# RumorDecay: Rumor Dissemination Interruption for Target Recipients in Social Networks

Zhongyuan Jiang, Xianyu Chen, Jianfeng Ma, *Member, IEEE*, and Philip S. Yu, *Life Fellow, IEEE*

***Abstract*—Rumors (i.e., untrue emergence saying of COVID-19 in an area) that rapidly disseminate on the ubiquitous social media easily cause public panics and irrational behaviors (e.g., taking unnecessary medicine) of many very sensitive individuals referred to as *target recipients*. Thus, rumor controlling or blocking for these target recipients is very critical, which differs from the traditional way of protecting all individuals and remains an open and challenging problem so far. In this work, on the basis that rumors are disseminated from the given sources to target recipients via multiple paths which may be significantly interrupted by deleting a few key links referred to as *protectors*, we first mathematically define a general target information disseminating (TID) model and do theoretical proofs. Second, based on the TID model, we introduce a random walk algorithm to sample the paths of rumor dissemination for recipients. Third, aiming at deleting a budget-limited set of protectors efficiently in a large number of selected paths to reduce or weaken the negative rumor influences on the target recipients, we propose a heuristical strategy-based rumor influence decay mechanism referred to as *RumorDecay* (i.e., RumorDecay $k$ hop nearest neighbor method and RumorDecay $k$ hop random walk method in this work) which can locate the optimal protectors quickly and efficiently. Finally, we conduct extensive experiments on many real social networks and the results show that the RumorDecay strategy can significantly weaken the rumor dissemination ability with less time cost.**

***Index Terms*—Rumor blocking, rumor dissemination interruption, social network, target information dissemination.**

## I. Introduction

With the development of computer technology, more and more social media, such as Facebook and WeChat are emerging rapidly. On the one hand, it provides people with convenient communication services and shortens the information interaction time between two individuals. On the other hand, the convenient social media also accelerate the generation and dissemination of rumors (e.g., false messages or wrong statements) in social networks. Due to the high degree of freedom and activity of social network participants, the dissemination of rumors becomes more complex and difficult to control. Someone who has an ulterior motivation deliberately creates rumors to attract the attention of the general public, such as for money or interests. Some people who do not know the truth may believe the rumors that are wildly disseminated on the Internet without thinking. Furthermore, these people who are cheated by rumors forward and share the rumors in social media, which results in the rapid dissemination of rumors. In more serious cases, it will disturb the normal social stability. Therefore, the malicious dissemination of rumors on the Internet has a great negative influence and resistance on the healthy development of the whole Internet and society. The dissemination of rumor in networks may bring widespread panic or huge economic losses within a short time [1]. Hence, the reasonable modeling and analysis of rumor dissemination processes have been a long-standing area of research [2]–[4].

However, most current rumor controlling studies assume that all users are recipients in the social networks, which is a very strong assumption. As we know, many social network users are immune to the rumors, who may know the truth or do not care about rumors, and the rumors are meaningless to them. Only those ones who received the rumors and took irrational behaviors (e.g., taking unnecessary medicine for COVID-19) are the victims that are the intended recipients (namely, the target recipients) of the rumors. It is different from the traditional all-recipients rumor dissemination models. First, traditional models are a very time-consuming process of stopping rumor dissemination, because the model considers all social users to be target recipients. Meanwhile, the research goal of this article is more specific and it only focuses on the target recipients. Therefore, the research method in this article is bound to consume less time. Second, there are no available solutions for the target-recipients rumor blocking problem for which there is also a lack of mathematical problem definition, model formulation and theoretical proofs. To the best of our knowledge, the target-recipients rumor blocking issue is still challenging and open so far.

We aim to interrupt the rumor dissemination paths for the given target recipients in social networks, and the negative rumor influence on target recipients will decay with the increasing budget. First, we define a reasonable mathematical model of target influence dissemination (TID) in which the rumor information arrives at a recipient along multiple paths

Manuscript received 22 November 2021; accepted 11 January 2022. Date of publication 3 February 2022; date of current version 16 September 2022. This work was supported in part by the Fundamental Research Funds for the Central Universities of China under Grant JB211501; in part by the Natural Science Basis Research Plan in Shaanxi Province of China under Grant 2020JM-203; in part by the National Natural Science Foundation of China under Grant 61502375 and Grant 62076191; and in part by the National Science Foundation under Grant III-1763325, Grant III-1909323, Grant III-2106758, and Grant SaTC-1930941. This article was recommended by Associate Editor M. Perc. *(Corresponding author: Zhongyuan Jiang.)*

Zhongyuan Jiang, Xianyu Chen, and Jianfeng Ma are with the School of Cyber Engineering, Xidian University, Xi'an 710071, Shaanxi, China (e-mail: zyjiang@xidian.edu.cn; xychen_3@stu.xidian.edu.cn; jfma@mail.xidian.edu.cn).

Philip S. Yu is with the Department of Computer Science, University of Illinois at Chicago, Chicago, IL 60607 USA (e-mail: psyu@uic.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSMC.2022.3144141.

Digital Object Identifier 10.1109/TSMC.2022.3144141

with a probability, and we do theoretical proofs for the models. Second, considering that a link may participate in multiple rumor propagating paths, we aim to use the greedy algorithm to find the link whose deletion will interrupt the most rumor disseminating paths at the cost of high computing time. Third, to reduce the time complexity, we propose to sample the rumor paths by employing the random walks with low time cost. Fourth, based on the sampled paths, the most suitable protectors are quickly deleted. Finally, we conduct extensive experiments on many real social graphs to demonstrate the effectiveness and efficiency of our methods.

Our work studies a new rumor dissemination model, and give feasible solutions to decay the negative rumor influence to target recipients. This work makes the following contributions:

1) *Mathematical Model of Target Influence Dissemination Is Established:* Aiming at decaying the rumor influence on the target recipients, we first define the target information dissemination problem, propose a mathematical model called TID describing that rumors are originally made by multiple source users and disseminated via multiple paths to multiple destination users, and theoretically prove the model.
2) *Rumor Dissemination Interruption Strategy Is Proposed:* We propose a rumor interruption strategy based on link deletion. The strategy first introduces random walk algorithm (RW) to sample the rumor dissemination paths between sources and target recipients. Then, we propose a heuristical strategy-based rumor influence decay mechanism referred to as RumorDecay (i.e., RumorDecay $k$ hop nearest neighbor method (RumorDecay-kHNN) and RumorDecay $k$ hop random walk method (RumorDecay-kHRW) in this work) to select the protectors. Under the given conditions, deleting these protectors can significantly weaken the dissemination of rumors.
3) *Extensive Controlled Experiments and Extended Experiments:* To illustrate the effectiveness of RumorDecay for rumor dissemination interruption, we conduct extensive experiments on a set of widely used social networks. The results show that our RumorDecay method can efficiently and effectively block the rumor influence on recipients with less time cost and limited graph utility loss.

The remainder of this article is organized as follows: Section II reviews the related research on rumor controlling. Section III introduces the TID model in social networks in detail. We analyze and demonstrate the effectiveness of the rumor dissemination interruption solution RumorDecay based on link deletion in Section IV. Section V conducts a lot of controlled experiments and graphically display the experimental results. Section VI is the extended experiments. Finally, we summarize the whole article and describe the potential future work in Section VII.

## II. Related Work

It is extremely important and urgent to construct the rumor dissemination interruption mechanism for major mass incidents. How to effectively prevent the rumors or false information from disseminating in the network and life has become an important study topic for scientists. In terms of rumor interruption, there are many considerable research results. These results mainly improve the early information dissemination model. These improved dissemination models are more in line with the dissemination characteristics of rumors in social networks. Based on these dissemination models, researchers have explored many effective methods to control and stop the dissemination of rumors. The current methods of rumor controlling or rumor blocking can be summarized into the following five categories: 1) information dissemination model; 2) rumor blocking method based on maximizing the influence of nodes; 3) rumor blocking method based on resisting information dissemination; 4) rumor dissemination path control based on link disturbance; and 5) rumor control strategy based on rumor community identification.

### A. Information Dissemination Model

Budak *et al.* [6] were among the first who study the false information controlling problem. In particular, they consider the multicampaign-independent cascade model and investigate the problem of identifying a subset of individuals that needs to be convinced to adopt the good campaign so as to minimize the number of people that adopt the rumor. He *et al.* [7] and Fan *et al.* [8] further studied this problem under the competitive linear threshold model and the OPOAO model, respectively.

### B. Maximizing the Influence of Nodes

In this part of the research, the main idea is to maximize the influence of information. That is, researchers use several influential important nodes to refute rumors, so that rumors in social networks cannot disseminate in the network. Wang *et al.* [9] proposed a fake information diffusion control method Fidic. In this article, considering the sequences of users during the diffusion of fake information in social networks, the PageRank-based method ranks the users according to the links of their diffusion behaviors. Bao *et al.* [10] established a trust network based on the trust relationship between users, then they used an immunization strategy to control the rumor dissemination. Kandhway and Kuri [11] used node centrality and optimal control to maximize information diffusion in social networks.

### C. Resisting Information Dissemination

When rumors disseminate in a social network, we release positive information to some nodes that are not affected by rumors. In this way, these positive nodes are free from the interference of rumors. In addition, we use these immunized nodes to immunize their neighbors, which can eliminate the negative impact of rumors and even eliminate rumors. Wen *et al.* [1] studied a feasible method to clarify rumors by disseminating the truth. It combines truth dissemination with the method of blocking rumors by influential nodes or communities. It proves that their method is more effective than only considering a single blocking strategy. Ding *et al.* [12] discussed a hybrid control strategy, which combines the

continuous truth dissemination method and the directional rumor screening method. It solves the imbalance problem between rumor suppression and control cost minimization.

### D. Rumor Dissemination Path Control

Rumor dissemination path controlling based on link perturbation or node perturbation is one of the important methods to prevent rumor dissemination. This method aims to delete some links or nodes that play a key role in rumor dissemination [13]–[17]. Therefore, in this study, the problem of rumor blocking is transformed into the problem of link selection.

Yan *et al.* [18] studied the problem of preventing rumor dissemination by deleting links in social networks. They try to remove a set of links from the network to minimize the influence of rumor dissemination. In [19], to minimize the dissemination of rumors, two heuristic algorithms based on betweenness and outdegree are proposed to perturb a limited number of links in social network.

### E. Rumor Community Identification

Recently, the research on the identification of rumor communities is one of the hot research topic in the field of rumor blocking. The close relationship between multiple rumor dissemination members leads to the emergence of rumor communities. How to identify these rumor communities and control the dissemination of them is particularly important.

Ping *et al.* pointed out that the presence of witch users can significantly reduce the rumor blocking effect by 30% and the existing rumor blocking methods fail to consider the influence of witch attacks on rumor blocking. Based on this, a rumor blocking framework (SLCRB) with minimum cost based on witch perception is proposed [20]. Zheng and Li [21] proposed a method for finding rumor communities, first, he identified the minimal subset of rumor nodes. Then, he deleted all nodes in the subset and their corresponding out degree edge and in degree edge from the network. Zhang *et al.* [22] proposed a rumor masking method (RBMTU) based on user tolerance degree. Wang *et al.* [23] proposed that the existing rumor blocking methods are either continuous blocking or permanent blocking. However, long time blocking may lead to complaints and dissatisfaction of users, or even withdrawal from social activities. Therefore, the author studied the feasibility of temporary rumor blocking in rumor control problem. Wu *et al.* [24] proposed a dynamic rumor influence minimization algorithm based on the group structure. Fang *et al.* [25] proposed a general rumor blocking method based on the nonpredictive rumor seed set.

## III. Target Information Disseminating in Social Networks

In this section, before studying the dissemination of rumors, we first study the TID problem in which a set of nodes is the target recipients. We focus on the information dissemination dynamics on information controlling and try to find a model that expresses rumor dissemination and rumor controlling.

Here, we first give the general representation symbol in the following research. Given a network $G = (V, E)$, where $V$ is the node set and $E$ is the link set. The number of nodes in $G$ is $N = |V|$ and the number of links in $G$ is $M = |E|$. A set of target recipients is denoted by $T$ and a set of information source nodes is denoted by $S$. The number of nodes in $S$ and $T$ are denoted by $n$ and $m$, respectively. All possible target information flow paths between $S$ and $T$ are defined as $R_{ST}$. For any source node $s$ ($s \in S$) and any target node $t$ ($t \in T$), the dissemination path between $s$ and $t$ is denoted by $R_{st}$. That is, $R_{ST} = \bigcup R_{s_i t_j}$, where $i \in [1, n]$ and $j \in [1, m]$. A path is denoted by $e = \{s \equiv v_0, v_1, \ldots, v_k \equiv t\}$, where $k$ is the path length (i.e., the hop number).

### A. Single-Source-Single-Target Problem

Suppose that there is only one source node $s$ and one target node $t$ in the network. The target information can be disseminated from $s$ to $t$ by any path in $R_{st}$. The probability that the target node $t$ is the final recipient of the information from $s$ is denoted by (1), where $p^e_{st}$ is the probability that the information flows from $s$ to $t$ through path $e$. We define (2) as the probability that $t$ is not the recipient of information from $s$ via all paths in $R_{st}$

$$H_{st} = 1 - \prod_{e \in R_{st}} \left(1 - p^e_{st}\right) \tag{1}$$

$$F_{st} = \prod_{e \in R_{st}} \left(1 - p^e_{st}\right). \tag{2}$$

*Problem 1 [Target Information Dissemination Minimization for Single Source Node Single Target Node (SS-MinTID)]:* The goal of SS-MinTID is to minimize the objective function $H_{st}$ by disabling or deleting a budget-limited number of links. The disabled or deleted links are denoted by set $E^-$ and the corresponding objective function is $H^{\#}_{st}(E^-)$. Equally, the goal is to maximize the function $F^{\#}_{st}(E^-) = 1 - H^{\#}_{st}(E^-)$.

*Theorem 1:* The SS-MinTID problem is monotone. Given two link sets $E_1$ and $E_2$, where $E_1 \subset E_2 \subset E$. The conclusion is that $F^{\#}_{st}(E_1) \leq F^{\#}_{st}(E_2)$ always holds.

*Proof:* Suppose that $E_2 = E_1 \cup x$, where $x \in E$. There are two cases for the link $x$: 1) $x$ is not in any path of the set $R_{st}$, so we have $F^{\#}_{st}(E_1) = F^{\#}_{st}(E_2)$ and 2) at least one path in $R_{st}$ contains link $x$. When link $x$ is deleted, the disappearing path set in $R_{st}$ is defined as $R_x$, where $R_x \subset R_{st}$. Thus, $F^{\#}_{st}(E_2)$ can be replaced by (3). Because it satisfies $0 < \prod_{e \in R_x}(1 - p^e_{st}) < 1$, so inequality $F^{\#}_{st}(E_1) < F^{\#}_{st}(E_2)$ always holds

$$F^{\#}_{st}(E_2) = \frac{F^{\#}_{st}(E_1)}{\prod_{e \in R_x}\left(1 - p^e_{st}\right)}. \tag{3}$$

■

*Theorem 2:* The SS-MinTID problem is not submodular. Given two link sets $E_1 \subset E_2 \subset E$ and a link $x$, where $x \in E \backslash E_2$. We have $F^{\#}_{st}(E_1 \cup \{x\}) - F^{\#}_{st}(E_1) < F^{\#}_{st}(E_2 \cup \{x\}) - F^{\#}_{st}(E_2)$.

*Proof:* As shown in Fig. 1, there are three paths with different colors between $s$ and $t$. That is, $R_{st} = \{e_1, e_2, e_3\}$. Let $E_1 = \emptyset$, $E_2 = \{(v_5, v_6)\}$ and $x = (v_1, v_2)$. The result of $F^{\#}_{st}(E_1 \cup \{x\}) - F^{\#}_{st}(E_1)$ is shown in (4) and the result of $F^{\#}_{st}(E_2 \cup \{x\}) - F^{\#}_{st}(E_2)$ is shown in (5). Therefore, $F^{\#}_{st}(E_1 \cup \{x\}) - F^{\#}_{st}(E_1) = \{F^{\#}_{st}(E_2 \cup \{x\}) - F^{\#}_{st}(E_2)\} * \{1 - p^{e_3}_{st}\}$.

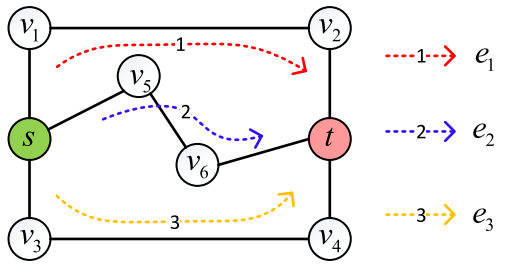



Fig. 1. Case to illustrate the SS-MinTID problem is not submodular.

So, Theorem 2 always holds because of $0 < (1 - p_{st}^{e_3}) < 1$

$$F_{st}^{\#}(E_1 \cup \{x\}) - F_{st}^{\#}(E_1) = \left\{\left(1 - p_{st}^{e_2}\right) * \left(1 - p_{st}^{e_3}\right)\right\} - \left\{\left(1 - p_{st}^{e_1}\right) * \left(1 - p_{st}^{e_2}\right) * \left(1 - p_{st}^{e_3}\right)\right\} = p_{st}^{e_1} * \left(1 - p_{st}^{e_2}\right) * \left(1 - p_{st}^{e_3}\right) \quad (4)$$

$$F_{st}^{\#}(E_2 \cup \{x\}) - F_{st}^{\#}(E_2) = \left\{\left(1 - p_{st}^{e_2}\right)\right\} - \left\{\left(1 - p_{st}^{e_1}\right) * \left(1 - p_{st}^{e_2}\right)\right\} = p_{st}^{e_1} * \left(1 - p_{st}^{e_2}\right). \quad (5)$$

■

*Greedy Solution for the SS-MinTID:* For information flow blocking problem, every time, we select a link to be deleted that maximally increases $F_{st}^{\#}$. Then, we can calculate the gain $G_x^{\#}$ of the function $F_{st}^{\#}$ when deleting a link $x$. We assume the link set $E^-$ has been deleted and the disappearing paths in $R_{st}$ is denoted by $R_{E^-}$. Equation (6) gives the definition of $G_x^{\#}$

$$\begin{aligned} G_x^{\#} &= F_{st}^{\#}\left(E^- \cup \{x\}\right) - F_{st}^{\#}\left(E^-\right) \\ &= \frac{\prod_{e \in R_{st}}\left(1 - p_{st}^e\right)}{\prod_{e \in \left\{R_{\left(E^- \cup \{x\}\right)} \cap R_{st}\right\}}\left(1 - p_{st}^e\right)} - \frac{\prod_{e \in R_{st}}\left(1 - p_{st}^e\right)}{\prod_{e \in \{R_{E^-} \cap R_{st}\}}\left(1 - p_{st}^e\right)} \\ &= \frac{F_{st}}{\prod_{e \in \{R_{E^-} \cap R_{st}\}}\left(1 - p_{st}^e\right)} \\ &\quad \times \left[\frac{1}{\prod_{e \in \{R_x \cap (R_{st} \setminus (R_{st} \cap R_{E^-}))\}}\left(1 - p_{st}^e\right)} - 1\right]. \end{aligned} \quad (6)$$

For any link $x$ to be deleted, the parts $F_{st}$ and $\prod_{e \in \{R_{E^-} \cap R_{st}\}}(1 - p_{st}^e)$ are fixed. As shown in (7), in order to get the maximum gain, we need to find a link $x$ when $G_x^{\#}$ is maximum and delete it from network $G$

$$\max \quad \left(G_x^{\#}\right) \sim \min\left\{\prod_{e \in \{R_x \cap (R_{st} \setminus (R_{st} \cap R_{E^-}))\}}\left(1 - p_{st}^e\right)\right\}. \quad (7)$$

### B. Multisources–Multitargets Problem

In this section, we will study the problem of multisource nodes and multitarget recipients for target information flow in networks. The source nodes of information are denoted by $S = \{s_1, s_2, \ldots, s_n\}$, where $n$ is the number of source nodes. Similarly, the target recipients of information are denoted by $T = \{t_1, t_2, \ldots, t_m\}$, where $m$ is the number of target recipients. All paths of information flow between $S$ and $T$ are defined as $R_{ST}$, where $R_{ST} = R_{St_1} \cup R_{St_2} \cup \cdots \cup R_{St_m} = R_{s_1t_1} \cup R_{s_2t_1} \cup \cdots \cup R_{s_nt_m}$. Similar to the (single-source–single-target) SS problem, we first give two concepts: $F_{ST}$ and $H_{ST}$, which are shown in (8) and (9), respectively. $F_{ST}$ is the probability that $T$ is not the recipients of target information from $S$ via all paths in $R_{ST}$. $H_{ST}$ is the probability that $T$ are the final recipients of the target information from $S$ via all paths in $R_{ST}$

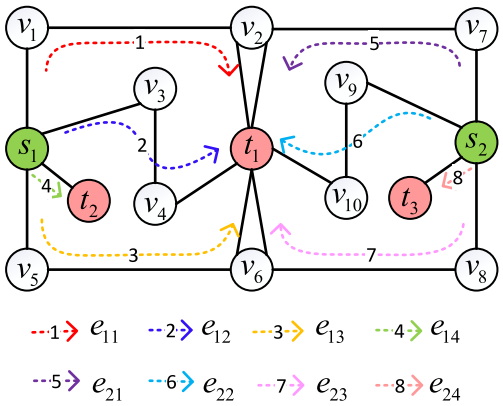



Fig. 2. Case to illustrate the MM-MinTID problem is not submodular.

$$F_{ST} = \prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e \in R_{s_it_j}}\left(1 - p_{s_it_j}^e\right) = F_{s_1t_1} * \cdots * F_{s_nt_m} \quad (8)$$

$$H_{ST} = 1 - \prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e \in R_{s_it_j}}\left(1 - p_{s_it_j}^e\right) = 1 - F_{ST}. \quad (9)$$

*Problem 2 [Target Information Dissemination Minimization for Multisource Nodes Multitarget Nodes (MM-MinTID)]:* The goal of MM-MinTID is to minimize the objective function $H_{ST}$ by disabling or deleting a budget-limited number of links. The disabled or deleted links are denoted by set $E^-$ and the corresponding objective function is $H_{ST}^{\#}(E^-)$. Equally, the goal is to maximize the function $F_{ST}^{\#}(E^-) = 1 - H_{ST}^{\#}(E^-)$.

*Theorem 3:* The MM-MinTID problem is monotone. Given two link sets $E_1$ and $E_2$, where $E_1 \subset E_2 \subset E$. Then, $F_{ST}^{\#}(E_1) \leq F_{ST}^{\#}(E_2)$ always holds.

*Proof:* Suppose that $E_2 = E_1 \cup x$, where $x \in E$. There are two cases for the link $x$: 1) $x$ is not in any path of the $R_{ST}$, so we have $F_{ST}^{\#}(E_1) = F_{ST}^{\#}(E_2)$ and 2) at least one path in $R_{ST}$ contains link $x$. When link $x$ is deleted, the disappearing path set in $R_{ST}$ is defined as $R_x$, where $R_x \subset R_{ST}$. Thus, $F_{ST}^{\#}(E_2)$ can be replaced by (10). Because it satisfies $0 < \prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e \in \{R_x \cap R_{s_it_j}\}}(1 - p_{s_it_j}^e) < 1$, so inequality $F_{ST}^{\#}(E_1) < F_{ST}^{\#}(E_2)$ always holds. That is to say, the MM-MinTID problem is monotone

$$\begin{aligned} F_{ST}^{\#}(E_2) &= \frac{F_{ST}^{\#}(E_1)}{\prod_{e \in R_x}\left(1 - p_{ST}^e\right)} \\ &= \frac{F_{ST}^{\#}(E_1)}{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e \in \left\{R_x \cap R_{s_it_j}\right\}}\left(1 - p_{s_it_j}^e\right)}. \end{aligned} \quad (10)$$

■

*Theorem 4:* The MM-MinTID problem is not submodular. Given two link sets $E_1 \subset E_2 \subset E$ and a link $x$, where $x \in E \setminus E_2$. Then, we have $F_{ST}^{\#}(E_1 \cup \{x\}) - F_{ST}^{\#}(E_1) < F_{ST}^{\#}(E_2 \cup \{x\}) - F_{ST}^{\#}(E_2)$.

*Proof:* Fig. 2 is the case to illustrate the MM-TID problem is not submodular. According to this figure, there are eight information flow paths with different colors between

$S = \{s_1, s_2\}$ and $T = \{t_1, t_2, t_3\}$. There is no path between $s_1$ and $t_3$. Similarly, there is no path between $s_2$ and $t_2$. $R_{ST} = \{e_{11}, e_{12}, e_{13}, e_{14}, e_{21}, e_{22}, e_{23}, e_{24}\}$. Suppose that there are two link sets $E_1 = \emptyset$ and $E_2 = \{(v_9, v_{10})\}$. A link $x = (v_1, v_2)$. Equation (11) gives the result of $F^{\#}_{ST}(E_1 \cup \{x\}) - F^{\#}_{ST}(E_1)$ and the result of $F^{\#}_{ST}(E_2 \cup \{x\}) - F^{\#}_{ST}(E_2)$ is shown in (12). Theorem 6 has been proved because of $0 < (1 - p^{e_{22}}_{s_2t}) < 1$

$$\begin{aligned} &F^{\#}_{ST}(E_1 \cup \{x\}) - F^{\#}_{ST}(E_1) \\ &= \frac{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in R_{s_it_j}}\left(1 - p^{e}_{s_it_j}\right)}{1 - p^{e_{11}}_{s_1t_1}} \\ &\quad - \prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in R_{s_it_j}}\left(1 - p^{e}_{s_it_j}\right) \\ &= \frac{F_{ST}}{1 - p^{e_{11}}_{s_1t_1}} - F_{ST} \\ &= \frac{p^{e_{11}}_{s_1t_1} * F_{s_1t_1} * F_{s_1t_2} * F_{s_2t_1} * F_{s_2t_3}}{1 - p^{e_{11}}_{s_1t_1}} \end{aligned} \tag{11}$$

$$\begin{aligned} &F^{\#}_{ST}(E_2 \cup \{x\}) - F^{\#}_{ST}(E_2) \\ &= \frac{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in R_{s_it_j}}\left(1 - p^{e}_{s_it_j}\right)}{(1 - p^{e_{11}}_{s_1t_1}) * (1 - p^{e_{22}}_{s_2t_1})} \\ &\quad - \frac{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in R_{s_it_j}}\left(1 - p^{e}_{s_it_j}\right)}{1 - p^{e_{22}}_{s_2t_1}} \\ &= \frac{F_{ST}}{(1 - p^{e_{11}}_{s_1t_1}) * (1 - p^{e_{22}}_{s_2t_1})} - \frac{F_{ST}}{1 - p^{e_{22}}_{s_2t_1}} \\ &= \frac{F^{\#}_{ST}(E_1 \cup \{x\}) - F^{\#}_{ST}(E_1)}{1 - p^{e_{22}}_{s_2t_1}}. \end{aligned} \tag{12}$$

■

*Greedy Solution for the MM-MinTID:* For the information flow blocking problem, every time, we select a link to be deleted that maximally increase $F^{\#}_{SF}$. Then, we can calculate the gain $G^{\#}_x$ of the function $F^{\#}_{ST}$ when deleting a link $x$. Suppose that the link set $E^-$ has been deleted and the disappearing paths in $R_{ST}$ is denoted by $R_{E^-}$. Equation (13) gives the definition of $G^{\#}_x$

$$\begin{aligned} G^{\#}_x &= F^{\#}_{ST}\left(E^- \cup \{x\}\right) - F^{\#}_{ST}\left(E^-\right) \\ &= \frac{F_{ST}}{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_{\left(E^-\cup\{x\}\right)}\cap R_{s_it_j}\right\}}\left(1 - p^{e}_{s_it_j}\right)} \\ &\quad - \frac{F_{ST}}{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_{E^-}\cap R_{s_it_j}\right\}}\left(1 - p^{e}_{s_it_j}\right)} \\ &= \frac{F_{ST}}{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_{E^-}\cap R_{s_it_j}\right\}}\left(1 - p^{e}_{s_it_j}\right)} \\ &\quad \times \frac{1 - \prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_x\cap\left(R_{s_it_j}\setminus\left(R_{s_it_j}\cap R_{E^-}\right)\right)\right\}}\left(1 - p^{e}_{s_it_j}\right)}{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_x\cap\left(R_{s_it_j}\setminus\left(R_{s_it_j}\cap R_{E^-}\right)\right)\right\}}\left(1 - p^{e}_{s_it_j}\right)}. \end{aligned} \tag{13}$$

To get the maximum gain, we need to find a link $x$ when $G^{\#}_x$ is maximum and delete it in network $G$. The specific solution is shown in

$$\max\ \left(G^{\#}_x\right) \sim \min\left[\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_x\cap\left(R_{s_it_j}\setminus\left(R_{s_it_j}\cap R_{E^-}\right)\right)\right\}}\left(1 - p^{e}_{s_it_j}\right)\right]. \tag{14}$$

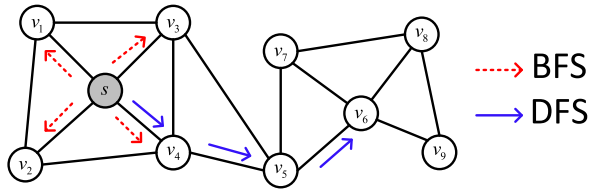



Fig. 3. Schematic of DFS and BFS algorithms.

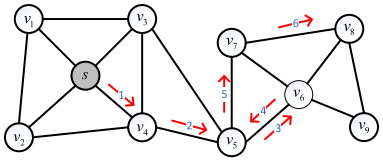


Fig. 4. Schematic of RW algorithm.

## IV. Rumor Dissemination Interruption

### A. Path Sampling

According to the TID model, first, we need to compute all the rumor dissemination paths $R_{ST}$ between the source nodes $S$ of rumor producers and the target recipients $T$. Hence, the solution to get all the paths $R_{ST}$ is the first problem that needs to be solved. Enumeration is one of the important methods to solve the problem of computation of set $R_{ST}$. We can enumerate all the paths between $S$ and $T$ and define the enumeration paths as $R^1_{ST}$. Because the core idea of the enumeration method is to list all the cases that meet the given conditions. Hence, set $R^1_{ST}$ contains all the rumor dissemination paths, that is, $R^1_{ST} = R_{ST}$. The main implementation methods of the enumeration method include depth-first search algorithm (DFS) and breadth-first search algorithm (BFS). Fig. 3 gives the search path of DFS and BFS algorithms. According to the figure, BFS focuses on local information and DFS focuses on global information. We will apply the DFS algorithm to compute the path set $R^1_{ST}$. The enumeration method seems to be a good solution for small-scale networks and we can even observe the paths in some small social networks with the naked eye.

However, with the increase of network scale, enumerating all the paths in social networks becomes a very time-consuming task. To speed up the path computing process of the greedy solution, we introduce the RW algorithm to sample the paths. RW is an irregular and completely random movement. As shown in Fig. 4, we cannot control the direction of the random walk. Particularly, there may be loops in some paths, such as steps 3 and 4 in Fig. 4. The only parameter that we can set is the frequency ($\tau$) of RW, such as $\tau = 1 \times 10^4$, $\tau = 2 \times 10^5$, or even bigger. Here, we apply the RW algorithm to sample the paths between $S$ and $T$ and define the sampling paths as $R^2_{ST}$. Since the RW algorithm has a random

**Algorithm 1** Path Sampling Algorithm: *DFS*

**Require:** $G = (V, E), s, t, r, R^1_{ST} = \{\}, e = [], l$
**Ensure:** $R^1_{ST}, F^{\#}_{ST}$
1: $e = e \longrightarrow s$;
2: **if** $s == t$ & $getHopNum(e) <= l$
3: $R^1_{ST} = R^1_{ST} \cup e$;
4: **else**
5: **if** $getHopNum(e) <= l$
6: $U = getNeighborNodes(G, s)$;
7: **for each** $u \in U$
8: **if** $u \notin e$
9: $R^1_{ST} = DFS(G, u, t, r, R^1_{ST}, e, l)$;
10: **end if**
11: **end for**
12: **end if**
13: **end if**
14: $F^{\#}_{ST} = getFSet(R^1_{ST}, r)$;

**Algorithm 2** Path Sampling Algorithm: *RW*

**Require:** $G = (V, E), s, t, r, l, \tau$
**Ensure:** $R^2_{ST}, F^{\#}_{ST}$
1: **do**
2: $cur = s; e = []; e = e \longrightarrow cur; R^2_{ST} = \{\}$;
3: **do**
4: $U = getNeighborNodes(G, cur)$;
5: $flag = 0$;
6: **if** $U \neq \varnothing$
7: $cur = getARandomNeighborNode(U)$;
8: $e = e \longrightarrow cur; flag = cur$;
9: **end if**
10: **if** $flag == 0$
11: *break*;
12: **else if** $cur == t$
13: $R^2_{ST} = R^2_{ST} \cup e$;
14: **end if**
15: $l = l - 1$;
16: **while** $l > 0$
17: $\tau = \tau - 1$;
18: **while** $\tau > 0$
19: $R^2_{ST} = getUniquePaths(R^2_{ST})$;
20: $F^{\#}_{ST} = getFSet(R^2_{ST}, r)$;

nature, so we have $R^2_{ST} \subset R_{ST}$. Although $R^2_{ST}$ is not as accurate as $R^1_{ST}$, it is worthy for us to lose a little bit of accuracy in exchange for the reduction of the time consumed for path sampling.

In our experiments, suppose that the rumor disseminates throughout the network and the dissemination rate (i.e., the dissemination probability of rumor between node pairs) of rumor in $G$ is $r$. Besides, we hold that the longest length (i.e., the hop number) of rumor dissemination path cannot exceed a certain threshold $l$. In our experiments, $l$ equals the diameter of $G$. For a feasible path $e = \{s \equiv v_0, v_1, \ldots, v_k \equiv t\}$ between a source node $s$ ($s \in S$) and a target recipient $t$ ($t \in T$), where $k$ is the hop number of path $e$ and $k \leq l$. Hence, when only $e$ is considered, we define the rumor acceptance probability of $t$ as $p^e_{st} = r^k$. Now, the purpose of path sampling is to find all qualified rumor dissemination paths $R_{ST}$ between $S$ and $T$. Meanwhile, we visit each path $e$ in $R_{ST}$ in turn and get all the links that exist in $e$. This set of links is denoted by $E_{ST}$. For each link $(u, v) \in E_{ST}$, we can get $F^{\#}_{ST}(\{(u, v)\})$ based on the knowledge of the previous section. Then, we get a set $F^{\#}_{ST} = \cup F^{\#}_{ST}(\{x_i\})$, where link $x_i \in E_{ST}$. Next, we will introduce the DFS and RW methods in detail, respectively.

*1) Depth-First Search:* DFS is a basic enumeration process, which traverses the whole graph in the form of depth first. Algorithm 1 gives the core steps of the DFS algorithm used in path sampling. Line 9 of Algorithm 1 is a recursive process. In the worst case, the time complexity of Algorithm 1 is $O(N^2)$.

*2) Random Walk:* Random walk is a statistical model, which is composed of a series of tracks. For the path sampling of rumor dissemination, random walk is a pretty good solution. Algorithm 2 gives the specific steps of random walk and the time complexity of the RW algorithm is $O(N)$.

### B. Important Links Detection

To interrupt the dissemination of rumors in $G$, we can destroy the dissemination path of rumors by deleting some of the most critical social links, so as to prevent the effective dissemination of rumors. Of course, it is not advisable to delete all links in $E_{ST}$, because it will cause serious damage to the utility of $G$. The operation we can accept is to delete a certain number of critical links and we define the number as $\beta$.

To interrupt the dissemination of rumor, our main work is to delete the $\beta$ most critical links in $G$ and these critical links are part of $E_{ST}$. So, how to find the most suitable $\beta$ links in a large number of links of $E_{ST}$? First, we find a small number of critical links between $S$ and $T$ that have information flow and we denote these links as $E'$, where $E' \subseteq E$. Here, we introduce two methods to compute $E'$: 1) RumorDecay-kHNN and 2) RumorDecay-kHRW. Then, we introduce a concept of link weight. For a link $(u, v) \in E'$, the weight of it equals $k_u \times k_v$. We denoted it by $W'_{(u,v)} = k_u \times k_v$, where $k_u$ is the degree of node $u$. So, we can calculate the link weight of each link in $E'$ and get weight set $W'$. For convenience, we normalize $W'$, and $W'_{(u,v)} = W'_{(u,v)} / \sum W'$. Next, we will explain the RumorDecay-kHNN and RumorDecay-kHRW in detail, respectively.

*1) RumorDecay k Hop Nearest Neighbor:* RumorDecay-kHNN method is a link selection algorithm and can select some most important links for rumor dissemination in $G$. These selected links form a bridge of information dissemination between $S$ and $T$ in $G$. The main idea of RumorDecay-kHNN algorithm is as follows: given source node $s$ and target recipient $t$ in $G$ and the longest length (i.e., the hop number) $l$ of rumor dissemination path. First, we find 1 to $k = \lceil l/2 \rceil$ hop neighbors of $s$ and define it as *s_kneigh*. Similarly, we find 1 to $k = \lceil l/2 \rceil$ hop neighbors (*t_kneigh*) of $t$. *s_kneigh* and *t_kneigh* will produce an intersection $V'$, where $V' \subseteq V$. Extracting the node set $V'$ in $G$ and the links corresponding to $V'$, we get a subgraph $G'$. Finally, based on $G'$, the link set $E'$ and the weight set $W'$ can be easily calculated. The specific

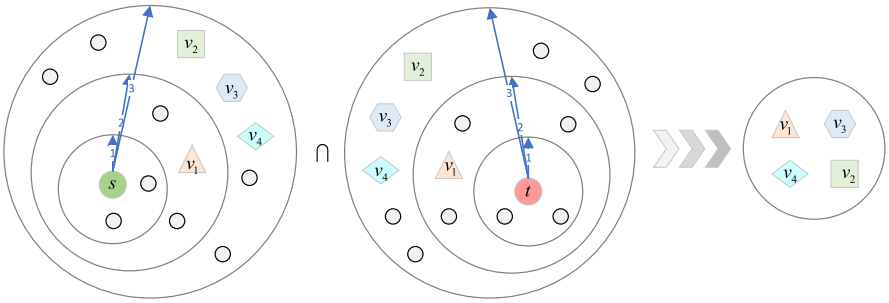



Fig. 5. Schematic of the RumorDecay-kHNN algorithm.

**Algorithm 3** *RumorDecay-kHNN*

**Require:** $G = (V, E), s, t, l$
**Ensure:** $E', W'$

$k = \lceil l/2 \rceil$;
$s\_kneigh = getkHopNearestNeigh(G, s, k)$;
$t\_kneigh = getkHopNearestNeigh(G, t, k)$;
$V' = getIntersection(s\_kneigh, t\_kneigh)$;
$G' = getSubGraph(G, V')$;
$(E', W') = getLinkWeight(G')$;

**Algorithm 4** *RumorDecay-kHRW*

**Require:** $G = (V, E), s, t, l, \tau$
**Ensure:** $E', W'$

$k = \lceil l/2 \rceil$;
$s\_knodes = getkHopRandWalk(G, s, k, \tau)$;
$t\_knodes = getkHopRandWalk(G, t, k, \tau)$;
$V' = getIntersection(s\_knodes, t\_knodes)$;
$G' = getSubGraph(G, V')$;
$(E', W') = getLinkWeight(G')$;

steps of RumorDecay-kHNN are shown in Algorithm 3. The time complexity of RumorDecay-kHNN is $O(N)$.

Fig. 5 is a case to illustrate the idea of the RumorDecay-kHNN algorithm. In this case, $k = 3$. There are 12 nearest neighbors (*s_kneigh*) of $s$ and 13 nearest neighbors (*t_kneigh*) of $t$. Set $\{v_1, v_2, v_3, v_4\}$ exists in both *s_kneigh* and *t_kneigh*. Hence, $V' = \{v_1, v_2, v_3, v_4\}$.

*2) RumorDecay k Hop Random Walk:* RumorDecay-kHRW is also a link selection algorithm and can select some most important links for rumor dissemination in $G$. RumorDecay-kHRW is similar to RumorDecay-kHNN, the only difference is the method to find 1 to $k$ hop neighbors. RumorDecay-kHRW finds the 1 to $k$ hop neighbors of node $s$ by random walk. So, given two nodes $s$ and $t$ in $G$ and the longest path length (i.e., the hop number) $l$. First, based on random walk, we find the 1 to $k = \lceil l/2 \rceil$ hop neighbors (*s_knodes*) of node $s$. Similarly, we find the 1 to $k = \lceil l/2 \rceil$ hop neighbors (*t_knodes*) of target recipient $t$. For *s_knodes* and *t_knodes*, the subsequent processes are consistent with method RumorDecay-kHNN. The specific steps of RumorDecay-kHRW are shown in Algorithm 4. The time complexity of RumorDecay-kHRW is $O(N)$.

**Algorithm 5** Rumor Dissemination Interruption: *RumorDecay*

**Require:** $G = (V, E), S, T, r, \beta, l, \tau$
**Ensure:** $G^-$

$(R^2_{ST}, F^{\#}_{ST}) = getPaths(G, S, T, r, l, \tau)$; /*Algorithm 2*/
$(E', W') = getWeight(G, S, T, l, \tau)$; /*Algorithm 3 or 4*/
**for** each $(u, v) \in E'$ **do**
  $I_{(u,v)} = getLinkImportance(F^{\#}_{ST}, W')$;
**end for**
$I^{sorted} = sortByLinkImportance(I)$;
$E^- = getDeletedLinks(I^{sorted}, \beta)$;
$G^- = getUpdatedNetwork(G, E^-)$;

*3) Rumor Dissemination Interruption:* Before explaining the rumor dissemination interruption method, this section first introduces the concept of link importance for rumor dissemination. $F^{\#}_{ST}$ and $W'$ are two different indicators to evaluate the influence of links on rumor dissemination, we use these two evaluation methods synthetically and propose a calculation method to evaluate the link importance for rumor dissemination. Thus, based on $F^{\#}_{ST}$ and $W'$, an evaluation index called $I_{(u,v)}$ to measure the importance of link $(u, v)$ in $E'$ is proposed. The calculation process of $I_{(u,v)}$ is shown in (15), where $(u, v) \in E'$. The value range of $I_{(u,v)}$ is [0, 1]. The larger $I$ of a link, the greater the effect of this link on the dissemination of rumor. Therefore, we can effectively interrupt the dissemination of rumor by deleting several links with larger $I$ within the given threshold range

$$I_{(u,v)} = \frac{F^{\#}_{ST}(\{(u,v)\}) + W'_{(u,v)} - \min\left(F^{\#}_{ST}\right) - \min\left(W'\right)}{\max\left(F^{\#}_{ST}\right) + \max(W') - \min\left(F^{\#}_{ST}\right) - \min(W')}. \quad (15)$$

In summary, we put forward a strategy RumorDecay based on link deletion in social networks to weaken the rumor dissemination. Algorithm 5 gives the specific steps of RumorDecay and the time complexity is $O(N)$. Specifically, given a social network $G = (V, E)$ and the diameter of $G$ is $l$. There are some nodes $S$ of rumor producers and some target recipients $T$ in $G$. Suppose that the probability of rumor dissemination between any two nodes is $r$. Based on these known conditions, RumorDecay aims to delete the optimal $\beta$ social links in order to weaken the dissemination of rumor. First, we use Algorithm 2 to sample the qualified paths $R^2_{ST}$ between $S$ and $T$ and the corresponding set $F^{\#}_{ST}$. Then, according to Algorithms 3 or 4, we can obtain a smaller link set

TABLE I
SOCIAL NETWORKS CONSIDERED IN THIS WORK

| Net | $N$ | $M$ | $l$ | Description |
|---|---|---|---|---|
| kar | 34 | 78 | 5 | Zachary karate's club [26] |
| dol | 62 | 159 | 8 | Dolphins association [27] |
| polb | 105 | 441 | 7 | Books about US politics [28] |
| footb | 115 | 613 | 4 | American football games [29] |
| email | 1005 | 16064 | 7 | email-Eu-core network [30] |
| ranGraph1 | 2100 | 9862 | 6 | an random graph |

$E'$ and the corresponding weight set $W'$. Next, for each link $(u, v) \in E'$, the link importance $I$ can be calculated. Sort $I$ in descending order to get $I^{\text{sorted}}$. At this time, the $\beta$ links with the largest $I$ are the links that need to be deleted, which is defined as $E^-$. Finally, the updated network $G^- = (V, E \backslash E^-)$ can effectively weaken the dissemination of rumor under the given constraints.

## V. EXPERIMENTS

### A. Elementary Knowledge

In this section, we briefly introduce some elementary knowledge of all the experiments in our research. They mainly include experimental environment, datasets, and experimental evaluation.

*1) Experimental Environment:* All the experiments are run on a windows10 server of 3.50-GHz CPU and 16.00-GB RAM. The programming simulation language is MATLAB.

*2) Datasets:* The datasets used in this work are many social networks which are free for scientific use and can be downloaded online.[1,2,3] In addition, we also generate a random networks as the experimental dataset. Table I gives an overview of the networks considered.

*3) Evaluation Methodology:* To prove the effectiveness of rumor dissemination path interruption algorithm RumorDecay, we mainly evaluate the experimental results in two aspects: 1) network utility and 2) the scope of rumor dissemination.

Network utility describes the change degree of the structure of original network $G$. For the study of rumor dissemination interruption, we just need to delete a certain number of social links in $G$ and the network utility is bound to decline. If the network utility has a big change, it means that the RumorDecay algorithm destroys the original network too much. Therefore, we not only hope to weaken the dissemination of rumor in $G$ but also hope to damage the original network as little as possible. In this article, we introduce the average clustering coefficient $C$ and the average path length $L$ to evaluate the network utility. The average clustering coefficient describes the coefficient of the degree of clustering of nodes in networks. Equation (16) gives the definition of $C$, where $E_i$ is the number of links that actually exist between the $k_i$ neighbor nodes of node $i$ and $C_i$ is the clustering coefficient of node $i$. The average path length is defined as the average distance between any two nodes in the network. Equation (17) gives the definition of $L$, where $d_{ij}$ is the number of links on the shortest path between nodes $i$ and $j$

$$C = \frac{1}{N}\sum_{i=1}^{N} C_i = \frac{1}{N}\sum_{i=1}^{N} \frac{2E_i}{k_i(k_i-1)} \tag{16}$$

$$L = \frac{1}{\frac{1}{2}N(N-1)}\sum_{i \geq j} d_{ij}. \tag{17}$$

The infectious disease model is a mathematical model to predict the trend of infectious diseases. According to the types of infectious diseases, common infectious disease models can be divided into SIS [31], SIR [32], SIRS [33], and so on. In this article, the SIR model is used to predict the dissemination scope of rumors. Before and after rumor dissemination interruption, we apply the SIR algorithm in $G$ and $G^-$, respectively, to calculate the scope of rumor dissemination. For an experiment, under the evaluation of SIR model, suppose that $m'$ nodes in $T$ finally accept the rumor. Here, the scope of rumor dissemination is denoted by $\text{sir} = m'/m$. In this work, for each rumor interruption result $G^-$, we perform the SIR algorithm for 50 times and get the average value.

In addition, according to Section III, $H_{ST}$ describes the probability that $T$ is affected by $S$. $H_{ST}$ is also an index to evaluate the scope of rumor dissemination. Therefore, by calculating the difference of $H_{ST}$ before and after deleting links, we can know whether our method effectively weakens the dissemination of rumor.

### B. Experimental Results

In this work, the different controlled experiments are designed in two directions: one is to study the influence of the threshold ($\beta$) of links to be deleted on the experimental results and the other is to study the influence of different dissemination rate ($r$) on the experimental results.

In our experiments, we observe that when the frequency of RW exceeds a certain threshold, it will have a small effect on the final results. Therefore, in our experiments, we set $\tau = 2 \times 10^5$ and it is big enough to deal with the sampling of paths. The most important point is that $2 \times 10^5$ rounds of RW can be completed in a relatively short period of time and it is much less than the time which is required of DFS algorithm.

*1) Threshold of Link Deletion:* For threshold $\beta$, we let $\beta = 0.02 \times N, 0.04 \times N, \ldots, 0.2 \times N$, respectively, and $r = 0.16$. Figs. 6 and 7 clearly describe the experimental results ($C$, $L$, $H$, and $sir$) corresponding to different rumor interruption algorithms with the change of threshold $\beta$. Fig. 6 is for networks kar, dol, and polb, while Fig. 7 is for networks footb, e-mail, and ranGraph1. In this work, all the same experiments are repeated for six times. In addition, both in Figs. 6 and 7, the green curve with the star symbol represents the standard curve corresponding to the sampling result of algorithm RW. The red curve with the circle symbol is the result of RumorDecay-kHNN and the blue curve with cross symbol is the result of RumorDecay-kHRW.

Generally speaking, for network utilities $C$ and $L$, the change degree of network utilities increases with the increase of $\beta$. The larger $\beta$, the more links are deleted in the

[1] http://www-personal.umich.edu/~mejn/netdata/
[2] http://snap.stanford.edu/data/index.html
[3] http://www.orgnet.com/

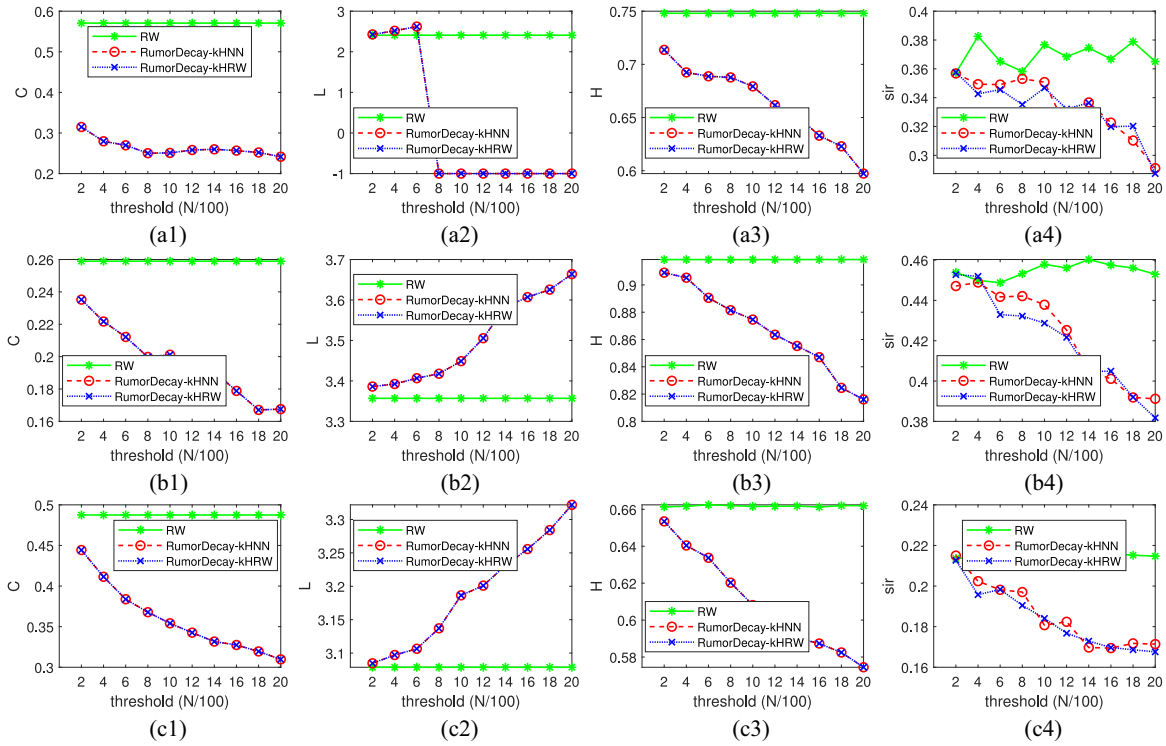



Fig. 6. Experimental results of networks (a1)–(a4) kar, (b1)–(b4) dol, and (c1)–(c4) polb with different threshold $\beta$.

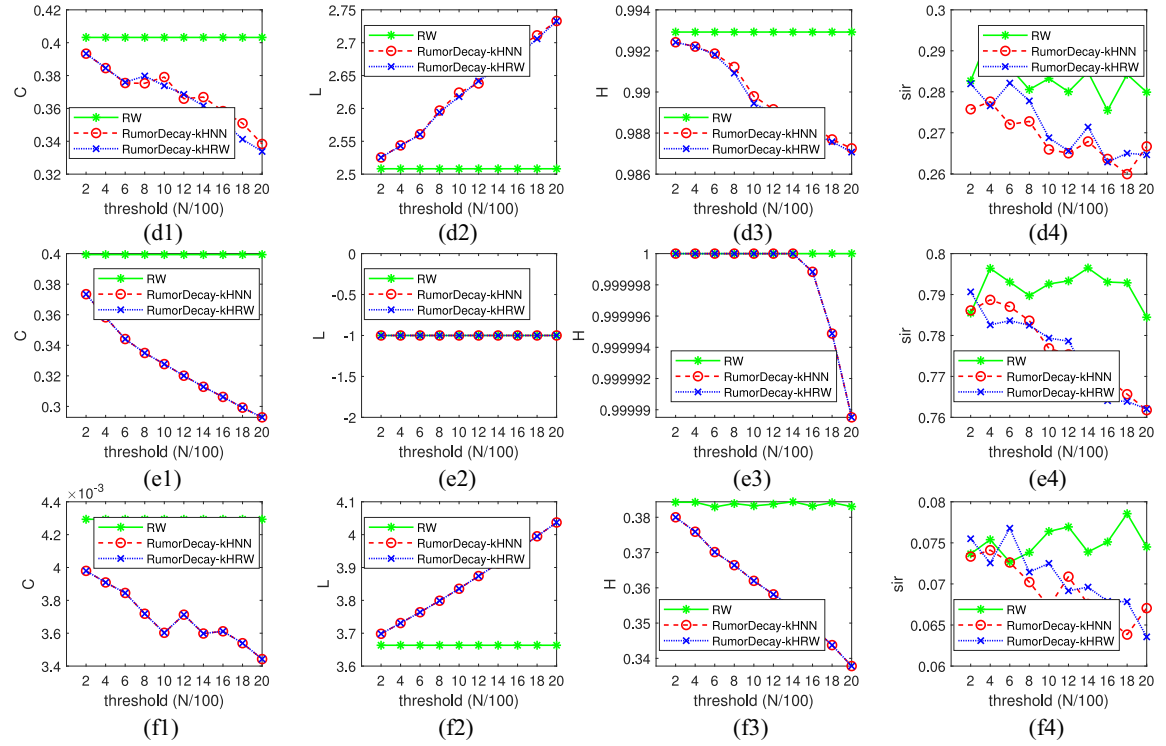



Fig. 7. Experimental results of networks (d1)–(d4) footb, (e1)–(e4) e-mail, and (f1)–(f4) ranGraph1 with different threshold $\beta$.

network and the greater the damage to the original network $G$. Therefore, as shown in Figs. 6 and 7, the curves of $C$ and $L$ corresponding to methods RumorDecay-kHNN and RumorDecay-kHRW are almost linear. Particularly, in Fig. 6(a2), we can observe that when the threshold exceeds $0.06 \times N$, the network kar will become a disconnected graph after deleting links. Thus, we cannot calculate the average path length of e-mail and replace the average path length of e-mail with $-1$. This is a reasonable phenomenon, deleting many links in network kar which has only 78 links will definitely have a greater impact on the structure of it.

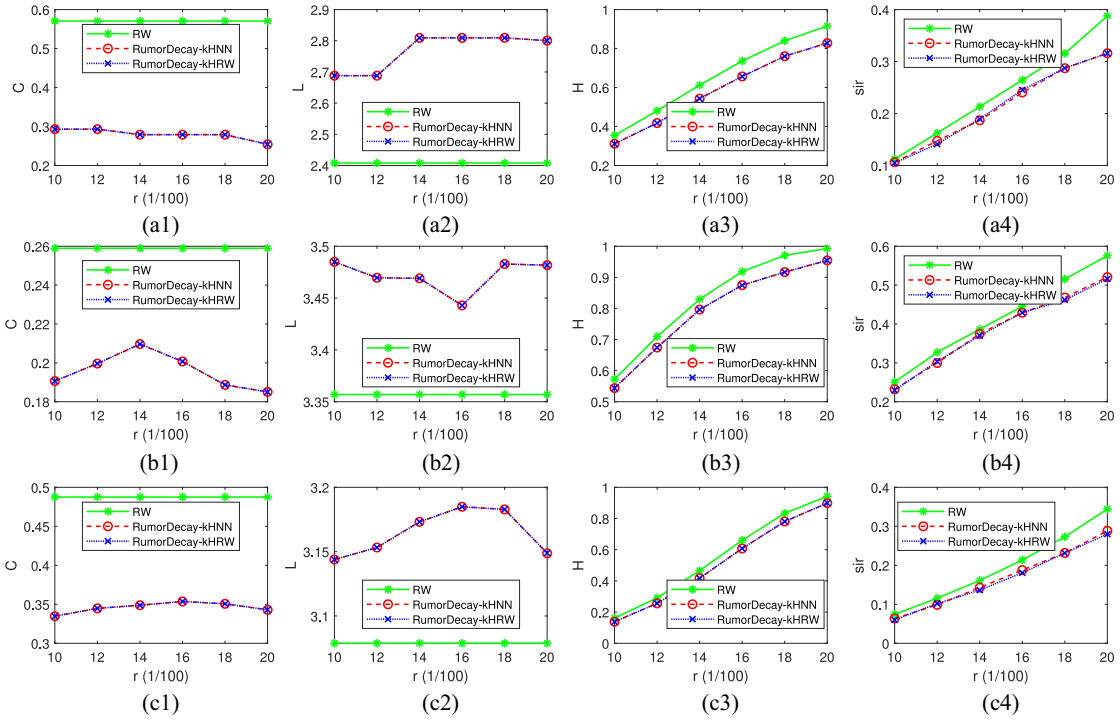



Fig. 8. Experimental results of networks (a1)–(a4) kar, (b1)–(b4) dol, and (c1)–(c4) polb with different rate $r$.

Now, we observe the influence of different $\beta$ on the scope of rumor dissemination. First, no matter how $\beta$ changes, the standard curves (the green curve with star symbol) of $H$ and *sir* should fluctuate in a smaller range because $r$ is a fixed value in all experiments. According to the observation, the standard curves of $H$ in Figs. 6 and 7 basically meet the above requirements. But the standard curves of *sir* show a sawtooth shape. The reason is that there are random modules in the SIR model, so it is difficult to get the same results for different experiments. On the whole, the large $\beta$, the better the effect of rumor dissemination interruption. The special case is shown in Fig. 7(e3), when $r$ is less than $0.16 \times N$, methods RumorDecay-kHNN and RumorDecay-kHRW cannot effectively weaken the dissemination of rumor. Besides, the experimental results show that the experimental performance of methods RumorDecay-kHNN and RumorDecay-kHRW is similar. In terms of *sir* in Figs. 6 and 7, methods RumorDecay-kHNN and RumorDecay-kHRW have their own advantages on different $\beta$. Therefore, we cannot clearly point out which method is better. Although the larger the threshold of link deletion, the better the interruption of rumor dissemination. However, deleting a large number of links in the network may completely destroy the availability of the network and the network utilities have also been greatly changed. Therefore, in our experiment, we suggest $\beta = 0.1 \times N$.

Table II shows the average running time of algorithms RW, RumorDecay-kHNN and RumorDecay-kHRW in different networks. First, in general, the path sampling time based on the RW algorithm will increase with the increase of network scale. Although the RW algorithm is repeated for $2 \times 10^5$ times in each experiment, the path sampling work can be completed in 1 min for all individual experiments. The

TABLE II
AVERAGE RUNNING TIME (IN SECONDS) OF THE CONTROLLED EXPERIMENTS

| | RW | RumorDecay-kHNN | RumorDecay-kHRW |
|---|---|---|---|
| kar | 10.1051 | 0.00040 | 20.8878 |
| dol | 13.2238 | 0.00155 | 27.5058 |
| polb | 19.8989 | 0.00272 | 31.8390 |
| footb | 13.7445 | 0.00188 | 21.0463 |
| email | 28.7753 | 0.28789 | 34.8930 |
| ranGraph1 | 33.1942 | 0.78104 | 34.6298 |

running process of algorithm RumorDecay-kHNN is very fast, because RumorDecay-kHNN only needs to find neighbors of some nodes. In our experiments, the fastest running time of RumorDecay-kHNN is about only 3 ms and the maximum running time is no more than 1 s. For RumorDecay-kHRW, both the source node and the target node need to perform a random walk process, which is different from the single random walk model of RW, so the running time of algorithm RumorDecay-kHRW is generally larger than the running time of RW.

*2) Rumor Dissemination Rate:* For rate $r$, we let $r = 0.1, 0.12, \ldots, 0.2$, respectively, and $\beta = 0.1 \times N$. Figs. 8 and 9 clearly describe the results corresponding to different rumor interruption algorithms with the change of rate $r$. Fig. 8 is for networks kar, dol, and polb, while Fig. 9 is for networks footb, e-mail, and ranGraph1. Similarly, all the same experiments are repeated six times.

First, as shown in Figs. 8 and 9, deleting a certain number ($\beta = 0.1 \times N$) of links in network $G$ will affect the clustering coefficient and the average path length of the network. All clustering coefficient $C$ of $G$ in different experiments show

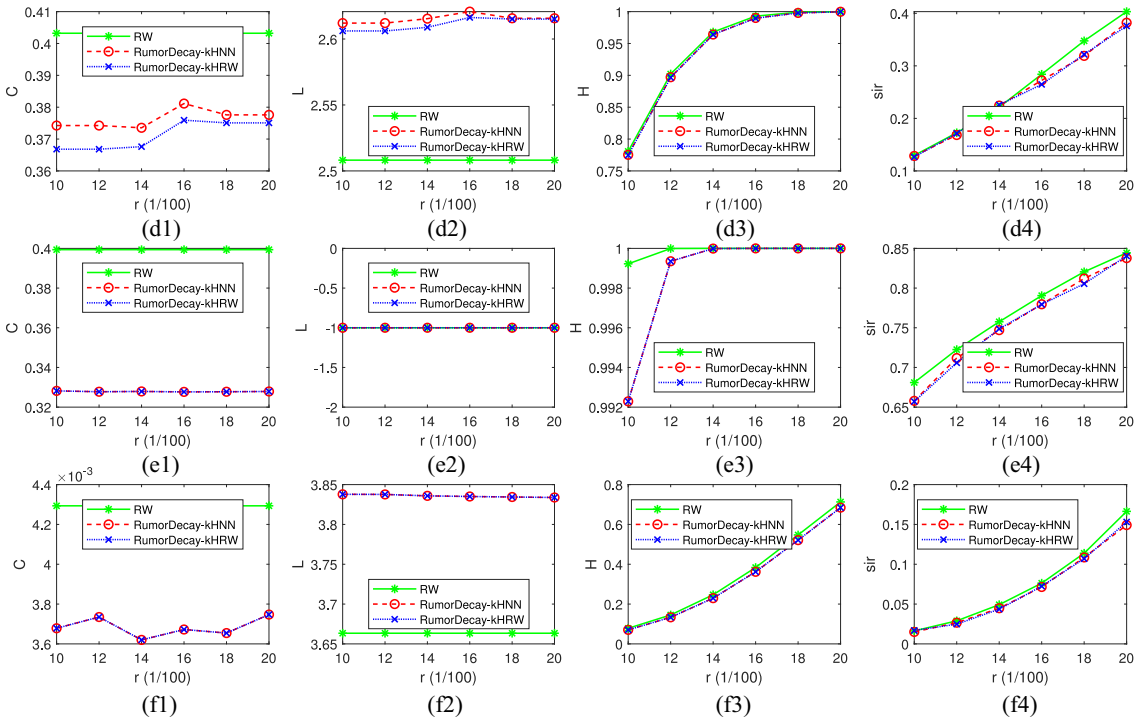



Fig. 9. Experimental results of networks (d1)–(d4) footb, (e1)–(e4) e-mail, and (f1)–(f4) ranGraph1 with different rate $r$.

different degrees of reduction and the average path length $L$ of $G$ generally increases in varying degrees. In addition, the change of network utilities is independent of rate $r$. Namely, the change of $r$ cannot affect the network utility directly. Particularly, look closely at Fig. 8(a1), only in this case, the change rate of network utilities exceeds 50%. In other experiments, the change rate of network utilities is between 10% and 30%. For our rumor dissemination interruption study, the change rate of network utilities is reasonable.

Now, we will discuss the influence of different $r$ on the scope of rumor dissemination. As shown in the third and fourth columns of Figs. 8 and 9, on the whole, both $H_{ST}$ and *sir* increase with the increase of $r$. In addition, methods RumorDecay-kHNN and RumorDecay-kHRW both have almost a same effect on the interruption of rumor dissemination. When $r$ is small, such as $r = 0.10$, deleting links with a certain threshold may only have a very weak effect on the final results. The main reason is that a small rate $r$ cannot have a large enough impact on the dissemination of rumor even in the original network $G$. When $r \geq 0.12$, $H$ and *sir* curves of RumorDecay-kHNN and RumorDecay-kHRW are lower than the standard curve in more than 95% of the experiments. That is, both methods RumorDecay-kHNN and RumorDecay-kHRW based on link deletion can effectively weaken the dissemination of rumor in most cases. Of course, in some rare cases, our methods may not be able to effectively weaken the dissemination of rumor, such as networks footb [Fig. 9(d3)] and e-mail [Fig. 9(e3)]. In fact, whether it is the dissemination of rumors or infectious diseases, the dissemination rate $r$ is generally not too large or too small. When $r$ is too small, we can believe that the rumors or infectious diseases are not harmful for our real life. When $r$ is too large,

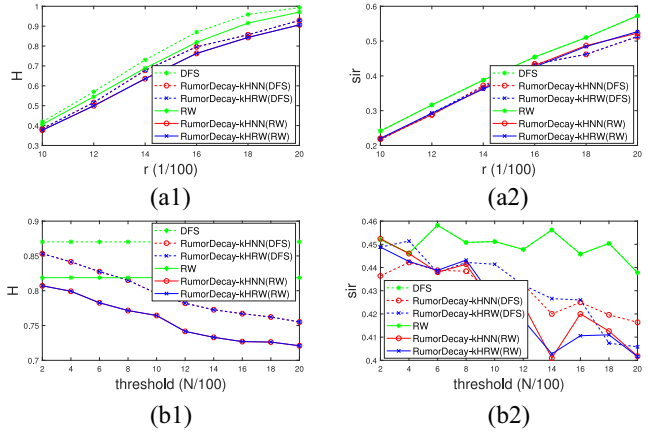



Fig. 10. Experimental results of (a1), (a2), (b1), and (b2) dol based on DFS and RW path sampling algorithms, respectively.

we will certainly take more effective measures to interrupt the dissemination of rumors or infectious diseases. For example, the release of truth information for rumor interruption and the intervention of medical drugs for infectious diseases.

## VI. Extended Experiments

In this section, we will compare the efficiency of algorithms DFS and RW in the path sampling work for network dol. Here, we do not discuss the network utilities and only discuss the dissemination scope of rumor for different rumor interruption methods. As shown in Fig. 10, the dotted lines represent the results of DFS-based rumor interruption methods and the solid lines represent the results of RW-based rumor interruption methods. Obviously, the standard of rumor

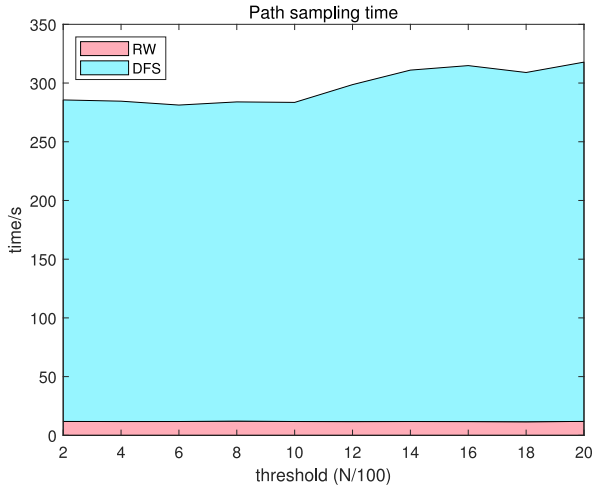



Fig. 11. Path sampling time of RW and DFS.

dissemination scope ($H$ and $sir$) based on the DFS algorithm will get the maximum and it is larger than the standard based on the RW algorithm. All in all, both algorithms RumorDecay-kHNN and RumorDecay-kHRW based on DFS or RW can weaken the dissemination of rumor to a certain extent.

Fig. 11 shows the path sampling time of DFS and RW. Because the change of $r$ or $\beta$ will not affect the process of all experiments, so we only compare the running time of the $\beta$-based controlled experiments. As shown in Fig. 11, for path sampling work, RW only needs a small amount of time to complete path sampling, while the running time of the DFS algorithm is much longer than RW. With the increase of network scale, DFS is obviously not suitable to solve the problem of path sampling.

## VII. Conclusion

This work studies the problem of rumor dissemination interruption for target recipients in social networks and proposes a rumor controlling strategy RumorDecay (i.e., RumorDecay-kHNN and RumorDecay-kHRW). The core idea of RumorDecay is link deletion with limited threshold.

First, this work built a mathematical model TID for the target information disseminating problem and theoretically proves its monotonicity and submodularity. Second, according to TID, in order to select the best candidate links to be deleted, we introduced the RW algorithm to sample quickly the rumor dissemination paths. Then, we proposed a heuristical strategy-based rumor influence decay mechanism referred to as RumorDecay. RumorDecay can find the critical links to be deleted in a short period of time. Finally, we designed and conducted a large number of controlled experiments in two input parameters: 1) rumor dissemination rate $r$ and 2) the threshold $\beta$ of links to be deleted. Experimental results showed that the RumorDecay strategy can effectively weaken rumor dissemination in social networks. Deleting key links in social networks can effectively weaken the spreading ability of rumors. In real life, both rumor controlling and infectious disease controlling can be achieved by the operation of critical link deletion. Therefore, the critical link selection method has become the main content of our work.

Limited by the immaturity of current technology, there are still some problems that have not been solved. There are several avenues for future research: 1) design a more comprehensive rumor dissemination interruption model. For example, we can consider the impact of the release of the truth or the intervention of official departments on the dissemination of rumors and 2) the community phenomenon of rumor dissemination. Specifically, the acceptance probability of a rumor by members of a community is consistent. So, we can consider the structure disturbance of the recipient community in the network to weaken or interrupt the dissemination of rumors. Furthermore, rumor producers often exist in the network in the form of community. Therefore, identifying the rumor community and controlling it are also a feasible idea.

## References

[1] S. Wen, J. Jiang, Y. Xiang, S. Yu, W. Zhou, and W. Jia, "To shut them up or to clarify: Restraining the spread of rumors in Online social networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 12, pp. 3306–3316, Dec. 2014.

[2] S. Wen, M. S. Haghighi, C. Chen, Y. Xiang, W. Zhou, and W. Jia, "A sword with two edges: Propagation studies on both positive and negative information in online social networks," *IEEE Trans. Comput*, vol. 64, no. 3, pp. 640–653, Mar. 2015.

[3] Y. Wang, A. V. Vasilakos, J. Ma, and N. Xiong, "On studying the impact of uncertainty on behavior diffusion in social networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 45, no. 2, pp. 185–197, Feb. 2015.

[4] Y. G. Lin *et al.*, "Dynamic control of fraud information spreading in mobile social networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 6, pp. 3725–3738, Jun. 2021, doi: 10.1109/TSMC.2019.2930908.

[5] L. J. Zhao *et al.*, "The impact of authorities' media and rumor dissemination on the evolution of emergency," *Physica A Stat. Mech. Appl.*, vol. 391, no. 15, pp. 3978–3987, 2012.

[6] C. Budak, D. Agrawal, and A. E. Abbadi, "Limiting the spread of misinformation in social networks," in *Proc. 20th WWW. ACM*, 2011, pp. 665–674.

[7] X. He, G. Song, W. Chen, and Q. Jiang, "Influence blocking maximization in social networks under the competitive linear threshold model," in *Proc. SDM. SIAM*, 2012, pp. 463–474.

[8] L. Fan, Z. Lu, W. Wu, B. Thuraisingham, H. Ma, and Y. Bi, "Least cost rumor blocking in social networks," in *Proc. IEEE 33rd ICDCS.*, 2013, pp. 540–549.

[9] Y. G. Wang, F. Z. Cai, E. K. Lang, J. B. Hu, and Z. Chen, "A diffusion control method of fake information in social networks," *J. Comput. Res. Develop.*, vol. 49, no. 2, pp. 131–137, 2012.

[10] Y. Bao, N. Yan, C. Yi, and Y. Xue, "Effective immunization strategy for rumor propagation based on maximum spanning tree," in *Proc. Int. Conf. Comput. Netw. Commun. (ICNC)*, 2014, pp. 11–15, doi: 10.1109/ICCNC.2014.6785296.

[11] K. Kandhway and J. Kuri, "Using node centrality and optimal control to maximize information diffusion in social networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 7, pp. 1099–1110, Jul. 2017.

[12] L. Ding, P. Hu, Z. H. Guan, and T. Li, "An efficient hybrid control strategy for restraining rumor spreading," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 11, pp. 6779–6791, Nov. 2021, doi: 10.1109/TSMC.2019.2963418.

[13] L.-L. Ma, C. Ma, H.-F. Zhang, and B.-H. Wang, "Identifying influential spreaders in complex networks based on gravity formula," *Physica A Stat. Mech. Appl.*, vol. 451, no. 1, pp. 205–212, 2016.

[14] S. Wang, X. Zhao, Y. Chen, Z, Li, K. Zhang, and J. Xia, "Negative influence minimizing by blocking nodes in social networks," in *Proc. 17th AAAI Conf. Artif. Intell.*, 2013, pp. 134–136.

[15] E. B. Khalil, D. Dilkina, and L. Song, "Scalable diffusion-aware optimization of network topology," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2014, pp. 1226–1235.

[16] M. Kimura, K. Saito, and H. Motoda, "Minimizing the spread of contamination by blocking links in a network," in *Proc. AAAI*, vol. 8, 2008, pp. 1175–1180.

[17] H. Tong, B. A. Prakash, T. E. Rad, M. Faloutsos, and C. Faloutsos, "Gelling, and melting, large graphs by edge manipulation," in *Proc. 21st ACM Int. Conf. Inf. Knowl. Manag.*, 2012, pp. 245–254.

[18] R. Yan, Y. Li, W. Wu, D. Li, and Y. Wang, "Rumor blocking through Online link deletion on social networks," *ACM Trans. Knowl. Disc. Data*, vol. 13, no. 2, pp. 1–26, 2019.

[19] Q. Yao, C. Zhou, L. Xiang, Y Yao, and G. Li, "Minimizing the negative influence by blocking links in social networks," in *Proc. Int. Conf. Trustworthy Comput. Services*, 2014, pp. 65–73, doi: 10.1007/978–3-662–47401-3_9.

[20] Y. Ping, Z. Cao, and H. Zhu, "Sybil-aware least cost rumor blocking in social networks," in *Proc. IEEE Global Commun. Conf.*, 2014, pp. 692–697, doi: 10.1109/GLOCOM.2014.7036888.

[21] J. Zheng and P. Li, "Least cost rumor community blocking optimization in social networks," in *Proc. 3rd Int. Conf. Secur. Smart Cities Ind. Control Syst. Commun. (SSIC)*, 2018, pp. 1–5, doi: 10.1109/SSIC.2018.8556739.

[22] M. Zhang, X. M. Wang, and X. Y. Qin, "A rumor blocking method based on tolerance of users," in *Proc. Int. Conf. Ubiquitous Comput. Commun. Data Sci. Comput. Intell. Comput. Netw. Serv.*, 2019, pp. 462–468, doi: 10.1109/IUCC/DSCI/SmartCNS.2019.00103.

[23] B. Wang, G. Chen, L. Fu, S. Li, and X. Wang, "DRIMUX: Dynamic rumor influence minimization with user experience in social networks," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 10, pp. 2168–2181, Oct. 2017.

[24] Q. Wu, X. Zhao, L. Zhou, Y. Wang, and Y. Wang, "Minimizing the influence of dynamic rumors based on community structure," *Int. J. Crowd Sci.*, vol. 3, no. 3, pp. 303–314, 2019.

[25] Q. Z. Fang *et al.*, "General rumor blocking: An efficient random algorithm with martingale approach," in *Proc. 12th Int. Conf. AAIM*, 2018, pp. 161–176. [Online]. Available: https://doi.org/10.1016/j.tcs.2019.05.044

[26] W. W. Zachary, "An information flow model for conflict and fission in small groups," *J. Anthropol. Res.*, vol. 33, no. 4, pp. 452–473, 1977.

[27] D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, and S. M. Dawson, "The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations," *Behav. Ecol. Sociobiol.*, vol. 54, no. 4, pp. 396–405, 2003.

[28] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 U.S. Election," in *Proc. WWW Workshop Weblogging Ecosyst.*, 2005, pp. 36–43.

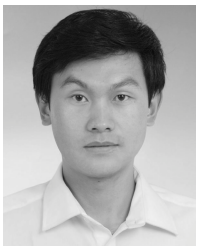
[29] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 U.S. election: Divided they blog," in *Proc. 3rd Int. Workshop Link Discov.*, 2005, pp. 36–43.

[30] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graph evolution: Densification and shrinking diameters," *ACM Trans. Knowl. Disc. Data*, vol. 1, no. 1, p. 2, 2007.

[31] J. O. Kephart and S. R. White, "Directed-graph epidemiological models of computer viruses," in *Proc. IEEE Comput. Soc. Symp. Res. Secur. Privacy*, 1991, pp. 343–361.

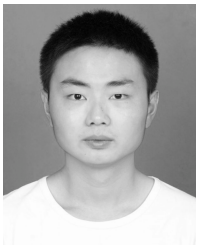
[32] J. H. Kim, S. Radhakrishnan, and S. K. Dhall, "Measurement and analysis of worm propagation on Internet network topology," in *Proc. 13th Int. Conf. Comput. Commun. Netw. (ICCCN)*, 2004, pp. 495–500.

[33] X. Han and Q. Tan, "Dynamical behavior of computer virus on Internet," *Appl. Math. Comput.*, vol. 217, no. 6, pp. 2520–2526, 2010.

**Zhongyuan Jiang** received the B.S. degree in computer science and technology and the Ph.D. degree in signal and information processing from Beijing Jiaotong University, Beijing, China, in 2009 and 2013, respectively.

He is currently an Associate Professor with the School of Cyber Engineering, Xidian University, Xi'an, China. His research interests include privacy preserving, social computing, urban computing, and network functions virtualization.

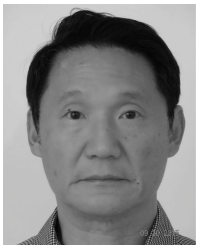


**Xianyu Chen** received the B.S. degree in computer science and technology from Anhui University, Hefei, China, in 2018. He is currently pursuing the M.S. degree in computer technology with the School of Cyber Engineering, Xidian University, Xi'an, China.

His current research interests include social network analysis and privacy protection.

**Jianfeng Ma** (Member, IEEE) received the B.S. degree in mathematics from Shaanxi Normal University, Xi'an, China, in 1985, and the Ph.D. degree in communication and electronic system from Xidian University, Xi'an, in 1995.

He is a Professor with the School of Cyber Engineering, Xidian University and the Director of the Key Laboratory of Network and System Security of Shaanxi Province. His research interests include cryptography, wireless network security, data security, and mobile intelligent system security.



**Philip S. Yu** (Life Fellow, IEEE) received the M.B.A. degree from New York University, New York, NY, USA, in 1982, and the Ph.D. degree in EE from Stanford University, Stanford, CA, USA, in 1978.

He is a Distinguished Professor and the Wexler Chair of Information Technology with the Department of Computer Science, University of Illinois at Chicago (UIC), Chicago, IL, USA. Before joining UIC, he was with the IBM Watson Research Center, Ossining, NY, USA, where he built a World-Renowned Data Mining and Database Department. His research interests include data mining, social computing, and privacy protection.