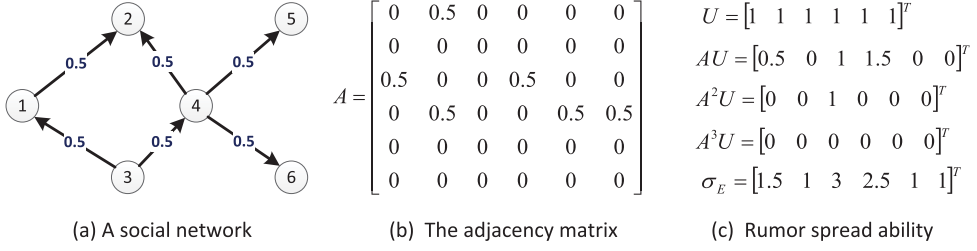


Fig. 4. An example to show the rumor spread ability of each node.

submodular functions using a variational framework, which is based on the concave-convex procedure. They also show that any set function can be expressed as a difference between two submodular functions and hence every set function optimization problem can be reduced to minimizing a difference between submodular functions. In this section, we first propose a heuristic algorithm to approximate the original objective function. Then, we discuss the difference between the upper-bound function and the original function.

### 5.1 Heuristic Algorithm

RSM problem aims at selecting a  $K$  edge set  $\mathcal{E}$  from the candidate edge set  $E'$  and removing from the original network such that the total rumor spread value  $\sum_{v \in V} \theta_{E \setminus \mathcal{E}}(v)$  is minimized on topology  $E \setminus \mathcal{E}$ . A feasible method is to delete or block edges between the seed nodes and the influential nodes who have strong ability to disseminate influence. Thus, how to measure the strength of a node to disseminate influence? This issue is very important.

Given a social network  $G = (V, E, p)$  with  $|V| = n$  and  $|E| = m$ , for  $v \in V$ , let  $\sigma_E(v)$  be the rumor spread ability when the  $v$  is a seed node on topology  $E$ . The  $n$  dimensional column vector  $\sigma_E = [\dots, \sigma_E(v), \dots]^T$  denotes the rumor spread abilities of all nodes on topology  $E$ . Let  $\mathcal{A} = (a_{uv})_{n \times n}$  be the adjacency matrix of  $G$  whose  $a_{uv} = p_{uv}$  if  $(u, v) \in E$  and  $a_{uv} = 0$  otherwise. The  $n$  dimensional column vector  $U = [1, \dots, 1]^T$  denotes unit vector and  $d$  is the network diameter. Then, we have  $\sigma_E = \sum_{0 \leq i \leq d} \mathcal{A}^i U$ . We adopt this method to evaluate the ability of each node to disseminate influence because the  $\mathcal{A}^l U$  indicates the  $l$ th hop propagation spread of all nodes on network. Figure 4 shows an example.

Figure 4(a) shows a simple social network with 6 nodes and 6 edges, where all the propagation probabilities are 0.5. Figure 4(b) shows the adjacency matrix of this network. Figure 4(c) shows how to compute the rumor spread ability of each node. More specifically, the  $AU = [0.5 \ 0 \ 1 \ 1.5 \ 0 \ 0]^T$  indicates the 1th hop propagation spread of all nodes. Similarly, the  $A^2U = [0 \ 0 \ 1 \ 0 \ 0 \ 0]^T$  indicates the 2-th hop propagation spread of all nodes. And the  $A^3U = [0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$  since the network diameter  $d$  is 2. Finally,  $\sigma_E = \sum_{0 \leq i \leq d} \mathcal{A}^i U = [1.5 \ 1 \ 3 \ 2.5 \ 1 \ 1]^T$ . Therefore, we say the node 3 has the strongest rumor spread ability.

Based on the  $\sigma_E$  and previous lemma, we design a heuristic strategy (see Algorithm 2), which iteratively selects  $K$  edges from the candidate edge set  $E'$ . Specifically, in  $k$ th iteration, we first determine which edge  $e_{st} = (s, t) \in E'$  has the maximum decrement spread, i.e.,  $e_{st} = \arg \max_{e_{st} \in E'} (1 - \theta_{E \setminus \mathcal{E}^k}(u))(\sigma_{E \setminus \mathcal{E}^k}(v) + \sigma_{E \setminus \mathcal{E}^k}(u))$ , where  $\theta_{E \setminus \mathcal{E}^k}(v)$  denotes the probability that node  $v$  is being activated by seed set  $S$  on topology  $E \setminus \mathcal{E}^k$  and  $\sigma_{E \setminus \mathcal{E}^k}(v)$  denotes the rumor spread ability of  $v$ . Then, we remove the edge  $e_{st} = (s, t)$  from the network. Finally, we update  $\sigma_{E \setminus \mathcal{E}^k}(v)$  of each node in network because the network topology changes when we remove an edge.

Note that it is not necessary to update  $\sigma_{E \setminus \mathcal{E}^k}(v)$  for all nodes. Three consequences will appear if we remove the edge  $e_{st} = (s, t)$  in  $k$ th iteration:

- The  $\sigma_{E \setminus \mathcal{E}^k}(s)$  should be updated. The reason is removing the edge  $(s, t)$  will have an impact on the rumor propagation ability of  $s$ .
- Let  $V_s$  be ancestors of the node  $s$ . For each node  $v \in V_s$ ,  $\sigma_{E \setminus \mathcal{E}^k}(v)$  should be reversely updated from node  $s$  because its rumor spread ability has changed when removing the edge  $(s, t)$ .
- For each node  $v \in V \setminus \{s \cup V_s\}$ ,  $\sigma_{E \setminus \mathcal{E}^k}(v)$  remain unchanged. This is because its network topology does not change. In other words, its rumor spread ability does not change when removing the edge  $(s, t)$ .

We propose a subroutine (Algorithm 1) to update the rumor spread ability of node  $v$  in  $k$ th iteration of Algorithm 2 if we remove the edge  $e_{st} = (s, t)$ , i.e.,  $\sigma_{E \setminus \mathcal{E}^k}(v)$ .

---

**ALGORITHM 1:** Updating the  $\sigma_{E \setminus \mathcal{E}^k}(v)$ 


---

**Input:** a social network  $G$ , an edge  $e_{st} = (s, t)$  and the  $(k - 1)$ -th  $\sigma_{E \setminus \mathcal{E}^{k-1}}(v)$

**Output:** the rumor spread ability  $\sigma_{E \setminus \mathcal{E}^k}(v)$

---

```

if the node  $v \in V_s \cup \{s\}$  then
    if the node  $v = s$  then
         $\sigma_{E \setminus \mathcal{E}^k}(s) \leftarrow \sigma_{E \setminus \mathcal{E}^{k-1}}(s) - p_{st} \cdot \sigma_{E \setminus \mathcal{E}^{k-1}}(t)$ ;
    else
        while each ancestor node  $v \in V_s$  do
             $\sigma_{E \setminus \mathcal{E}^k}(v) \leftarrow \sigma_{E \setminus \mathcal{E}^{k-1}}(v) - p_{vu} \cdot \sigma_{E \setminus \mathcal{E}^{k-1}}(u)$  for  $v, u \in V_s$  and  $u \in N_{E \setminus \mathcal{E}^{k-1}}^{out}(v)$ ;
        end
    end
else
     $\sigma_{E \setminus \mathcal{E}^k}(v) \leftarrow \sigma_{E \setminus \mathcal{E}^{k-1}}(v)$ ;
end

```

---

In Algorithm 1, first updating process starts from the node  $s$ . Second, we update the parent nodes of  $s$ . Finally, we recursively updating the parent nodes of parent nodes of  $s$  until  $v \in V_s$  have been updated where  $V_s$  denotes the ancestors of  $s$ . Based on this subroutine, we propose a heuristic algorithm (Algorithm 2) to get a high quality approximate solution of the original objective function.

Algorithm 2 consists of two steps: the initialization step and the iteration step. In the initialization step, we compute the rumor spread ability  $\sigma_E(v)$  and the probability that a node  $v$  is activated by the seed set  $S$ , that is,  $\theta_E(v)$ . In the iteration step, we select  $K$  edges which have the maximum marginal decrement as the output of algorithm.

Let's analyze the time complexity of Algorithm 2. It takes at most  $O(dm)$  time to compute  $\sigma_E(v)$  with adjacency matrix  $\mathcal{A}$ , where  $d$  is network diameter. For each  $v \in V$ , computing  $\theta_E(v)$  needs  $O(nm)$  time. The loop contains  $K$  iterations. In each iteration, selecting best edge  $e_{st}$  from candidate set needs  $|E'|$  comparisons. And, updating  $\sigma_{E \setminus \mathcal{E}^k}(v)$ ,  $\Delta_{e_{st}} \theta_{E \setminus \mathcal{E}^k}(v)$ ,  $\theta_{E \setminus \mathcal{E}^k}(v)$  takes  $O(2m)$  time. Therefore, the total time is  $O(dm + nm + K(|E'| + 2m))$ .

## 5.2 Difference of Submodular Functions

In this section, inspired by the results of [10, 11], we study the difference between original objective function and these two submodular bounds. The main results are Theorem 5.1 and Theorem 5.2, respectively.

**THEOREM 5.1.** *The function  $h(\mathcal{E}) = f(\mathcal{E}) - \underline{f}(\mathcal{E})$  is supermodular.*

**ALGORITHM 2:** Marginal Decrement Strategy (MDS) for RSM

**Input:** A social network  $G = (V, E, p)$ , a candidate edge set  $E'$ , a positive inter  $K$ , a rumor seed set  $S$  and the IC model  $\mathcal{M}$

**Output:** a rumor deletion edge set  $\mathcal{E}$

$\mathcal{E}^0 \leftarrow \emptyset$ ;

**for** each node  $v \in V$  **do**

    Compute the rumor spread ability  $\sigma_E(v)$ ;

    Compute the probability that  $v$  is activated by the seed set  $S$ , i.e.,  $\theta_E(v)$ ;

**end**

**while**  $|\mathcal{E}^k| < K$  **do**

$e_{st} = (s, t) \leftarrow \arg \max_{e_{vu} \in E'} (1 - \theta_{E \setminus \mathcal{E}^k}(u)) (\sigma_{E \setminus \mathcal{E}^k}(v) + \sigma_{E \setminus \mathcal{E}^k}(u))$ ;

    Update  $\sigma_{E \setminus \mathcal{E}^k}(v)$  with Algorithm 1;

    Update  $\Delta_{e_{st}} \theta_{E \setminus \mathcal{E}^k}(v), \theta_{E \setminus \mathcal{E}^k}(v)$  with Lemma 3.2;

$\mathcal{E}^k \leftarrow \mathcal{E}^k \cup \{e_{st}\}$ ;

**end**

**return**  $\mathcal{E} \leftarrow \mathcal{E}^k$ .

PROOF. According to Equations (10) and (11),

$$f(\mathcal{E}) = \sum_{v \in V} \theta_E(v) - g(\mathcal{E}), \quad (17)$$

$$\underline{f}(\mathcal{E}) = \sum_{v \in V} \theta_E(v) - \bar{g}(\mathcal{E}). \quad (18)$$

Let function

$$h(\mathcal{E}) = f(\mathcal{E}) - \underline{f}(\mathcal{E}) = \bar{g}(\mathcal{E}) - g(\mathcal{E}).$$

On one hand, according to Equations (12) and (14), the submodular upper-bound  $\bar{g}(\mathcal{E}) = \sum_{k=1}^K \Delta_{e_{s_k t_k}} \bar{g}(\mathcal{E}^k)$ , that is,

$$\begin{aligned} \bar{g}(\mathcal{E}) &= \sum_{k=1}^K (\Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(t_k)) \\ &+ \sum_{v_1 \in N_E^{out}(t_k)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_1) \\ &+ \sum_{v_2 \in N_E^{out}(v_1)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_2) \\ &\dots + \sum_{v_d \in N_E^{out}(v_{d-1})} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_d)). \end{aligned} \quad (19)$$

On the other hand,

$$\begin{aligned} g(\mathcal{E}) &= \sum_{k=1}^K (\Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(t_k)) \\ &+ \sum_{v_1 \in N_{E \setminus \mathcal{E}^k}^{out}(t_k)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_1) \end{aligned}$$

$$\begin{aligned}
& + \sum_{v_2 \in N_{E \setminus \mathcal{E}^k}^{out}(v_1)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_2) \\
\ldots & + \sum_{v_d \in N_{E \setminus \mathcal{E}^k}^{out}(v_{d-1})} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_d).
\end{aligned} \tag{20}$$

Combining Equations (19) and (20), we have

$$\begin{aligned}
h(\mathcal{E}) &= \bar{g}(\mathcal{E}) - g(\mathcal{E}) \\
&= \sum_{k=1}^K \left( \sum_{v_1 \in N_{\mathcal{E}^k}^{out}(t_k)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_1) \right. \\
&\quad + \sum_{v_2 \in N_{\mathcal{E}^k}^{out}(v_1)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_2) \\
&\quad \left. \ldots + \sum_{v_d \in N_{\mathcal{E}^k}^{out}(v_{d-1})} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_d) \right).
\end{aligned} \tag{21}$$

We have the following observations based on the above equation:

- The marginal decrement is monotone increasing with respect to the removed edge set  $\mathcal{E}^k$  on current topology.
- The number of terms such as  $v_1 \in N_{\mathcal{E}^k}^{out}$  is also monotone increasing with respect to the removed edge set  $\mathcal{E}^k$ .

Therefore, the function  $h(\mathcal{E})$  is supermodular.  $\square$

**THEOREM 5.2.** *The function  $l(\mathcal{E}) = \bar{f}(\mathcal{E}) - f(\mathcal{E})$  is submodular.*

**PROOF.** According to Equations (10) and (16),

$$f(\mathcal{E}) = \sum_{v \in V} \theta_E(v) - g(\mathcal{E}), \tag{22}$$

$$\bar{f}(\mathcal{E}) = \sum_{v \in V} \theta_E(v) - \underline{g}(\mathcal{E}). \tag{23}$$

Let function

$$l(\mathcal{E}) = \bar{f}(\mathcal{E}) - f(\mathcal{E}) = g(\mathcal{E}) - \underline{g}(\mathcal{E}).$$

On one hand, according to Equations (12) and (16), the submodular lower-bound  $\underline{g}(\mathcal{E}) = \sum_{k=1}^K \Delta_{e_{s_k t_k}} \underline{g}(\mathcal{E}^k)$ , that is,

$$\begin{aligned}
\underline{g}(\mathcal{E}) &= \sum_{k=1}^K \left( \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(t_k) \right. \\
&\quad \left. + \sum_{v_1 \in N_{E \setminus \mathcal{E}^k}^{out}(t_k)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_1) \right)
\end{aligned}$$

$$\begin{aligned}
& + \sum_{v_2 \in N_{E \setminus E'}^{out}(v_1)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_2) \\
& \dots + \sum_{v_d \in N_{E \setminus E'}^{out}(v_{d-1})} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_d) \Bigg). \tag{24}
\end{aligned}$$

On the other hand,

$$\begin{aligned}
g(\mathcal{E}) &= \sum_{k=1}^K \left( \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(t_k) \right. \\
& + \sum_{v_1 \in N_{E \setminus \mathcal{E}^k}^{out}(t_k)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_1) \\
& + \sum_{v_2 \in N_{E \setminus \mathcal{E}^k}^{out}(v_1)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_2) \\
& \dots + \sum_{v_d \in N_{E \setminus \mathcal{E}^k}^{out}(v_{d-1})} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_d) \Bigg). \tag{25}
\end{aligned}$$

Combining Equations (24) and (25), we have

$$\begin{aligned}
l(\mathcal{E}) &= g(\mathcal{E}) - \underline{g}(\mathcal{E}) \\
&= \sum_{k=1}^K \left( \sum_{v_1 \in N_{E' \setminus \mathcal{E}^k}^{out}(t_k)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_1) \right. \\
& + \sum_{v_2 \in N_{E' \setminus \mathcal{E}^k}^{out}(v_1)} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_2) \\
& \dots + \sum_{v_d \in N_{E' \setminus \mathcal{E}^k}^{out}(v_{d-1})} \Delta_{e_{s_k t_k}} \theta_{E \setminus \mathcal{E}^k}(v_d) \Bigg). \tag{26}
\end{aligned}$$

We have the following observations based on the above equation:

- The marginal decrement is monotone decreasing with respect to set  $E' \setminus \mathcal{E}^k$ ;
- The number of terms such as  $v_1 \in N_{E' \setminus \mathcal{E}^k}^{out}$  is also monotone decreasing with respect to  $E' \setminus \mathcal{E}^k$ .

Therefore, the function  $l(\mathcal{E})$  is submodular.  $\square$

According to Theorem 5.2,  $f(\mathcal{E})$  can be rewritten as the difference of two submodular functions, i.e.,  $f(\mathcal{E}) = \bar{f}(\mathcal{E}) - l(\mathcal{E})$  where  $\mathcal{E} \subseteq E'$ . In this setting, we can evaluate the quality of the boundaries. Iyer et al. in [10] propose algorithms for minimization of the difference between submodular functions. More specifically, given two submodular functions  $f'$  and  $g'$ , solve the optimization problem:  $\min_{X \subseteq E} [f'(X) - g'(X)]$  where  $E$  is a ground set. Their Modular-Modular procedure (MMP) is shown in Algorithm 3.

Table 2. The Statistics of Datasets

Dataset	Relationship	#Node	#Edge	Diameter
Wikipedia	Voting activity	7.1K	103.7K	7
Slashdot	Friendship	13.1K	30.9K	14
Google+	Friendship	23.6K	39.2K	8
HEP-TH	Citation	27.7K	352.8K	15

**ALGORITHM 3:** Modular-Modular Procedure (MMP)

Let  $X^0 = \emptyset$ ,  $t \leftarrow 0$ ;

**repeat**

Choose a permutation  $\gamma^t$  whose chain contains the set  $X^t$ ;

$X^{t+1} := \arg \min_X w_{X^t}^{f'}(X) - q_{X^t, \gamma^t}^{g'}(X)$ ;

$t \leftarrow t + 1$ ;

**until** converged (i.e.,  $X^{t+1} = X^t$ );

**return**  $X^t$

In order to apply MMP, we let  $f' = \bar{f}(\mathcal{E})$  and  $g' = l(\mathcal{E})$ . In Algorithm 3, the key is the iterative loop. In each iteration, we find the modular lower bound  $q_{X^t, \gamma^t}^{g'}(X)$  and modular upper bound  $w_{X^t}^{f'}(X)$  that can be done in  $O(n)$  time (see [10]), where  $n$  is the number of nodes. Furthermore, we are guaranteed to monotonically decrease the objective at every iteration and converge to a local minimal.

## 6 EXPERIMENTS

In this section, we conduct experiments on four real-world datasets to test and evaluate the performance of algorithms. We first give the description of each dataset and experiment setup. Then, we analyze and discuss the main results.

### 6.1 Datasets

Four real-world datasets represent a variety of relationships. Table 2 provides the details of these datasets.

- *Wikipedia*: This dataset is generated by a voting activity, which Wikipedia community discuss and vote for people who to become an administrator. There are 7,115 nodes and 103,689 edges. Each node in the network represents a user attend the voting procedure. Each directed edge denotes who vote for whom.
- *Slashdot*: This dataset represents the friendship of users in the Slashdot site. If one user tag the other user as “friends,” there is directed edges between them. It contains 13,182 nodes and 30,914 edges.
- *Google+*: This directed network contains Google+ user-user links. A node represents a user, and a directed edge denotes that one user has the other user in his circles. This network contains 23,628 nodes and 39,242 edges.
- *HEP-TH*: This is a citation network which from the e-print arXiv and covers all the citations within a dataset of 27,770 papers with 352,807 edges. If a paper  $v$  cites paper  $u$ , the network contains a directed edge from  $v$  to  $u$ .

## 6.2 Experiment Setup

We make the following experiment setups for our rumor propagation process: 1% of total nodes are selected randomly and uniformly as rumor seed nodes by a rate from each dataset. We generate a candidate edge set  $E'$  on each network by choosing the edges that could block the paths from seed nodes to the strong rumor spread ability nodes and the paths between the strong rumor spread ability nodes. For computing the rumor spread ability of nodes, we have described in Section 5.1. And, we adopt the IC model as the influence propagation model in all experiments. In addition, since all data sets lack propagation probability, we assign the propagation probability of each edge in the following two ways. One assigns a uniform probability  $p = 0.05$  for each edge in a network. Another assigns a probability with the trivalency model ( $p = TRI$ ). More specifically, for each edge, we uniformly select a value from  $\{0.1, 0.01, 0.001\}$  at random, which corresponds to high, medium, and low influences.

In experiments, we observe that the rumor spread will be very weak after 6 iterations. So we set the propagation hops from 1 to 6. And we update the rumor spread ability and the probability that a node is activated by the seed set after removing each edge until the number of total removed edges are  $0.8|E'|$ .

We use the basic greedy algorithm to compute the upper-bound and lower-bound. Since the bounds have been proved submodular thus we can achieve  $(1 - 1/e)$ -approximation guarantee. Here, we use the following two evaluation criteria:

- *Rumor spread value*: That is total probability of each node being activated by the seed nodes after removing  $K$  edges. The smaller the value, the better the performance of the algorithm.
- *Running time*: The running time of the algorithm indicates the scalability of the algorithm.

## 6.3 Comparative Methods

To evaluate proposed algorithm, we compare with existing popular greedy method and heuristic methods. For the greedy algorithms, we compare with the work [14, 23]. For the heuristic algorithms, we compare with the Out-degree [12], Pagerank [20] and Random. We briefly introduce these greedy and heuristic algorithms.

- *Bond Percolation Method (BPM)* [14]. Authors develop an estimation method on the basis of the BPM and greedily remove  $K$  edges from the network.
- *K edge deletion (KED)* [23]. Authors prove that deleting  $K$  edges on original graph is equivalent to removing  $K$  nodes on the line graph. And, they propose a greedy algorithm to delete  $K$  nodes on the line graph such that creates the largest decrease in terms of the leading eigenvalue of the line graph.
- *Out-degree (OD)* [12]. We adopt this method and remove  $K$  edges between the nodes with high out-degree. In other words, we select the top  $K$  edges with the sum of nodes out-degree from the candidate edge set.
- *Pagerank (PR)* [20]. The pagerank score indicates the importance of a node. There is a damping factor, and we set it to 0.9. And, we select  $K$  edges from the candidate edge set by pagerank score and remove them from the network.
- *Random (RAN)*, we randomly select  $K$  edges and remove them from the network.

## 6.4 Experiment Results

We evaluate the effectiveness of Algorithm 2 for RSM problem. First, we show the rumor spread value by choosing high rumor spread ability nodes in different propagation hops. Second, we compare the upper-bound, lower-bound, and Algorithm 2 by varying number of blocking edges and

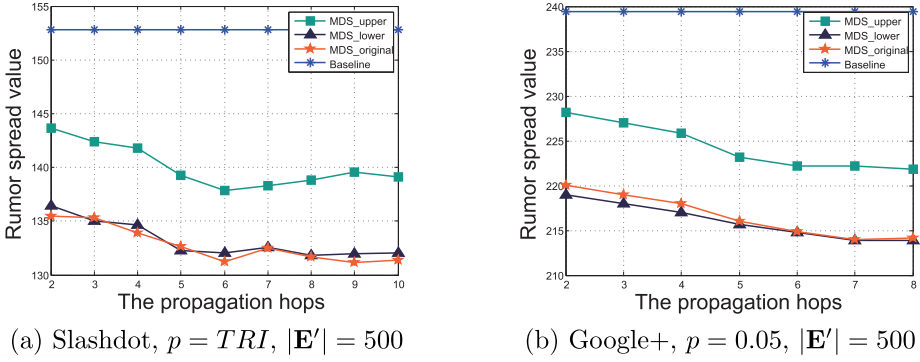


Fig. 5. Rumor spread value vs. propagation hops on Slashdot and Google+ networks.

seed nodes under two propagation probabilities:  $p = 0.05$  or  $p = TRI$ . Third, we compare with existing edge selection strategies. Finally, we show the running time of each algorithm. Notice all the experiments we run 10 times to take the average as a record to show in the figures.

**6.4.1 Rumor Spread Value vs. Varying Propagation Hops.** We show the rumor spread value by varying the hop from 2 to 10. We randomly and uniformly select 1% of total nodes as rumor seed nodes on Slashdot network and Google+ network. And, we set the propagation probability  $p = TRI$  and  $p = 0.05$  for each edge on Slashdot network and on Google+ network, respectively. In addition, we let  $|E'| = 500$  and remove  $0.8|E'|$  edges from each network. The results are shown in Figure 5. The horizontal axis holds the number of propagation hops and vertical axis holds the rumor spread value (total probability of each node being activated by seed nodes). We use Algorithm 2 to calculate MDS\_original, which is the original objective function. MDS\_upper and MDS\_lower are the upper bound and the lower bound, respectively. Baseline means that no edges are deleted from the network. We can observe as the hops increases, the rumor spread value is gradually going down until reach a certain number. Here, in both Slashdot and Google+ datasets, the rumor spread value tends to converge after 6 or 7 hops. It is easy to understand a node has large influence to the nodes within few hops but the influence is getting smaller to the nodes far away. So we set propagation hops to 6 in the following experiments since after 6 hops the influence will not have large impact.

**6.4.2 Rumor Spread Value vs. the Number of Seed Nodes.** We also simulate proposed Algorithm 2 by varying the percentage of seed nodes from 1% to 10%. We set the propagation probability  $p = 0.05$  and  $p = TRI$  for each edge on Wikipedia network and Slashdot network, respectively. And, we set  $|E'| = 500$  and remove  $0.8|E'|$  edges from networks. The results are shown in Figure 6. The horizontal axis denotes the seed nodes percentage from 1% to 10%. The vertical axis denotes the rumor spread value. From the figure, we have the following observations: (1) The rumor spread value grows with the seed nodes size increases in MDS\_original, MDS\_upper, MDS\_lower and Baseline. This phenomenon is reasonable because the more seed nodes are selected, the more the rumor spread value will be. (2) In Figure 6(a), the Wikipedia network shows the rumor spread value reduces almost half of the baseline in different percentage settings, which corresponds to Figure 7(a) that when we set seed nodes to 1% the rumor spread value also shows half of the Baseline. The same trend also happens in Slashdot datasets in Figures 6(b) and 7(b). (3) The MDS\_original, MDS\_upper, and MDS\_lower are very close in both subfigures. It indicates the high quality of bounds.

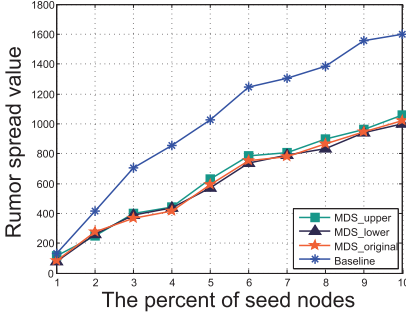
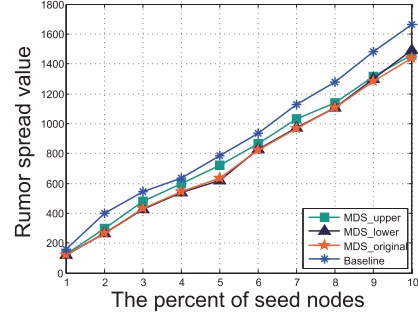
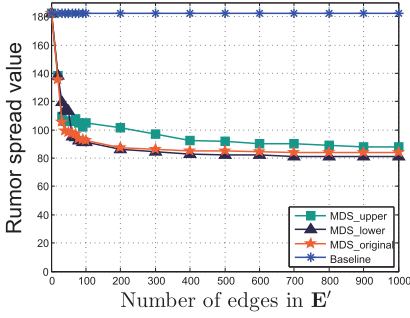
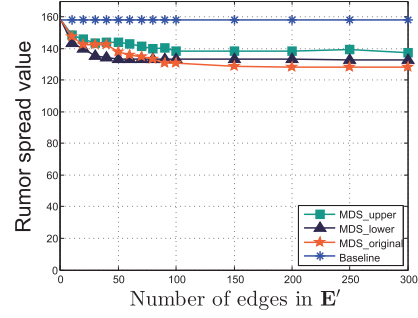
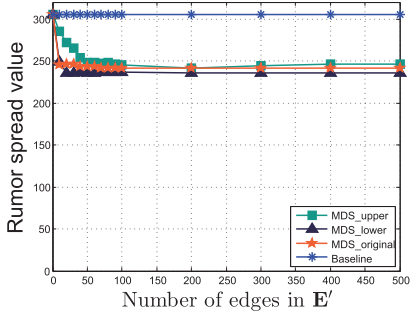
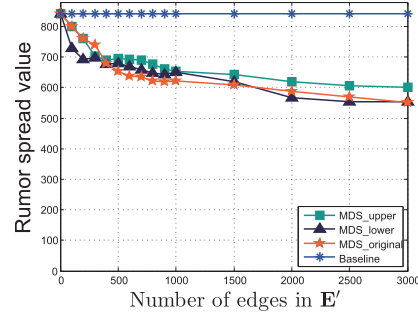
(a) Wikipedia,  $p = 0.05$ ,  $|\mathbf{E}'| = 500$ (b) Slashdot,  $p = TRI$ ,  $|\mathbf{E}'| = 500$ 

Fig. 6. Rumor spread value vs. increasing the seed nodes on Wikipedia and Slashdot networks.

(a) Wikipedia,  $p = 0.05$ , 1% seed nodes(b) Slashdot,  $p = 0.05$ , 1% seed nodes(c) Google+,  $p = TRI$ , 1% seed nodes(d) HEP-TH,  $p = TRI$ , 1% seed nodesFig. 7. Rumor spread value vs. removing number of edges in  $|\mathbf{E}'|$ .

**6.4.3 Rumor Spread Value vs. the Number of Blocking Edges.** We evaluate the rumor spread value by comparing upper-bound (MDS\_upper), lower-bound (MDS\_lower), and original objective function (MDS\_original) for RSM problem along with different number of removed edges. More specifically, we first randomly and uniformly select 1% of total nodes as rumor seed nodes on all networks. Then, we perform simulations on four networks under two propagation probabilities:  $p = 0.05$  or  $p = TRI$ , respectively. We vary the size of candidate edge set  $|\mathbf{E}'|$  from 0 to 3,000 and remove  $0.8|\mathbf{E}'|$  on each network. Finally, the results are shown in Figure 7. The horizontal axis denotes the number of edges in candidate edge set  $\mathbf{E}'$  and the vertical axis denotes the rumor spread value over the network after removing  $K = 0.8|\mathbf{E}'|$  edges. Baseline shows the rumor spread

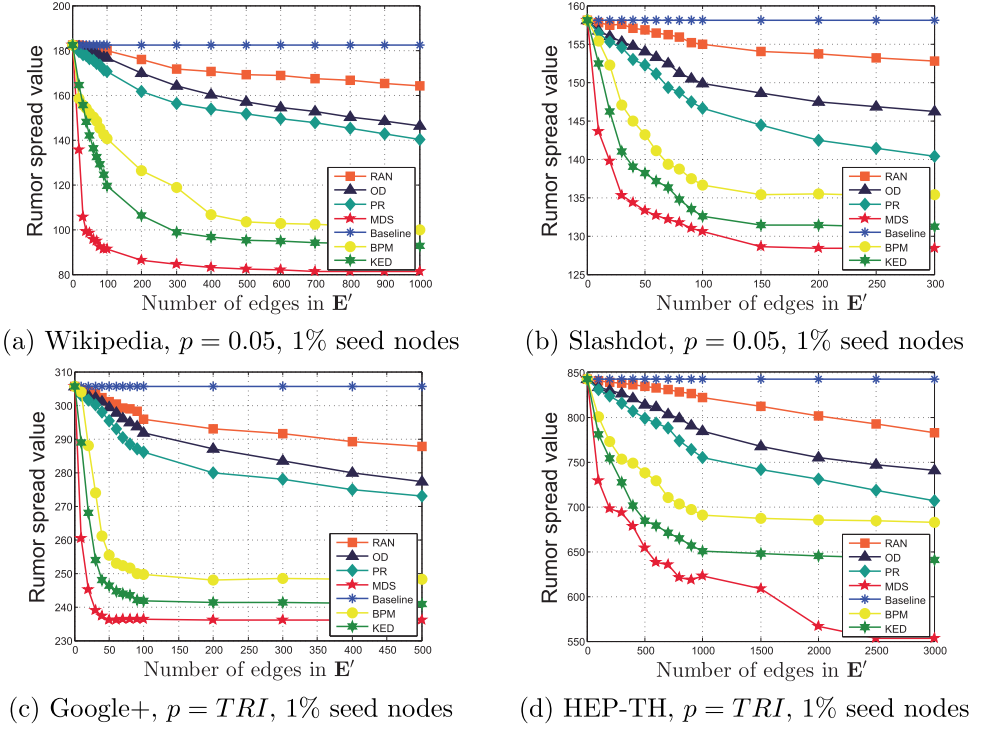


Fig. 8. The rumor spread vs. the number of edges in  $E'$  in each comparison method.

value without removing any edges. In Figure 7, the rumor spread value drops as the number of edges decreases. This indicates the edges we select have large impact on the rumor spreading over the networks. In other words, the rumor spreading can be weakened by removing these edges. Take Wikipedia network as an example, when we remove 80 edges (the total number of edges in candidate set is 100.) which is only 0.08% of total 100,000 edges in  $E$ , the rumor spread value drops almost half of the Baseline. This illustrates the effectiveness of the algorithm.

We also observe that the rumor spread value reduction shows a dramatically change at the beginning when removing a relative small number of edges. That is to say when dealing with RSM problem it just need to remove a small number of edges to achieve a good result. It is well explained although our RSM problem is not submodular theoretically, two submodular bounds could ensure the submodularity of RSM to a large extend. We also see that the difference between lower-bound and upper-bound is very close and also close to the MDS\_original, which indicates the good quality of the bounds we proposed.

**6.4.4 Comparing with Other Edge Deletion Strategies.** To fully evaluate MDS algorithm, we compare our algorithm with the existing popular greedy algorithms such as BPM [14] as well as KED [23] and heuristic algorithms such as Out-degree (OD), Pagerank (PR), as well as Random (RAN). Each method represents a strategy for removing edges.

We first randomly and uniformly select 1% of total nodes as rumor seed nodes on all networks. Then we respectively set the propagation probabilities:  $p = 0.05$  or  $p = TRI$  on networks. And, we vary the size of candidate edge set  $|E'|$  from 0 to 3,000 and remove  $0.8|E'|$  on each network. Finally, the experimental results are shown in Figure 8. The horizontal and vertical axes denote the number of edges in  $E'$  and rumor spread value, respectively.

Table 3. The Proportion of the Same Edges Appear in the Solutions of Algorithm 2 and Algorithm 3

Dataset	Wikipedia	Slashdot	Google+	HEP-TH
Proportion of same edges	15.4%	18.1%	17.2%	13.6%

Form the figure, we have the following observations: (1) In general, the rumor spread value of all methods decreases as the number of removed edges increases. (2) Greedy algorithms (BPM and KED) are better than heuristic algorithms (OD, PR, and RAN). The reason is that the greedy algorithm always reduces the rumor spread value at each iteration. However, for the heuristic algorithm, it does not have a strong effect on the rumor spread value such as RAN. Among the three heuristic methods, Pagerank performs better than the other two because high pagerank score means the nodes are more important than the others on networks. When we remove the edges between important nodes the rumor spread value will decrease. (3) The Maxdegree performs not as well as we expected. It is intuitively that high degree nodes have greater rumor spreading ability, but it is not always the case. The rumor spread ability of a node is also highly relied on the probability of being activated and the influence ranking. (4) Our proposed method outperforms the greedy algorithms since our method has an approximate guarantee based on submodular upper and lower bounds. Furthermore, the rumor spread value drops dramatically at the beginning of MDS while the other methods in a slow motion. From Figure 8(c), we can see the rumor spread decrement of MDS is three times of that of Pagerank when the removing 40 edges (candidate edge set size is 50) even though the Pagerank performs the best among the three heuristic methods. That is, to say MDS could reduce the rumor spread three times of Pagerank over the network. (5) Note that the difference between MDS and other methods grows rapidly in the beginning with a small number of removed edges. And, then the difference will not change too much. This is because two submodular bounds could ensure the submodularity of RSM to a large extent even though it is not submodular.

**6.4.5 The Difference between Algorithm 2 and Algorithm 3.** We implement the MMP (Algorithm 3) and compare with the solutions of Algorithm 2. We observe that their solutions are not exactly the same. The reason for this phenomenon is that the objective functions of these two algorithms are essentially different. In addition, we also find that some solutions appear in the results of Algorithm 2 and Algorithm 3 at the same time. This indicates these solutions (removed edges) play a key role in the dissemination of rumor. Table 3 shows the proportion of the same edges appear in the both solutions of Algorithm 2 and Algorithm 3.

**6.4.6 The Running Time of Each Methods.** We show the running time of all methods with the same parameter settings in Figure 9 on Slashdot network and HEP-TH network. In the figure, the horizontal and vertical axes denote the number of edges in  $E'$  and running time (minutes), respectively. From the figure, we have the following observations. (1) RAN method has the least running time. (2) The running time of the heuristic algorithm is less than the running time of the greedy algorithm. For example, on Slashdot network, the running time of the heuristic algorithm is in two minutes when we remove  $0.8|E'|$  edges. However, the running time of the greedy algorithm is more than 2.8 minutes. This is because heuristic algorithms only select edges based on certain heuristic criteria rather than performing Monte-Carlo simulations like greedy algorithms, which is time consuming. (3) For the greedy algorithm, our algorithm is slightly slower than BPM and KED. Although these two algorithms are slightly faster than MDS, they suffer in the quality of the total rumor spread value.

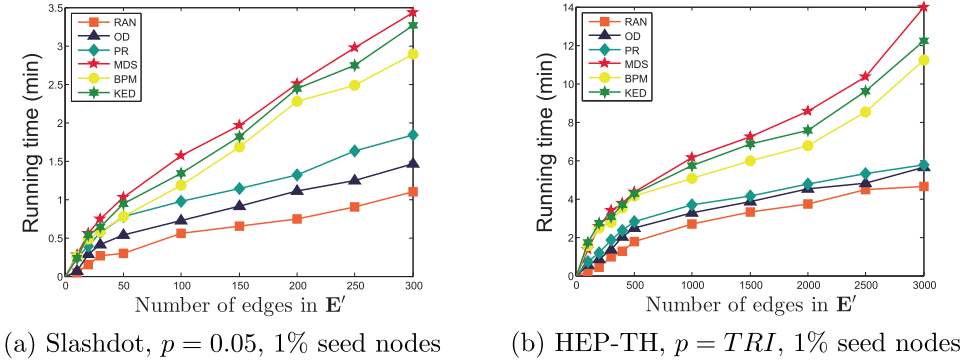


Fig. 9. Running time of each method on Slashdot network and HEP-TH network.

## 7 CONCLUSIONS

In this article, we study a RSM problem that removes an edge set from the network such that the rumor spread value (total probability of each node being activated by the seed nodes) is minimized. We first prove objective function is not submodular, which means this problem is challenging. Considering the objective function without submodularity, we propose a submodular lower-bound and an upper-bound. Furthermore, we develop a heuristic algorithm (MDS algorithm) to calculate original objective function value. We also discuss the difference between the original objective function and the upper as well as the lower bound. Then, we prove the following: (1) The original objective function minus the lower-bound function is a supermodular function. (2) The upper-bound function minus the original objective function is a submodular function. Based on the above conclusions, we observe that the original objective function can be expressed as the difference between the two submodular functions. Finally, in order to evaluate our proposed method, extensive experiments have been conducted on real-world networks. The experiment results show the good quality of upper and lower bounds and our method outperforms comparison methods with respect to the rumor spread value.

## REFERENCES

- [1] Béla Bollobás. 1984. *Graph Theory and Combinatorics: Proceedings of the Cambridge Combinatorial Conference in Honour of Paul Erdős, [Trinity College, Cambridge, 21–25 March 1983]*. Academic Press.
- [2] Ceren Budak, Divyakant Agrawal, and Amr El Abbadi. 2011. Limiting the spread of misinformation in social networks. In *Proceedings of the 20th International Conference on World Wide Web*. ACM, 665–674.
- [3] Shai Carmi, Shlomo Havlin, Scott Kirkpatrick, Yuval Shavitt, and Eran Shir. 2007. A model of internet topology using k-shell decomposition. *Proceedings of the National Academy of Sciences of the USA* 104, 27 (2007), 11150–11154.
- [4] Wei Chen, Chi Wang, and Yajun Wang. 2010. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1029–1038.
- [5] Wei Chen, Yajun Wang, and Siyu Yang. 2009. Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 199–208.
- [6] Cesar Henrique Comin and Luciano da Fontoura Costa. 2011. Identifying the starting point of a spreading process in complex networks. *Physical Review E* 84, 5 (2011), 056105.
- [7] Benjamin Doerr, Mahmoud Fouz, and Tobias Friedrich. 2012. Why rumors spread so quickly in social networks. *Communications of the ACM* 55, 6 (2012), 70–75.
- [8] Pedro Domingos and Matt Richardson. 2001. Mining the network value of customers. In *Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 57–66.
- [9] Lidian Fan, Zaixin Lu, Weili Wu, Bhavani Thuraisingham, Huan Ma, and Yuanjun Bi. 2013. Least cost rumor blocking in social networks. In *Proceedings of the 33rd International Conference on Distributed Computing Systems (ICDCS'03)*. IEEE, 540–549.

- [10] Rishabh Iyer and Jeff Bilmes. 2012. Algorithms for approximate minimization of the difference between submodular functions, with applications. In *Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*.
- [11] Rishabh Iyer, Stefanie Jegelka, and Jeff Bilmes. 2013. Fast semidifferential-based submodular function optimization. In *Proceedings of the International Conference on Machine Learning*. 855–863.
- [12] David Kempe, Jon Kleinberg, and Éva Tardos. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 137–146.
- [13] Elias Boutros Khalil, Bistra Dilkina, and Le Song. 2014. Scalable diffusion-aware optimization of network topology. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1226–1235.
- [14] Masahiro Kimura, Kazumi Saito, and Hiroshi Motoda. 2008. Minimizing the spread of contamination by blocking links in a network. In *Proceedings of the 23rd AAAI Conference of Artificial Intelligence*. AAAI, 1175–1180.
- [15] Maksim Kitsak, Lazaros K. Gallos, Shlomo Havlin, Fredrik Liljeros, Lev Muchnik, H. Eugene Stanley, and Hernán A. Makse. 2010. Identification of influential spreaders in complex networks. *Nature Physics* 6, 11 (2010), 888.
- [16] Wei Lu, Wei Chen, and Laks V. S. Lakshmanan. 2015. From competition to complementarity: Comparative influence diffusion and maximization. *Proceedings of the VLDB Endowment* 9, 2 (2015), 60–71.
- [17] Mukund Narasimhan and Jeff A. Bilmes. 2012. A submodular-supermodular procedure with applications to discriminative structure learning. arXiv:1207.1404 (2012).
- [18] M. E. J. Newman, Stephanie Forrest, and Justin Balthrop. 2002. Email networks and the spread of computer viruses. *Physical Review E* 66 (2002), 035101, 1–4.
- [19] Nam P. Nguyen, Guanhua Yan, My T. Thai, and Stephan Eidenbenz. 2012. Containment of misinformation spread in online social networks. In *Proceedings of the 4th Annual ACM Web Science Conference*. ACM, 213–222.
- [20] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1998. *The Pagerank Citation Ranking: Bringing Order to the Web*. Technical Report. Stanford InfoLab.
- [21] Guangmo Tong, Weili Wu, Ling Guo, Deying Li, Cong Liu, Bin Liu, and Ding-Zhu Du. 2017. An efficient randomized algorithm for rumor blocking in online social networks. *IEEE Transactions on Network Science and Engineering*.
- [22] Guangmo Amo Tong, Weili Wu, and Ding-Zhu Du. 2017. Distributed rumor blocking with multiple positive cascades. In *IEEE Transactions on Computational Social Systems PP* 99 (2017), 1–13.
- [23] Hanghang Tong, B. Aditya Prakash, Tina Eliassi-Rad, Michalis Faloutsos, and Christos Faloutsos. 2012. Gelling, and melting, large graphs by edge manipulation. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*. ACM, 245–254.
- [24] Senzhang Wang, Xiaojian Zhao, Yan Chen, Zhoujun Li, Kai Zhang, and Jiali Xia. 2013. Negative influence minimizing by blocking nodes in social networks. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence (Late-Breaking Developments) (AAAI'13)*. 134–136.
- [25] Ruidong Yan, Yuqing Zhu, Deying Li, and Zilong Ye. 2018. Minimum cost seed set for threshold influence problem under competitive models. *World Wide Web* (2018), 1–20.
- [26] Qipeng Yao, Chuan Zhou, Linbo Xiang, Yanan Cao, and Li Guo. 2014. Minimizing the negative influence by blocking links in social networks. In *Proceedings of the International Conference on Trustworthy Computing and Services (ISCTCS'14)*. CCIS, vol. 520, 65–73.
- [27] Arkaitz Zubiaga, Maria Liakata, Rob Procter, Geraldine Wong Sak Hoi, and Peter Tolmie. 2016. Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PLoS One* 11, 3 (2016) e0150989, 1–29. DOI: <https://doi.org/10.1371/journal.pone.0150989>

Received August 2018; revised November 2018; accepted December 2018