Information Sciences 651 (2023) 119718

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Online conflict resolution: Algorithm design and analysis ☆

Guoyao Rao [a], Deying Li [b,*], Yongcai Wang [b], Wenping Chen [b], Chunlai Zhou [b], Yuqing Zhu [c]

[a] Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou 324000, China
[b] School of Information, Renmin University of China, Beijing, 100872, China
[c] Research Center for Applied Mathematics and Machine Intelligence, Zhejiang Lab, Hangzhou, Zhejiang, China

ARTICLE INFO

Keywords:
Online conflicts
Influence maximization
Social networks

ABSTRACT

In today's digital age, the widespread dissemination of information through social media platforms has a profound impact on people's thoughts and actions. It fosters lively discussions and debates among individuals, but also sparks online conflicts that exacerbate social tensions and divisions. Therefore, it brings a challenge: how to eliminate these online conflicts. In this paper, we analyze and solve it from an algorithmic perspective. We introduce individuals who are willing to persuade their close friends to avoid conflicts with others. Our objective is to strategically select a group of influencers who can encourage others to act as mediators, thus mitigating online conflicts within the realm of social networks. Hence, an optimization problem is formulated, which is proved to be NP-hard and #P-hard. Subsequently, to solve the problem, an algorithm grounded in a reverse sampling technique is proposed, which achieves a performance bound of $(1 - 1/e)$. Finally, the result of experiments on real-world networks datasets to evaluate our algorithm and shows that our algorithm outperforms all comparisons.

1. Introduction

As the result of the development of Internet, nowadays social media spreads rich information at an exponential rate and is profoundly affecting people's opinions and communications on various issues. Meanwhile, social media brings many new online conflicts and deepens the existing online conflicts among people [5]. In the United States, people are deeply troubled by debates on various social issues such as racial discrimination and abortion, which further escalate into online conflicts involving the dissemination of violent messages, hate speech, and comments among individuals [39]. In the European Union, about 40% of people have been suffered from online conflicts by feeling attacked or threatened [9]. All of these irrational and rude online conflicts such as verbal violence make the social environment extremely noisy and chaotic, and further aggravate the division and opposition of the crowd [26]. Even worse, these online conflicts may spread to the offline world, resulting in exhaustion within real-world society and triggering physical violence [8]. In this paper, we specifically denote all the online conflicts as an directed graph with edges like e_{uv} between two person u and v, where the direction means an attack directed from an attacker to a sufferer. So having only one directed edge

☆ This work is partly supported by National Natural Science Foundation of China under grant 12071478, 61972404, 62102376, and Public Computing Cloud of Renmin University of China.

* Corresponding author.
E-mail addresses: gyr@ruc.edu.cn (G. Rao), deyingli@ruc.edu.cn (D. Li), ycw@ruc.edu.cn (Y. Wang), chenwenping@ruc.edu.cn (W. Chen), czhou@ruc.edu.cn (C. Zhou), yzhu@zhejianglab.com (Y. Zhu).

https://doi.org/10.1016/j.ins.2023.119718
Received 12 April 2023; Received in revised form 12 September 2023; Accepted 21 September 2023

Available online 27 September 2023

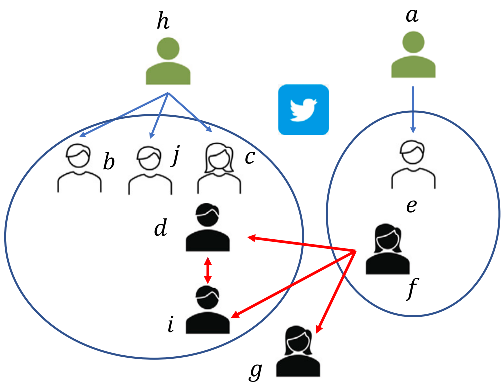


Fig. 1. Illustration of resolving a conflict in Twitter.

between two nodes means an unilateral conflict and having both directed edges between two nodes means an bilateral conflict. Then we pose questions: How are these online conflicts generated and how can we eliminate them?

Because of its importance and popularity, online conflicts are attracting studies in the fields of sociology, such as the works in [10][38][14] and so on. In reality, most conflicts arise because individuals accept incomplete, radical, or incorrect information within closed environments [37]. Factors such as rumors, misinformation, and outdated news can contribute to the escalation of conflicts. Therefore, timely and widely disseminated positive mediation can play a crucial role in resolving conflicts and correcting misperceptions. Providing authoritative and impartial information can help counteract the spread of misinformation and rumors, offering individuals a more accurate understanding of the situation [18]. Additionally, impulsive thinking and reactionary behavior also can further exacerbate the online conflict. Thus, promoting harmony and fostering a sense of unity can help alleviate negative emotions and create an environment conducive to peaceful resolution. Encouraging dialogue, empathy, and understanding can facilitate constructive conversations and bridge the divide between conflicting parties. When people feel heard and respected, it can reduce hostility and create opportunities for finding common ground and shared solutions [2].

So it is important for individuals, community leaders, organizations, and governments to actively engage in promoting positive mediation. By doing so, it becomes possible to mitigate conflicts and build a more harmonious and inclusive society. Note that the "Word-of-mouth" effect suggests that people usually trust the words from their close friends and can be easily persuaded by them [16]. Therefore, with this advantage of social influence, encouraging people to mediate their contentious online friends involved in the conflict also provides a way to help resolve the online conflicts [25]. We give a vivid scenario from an observant on Twitter to show how the conflict is resolved. As shown in Fig. 1, users b, j, and c are friends of users d and i, while user e is a friend of user f. Users a and f are attacking each other by verbal violence in the comment area of certain tweet. By the call of harmony from users a, h, users b, c and e may be influenced as mediators to persuade their close friends d, i, f to stop attacking.

Here, we adopt this idea of selecting a subset of users as seeds to influence all attackers' close friends to act as **mediators** who can persuade attackers to stop the conflicts and hence solve the problem of online conflicts effectively from an algorithmic perspective. Intuitively, we might believe the strategy that influencing more people to become mediators can eliminate conflicts, and it aligns with the optimization goal of traditional influence maximization [15]. However, the following examples demonstrate that such a simplistic optimization goal does not work well. In Fig. 1, we assume that the direct edges from nodes h and a represent a definite influence, where a node must be persuaded by its close friend who is already influenced to become a mediator.

When considering the most influential approach of encouraging more people to become mediators, selecting seed node h can influence three individuals b, j, c, to become mediators. This intervention can prevent attackers d and i from attacking each other, but it fails to completely resolve the conflict and none are freed from the conflict. But if we select user a as the seed, even though it can only influence one individual e to become a mediator, it has the advantage of freeing individuals f and g from the conflict completely, and reducing the level of attack on users d and i. From this example, it is evident that in order to better eliminate conflicts, it is necessary to design more sophisticated models and algorithms to provide more accurate strategies to select seeds.

At the modeling level, we first adopt a probability $p_a(u) \in [0, 1]$ to measure the likelihood of an attacker u being persuaded to stop the conflicts by its close friends to cease the conflicts. Here, we intuitively set this probability using the Bernoulli process, where u is persuaded by each close friend one by one, each with the same probability and we have

$$p_a(u) = 1 - (1 - b_u)^{|I(S) \cap \mathcal{F}_u|} \tag{1}$$

where $I(S)$ is the set of all nodes being influenced as mediators by a set of seeds S, $\mathcal{F}_u$ is the set of all close friends of an attacker u, and b_u is the Bernoulli probability of u being persuaded by one of its close friend. Then we adopt a weight $w_s^*(v) \in [0, 1]$ to measure the attack level on a sufferer v after the mediation, and we set it intuitively with the percentage of its attackers not being persuaded as follows.

$$w_s^*(v) = \frac{|\mathcal{A}_v^*|}{|\mathcal{A}_v|} \tag{2}$$

where $\mathcal{A}_v^*$ is the set of all attackers who are still attacking after the mediation influence and $\mathcal{A}_v$ is the set of all attackers attacking a sufferer v before the mediation influence. Finally, in this paper, we consider a discrete optimization of choosing seeds to minimize the attack level to all sufferers after spreading mediation influence over social networks. Our main contributions are summarized as follows:

- We first consider resolving online conflicts from an algorithmic perspective, and formulate it to a discrete optimization of choosing seeds to spread mediation influence over social networks to minimize the attack to all sufferers. We prove the problem is NP-hard and its objective is #P-hard to compute.
- Instead of heavily repeated Monte Carlo simulations which are time-consuming in large networks, we propose an estimation method for the objective function based on the Reverse Sampling technique. A near-optimal $(1-1/e)$ algorithm with theoretical analysis is then proposed.
- We conduct various experiments based on real-world databases to evaluate our algorithms.

In the rest of this paper, we firstly review some existing related works of influence maximization in section 2. In section 3, we introduce some preliminaries and formulations for our problem and analyze its hardness and related properties. In section 4, we introduce an estimation method for the target function and propose an algorithm for the problem. Finally, in section 5, we show our experiments and give the conclusion in section 6.

2. Related work

In this section, we introduce the related work from following two fields in influence maximization: (1) Variants of influence maximization, and (2) Related algorithms and techniques.

2.1. Variants of influence maximization

The research on the Influence Maximization problem (IM) has always been a focal point in the field of social network analysis. Kempe et al. [15] first formulated the IM problem as a discrete optimization task, aiming to select k nodes as seeds to maximize the expected number of influenced nodes through a stochastic diffusion process in social networks. They specifically analyzed the IM problem using the IC and LT models. This work attracted and inspired numerous researchers to delve into social influence, leading to the emergence of various variants and extensions.

Many studies focus on the spread problem in various specific scenarios, such as time-constrained scenarios [19,7], topic-aware scenarios [1,6], competition scenarios [3,20], rumor control scenarios [13], multi-round scenarios [32], group-fairness scenarios [30,36], and others. These works primarily address diffusion models to suit specific scenarios. However, they are all individual-based, meaning they do not consider the constraints imposed by additional relationships among the nodes. While some works, such as [27–29], have considered positive relationships like cooperation and matching, there is also a negative aspect where nodes may have conflicting relationships with each other. Ignoring this negative relationship can be detrimental in many scenarios. Therefore, in our variant, we take into account online conflicts and consider conflicting relationships among nodes. This necessitates the adoption of more precise seed selection strategies to address the challenges posed by the complex conflicting relationships among the nodes.

2.2. Related algorithms and techniques

The basic IM problem is proved to be NP-hard and the influence computation is #P-hard. With the good property of submodularity the basic greedy method can provide a $(1-1/e-\epsilon)$-approximation solution [22], where ϵ is the loss caused by influence estimation since it's #P-hard to get the accurate influence. However, such greedy-based methods cost too much time using the heavy Monte Carlo simulations to estimate the marginal gain of node's influence, and it's hard to apply to the large scale network. There are many improvements [17,11,24] to reduce the Monte Carlo simulations. Then Tang et al. [35] and Borgs et al. [4] proposed the reverse influence set (RIS) sampling technique to estimate the influence. The concept of Reverse Influence Sampling (RIS) involves reverse propagation from a node v to identify a random set of nodes that can influence v. This approach efficiently estimates influence compared to repetitive simulations. RIS has led to the transformation of the Influence Maximization (IM) problem into the classical set cover problem. Various extensions and enhancements based on RIS have been developed, including IMM [34], SSA and D-SSA [23], OPIM [33], SUBSIM [12].

3. Preliminaries

3.1. IC model

We use the classical Independent Cascade (IC) model to be the diffusion model of how people are influenced to be mediators. IC model draws the influence relationships among nodes by a directed graph as $G(V,E,p)$, where V is the set of all nodes in the social network, E is the set of all influence edges, and $p \in [0,1]$ is the set of edge weights. Specially an influence edge $\vec{e}_{uv}$ with the weight $p_{uv} \in [0,1]$ represents that a node u can influence another node v with a probability p_{uv}. The dynamic process of influence being spread from a set of seeds S in discrete rounds is as following: (1) in the first round, we mark all seeds as *active*; (2) in round afterwards, all nodes marked as *active* newly in last round have one chance to influence its uninfluenced neighbors and mark them as *active* following the influence edges with the probability of edge weight; (3) the process terminates when no new nodes can be marked as *active*.

In the stochastic dynamic process above, marking the set of all nodes as *active* is a possible realization of influence spread. Alternatively, there is an equivalent formulation in live-edge theory [1], which generates the live-edge graph to produce a final instance of influence spread. A live-edge graph $\mathcal{G}$ is randomly realized from $\mathcal{G}$ by removing each influence edge $\vec{e}_{uv}$ with a Bernoulli probability of $(1-p_{uv})$. We use Ω to denote the sample space of all possible live-edge graphs, then we have the probability distribution of $g \in \Omega$ as following:

$$Pr[\mathcal{G}] = \prod_{\vec{e}_{uv} \in g} p_{uv} \prod_{\vec{e}_{uv} \notin g} (1 - p_{uv}) \tag{3}$$

We denote $\mathcal{G}_S$ as the set of all nodes reachable within the graph $\mathcal{G}$ by any seed node in seed set S. Then $\mathcal{G}_S$ represents a potential instance of influence spread.

3.2. Submodular optimization

Let V be a ground set. Given a set function $f : 2^V \to \mathbb{R}$, we claim f is nonnegative, submodular and monotone non-decreasing as following:

- **Nonnegative**: For every $S \subseteq T$, we have that $f(S) \geq 0$.
- **Submodular**: For every $S \subseteq T$ and any $x \in V/T$, we have that $f(S \cup \{x\}) - f(S) \geq f(T \cup \{x\}) - f(T)$.
- **Monotonic non-decreasing** : For every $S \subseteq T$, we have that $f(S) \leq f(T)$.

The maximization optimization problem constrained by set size for a set function with all three aforementioned properties can be solved in designing of approximation algorithms. The basic greedy method of repeatedly adding the element with the maximum marginal gain into the solution set can yield a $(1-1/e)$-approximation [22]. In this paper, we will prove that the objective function in the maximization version of original problem has such three good properties, hence we can solve the optimization problem through the greedy framework.

4. Problem statement

Before claiming our problem, we first declare a list of prerequisites clearly as follows:

- There is an influence graph $G^i(V^i, E^i, p^i)$ on IC model to model how the latent mediators are influenced to be real mediators.
- There is a conflict graph $G^c(V^c, E^c)$ to illustrate the online conflicts, where V^c is the set of all nodes involved in the conflict, which consists of both attackers denoted by a set $V^{c,a}$ and sufferers denoted by a set $V^{c,s}$, and each direct edge in E^c represents an attack from an attacker to a sufferer.
- There is a close friend graph $G^f(V^f, E^f)$ which is a bipartite graph to illustrate all attackers' close friends who may be influenced to be mediators, where $V^f = V^i \cup V^{c,a}$, and each direct edge in E^c represents a close friendship between an attacker in $V^{c,a}$ and a latent mediator in V^i.
- For each attacker u, there is a weight b_u to measure the Bernoulli probability of u being persuaded by his/her any close friend. This probability is a premise parameter in our paper, and it can be measured from history record or predicted by machine learning.
- There is a set of all alternative seeds $\mathbb{S} \subseteq V^i$, and we consider choosing budgeted seeds from $\mathbb{S}$.

We combine the aforementioned three graphs into a heterogeneous graph called the **influence-conflict-close** graph, denoted by $G(V, E, p)$. Here, $V^f = V^i \cup V^c$ represents the set of nodes, $E = E^i \cup E^c \cup E^f$ represents the set of edges, and $p = p^i \cup p^c$ represents the set of probabilities.

Fig. 2 illustrates an instance of this graph. In the figure, the green nodes numbered from 1 to 11 represent latent mediators in the influence network. The green weighted arrows indicate directed influences among them with corresponding influence probabilities. Notably, nodes 1, 2, and 3 serve as alternative seeds. Nodes 12 to 18 represent individuals involved in the conflict, while the red arrows among them represent directed conflicts from attackers to sufferers. Specifically, nodes 12 to 16 are attackers, while nodes 14 to 18 are sufferers. The weighted black nodes from 14 to 18 represent the probabilities of sufferers being persuaded by a close friend, and the red-black nodes (14 to 16) indicate individuals who are both attackers and sufferers. Additionally, the undirected green lines represent close friendships among latent mediators and attackers.

Under the premise of this preparation, we give our final problem definition as following:

Problem 1. Minimizing the attacks to sufferers (MAS): The MAS problem is to choose a set S^* of at most k nodes from $\mathbb{S}$ as seeds to spread the influence to enable individuals in V^i to act as mediators to persuade her/his friend in $V^{c,a}$ to stop attacking sufferers in $V^{c,s}$ and further minimize the expected attacks to all sufferers in $V^{c,s}$, i.e.,

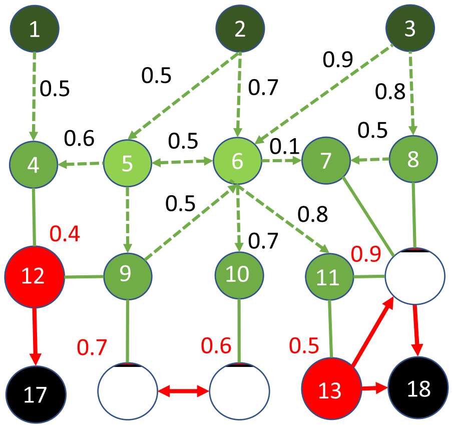


Fig. 2. An instance of influence-conflict-close heterogeneous graph.

$$S^* := \underset{S\subseteq\mathbb{S},|S|\le k}{argmin} \sum_{v\in V^{c,s}} E_S[w^*(v)], \tag{4}$$

where $V^{c,s}$ denotes all sufferers in the conflict, $E_S[\cdot]$ denotes the expected operator under the premise of a given seed set S, and $w^*(v)$ is computed as Equation (2).

Note that $w^*(v) = 1 - w(v)$, where $w(v) = \frac{|\mathcal{A}_v^\star|}{|\mathcal{A}_v|}$ and $\mathcal{A}_v^\star$ is the set of attackers in $\mathcal{A}_v$ who are persuaded to stop attacking v. Then we have $E_S[w^*(v)] = 1 - E_S[w(v)]$, and the MAS problem is equivalent to the following maximization problem:

Problem 2. The maximization optimization version of MAS (M-MAS) is defined as below:

$$S^* := \underset{S\subseteq\mathbb{S},|S|\le k}{argmax} \sum_{v\in V^{c,s}} E_S[w(v)] \tag{5}$$

According to the live-edge theory for the IC model, by generating a live-edge graph $\mathcal{G}$ from the live-edge graphs sample space Ω, we can capture $\mathcal{G}_S$ as a possible influence realization which represents a set of nodes being influenced to act as mediators by seeds set S. Let $E(|\mathcal{A}_v^\star|)$ be the set of nodes in $\mathcal{A}_v$ who are persuaded to stop attacking v in $\mathcal{G}_S$. So we can unfold $E_S[w(v)]$ as follows:

$$\begin{aligned} E_S[w(v)] &= \sum_{\mathcal{G}\in\Omega} Pr[\mathcal{G}]\frac{E(|\mathcal{A}_v^\star|)}{|\mathcal{A}_v|} \\ &= \sum_{\mathcal{G}\in\Omega} Pr[\mathcal{G}]\frac{\sum_{u\in\mathcal{A}_v} p_a(u)}{|\mathcal{A}_v|} \\ &= \sum_{\mathcal{G}\in\Omega} Pr[\mathcal{G}]\frac{\sum_{u\in\mathcal{A}_v} 1-(1-b_u)^{|\mathcal{G}_S\cap f_u|}}{|\mathcal{A}_v|} \\ &= \sum_{\mathcal{G}\in\Omega} Pr[\mathcal{G}]\frac{\sum_{u\in\mathcal{A}_v} 1-b_u^{*f_u(S)}}{|\mathcal{A}_v|}. \end{aligned} \tag{6}$$

Specially we mark $b_u^* := 1 - b_u$ and $f_u(S) := |\mathcal{G}_S \cap \mathcal{F}_u|$. Further let $\sigma(S) := \sum_{v\in V^{c,s}} E_S[Ws(v)]$ denote the objective function. For its maximization problem in the constrained setting of set size, we encounter the following hardness results:

Theorem 1. *The M-MAS problem is NP-hard and the computation problem of objective function σ is #P-hard.*

Proof. Consider a special case of MAS as shown in Fig. 3 where each node v in influence graph G^i corresponds to a close friend u which is attacking a sufferer s and we set the Bernoulli probability of u being persuaded by a close friend $b_u := 1$. We have that for each sufferer s, its attack-pacification $w(s)$ is either 1 or 0 and we have that the probability of $w(s) = 1$ corresponds to the probability that v is influenced. Then the objective function is $\sigma(S) := \sum_{s\in V^{c,a}} E_S[w(s)] = \sum_{v\in V^i} Pr\{g|v \in \mathcal{G}_S\} = E[|I(S)|]$, where $E[|I(S)|]$ is the expected number of nodes being influenced by seeds set S. Therefore this special case is equal to the influence maximization problem which is NP-hard and the computation of influence $I(S)$ is #P-hard. □

Theorem 2. *The objective function σ is nonnegative, monotonic non-decreasing and submodular.*

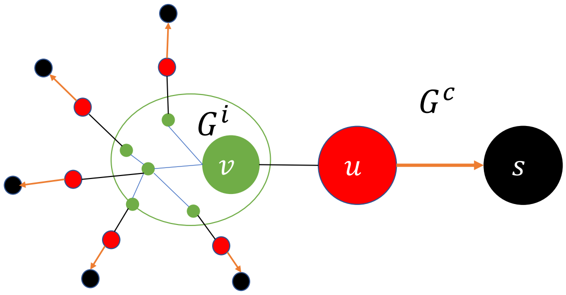


Fig. 3. A special case of MAS to illustrate its hardness.

Proof. The objective function σ is nonnegative obviously. To prove it's monotone and submodular, we mark $f_v(S) := E_S[w(v)]$, and we just need to prove $f_v(\cdot)$ is monotone and submodular.
Monotonic non-decreasing: Note the fact that for any $S \subseteq T$, we have $\mathcal{G}_S \subseteq \mathcal{G}_T$. Therefore we have $\mathcal{G}_S \cap \mathcal{F}_u \subseteq \mathcal{G}_T \cap \mathcal{F}_u$ and $f_u(S) \leq f_u(T)$. So $E_S(w(v)) \leq E_T(w(v))$ and the set function $f_v(\cdot)$ is monotonic non-decreasing.
Submodular: Let $\Delta_x f_v(S)$ denote the incremental value of objective $f_v(S)$ after adding a single element x to S and we have:

$$\Delta_x f_v(S) = \sum_{\mathcal{G}\in\Omega} Pr[\mathcal{G}] \frac{\sum_{u\in\mathcal{A}_v} b_u^{*f_u(S)} - b_u^{*f_u(S\cup\{x\})}}{\mathcal{A}_v}. \tag{7}$$

We aim to prove $\Delta_x f_v(S) \geq \Delta_x f_v(T)$ and then $f_v(\cdot)$ is submodular. It will be a direct conclusion if we prove the claim: $b_u^{*f_u(S)} - b_u^{*f_u(S\cup\{x\})} \geq b_u^{*f_u(T)} - b_u^{*f_u(T\cup\{x\})}$. Note that since $f_u(T\cup\{x\}) - f_u(T) = |(\mathcal{G}_{\{x\}}/\mathcal{G}_T) \cap \mathcal{F}_u|$, $f_u(S\cup\{x\}) - f_u(S) = |(\mathcal{G}_{\{x\}}/\mathcal{G}_S) \cap \mathcal{F}_u|$ and $\mathcal{G}_{\{x\}}/\mathcal{G}_T \subseteq \mathcal{G}_{\{x\}}/\mathcal{G}_S$, we have a fact that $f_u(T\cup\{x\}) - f_u(T) \leq f_u(S\cup\{x\}) - f_u(S)$, i.e., $f_u(\cdot)$ is submodular. Nextly with this fact, we can prove the aforementioned claim as following:

$$\begin{aligned}
& b_u^{*f_u(S)} - b_u^{*f_u(S\cup\{x\})} - (b_u^{*f_u(T\cup\{x\})} - b_u^{*f_u(T)}) \\
=& b_u^{*f_u(S)}(1 + b_u^{*f_u(T)-f_u(S)} - b_u^{*f_u(S\cup\{x\})-f_u(S)} - b_u^{*f_u(T\cup\{x\})-f_u(S)}) \\
=& b_u^{*f_u(S)}(1 + b_u^{*f_u(T)-f_u(S)}(1 - b_u^{*f_u(T\cup\{x\})-f_u(T)}) - b_u^{*f_u(S\cup\{x\})-f_u(S)}) \\
\geq& b_u^{*f_u(S)}(1 + b_u^{*f_u(T)-f_u(S)}(1 - b_u^{*f_u(T\cup\{x\})-f_u(T)}) - b_u^{*f_u(T\cup\{x\})-f_u(T)}) \\
\geq& b_u^{*f_u(S)}(1 + b_u^{*f_u(T)-f_u(S)})(1 - b_u^{*f_u(T\cup\{x\})-f_u(T)}) \geq 0
\end{aligned}$$

So we have proved $f_v(\cdot)$ is submodular. □

5. Proposed algorithm

In this section, since computing the objective function is #P-hard, obtaining the exact value of the objective function is challenging. Therefore, to effectively estimate the objective function, we first consider the Reverse Sampling Technique and introduce a Reverse Sample, which is a collection of sets. We then propose our algorithm.

5.1. Reverse sampling technique

We first introduce a reverse sampling technique as follows to generate a reverse sample for solving the estimation problem related to the objective function.

Definition 1 *(Reverse Sample)*. Given an **influence-conflict-close** heterogeneous graph, a reverse sample is generated using the following random process:

- Step (1): Sample a sufferer v from all sufferers by the probability distribution of $p_v := \frac{|\mathcal{A}_v|^{-1}}{\sum_{v\in V^{c,s}} |\mathcal{A}_v|^{-1}}$.
- Step (2): For each attacker u who is attacking the sufferer v, we select a set of u's close friends one by one randomly following the Bernoulli probability b_u and mark the result of set as $\mathcal{P}_u$.
- Step (3): Do a randomly multi-sources breadth-first search from source nodes in $\cup_{u\in\mathcal{A}_v} \mathcal{P}_u$ reversely and randomly over the influence graph $G^i(V^i, E^i, p)$.
- Step (4): We denote the set of all alternative seeds searched by source nodes in $\mathcal{P}_u$ denoted by R_u and call the collection $\mathcal{R} := \{R_u | u \in A_u\}$ as a reverse sample.

Specifically, we put step (1) and (2) in Algorithm 1, step (3) in Algorithm 2. and step (4) in Algorithm 3. For the sake of expla-

Algorithm 1: Selecting Source Nodes.

Pick a node v from all sufferers with the probability of p_v;
for *each attacker* u *who is attacking* v **do**
  $\mathcal{P}_u \leftarrow \emptyset$;
  for *each* u*'s close friend* s **do**
    add s into $\mathcal{P}_u$ with the probability of b_u;
return $\mathbb{P}_v := \{\mathcal{P}_u, u \in \mathcal{A}_v\}$

Algorithm 2: Multi-sourced breadth-first searching ($\mathbb{P}_v$).

Execute Algorithm 1 and get the result of a family of sourced sets $\mathbb{P}_v$;
$Seen_s \leftarrow \{u | s \in \mathcal{P}_u\}$;
$Visit = \{(s, \{b_u\}) | s \in \mathcal{P}_u\}$;
while $Visit \neq \emptyset$ **do**
  $VisitNext \leftarrow \emptyset$;
  for *each* $s \in \bigcup_{(o,B_o)\in Visit}\{o\}$ **do**
    $B_s^* = \bigcup_{(s,B_s')\in Visit} B_s'$;
    for *each* s*'s in-neighbor* t *in* G^i **do**
      if *edge* $\vec{e}_{ts}$ *has not been flipped* **then**
        Flip $\vec{e}_{ts}$ with probability p_{ts}^i;
      if e_{ts} *has been "on"* **then**
        $B_t \leftarrow B_s^* / \{b_o | o \in Seen_t\}$;
        if $B_t \neq \emptyset$ **then**
          $VisitNext \leftarrow VisitNext \cup (t, B_t)$;
          $Seen_t \leftarrow Seen_t \cup \{o | b_o \in B_t\}$;
  $Visit \leftarrow VisitNext$;
return $\{Seen_s | s \in \mathbb{S}\}$

Algorithm 3: Sampling a reverse sample.

Execute Algorithm 1 and get the result of a family of sourced sets $\mathbb{P}_v^*$;
Execute Algorithm 2 with $\mathbb{P}_v$ and get the result of searching $\{Seen_s | s \in \mathbb{S}\}$;
$r_u \leftarrow \emptyset$ for each $\mathcal{P}_u \neq \emptyset$;
for *each* $s \in \mathbb{S}$ **do**
  for *each* $u \in Seen_s$ **do**
    $r_u \leftarrow r_u \cup \{s\}$;
return $\{r_u\}$

nation and better understanding, we provide an example of sampling a reverse sample in Fig. 4 based on the instance shown in Fig. 2. In step (1), we choose a sufferer randomly (node 8) as drawn in Fig. 4(a). In step (2), we choose close friends for attackers 13,16 randomly and pick nodes $\{11\}, \{8, 11\}$ respectively as shown in Fig. 4(b). Let b_u represent a BFS process rooted from u, $Seen_s$ record the set of indices of BFS processes that have searched s, $VisitNext$ record the set of starting nodes in the next BFS level. As shown in Fig. 4(c), after initializing the multi-sourced BFS randomly reversely, we have $Seen_{11} = \{13, 16\}$, $Seen_8 = \{16\}$ and $VisitNext = \{(11, \{b_{13}, b_{16}\}), (8, \{b_{16}\})\}$. Fig. 4(d) draws the first level of BFS with a result of $Seen_6 = \{13, 16\}$, $Seen_3 = \{16\}$ and $VisitNext = \{(6, \{b_{13}, b_{16}\}), (3, \{b_{16}\})\}$. Fig. 4(e) draws the second level of BFS with a result of $Seen_9 = \{13, 16\}$, $Seen_2 = \{16\}$ and updated $Seen_3 = \{13, 16\}$, $VisitNext = \{(9, \{b_{13}, b_{16}\}), (2, \{b_{16}\}), (3, \{b_{13}\})\}$. Fig. 4(f) draws the last level of BFS with a result of $Seen_5 = \{13, 16\}$, $VisitNext = \emptyset$. Finally for step (4), we get a reverse sample $\{\{2, 3\}, \{2, 3\}\}$.

Note that performing a multi-sources BFS from source nodes in reverse and randomly on the influence graph G^i is equivalent to the process of first generating a live-edge graph $\mathcal{G}$ from G^i firstly and then conducting a multi-sources breadth-first searching from source nodes definitely over the reverse of $\mathcal{G}$, denoted as $\mathcal{G}^T$ to obtain a reverse sample $\mathcal{R} = \{\mathcal{G}^T_{\mathcal{P}_u} \cap \mathbb{S} | u \in A_u\}$. We can compute the union probability distribution for a reverse sample set $\mathcal{R}$ generated above with $v, \mathbb{P}_v, \mathcal{G}$ as follows:

$$Pr[\mathcal{R}] = Pr[v]Pr[\mathbb{P}_v]Pr[\mathcal{G}] = p_v \prod_{u\in\mathcal{A}_v} b_u^{|\mathcal{P}_u|}(1-b_u)^{|\mathcal{F}_u/\mathcal{P}_u|} Pr[\mathcal{G}] \tag{8}$$

Next, given the seed set S and a reverse sample $\mathcal{R}$, we define:

$$\xi_{\mathcal{R}}(S) := |\{R | R \cap S \neq \emptyset, R \in \mathcal{R}\}| \tag{9}$$

and we can get the following theorem:

Theorem 3. *Let* $E[\xi_{\mathcal{R}}(S)]$ *be the expectation of* $\xi_{\mathcal{R}}(S)$*, then we have* $\sigma(S) = \tau \cdot E[\xi_{\mathcal{R}}(S)]$*, where* $\tau = \sum_{v\in V^{c,s}} |\mathcal{A}_v|^{-1}$.

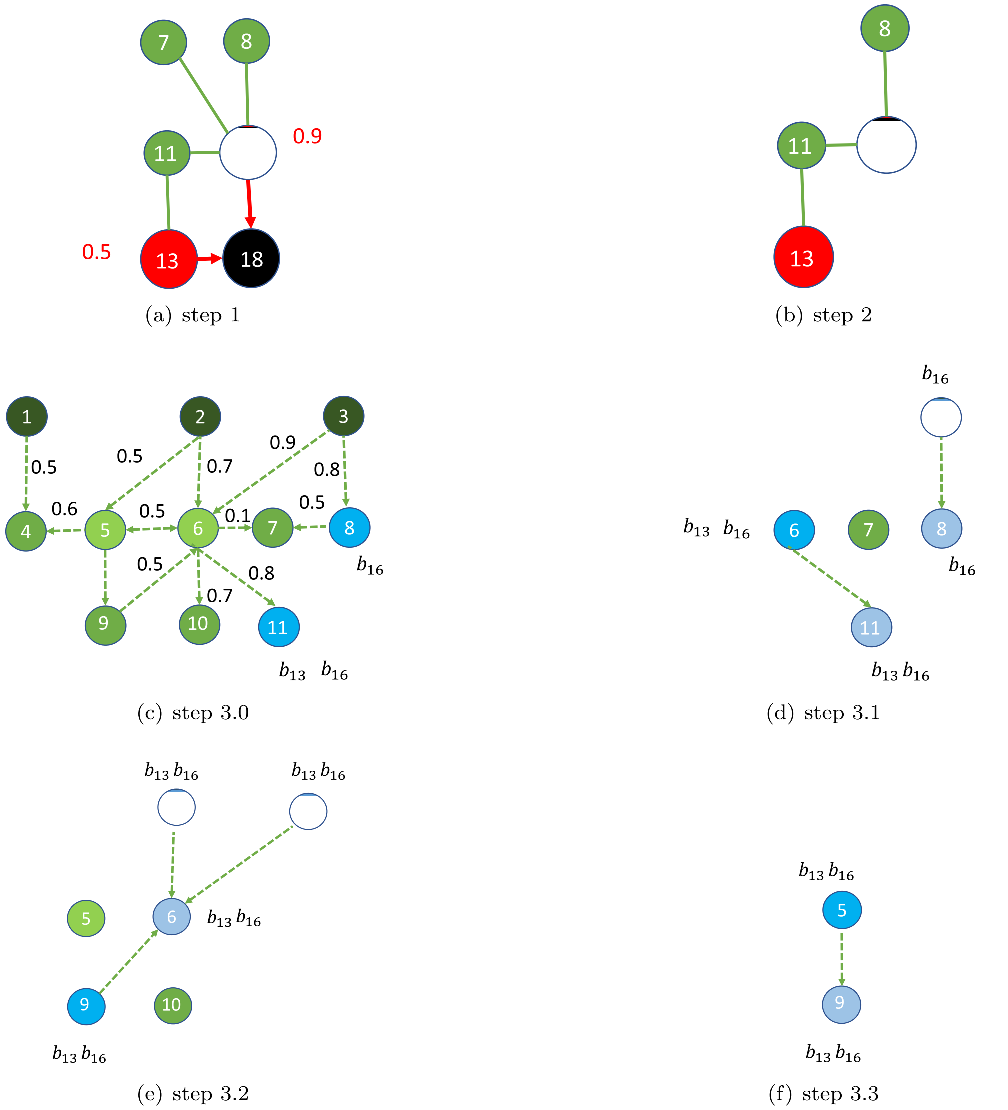


Fig. 4. An illustration of sampling algorithm.

Proof. Given a sufferer v and an attacker u who is attacking v, we have $\mathcal{G}_S \cap \mathcal{F}_u$ is an influence realization of set of u's close friends influenced to act as mediators over the live-edge graph $\mathcal{G}_S$ by seeds set S, then let $\mathcal{P}_u^*$ denote the set of nodes in $\mathcal{G}_S \cap \mathcal{F}_u$ that persuade u successfully. Then we can compute the probability distribution of the set family $\mathbb{P}_v^* := \{\mathcal{P}_u^*, u \in \mathcal{A}_v\}$ as $Pr[\mathbb{P}_v^*] = \prod_{u\in\mathcal{A}_v} b_u^{|\mathcal{P}_u^*|}(1-b_u)^{|\mathcal{F}_u \cap \mathcal{G}_S / \mathcal{P}_u^*|}$. Count the number of non-empty sets in $\mathbb{P}_v^*$ and denote it by $\alpha(\mathbb{P}_v^*)$ which is the number of attackers persuaded to stop attacking v. We have the expectation computed as follows:

$$\sigma(S) = \sum_{v\in V^{c,s}, \mathcal{G}\in\Omega} Pr[\mathcal{G}] \frac{\sum_{\mathbb{P}_v^*} Pr[\mathbb{P}_v^*]\alpha(\mathbb{P}_v^*)}{|\mathcal{A}_v|} \tag{10}$$

Next, we prove the following equation:

$$\sum_{\mathbb{P}_v^*} Pr[\mathbb{P}_v^*]\alpha(\mathbb{P}_v^*) = \sum_{\mathbb{P}_v} Pr[\mathbb{P}_v]\xi_{\mathcal{R}}(S), \tag{11}$$

where $\mathbb{P}_v := \{\mathcal{P}_u, u \in \mathcal{A}_v\}$ is the family of all sets like $\mathcal{P}_u$, which consists of all nodes picked with probability b_u one by one from the set of u's close friends $\mathcal{F}_u$. We have its probability distribution $Pr[\mathbb{P}_v] = \prod_{u\in\mathcal{A}_v} b_u^{|\mathcal{P}_u|}(1-b_u)^{|\mathcal{F}_u/\mathcal{P}_u|}$. Let $\mathcal{A}_v^*$ denote a subset of A_v and $Pr[\mathbb{P}_v|\mathcal{A}_v^*]$ denote the probability of $\mathbb{P}_v$ which satisfies following condition (1): $\mathcal{G}_{\mathcal{P}_u}^T \cap \mathbb{S} \cap S \neq \emptyset$ if $u \in \mathcal{A}_v^*$, otherwise $\mathcal{G}_{\mathcal{P}_u}^T \cap \mathbb{S} \cap S = \emptyset$. Hence we have the corresponding $\xi_{\mathcal{R}}(S) = |\mathcal{A}_v^*|$ and

$$\sum_{\mathbb{P}_v} Pr[\mathbb{P}_v]\xi_{\mathcal{R}}(S) = \sum_{\mathcal{A}_v^* \subseteq \mathcal{A}_v} Pr[\mathbb{P}_v|\mathcal{A}_v^*]|\mathcal{A}_v^*|. \tag{12}$$

Let $Pr[\mathbb{P}_v^*|\mathcal{A}_v^*]$ denote the probability of $\mathbb{P}_v^*$ which satisfies following condition (2): $\mathcal{P}_u^* \neq \emptyset$ if $u \in \mathcal{A}_v^*$, otherwise $\mathcal{P}_u^* = \emptyset$ and hence $\alpha(\mathbb{P}_v^*) = |\mathcal{A}_v^*|$. Then we have

$$\sum_{\mathbb{P}_v^*} Pr[\mathbb{P}_v^*]\alpha(\mathbb{P}_v^*) = \sum_{\mathcal{A}_v^* \subseteq \mathcal{A}_v} Pr[\mathbb{P}_v^*|\mathcal{A}_v^*]|\mathcal{A}_v^*|. \tag{13}$$

Actually, we can compute $Pr[\mathbb{P}_v^*|\mathcal{A}_v^*]$ as following:

$$Pr[\mathbb{P}_v^*|\mathcal{A}_v^*] = \prod_{u \in A_v^*} (1-(1-b_u)^{|\mathcal{G}_S \cap \mathcal{F}_u|}) \prod_{u \in A_v / A_v^*} b_u^{*|\mathcal{G}_S \cap \mathcal{F}_u|} = Pr[\mathbb{P}_v|\mathcal{A}_v^*]. \tag{14}$$

Since all families of $\mathbb{P}_v$ satisfying condition (1) equal to ones satisfying following condition (3): $\mathcal{P}_u \cap \mathcal{G}_S \cap \mathcal{F}_u \neq \emptyset$ if $u \in \mathcal{A}_v^*$, otherwise $\mathcal{P}_u \cap \mathcal{G}_S \cap \mathcal{F}_u = \emptyset$ if $u \notin \mathcal{A}_v^*$.

We have proved Equation (11), therefore, returning to Equation (10), we have

$$\begin{aligned}
\sigma(S) = & \sum_{v \in V^{c,s}, \mathcal{G} \in \Omega} Pr[\mathcal{G}] \frac{\sum_{\mathbb{P}_v} Pr[\mathbb{P}_v]\xi_{\mathcal{R}}(S)}{|\mathcal{A}_v|} \\
= & \tau \sum_{v,\mathcal{G}} \sum_{\mathbb{P}_v} Pr[\mathcal{R}]Pr[\mathbb{P}_v]Pr[v]\xi_{\mathcal{R}}(S) \\
= & \tau \sum_{v,\mathcal{G},\mathbb{P}_v} Pr[\mathcal{R}]\xi_{\mathcal{R}}(S) \\
= & \tau E(\xi_{\mathcal{R}}(S))
\end{aligned}$$

We proved it. □

5.2. Algorithms based on sampling

According to Theorem 3, we can sample a sufficient number of reverse samples to statistically estimate $E[\xi_{\mathcal{R}}(S)]$. Let λ be the total number of times the sampling Algorithm 3 is run, resulting in the set of all reverse samples $\mathbb{R} = \{\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_\lambda\}$. For a seed set $S \subseteq \mathbb{S}$, we compute $\eta(\mathbb{R}, S) := \sum_{\mathcal{R} \in \mathbb{R}} \xi_{\mathcal{R}}(S)$. Hence $\frac{\eta(\mathbb{R},S)}{\lambda}$ is an estimation for $E(\xi_{\mathcal{R}}(S))$, and $\frac{\tau\eta(\mathbb{R},S)}{\lambda}$ is an estimation for $\sigma(S)$, the original objective function in Problem 2. Hence we consider following problem:

Problem 3. Given a family of reverse samples $\mathbb{R}$, find a seed set $S^\star \subseteq \mathbb{S}$, s.t., $S^\star = argmax_{S \subseteq \mathbb{S}, |S| \leq k} \eta(\mathbb{R}, S)$.

For two sets, we say that a set hits another set if the intersection set of them is not empty. Since $\xi_{\mathcal{R}}(S)$ is the number of sets hit by S in a set family $\mathcal{R}$, we have an equivalent computation for $\eta(\mathbb{R}, S)$ by counting the number of sets in the set family $\mathcal{R}^*$ hit by S, where $\mathcal{R}^* := \{R | R \in \mathcal{R}, \mathcal{R} \in \mathbb{R}\}$ is the collection of all the sets from all reverse samples we sampled. Thus, Problem 3 is a classical set cover problem [22]. With respect of seeds set S, $\eta(\mathbb{R}, S)$ is submodular, nonnegative and monotonic non-decreasing. The general greedy selecting as Algorithm 4 can provide a solution with $(1-1/e)$-approximation.

Algorithm 4: Seeds selection $(k, \mathcal{R}^*)$.

```
Let S ← ∅;
for i from 1 to k do
    Get a node s_i := argmax_{s∈𝕊/S} |{R | s ∈ R, R ∈ R*}|;
    Update S ← S ∪ {s_i};
    Remove all sets nit by S from R*;
return S
```

Let $m = max\{|\mathcal{A}_v| | v \in V^{c,s}\}$, which is the maximum number of attackers among sufferers. Then the reverse sample has at most m sets. Next, we analyze the quality of the solution $S^\star$ for Problem 2, and we can state a theorem as follows:

Theorem 4. *let S^* be the optimal solution of Problem 2 with the optimal value OPT, λ be the number of reverse samples, and $S^\star$ be the solution provided by Algorithm 4. If any one of following inequalities holds:*

$$\exp\left(-\frac{2\epsilon_1^2 \eta(\mathbb{R}, S^\star)}{(1+\epsilon_1)m^2}\right) + \exp\left(-\frac{\epsilon_2^2 \eta(\mathbb{R}, S^\star)}{\tau(1+\epsilon_1)m^2}\right) \leq \delta, \tag{15}$$

$$\exp\left(-\frac{2\epsilon_1^2 \lambda}{\tau^2 m^2}\right) + \exp\left(-\frac{\epsilon_2^2 \lambda}{\tau^2 m^2}\right) \leq \delta, \tag{16}$$

then we have $\sigma(S^\star) \geq \frac{(1-\epsilon_2)}{(1+\epsilon_1)}(1-1/e)OPT$ *with a probability at least* $1-\delta$*, where* $\delta, \epsilon_1, \epsilon_2 \in (0,1)$.

Proof. Firstly, we introduce Chernoff Bounds [21]:

Let $X_1, X_2, \cdots, X_\lambda$ be random variables such that $a \leq X_i \leq b$ for all i. Let $X = \sum_{i=1}^{\lambda} X_i$ and $\mu = E(X)$. Then for any $\epsilon \geq 0$, we have:

$$Pr[X-\mu \geq \epsilon\mu] \leq \exp(-\frac{2\epsilon^2\mu^2}{\lambda(b-a)^2}), \tag{17}$$

$$Pr[X-\mu \leq -\epsilon\mu] \leq \exp(-\frac{\epsilon^2\mu^2}{\lambda(b-a)^2}). \tag{18}$$

Suppose that the random variable $X(S) := \xi_{\mathcal{R}}(S)$ is in a range $[l(S), u(S)]$, $0 \leq l(S) \leq u(S)$. We have $u(S) \leq m$ since each reverse sample has at most of $|\mathcal{A}_v|$ sets. Hence $u(S) - l(S) \leq m$. Based on the λ times of sampling reverse sample, we have two series of λ independent and identically distributed random variables $\{X_i(S^\star), i \in [\lambda]\}$, and $\{X_i(S^*), i \in [\lambda]\}$ respectively. Then $X(S^\star) := \sum_{i=1}^{\lambda} X_i(S^\star) = \eta(\mathbb{R}, S^\star)$, $\mu(S^\star) = E(X(S^\star)) = \sum_{i=1}^{\lambda} E(X_i(S^\star)) = \lambda\tau^{-1}\sigma(S^\star)$, and similarly $X(S^*) = \eta(\mathbb{R}, S^*)$ and $\mu(S^*) = \lambda\tau^{-1}\sigma(S^*)$. We naturally have $\sigma(S^\star), \sigma(S^\star) \geq 1$. Then by Inequality (17),

$$Pr[\eta(\mathbb{R}, S^\star) \geq (1+\epsilon_1)\lambda\tau^{-1}\sigma(S^\star)] \leq \exp(-\frac{2\epsilon_1^2\lambda\tau^{-2}\sigma(S^\star)^2}{(u(S^\star)-l(S^\star))^2}) \leq \exp(-\frac{2\epsilon_1^2\lambda\tau^{-2}\sigma(S^\star)}{m^2})$$
$$= \delta_1 \leq \exp(-\frac{2\epsilon_1^2\lambda\tau^{-2}}{m^2}) = \delta_3$$

and similarly by Inequality (18),

$$Pr[\eta(\mathbb{R}, S^*) \leq (1-\epsilon_2)\lambda\tau^{-1}\sigma(S^*)] \leq \exp(-\frac{\epsilon_2\lambda\tau^{-2}\sigma(S^*)^2}{(u(S^*)-l(S^*))^2}) \leq \exp(-\frac{\epsilon_2^2\lambda\tau^{-2}\sigma(S^*)^2}{m^2})$$
$$\leq \exp(-\frac{\epsilon_2^2\lambda\tau^{-2}\sigma(S^\star)}{m^2})$$
$$= \delta_2 \leq \exp(-\frac{\epsilon_2^2\lambda\tau^{-2}}{m^2}) = \delta_4$$

Suppose that inequalities $\eta(\mathbb{R}, S^\star) \leq (1+\epsilon_1)\lambda\tau^{-1}\sigma(S^\star)$ and $\eta(\mathbb{R}, S^*) \geq (1-\epsilon_2)\lambda\tau^{-1}\sigma(S^*)$ hold at the same time, and it occurs with a probability more than $1-(\delta_1+\delta_2)$ and $1-(\delta_3+\delta_4)$. Since Algorithm 4 guarantees that $\eta(\mathbb{R}, S^\star) \geq (1-1/e)OPT^1$, where OPT^1 is the optimal value of Problem 3, and $OPT^1 \geq \eta(\mathbb{R}, S^*)$, we have

$$\sigma(S^\star) \geq \frac{\tau}{(1+\epsilon_1)\lambda}\eta(\mathbb{R}, S^\star) \geq \frac{\tau(1-1/e)}{(1+\epsilon_1)\lambda}OPT^1 \geq \frac{\tau(1-1/e)}{(1+\epsilon_1)\lambda}\eta(\mathbb{R}, S^*) \geq \frac{(1-\epsilon_2)(1-1/e)}{(1+\epsilon_1)}OPT \tag{19}$$

holds with a probability more than $1-(\delta_1+\delta_2)$ and $1-(\delta_3+\delta_4)$.

- If Inequality (15) holds, by $\eta(\mathbb{R}, S^\star) \leq (1+\epsilon_1)\lambda\tau^{-1}\sigma(S^\star)$, i.e., $\frac{1}{1+\epsilon_1}\eta(\mathbb{R}, S^\star) \leq \lambda\tau^{-1}\sigma(S^\star)$, we have $\delta_1 \leq exp(-\frac{2\epsilon_1^2\eta(\mathbb{R},S^\star)}{\tau(1+\epsilon_1)m^2})$ and $\delta_2 \leq exp(-\frac{2\epsilon_2^2\eta(\mathbb{R},S^\star)}{\tau(1+\epsilon_1)m^2})$. Hence $1-(\delta_1+\delta_2) \geq 1-(exp(-\frac{2\epsilon_1^2\eta(\mathbb{R},S^\star)}{\tau(1+\epsilon_1)m^2}) + exp(-\frac{\epsilon_2^2\eta(\mathbb{R},S^\star)}{\tau(1+\epsilon_1)m^2})) \geq 1-\delta$.
- If Inequality (16) holds, we have $1-(\delta_3+\delta 4) = (1-(\exp(-\frac{2\epsilon_1^2\lambda}{\tau^2 m^2}) + \exp(-\frac{\epsilon_2^2\lambda}{\tau^2 m^2})) \geq 1-\delta$.

Thus as long as Inequality (15) or Inequality (16) holds, we can guarantee that Inequality (19) holds with a probability at least $1-\delta$. □

We propose an algorithm shown in Algorithm 5 that utilizes an incremental sampling approach for progressive seed selection, allowing for early termination to reduce the number of samples. Within a given number of iterations which can guarantee Inequality (15), in each iteration, we check whether it satisfies Inequality (16) to decide whether to continue doubling the number of sampling iterations. We can have the following theorem:

Theorem 5. *The Algorithm 5 provides a solution* $S^\star$ *for Problem 2 satisfying:* $\sigma(S^\star) \geq \epsilon(1-1/e)OPT$ *with a probability at least* $1-\delta$*, where* $\delta, \epsilon \in (0,1)$*, and* OPT *is the optimal value of Problem 2.*

Proof. Let $\epsilon_1 = \epsilon^*, \epsilon_2 = \sqrt{2}\epsilon^*$. By Theorem 4, if algorithm returns early in the iteration, i.e., $\eta(\mathbb{R}, S^\star) \geq \frac{\tau(\ln 2 - \ln\delta)m^2(1+\epsilon^*)}{2\epsilon^{*2}}$, then we can infer the Inequality (15) as

$$\exp(-\frac{2\epsilon_1^2\eta(\mathbb{R}, S^\star)}{\tau(1+\epsilon_1)m^2}) + \exp(-\frac{\epsilon_2^2\eta(\mathbb{R}, S^\star)}{\tau(1+\epsilon_1)m^2}) = 2exp(-\frac{2\epsilon_*^2\eta(\mathbb{R}, S^\star)}{\tau(1+\epsilon^*)m^2}) \leq \delta$$

holds. If algorithm returns after all iterations, we have the number of reverse samples are at least λ^*, and we have Inequality (16) as

Algorithm 5: Flying Sampling-Selecting (ϵ, δ).

$\epsilon^* = \frac{1-\epsilon}{\sqrt{2}+\epsilon}$;
$\lambda^* = -\frac{\tau^2 m^2 \ln \delta}{2\epsilon^{*2}}$;
$T = \lceil log_2(\lambda) \rceil$;
$\mathcal{R} \leftarrow \emptyset$;
for *each i from 1 to T-1* **do**
  Sample new 2^i reverse samples into R;
  Select seeds $S_i^\star$ by Algorithm 4;
  if $\eta(\mathbb{R}, S^\star) \geq \frac{\tau(\ln 2 - \ln \delta) m^2 (1+\epsilon^*)}{2\epsilon^{*2}}$ **then**
    return $S_i^\star$;
return $S_{T-1}^\star$

Table 1
Datasets.

Dataset	#Nodes	#Edges	(#Attackers,#Sufferers)		
			5%	10%	20%
Facebook	4K	80K	(0.07K,0.13K)	(0.15K,0.25K)	(0.3K,0.5K)
Flickr	80K	590K	(2K,2K)	(3K,5K)	(9K,7K)
DBLP	203K	382K	(4K,6K)	(10K,10K)	(25K,15K)
Twitter	580K	717K	(15K,14K)	(30K,28K)	(50K,66K)

$$\exp(-\frac{2\epsilon_1^2 \lambda}{\tau^2 m^2}) + \exp(-\frac{\epsilon_2^2 \lambda}{\tau^2 m^2}) = 2\exp(-\frac{2\epsilon^{*2}\lambda}{\tau^2 m^2}) \leq \delta$$

holds. So we have $\sigma(S^\star) \geq \epsilon(1 - 1/e)OPT$ with a probability at least $1 - \delta$, where $\epsilon = \frac{(1-\epsilon_2)}{(1+\epsilon_1)}$. □

6. Experiments

In this section, we conduct experimental studies to evaluate the performance of our proposed methods over 4 real-world datasets[1] (facebook, Flickr, DBLP, Twitter) and we provide the details in Table 1. First, we introduce the experiment setup, then we compare our method with other existing approaches, after that, we analyze and discuss the result from different perspectives.

6.1. Setup

For each network in the datasets, we randomly and uniformly select 5%, 10%, and 20% of the nodes to form the conflict graph. The conflict edges are generated randomly to create a directed graph representing the directed conflicts. The specific details are outlined in Table 1. We then set the vertex-induced subgraph with remaining 90% nodes in the original network as the influence graph.

Following general settings, we set the influence probability i.e. the weight of each direct influence edge from u to v with $D^{-1}(v)$, where $D(v)$ is the in-going degree of node v. For each attacker, we set all its neighbors in original network without conflicts as its close friends and assume a Bernoulli probability of attacker being persuaded by its close friend to be 0.5.

In particular, we designate all nodes in the influence graph as alternative seeds. To evaluate the quality of a seed set, we perform 10 000 simulations of Monte Carlo to simulate the stochastic process of influencing and mediating. We record the attack-level for each instance and calculate the average as the quality measure for a seed set. It's important to note that a smaller value indicates better quality.

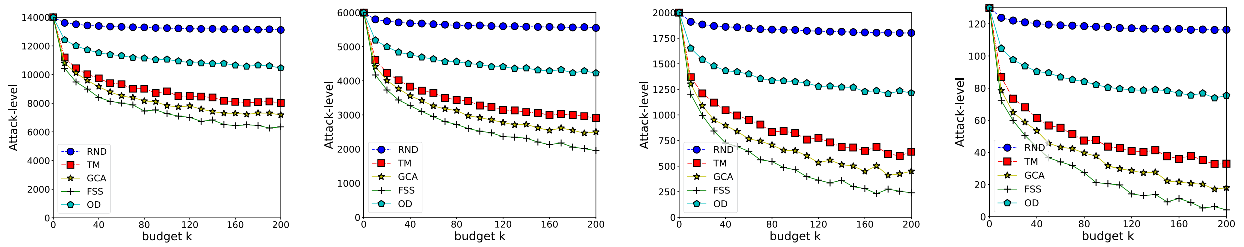
We compare our algorithms with some methods as following.

- **TM** ([31]): The algorithm proposed to solve the targeted influence, we set the targeted nodes to be all attackers' close friends.
- **GCA**: The general greedy algorithm to add seeds by the Monte Carlo simulations to estimate the objective function.
- **FSS**: Algorithm 5 we proposed based on the reverse sample and specially, to guarantee the quality, we set $\epsilon = 0.9$ and $\delta = 0.1$.
- **OD**: Select the top k nodes with the maximum out degree in influence graph.
- **RND**: The basic method by choosing seeds randomly.

6.2. Result

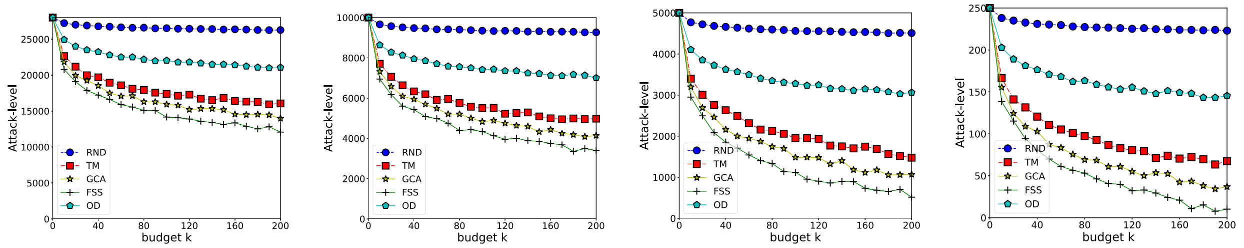
Effectiveness Comparison: Firstly, we compare the quality of seeds. For each dataset, we set the budget, i.e., the number of seeds k, in the range of $[0, 200]$, to evaluate the relationship between seed quality and budget. To evaluate the effectiveness of

[1] http://networkrepository.com.



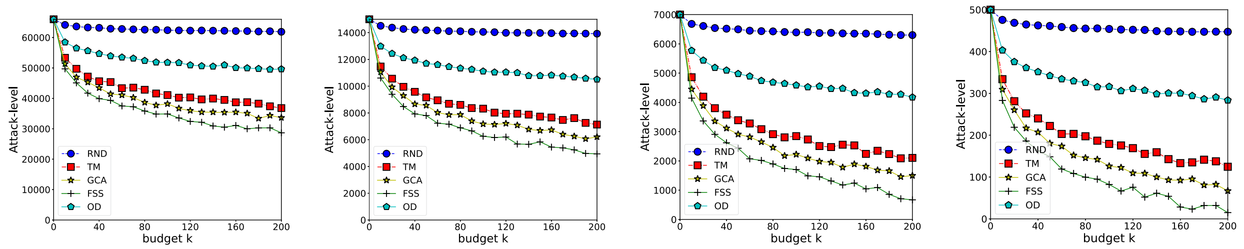
(a) Twitter (b) DBLP (c) Flickr (d) Facebook

Fig. 5. The comparison of the effectiveness of seed selection achieved by different algorithms on the Twitter, DBLP, Flickr, and Facebook datasets, considering 5% of the nodes involved in the conflict.



(a) Twitter (b) DBLP (c) Flickr (d) Facebook

Fig. 6. The comparison of the effectiveness of seed selection achieved by different algorithms on the Twitter, DBLP, Flickr, and Facebook datasets, considering 10% of the nodes involved in the conflict.

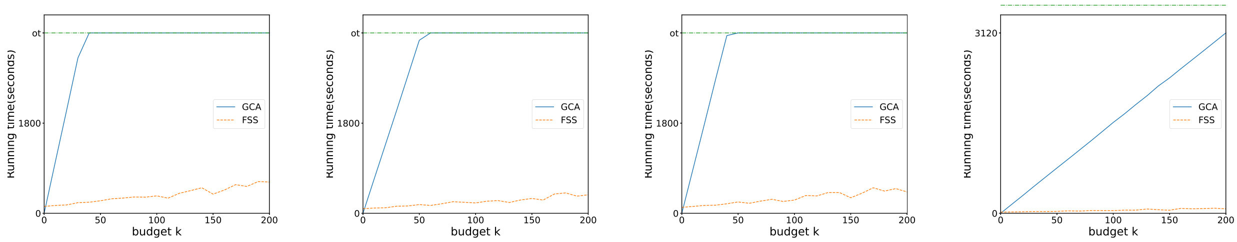


(a) Twitter (b) DBLP (c) Flickr (d) Facebook

Fig. 7. The comparison of the effectiveness of seed selection achieved by different algorithms on the Twitter, DBLP, Flickr, and Facebook datasets, considering 20% of the nodes involved in the conflict.

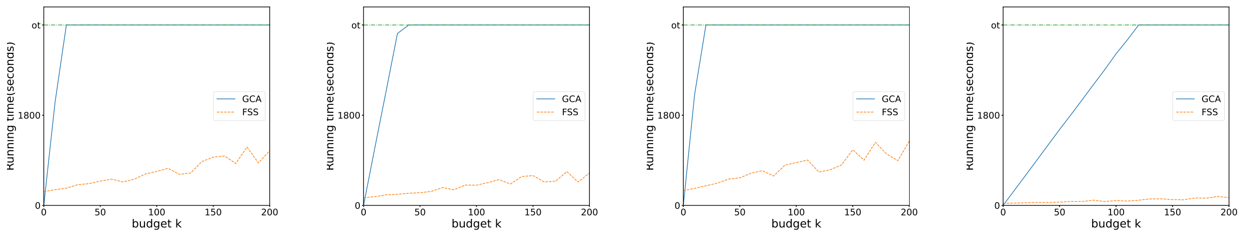
Algorithm 5, we compare it with TM, GCA, OD, and RND by executing each algorithm 10 times. Fig. 5, 6, and 7 depict the average results for varying numbers of sufferers and attackers at 5%, 10%, and 20% involvement rates, respectively. In each subfigure, the horizontal axis represents the budget k, while the vertical axis represents the attack level after spreading the influence of mediating with seeds provided by these algorithms for comparison. Specially, the value on the vertical axis represents the original attack level without any mediation at the horizontal axis, where $k = 0$. Our observations are as follows: (1) As the budget increases, the attack level σ^* decreases, indicating that more sufferers are protected from attacks. This is expected, as we have proven that the objective function $\sigma = 1 - \sigma^*$ in Problem (2) is non-decreasing. (2) When dealing with real network data involving conflict at different scales, our algorithm can provide the best seed strategy within the given budget. Greedy algorithms (GCA and FSS) significantly outperform heuristic algorithms (TM, OD) because the greedy method guarantees a solution of high quality with a $(1 - 1/e)$ approximation when the estimation of the objective function is accurate enough. Specifically, our proposed FSS performs better than GCA due to the superior performance of reverse sampling compared to Monte Carlo simulations when their costs are equal. TM is not better than GCA and FSS because it only focuses on influencing more attackers, which may result in a poor selection where many attackers target a common sufferer while neglecting other sufferers with fewer attackers. RND is the worst-performing algorithm as it does not identify any seeds. OD is better than RND in terms of identification ability, but it is still inferior to FSS and GCA due to its weaker identification capability.

Efficiency comparison: We compare the efficiency between FSS and GCA since they have similar effectiveness. Additionally, the time cost of other algorithms for OD and RND is negligible. For TTM algorithm, both in terms of time efficiency and solution quality, is theoretically not as good as our proposed algorithm. The FSS algorithm we propose is based on reverse sampling estimation, so its running time is primarily determined by the sampling algorithm and the number of samples. The worst-case scenario is when it



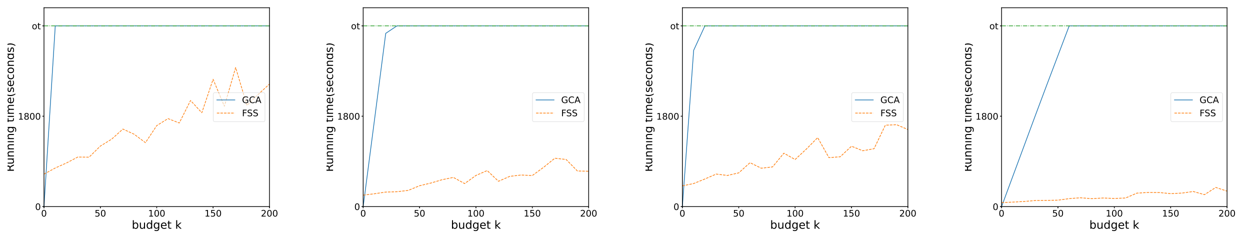
(a) Twitter (b) DBLP (c) Flickr (d) Facebook

Fig. 8. Efficiency comparisons for seed selection achieved by different estimation methods on the Twitter, DBLP, Flickr, and Facebook datasets, considering 5% of the nodes involved in the conflict.



(a) Twitter (b) DBLP (c) Flickr (d) Facebook

Fig. 9. Efficiency comparisons for seed selection achieved by different estimation methods on the Twitter, DBLP, Flickr, and Facebook datasets, considering 10% of the nodes involved in the conflict.



(a) Twitter (b) DBLP (c) Flickr (d) Facebook

Fig. 10. Efficiency comparisons for seed selection achieved by different estimation methods on the Twitter, DBLP, Flickr, and Facebook datasets, considering 20% of the nodes involved in the conflict.

returns after all iterations with a time complexity of $O(\tau^2 m^2)$ where τ and m are related to the degree distribution of the conflict graph among the conflict-involved nodes in the network and the maximum node degree. Since GCA utilizes Monte Carlo simulations to estimate the marginal gain for the objective function, we consider increasing the number of simulations to compare its efficiency when its effectiveness is close to that of FSS with an error of less than 10%. Fig. 8, 9, and 10 illustrate the running time comparison for varying numbers of sufferers and attackers at 5%, 10%, and 20% involvement rates, respectively. The horizontal axis represents the budget, and the value on the vertical axis represents the lower bound of the running time (within one hour) for GCA to achieve the same effectiveness as FSS. The "ot" (overtime) line represents running times that exceed one hour. As shown in Fig. 8, 9, and 10, in order to achieve similar effectiveness, GCA runs significantly slower than FSS in larger networks such as Twitter, DBLP, and Flicker. Through experiments, we have found that FSS often terminates early during iterations, which allows it to outperform GCA significantly. GCA requires numerous repeated simulations with a time complexity of $O(kt)$, where t is the number of simulations needed to estimate the marginal gain.

7. Conclusions

In this paper, we approach the resolution of online conflict from an algorithmic perspective. By the method of influence propagation on social media platforms, we propose the problem of minimizing attacks on sufferers (MAS) and further consider its equivalent maximization optimization problem. We propose reverse sampling method in providing accurate estimations for seed selection. By leveraging reverse samples, we are able to estimate the objective function accurately and present an algorithm with a $1-1/e$ approximation, which offers high-quality seed node selection strategies. Experimental results on real network data show that our algorithm can provide the best seed strategy within a given budget while maintaining low time complexity and computational cost. By adopting our proposed approach and algorithm, it becomes possible to reduce the level of attacks on sufferers, improve user experiences on social media platforms, and facilitate positive interactions and information dissemination. This is crucial for building and maintaining a healthy online social environment.

CRediT authorship contribution statement

Guoyao Rao: Conceptualization, Methodology, Software, Writing – original draft. **Deying Li:** Funding acquisition, Supervision, Writing – review & editing. **Yongcai Wang:** Writing – review & editing. **Wenping Chen:** Writing – review & editing. **Chunlai Zhou:** Writing – review & editing. **Yuqing Zhu:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] N. Barbieri, F. Bonchi, G. Manco, Topic-aware social influence propagation models, Knowl. Inf. Syst. 37 (2013) 555–584.
[2] J. Bercovitch, Social Conflicts and Third Parties: Strategies of Conflict Resolution, Routledge, 2019.
[3] S. Bharathi, D. Kempe, M. Salek, Competitive influence maximization in social networks, in: International Workshop on Web and Internet Economics, Springer, 2007, pp. 306–311.
[4] C. Borgs, M. Brautbar, J. Chayes, B. Lucier, Maximizing social influence in nearly optimal time, in: Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2014, pp. 946–957.
[5] S.A. Castaño-Pulgarín, N. Suárez-Betancur, L.M.T. Vega, H.M.H. López, Internet, social media and online hate speech. Systematic review, Aggress. Violent Behav. 58 (2021) 101608.
[6] S. Chen, J. Fan, G. Li, J. Feng, K.I. Tan, J. Tang, Online topic-aware influence maximization, Proc. VLDB Endow. 8 (2015) 666–677.
[7] E. Cohen, D. Delling, T. Pajor, R.F. Werneck, Timed influence: computation and maximization, arXiv preprint, arXiv:1410.6976, 2014.
[8] C. Elsaesser, D.U. Patton, E. Weinstein, J. Santiago, A. Clarke, R. Eschmann, Small becomes big, fast: adolescent perceptions of how social media features escalate online conflict to offline violence, Child. Youth Serv. Rev. 122 (2021) 105898.
[9] I. Gagliardone, D. Gal, T. Alves, G. Martinez, Countering Online Hate Speech, Unesco Publishing, 2015.
[10] A.R. Gohdes, Studying the Internet and violent conflict, Confl. Manage. Peace Sci. 35 (2018) 89–106.
[11] A. Goyal, W. Lu, L.V. Lakshmanan, Celf++: optimizing the greedy algorithm for influence maximization in social networks, in: Proceedings of the 20th International Conference Companion on World Wide Web, ACM, 2011, pp. 47–48.
[12] Q. Guo, S. Wang, Z. Wei, W. Lin, J. Tang, Influence maximization revisited: efficient sampling with bound tightened, ACM Trans. Database Syst. 47 (2022) 1–45.
[13] X. He, G. Song, W. Chen, Q. Jiang, Influence blocking maximization in social networks under the competitive linear threshold model, in: Proceedings of the 2012 Siam International Conference on Data Mining, SIAM, 2012, pp. 463–474.
[14] L.Y. Hunter, G. Biglaiser, The effects of social media, elites, and political polarization on civil conflict, Stud. Confl. Terrorism (2022) 1–28.
[15] D. Kempe, J. Kleinberg, É. Tardos, Maximizing the spread of influence through a social network, in: Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2003, pp. 137–146.
[16] P. Kiecker, D. Cowles, Interpersonal communication and personal influence on the Internet: a framework for examining online word-of-mouth, J. Euromark. 11 (2002) 71–88.
[17] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, N. Glance, Cost-effective outbreak detection in networks, in: Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2007, pp. 420–429.
[18] S. Lewandowsky, U.K. Ecker, C.M. Seifert, N. Schwarz, J. Cook, Misinformation and its correction: continued influence and successful debiasing, Psychol. Sci. Public Interest 13 (2012) 106–131.
[19] B. Liu, G. Cong, D. Xu, Y. Zeng, Time constrained influence maximization in social networks, in: 2012 IEEE 12th International Conference on Data Mining, IEEE, 2012, pp. 439–448.
[20] W. Lu, W. Chen, L.V. Lakshmanan, From competition to complementarity: comparative influence diffusion and maximization, Proc. VLDB Endow. 9 (2015) 60–71.
[21] R. Motwani, P. Raghavan, Randomized Algorithms, Cambridge University Press, 1995.
[22] G.L. Nemhauser, L.A. Wolsey, M.L. Fisher, An analysis of approximations for maximizing submodular set functions—I, Math. Program. 14 (1978) 265–294.
[23] H.T. Nguyen, M.T. Thai, T.N. Dinh, Stop-and-stare: optimal sampling algorithms for viral marketing in billion-scale networks, in: Proceedings of the 2016 International Conference on Management of Data, ACM, 2016, pp. 695–710.
[24] N. Ohsaka, T. Akiba, Y. Yoshida, K.i. Kawarabayashi, Fast and accurate influence maximization on large networks with pruned Monte-Carlo simulations, in: Twenty-Eighth AAAI Conference on Artificial Intelligence, 2014.
[25] E.L. Paluck, Reducing intergroup prejudice and conflict using the media: a field experiment in Rwanda, J. Pers. Soc. Psychol. 96 (2009) 574.
[26] M. Pérez-Escolar, J.M. Noguera-Vivo, Hate Speech and Polarization in Participatory Society, Taylor & Francis, 2022.
[27] G. Rao, D. Li, Y. Wang, W. Chen, C. Zhou, Y. Zhu, Maximizing the influence with κ-grouping constraint, Inf. Sci. (2023), https://doi.org/10.1016/j.ins.2023.01.139.
[28] G. Rao, Y. Wang, W. Chen, D. Li, W. Wu, Matching influence maximization in social networks, Theor. Comput. Sci. 857 (2021) 71–86.
[29] G. Rao, Y. Wang, W. Chen, D. Li, W. Wu, Union acceptable profit maximization in social networks, Theor. Comput. Sci. 917 (2022) 107–121.
[30] B. Razaghi, M. Roayaei, N.M. Charkari, On the group-fairness-aware influence maximization in social networks, IEEE Trans. Comput. Soc. Syst. (2022).
[31] C. Song, W. Hsu, M.L. Lee, Targeted influence maximization in social networks, in: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, ACM, 2016, pp. 1683–1692.
[32] L. Sun, W. Huang, P.S. Yu, W. Chen, Multi-round influence maximization, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, ACM, 2018, pp. 2249–2258.

[33] J. Tang, X. Tang, X. Xiao, J. Yuan, Online processing algorithms for influence maximization, in: Proceedings of the 2018 International Conference on Management of Data, 2018, pp. 991–1005.
[34] Y. Tang, Y. Shi, X. Xiao, Influence maximization in near-linear time: a martingale approach, in: Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, ACM, 2015, pp. 1539–1554.
[35] Y. Tang, X. Xiao, Y. Shi, Influence maximization: near-optimal time complexity meets practical efficiency, in: Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data, ACM, 2014, pp. 75–86.
[36] A. Tsang, B. Wilder, E. Rice, M. Tambe, Y. Zick, Group-fairness in influence maximization, arXiv preprint, arXiv:1903.00967, 2019.
[37] J.A. Tucker, A. Guess, P. Barberá, C. Vaccari, A. Siegel, S. Sanovich, D. Stukal, B. Nyhan, Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature, March 19, 2018.
[38] H.A. Unver, Using social media to monitor conflict-related migration: a review of implications for AI forecasting, Soc. Sci. 11 (2022) 395.
[39] Z. Zhang, L. Luo, Hate speech detection: a solved problem? The challenging case of long tail on Twitter, Semant. Web 10 (2019) 925–945.