

DIFFIM: Differentiable Influence Minimization with Surrogate Modeling and Continuous Relaxation

Junghun Lee, Hyunju Kim, Fanchen Bu, Jihoon Ko, Kijung Shin

KAIST, Daejeon, South Korea

{junghun.lee, hyunju.kim, boqvezen97, jihoonko, kijungs}@kaist.ac.kr

Abstract

In social networks, people *influence* each other through social links, which can be represented as propagation among nodes in graphs. *Influence minimization* (IMIN) is the problem of manipulating the structures of an input graph (e.g., removing edges) to reduce the propagation among nodes. IMIN can represent time-critical real-world applications, such as rumor blocking, but IMIN is theoretically difficult and computationally expensive. Moreover, the *discrete* nature of IMIN hinders the usage of powerful machine learning techniques, which requires *differentiable* computation. In this work, we propose DIFFIM, a novel method for IMIN with two *differentiable* schemes for acceleration: (1) *surrogate modeling* for efficient influence estimation, which avoids time-consuming simulations (e.g., Monte Carlo), and (2) the *continuous relaxation* of decisions, which avoids the evaluation of individual discrete decisions (e.g., removing an edge). We further propose a third accelerating scheme, *gradient-driven selection*, that chooses edges *instantly based on gradients* without optimization (spec., gradient descent iterations) on each test instance. Through extensive experiments on real-world graphs, we show that each proposed scheme significantly improves speed with little (or even no) IMIN performance degradation. Our method is Pareto-optimal (i.e., no baseline is faster *and* more effective than it) and typically several orders of magnitude (spec., up to 15,160 \times) faster than the most effective baseline, while being more effective.

Code, datasets and online appendix —
<https://github.com/junghun/DiffIM>

1 Introduction

In both online and offline social networks, a common phenomenon is *influence*. That is, people influence other people through social links. Typical examples include the spread of information (e.g., rumors) and the contagion of a disease (e.g., COVID-19). We can model social networks as graphs and use the propagation among nodes to simulate such processes (Kempe, Kleinberg, and Tardos 2005), and several *diffusion models* mathematically model such propagation.

While the problem of influence maximization has been widely studied, prior research has also explored *influence minimization* (IMIN), where one aims to manipulate graph

structures to reduce the propagation among nodes. IMIN is relevant to real-world scenarios, such as blocking the spread of rumors or diseases, and it has been studied with several different formulations (Yan et al. 2019; Ni, Zhu, and Wang 2023). In this work, we mainly focus on a formulation with edge removal under the independent cascade (IC) model (Kempe, Kleinberg, and Tardos 2003), due to its realism and generality; and we provide discussions and experiments on other models, spec., the linear threshold (LT) model (Kempe, Kleinberg, and Tardos 2003) and the general Markov chain susceptible-infected-recovered (G-SIR) model (Yi et al. 2022), in Appendix B. Specifically, the IC model is realistic and widely considered for modeling the spread of information (Tripathy, Bagchi, and Mehta 2010) and diseases (Borgs et al. 2014); and edge removal is general, including node removal, another widely considered graph manipulation, as a special case. For such real-world scenarios, we need timely actions since any delay could witness an exponential explosion in the spread of information (Jin et al. 2013) and diseases (Platto et al. 2021).

However, this problem is NP-hard, and even simply computing influence under the IC model is computationally expensive (spec., #P-hard; Chen, Wang, and Wang 2010). Therefore, several ideas, including Monte Carlo (MC) simulation, bond percolation (Kimura, Saito, and Motoda 2009), reverse influence sampling (Borgs et al. 2014; Yi et al. 2022), and marginal-decrement heuristics (Yan et al. 2019), are available for influence estimation. However, they offer limited speed improvements (e.g., due to extensive sampling requirements) and/or rely on assumptions (e.g., acyclic input graphs) that do not generally hold in practice.

Notably, existing methods approach the IMIN problem as a *discrete* optimization problem, which hinders the application of powerful continuous-optimization (e.g., gradient descent) and machine-learning (e.g., neural networks) techniques.

Instead, we propose a novel method for IMIN, called DIFFIM, with two main *differentiable* schemes: (1) *surrogate modeling* for efficient influence estimation, and (2) *continuous relaxation* of edge removal. To the best of our knowledge, we are the first to approach influence minimization using *differentiable learning* instead of discrete optimization.

Surrogate modeling. First, we propose an efficient scheme for influence estimation. Inspired by Ko et al. (2020), we pro-

Symbol	Definition
$G = (V, E)$	a graph with a node set V and an edge set E
$p : E \rightarrow [0, 1]$	activation probabilities
$S \subseteq V$	a seed set
$\pi(v; G, p, S)$	the influenced probability of v under $\text{IC}(G, p, S)$
$\sigma(S; G, p)$	the expected influence of S under $\text{IC}(G, p, S)$
$\text{GNN}_\theta(v; G, p, S)$	the predicted influenced probability of v by GNN_θ

Table 1: Frequently-used notations

pose to train graph neural networks (GNNs) to “predict” the influence when given an input graph and seed nodes (i.e., the nodes where the propagation starts). Such surrogate modeling leverages the efficiency of GNNs and avoids time-consuming MC simulations or other estimation methods. Although GNN training introduces additional overhead, it is affordable because once trained in advance, the GNN efficiently estimates the influence of unseen graphs and/or seed nodes.

Continuous relaxation. Second, to further enhance speed, we propose to relax the edge removal decisions, so that we can directly optimize continuous (i.e., probabilistic) decisions without evaluating individual edge removal. Specifically, for each edge, instead of considering a binary decision (to remove it or not), we consider a probabilistic decision representing the probability of removing it. Such relaxation also allows us to incorporate powerful machine-learning techniques, which typically use gradient descent and thus cannot be naively applied to the original discrete problem.

Gradient-driven selection. These two schemes enable us to compute gradients w.r.t. the probabilistic decisions on removing edges, and the gradient of each edge is naturally interpreted as its “sensitivity”. Specifically, with influence as the objective, the influence is more sensitive to the edges with higher gradients, and removing such edges is expected to reduce the influence more effectively. Hence, we further propose our third speed-up scheme, *gradient-driven selection*, that removes edges based on their gradients, without further optimization (spec., gradient descent iterations).

With the three proposed schemes, we propose three versions, DIFFIM, DIFFIM+, and DIFFIM++, equipped with the first scheme, the first two schemes, and all three schemes, respectively. More schemes result in faster speed, with little (or even no) IMIN performance degradation.

Through extensive experiments on three real-world graphs, we show the superiority of the DIFFIM family over baselines. Specifically, all the versions of DIFFIM are Pareto-optimal (i.e., no baseline is faster *and* more effective than any version of DIFFIM), and they are typically orders of magnitude faster (spec., up to 15,160 \times) than the most effective baseline, while also being more effective. We also show their ability to perform well when trained and tested on different graphs.

In short, our main contributions are three-fold:

- **Differentiable learning based approach:** To the best of our knowledge, we are the first to tackle IMIN using differentiable learning, instead of discrete combinatorial optimization. We make this approach feasible by leveraging GNNs as surrogate models and employing the continuous

relaxation of edge removal.

- **Gradient-driven acceleration:** We propose another speed-up scheme, *gradient-driven selection*, that selects edges based on their gradients without additional test-time optimization (spec., gradient descent iterations).
- **Extensive experiments:** We demonstrate the empirical superiority of our methods over baselines in IMIN on real-world graphs, in terms of both speed and effectiveness.

2 Preliminaries

Basic concepts. Refer to Table 1 for frequently-used notations. Let \mathbb{N} be the set of positive integers, and let $[n]$ be $\{1, \dots, n\}$. A graph $G = (V, E)$ is defined by a node set V and an edge set E . We consider *directed* edges, i.e., each edge $e = (u, v) \in E$ is a directed link from $u \in V$ to $v \in V$.

Definition 1 (Independent cascade (IC) model). *Given (1) a graph $G = (V, E)$, (2) activation probabilities $p : E \rightarrow [0, 1]$, (3) a seed set $S \subseteq V$, the IC model $\text{IC}(G, p, S)$ is a stochastic process defined as follows:*

- **Initialization:** At time step $t = 0$, each seed node $v_S \in S$ is activated and each non-seed node remains inactive.
- **Diffusion steps:** At each step $t \geq 1$, each node v that is activated in the previous step $t - 1$ activates each of its inactive out-neighbor u with activation probability $p(v, u)$.¹ That is, each activated node remains active for the whole process but can only activate other nodes one step after its activation. The process terminates when no node is activated in the previous step.

As shown in Def. 1, the process is stochastic and the states of nodes at the termination are thus probabilistic.

Definition 2 (Influenced probabilities and expected influence). *Given $G = (V, E)$, $p : E \rightarrow [0, 1]$, and $S \subseteq V$, for each node $v \in V$, the influenced probability of v , denoted by $\pi(v; G, p, S)$, is the probability of v being influenced (i.e., active) when the process of $\text{IC}(G, p, S)$ terminates, and the expected influence of the seed set S , denoted by $\sigma(S; G, p)$, is defined as the expected number of finally influenced nodes, i.e., $\sigma(S; G, p) := \sum_{v \in V} \pi(v; G, p, S)$.*

Graph neural networks (GNNs). In GNNs, there are two main types of operators on the node features: feature transformation and propagation (Zhu et al. 2021b). A feature transformation operator transforms node features at each layer into the next layer via nonlinear transformation. A propagation operator passes the features of a node to its neighbors, and updates the feature of each node by aggregating the features of its neighbors, typically in the form of $H \leftarrow \tilde{A}H$, where H is a node feature matrix, and \tilde{A} is a normalized adjacency matrix. The strength of different edges (i.e., edge weights or edge probabilities) can be incorporated in the entries of \tilde{A} .

3 Related Work

Influence estimation. As mentioned in Sec. 1, exactly computing the influence under the IC model is costly, and thus several methods have been proposed for influence estimation.

¹The set of *out-neighbors* of a node v is $\{u \in V : (v, u) \in E\}$.

The most related approach is by Ko et al. (2020), where graph neural networks (GNNs) are used to learn the influence under the IC model. The Monte Carlo (MC) simulation (i.e., taking the mean value of samplings) has been a common practice for influence computation (Zhou et al. 2013; Yang, Brenner, and Giua 2019; Manouchehri, Helfroush, and Danyali 2021). Kimura, Saito, and Motoda (2009) used the bond percolation method, and Yi et al. (2022) used the reverse influence sampling instead of MC, but samplings are still required.

Influence minimization. Influence minimization (IMIN) has been widely studied, and it has several different variants. Researchers have considered node removal (Zhu et al. 2021a; Ni, Zhu, and Wang 2023) and edge removal (Kimura, Saito, and Motoda 2009; Tong et al. 2012); and different models other than IC (Dai et al. 2022) have also been considered (note that we also considered other models; see Appendix B). Moreover, prior studies have explored blocking propagation to specific targets (Jiang et al. 2022; Wang et al. 2020) and/or without specific seed sets (Zareie and Sakellariou 2022), as well as active defense by propagating opposite information (Budak, Agrawal, and El Abbadi 2011; Luo et al. 2014). In this work, we consider edge removal under the IC model. While Yan et al. (2019) used the same problem formulation, their analysis, and proposed method were limited to acyclic graphs. Due to the difficulty of influence computation or even estimation, different heuristics without direct influence estimation have also been considered. They proposed considering the incremental differences when removing each edge, assuming that the input graph is acyclic. Tong et al. (2012) proposed to choose the edges according to the leading eigenvalues of the adjacency matrix, but the effect of seed nodes is not considered in the method. See Appendix D for more details on existing IMIN methods. Note that no existing method for IMIN has considered a differentiable learning scheme, which is a novel approach introduced in this work.

4 Problem Statement and Hardness

As mentioned in Sec. 3, there are different problem formulations for influence minimization (IMIN). One can consider different graph manipulation (edge removal or node removal) and different diffusion models. In this work, we consider the formulation with *edge removal* under the *independent cascade* (IC) model (see Sec. 2) due to the following reasons:

- The formulation with edge removal is more *general* than that with node removal. Specifically, node removal can be seen as edge removal with additional constraints that the edges incident to a node should be all kept or all removed.
- Blocking spread between users (i.e., edge removal), such as through contact restriction, is often more feasible than completely removing a user (i.e., node removal).
- Regarding the diffusion model, the IC model has been widely considered for the spread of information (e.g., rumors) (Tripathy, Bagchi, and Mehta 2010; Xu and Chen 2015; Shelke and Attar 2019) due to its simple yet realistic nature. However, note that **our proposed approach is not limited to the IC model** but can be applied to more diffusion models, spec., the linear threshold (LT)

model (Kempe, Kleinberg, and Tardos 2003) and the general Markov chain susceptible-infected-recovered (G-SIR) model (Yi et al. 2022), as explored in Appendix B.

Hereafter, we simply call the considered problem *influence minimization* when no confusion is likely.

Problem 1 (influence minimization).

- **Given:** a graph $G = (V, E)$, activation probabilities $p : E \rightarrow [0, 1]$, a seed set $S \subseteq V$, and a budget $b \in \mathbb{N}$,
- **Find:** a set \mathcal{E} of b edges, i.e., $\mathcal{E} \subseteq E$ and $|\mathcal{E}| = b$,
- **to Minimize:** the expected influence of S after removing the edges in \mathcal{E} from G , i.e., $\sigma(S; G_{\setminus \mathcal{E}}, p_{\setminus \mathcal{E}})$ with $G_{\setminus \mathcal{E}} := (V, E \setminus \mathcal{E})$ and $p_{\setminus \mathcal{E}}(e) = p(e), \forall e \in E \setminus \mathcal{E}$.

We show the NP-hardness of influence minimization (see Appendix A.1), and Yan et al. (2019) proved that influence minimization is non-submodular.

Theorem 1. *Influence minimization (Problem 1) is NP-hard.*

Theorem 2 (Yan et al. 2019). *Influence minimization is non-submodular, i.e., $f(\mathcal{E}; G, p, S) := \sigma(S; G, p) - \sigma(S; G_{\setminus \mathcal{E}}, p_{\setminus \mathcal{E}})$ is not submodular w.r.t. \mathcal{E} .*

5 Proposed Method

Thms. 1-2 show the non-triviality of influence minimization (IMIN; Problem 1). Below, we further analyze the challenges in IMIN and propose our method, DIFFIM, to address them.

5.1 Naive algorithms and their problems

Before introducing our method DIFFIM, we discuss some naive algorithms, and analyze their problems.

A naive enumeration algorithm evaluates all possible combinations of b edges and chooses the best combination. This requires computing influence for $\binom{|E|}{b} = \Theta(|E|^b)$ times, which is computationally prohibitive. One can reduce the frequency of influence computation by adapting it in an incremental manner. There are b rounds in total, and in each round, we choose an edge whose removal reduces the expected influence of S most. See Alg. 2 in Appendix C for the pseudo-code of such an incremental greedy algorithm. Although such an idea requires computing influence for only $O(b|E|)$ times, as discussed in Secs. 1 & 3, the exact computation of expected influence is computationally prohibitive, and the existing estimation methods, e.g., Monte Carlo (MC) simulation, are still time-consuming because they require extensive sampling. Below, we shall propose multiple schemes to speed up the process of choosing each edge.

5.2 DIFFIM: Surrogate modeling for efficient influence estimation without simulation

We shall first address the problem of time-consuming estimation of influence. We propose to use graph neural networks (GNNs), which are computationally efficient.

The high-level idea is to use a GNN as a surrogate model (neural approximation) of the influenced probabilities $\pi(\cdot)$ (see Def. 2). That is, we see π as a black-box function, and we aim to train a GNN $_{\theta}$ parameterized by θ such that $\text{GNN}_{\theta}(v; G, p, S) \approx \pi(v; G, p, S), \forall G, p, S, v$. Specifically,

p is used as edge weights and S is represented by one-dimensional binary node features. Although GNN training introduces additional overhead, it is affordable since it can be done once in advance on existing or randomly generated data. Once trained, GNN_θ efficiently estimates the influenced probabilities $\text{GNN}_\theta(v'; G', p', S') \approx \pi(v'; G', p', S')$ of new unseen cases (see Sec. 5.5 for complexity analysis), avoiding time-consuming MC simulation or other estimation methods (see Sec. 3 for examples).

Notably, even with a single input graph, we are able to generate multiple data points by generating different seed sets, and the training process can be easily extended to multiple graphs. We first obtain the influence of each training seed set by MC simulation and use it as the ‘‘ground-truth’’ influence, and then update the parameters of the GNN w.r.t. the difference between the predicted influence by the GNN and the ‘‘ground-truth’’ influence. Specifically, for each seed set S_i , the L2-loss is used, i.e.,

$$\begin{aligned} & \mathcal{L}(\text{GNN}_\theta; \tilde{\pi}, \{S_i\}, G, p) \\ & := \sqrt{\sum_{v \in V} (\text{GNN}_\theta(v; G, p, S_i) - \tilde{\pi}(v; G, p, S_i))^2}. \end{aligned}$$

The final loss function is averaged over seed sets, i.e.,

$$\mathcal{L}(\text{GNN}_\theta; \tilde{\pi}, S, G, p) := \frac{1}{|S|} \sum_{S_i \in S} \mathcal{L}(\text{GNN}_\theta; \tilde{\pi}, \{S_i\}, G, p).$$

After training, we can use the trained GNN as a surrogate model for each test instance $\mathcal{T} = (G, p, S, b)$ (see Problem 1), while still following an incremental greedy scheme, i.e., choosing the edges one by one with the highest effect estimated by the trained GNN, which results in DIFFIM. See Alg. 1 for pseudo-code of DIFFIM, and see Alg. 3 in Appendix C for that of GNN training.

5.3 DIFFIM+: Continuous relaxation of edge removal without individual removal evaluation

Although DIFFIM accelerates influence estimation using GNNs as a surrogate model, DIFFIM still needs to evaluate each individual edge removal (spec., compute the estimated influence when each edge is removed), which can still take considerable time even with efficient influence estimation. To this end, we propose to use *continuous relaxation* of edge removal to avoid individual edge-removal evaluation.

Continuous relaxation. The high-level idea is that, for each edge $e = (v, u)$, instead of a binary decision $r(v, u) = \mathbb{1}((v, u) \notin \mathcal{E}) \in \{0, 1\}$ (i.e., to keep it or not; recall \mathcal{E} is the set of edges to be removed), we consider a *probabilistic* decision $\tilde{r}(v, u) = \Pr[(v, u) \notin \mathcal{E}] \in [0, 1]$, i.e., the probability of keeping e . Such relaxation can be readily incorporated into our surrogate-model-based influence estimation. Specifically, by the multiplication rule: $\Pr[v \text{ activates } u \text{ via } (v, u)] = \Pr[v \text{ activates } u \wedge (v, u) \text{ is kept}] = p(v, u)\tilde{r}(v, u)$.

Therefore, given activation probabilities $p : E \rightarrow [0, 1]$ and probabilistic decisions $\tilde{r} : E \rightarrow [0, 1]$ on the edges, we obtain the *modified activation probabilities* $\tilde{p}_{\tilde{r}} : E \rightarrow [0, 1]$ by $\tilde{p}_{\tilde{r}}(v, u) := p(v, u)\tilde{r}(v, u), \forall (v, u) \in E$. Specifically, given a graph $G = (V, E)$, activation probabilities p , a seed set $S \subseteq V$, a trained GNN_θ , and probabilistic edge-removal

Algorithm 1: DIFFIM / DIFFIM+ / DIFFIM++

Input: (1) $G = (V, E)$: an input graph
(2) $p : E \rightarrow [0, 1]$: activation probabilities
(3) $S \subseteq V$: a seed set
(4) b : an edge-removal budget
(5) GNN_θ : a trained GNN
(6) $\tilde{r} : E \rightarrow [0, 1]$: initial probabilistic decisions
 \triangleright For DIFFIM+ and DIFFIM++
(7) n_{ep} : the number of epochs for each removal
 \triangleright For DIFFIM+ only

Output: $\mathcal{E} \subseteq E$: a set of edges chosen to be removed

```

1  $\mathcal{T} \leftarrow (G, p, S, b)$   $\triangleright$  Initialize the IMIN problem instance  $\mathcal{T}$ 
2  $\mathcal{E} \leftarrow \emptyset$   $\triangleright$  Initialize the set of edges to be removed
3 for  $i = 1, 2, \dots, b$  do
4    $e = \text{EdgeSelection}()$  or  $\text{EdgeSelection}+$   $()$ 
   or  $\text{EdgeSelection}++()$   $\triangleright$  Select an edge
5    $E \leftarrow E \setminus \{e\}; \mathcal{E} \leftarrow \mathcal{E} \cup \{e\}$   $\triangleright$  Remove the edge
6    $G \leftarrow (V, E); \mathcal{T} \leftarrow (G, p, S, b)$   $\triangleright$  Update the graph
7 return  $\mathcal{E}$ 
8 Function  $\text{EdgeSelection}()$ :
9   return  $\arg \min_{e \in E} \sum_{v \in V} \text{GNN}_\theta(v; G, p_{\setminus \mathcal{E}}, S)$ 
10 Function  $\text{EdgeSelection}+$   $()$ :
11   for  $j = 1, 2, \dots, n_{ep}$  do
12     Compute the derivative  $\nabla_{\tilde{r}} \mathcal{L} := \frac{\partial \mathcal{L}_O(\tilde{r}; \mathcal{T}, \text{GNN}_\theta)}{\partial \tilde{r}}$ 
13     Update  $\tilde{r}$  via gradient descent w.r.t.  $\nabla_{\tilde{r}} \mathcal{L}$ 
14   return  $\arg \min_{e \in E} \tilde{r}(e)$ 
15 Function  $\text{EdgeSelection}++()$ :
16    $\nabla_{\tilde{r}}(e) = \frac{\partial \sum_{v \in V} \text{GNN}_\theta(v; G, p_{\setminus \mathcal{E}}, S)}{\partial \tilde{r}(e)}, \forall e \in E$ 
17   return  $\arg \max_{e \in E} \nabla_{\tilde{r}}(e)$ 

```

decisions \tilde{r} , the influenced probability of each node $v \in V$ is estimated as $\text{GNN}_\theta(v; G, \tilde{p}_{\tilde{r}}, S)$. Notably, such an estimated influenced probability is differentiable w.r.t. \tilde{r} .

Lemma 1. $\text{GNN}_\theta(v; G, \tilde{p}_{\tilde{r}}, S)$ is differentiable w.r.t. \tilde{r} .

Proof. See Appendix A.2. \square

Differentiable optimization. After training GNN_θ , by Lemma 1, we can now conduct *differentiable* optimization on the probabilistic edge-removal decisions for each test instance $\mathcal{T} = (G, p, S, b)$ of the IMIN problem (see Problem 1), which results in DIFFIM+ (see Alg. 1). The high-level process is as follows: we fix GNN_θ after training, and only update \tilde{r} to minimize the loss \mathcal{L}_O , which has three parts:

$$\mathcal{L}_O := \mathcal{L}_{\text{obj}} + \alpha \mathcal{L}_{\text{budget}} + \beta \mathcal{L}_{\text{certainty}}, \quad (1)$$

where the loss coefficients, α and β , are hyperparameters.

The first part \mathcal{L}_{obj} is regarding the main optimization objective of IMIN (Problem 1), defined as

$$\begin{aligned} & \mathcal{L}_{\text{obj}}(\tilde{r}; \mathcal{T}, \text{GNN}_\theta) \\ & := - \frac{\sum_{v \in V} (\text{GNN}_\theta(v; G, p, S) - \text{GNN}_\theta(v; G, \tilde{p}_{\tilde{r}}, S))}{\sum_{v \in V} \text{GNN}_\theta(v; G, p, S) - |S|}, \end{aligned}$$

which is the *reduction ratio* in the estimated number of influenced non-seed nodes. Specifically, the numerator is the estimated influence reduction after the probabilistic edge removal \tilde{r} is applied, and the denominator is the estimated number of influenced non-seed nodes before the edge removal.

When the surrogate model is perfectly accurate, minimizing \mathcal{L}_{obj} is equivalent to optimizing the objective of IMIN.

The second part $\mathcal{L}_{\text{budget}}$ is regarding the budget constraint:

$$\mathcal{L}_{\text{budget}}(\tilde{r}; \mathcal{T}, \text{GNN}_{\theta}) := (|E| - \sum_{e \in E} \tilde{r}(e) - b)^2,$$

which is the squared difference between the expected number of removed edges and the required budget. When the budget is exactly used, \mathcal{L}_{obj} is 0, i.e., minimized.

The third part $\mathcal{L}_{\text{certainty}}$ is regarding the certainty (i.e., closeness to binary) of the probabilistic decisions \tilde{r} :

$$\begin{aligned} \mathcal{L}_{\text{certainty}}(\tilde{r}; \mathcal{T}, \text{GNN}_{\theta}) \\ := \frac{\sum_{e \in E} (\tilde{r}(e) \log r(e) - (1 - \tilde{r}(e)) \log(1 - \tilde{r}(e)))}{|E|}, \end{aligned}$$

which is inspired by the Shannon entropy, and $\mathcal{L}_{\text{certainty}}$ is smaller when each $\tilde{r}(e)$ is closer to 0 or 1.

Even with continuous relaxation, optimizing \mathcal{L}_O remains closely aligned with the IMIN problem, as discussed in Appendix A.3. In Alg. 1, we show pseudo-code of DIFFIM+, where some details (e.g., how we initialize and normalize \tilde{r}) are omitted and will be deferred to Sec. 6 when we describe the detailed experimental settings. Given initial probabilistic decisions \tilde{r} , in each iteration, we update it via gradient descent according to the loss function \tilde{L}_O (Eq. (1)) and its derivative. For each specified number n_{ep} of epochs, the edge with the smallest $\tilde{r}(e)$ value is removed.

5.4 DIFFIM++: Gradient-driven selection without test-time gradient-descent optimization

The previous two schemes entail differentiability, enabling us to compute gradients w.r.t. the probabilistic decisions on edges. The gradient of each edge can be naturally interpreted as its ‘‘sensitivity’’. Specifically, the influence is more sensitive to the edges with higher gradients, and removing such edges is expected to reduce the influence more effectively. Hence, we propose DIFFIM++ with *gradient-driven edge selection*, which *instantly* removes the edge with the largest gradient in each round, instead of performing optimization (spec., gradient descent) over many epochs.

Alg. 1 shows pseudo-code of DIFFIM++. After training GNN_{θ} , for each test instance $\mathcal{T} = (G = (V, E), p, S, b)$ of the IMIN problem, we compute the derivatives on all edges,

$$\nabla_{\tilde{r}}(e; \mathcal{T}) := \frac{\partial \sum_{v \in V} \text{GNN}_{\theta}(v; G, p, S)}{\partial \tilde{r}(e)}, \forall e \in E,$$

and remove the edge with the largest $\nabla_{\tilde{r}}(e; \mathcal{T})$ in each round.

Discussion. For DIFFIM+, instead of removing edges one by one after n_{ep} epochs, one can choose the bottom- b edges with the lowest \tilde{r} values at once. Similarly, for DIFFIM++, one can choose the top- b edges with the largest gradient at once. Empirically, we observe that choosing edges at once in such a way achieves similar or worse performance. See Appendix G.3 for more detailed results and discussions.

5.5 Time and space complexities

Given a test instance $\mathcal{T} = (G = (V, E), p, S, b)$ of the IMIN problem, assume that (1) we use graph convolutional networks (GCNs; Kipf and Welling 2017) with a constant number of layers, as in our experiments, (2) the dimensions of

dataset	abbr.	training graph		test graph	
		$ V $	$ E $	$ V $	$ E $
WannaCry	WC	16, 246	84, 217	19, 381	85, 202
Celebrity	CL	7, 848	28, 839	7, 336	27, 699
Extended	ET	5, 636	31, 826	5, 413	27, 146

Table 2: Basic statistics of the real-world datasets.

hidden features are fixed as constant, which are indeed fixed and much smaller than the size of graphs in our experiments, and (3) the input graph is sparse (specifically, $|E| = \Theta(|V|)$), which is indeed true for the datasets in our experiments (see Table 2), we can derive the time and space complexity of DIFFIM based on existing results (Chiang et al. 2019; Blakely, Lanchantin, and Qi 2021).

DIFFIM. A forward pass of GCN takes $O(|E|)$ time. In each round, we conduct $O(|E|)$ forward passes, with b rounds in total. Thus, the time complexity of DIFFIM is $O(b|E|^2)$.

DIFFIM+. A backward pass of GCN takes $O(|E|)$ time, and there are b rounds each with n_{ep} epochs (see Alg. 1). Therefore, the time complexity of DIFFIM+ is $O(n_{ep}b|E|)$.

DIFFIM++. We conduct one backward pass in each of the b rounds, so the time complexity of DIFFIM++ is $O(b|E|)$.

Note. Regarding the time complexity, $\text{DIFFIM} > \text{DIFFIM+} > \text{DIFFIM++}$, as intended.

Space complexity. For each version, it is $O(|E| + |V|) = O(|E|)$, dominated by the space complexity of GCN.

6 Experiments

We performed experiments on real-world graphs, aiming to answer the following questions:

- **Q1. Performance:** How effectively and quickly does DIFFIM minimize influence?
- **Q2. Scalability:** How does the running time of DIFFIM grow with the budget b ?
- **Q3. Influence estimation quality:** How well does the surrogate GNN model in DIFFIM estimate the influence?
- **Q4. Inductivity:** How does DIFFIM perform when trained and tested on different/same graphs?
- **Q5. Ablation studies:** How does each component or algorithmic design affect the performance of DIFFIM?

6.1 Experimental settings

Datasets. We used three real-world social-network datasets from Ko et al. (2020): WC, CL, and ET, consisting of interactions logs, e.g., retweets among users (Sabotke, Suciu, and Dumitras 2015). Such interactions naturally represent people’s impact on others, i.e., influence. We provided their basic statistics in Table 2. Each dataset was split into training and test graphs based on a time threshold t_{th} : edges before t_{th} were used for training, and those after were used for testing.

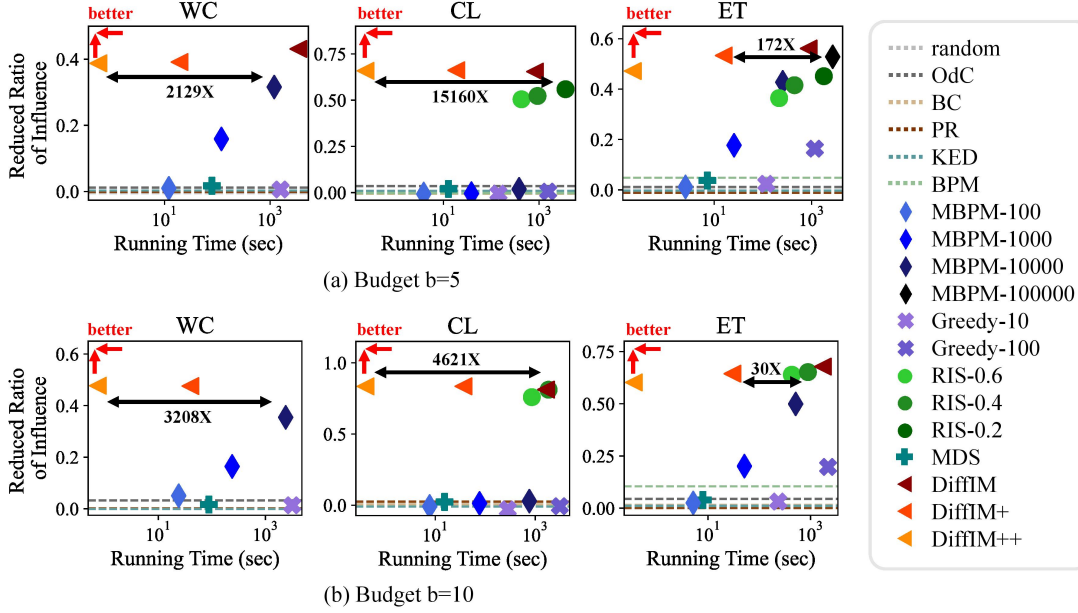


Figure 1: The effectiveness (the reduced ratio of influence) and running time of each method, with budget $b = 5$ (top) and $b = 10$ (bottom). The baselines with outputs independent of seed sets were represented as horizontal lines. We compared the running time of the best baseline to one of our methods with the most similar reduced ratio, and in all cases, DIFFIM was $30 \times$ to $15,160 \times$ faster. See Appendix G.1 for full results with standard deviations.

Baselines. We considered the following baselines:

1. **RANDOM** removes b edges chosen uniformly at random.
2. **ODC** (Out-degree Centrality; Kempe, Kleinberg, and Tardos 2003) removes the top- b edges w.r.t. the sum of out-degrees of their two endpoints.²
3. **BC** (Betweenness Centrality; Schneider et al. 2011) removes the top- b edges w.r.t edge betweenness.
4. **PR** (PageRank; Page et al. 1998) removes the top- b edges w.r.t. the sum of the PageRank scores of their endpoints.
5. **KED** (Tong et al. 2012) removes the b edges to minimize the leading eigenvalue of the adjacency matrix.
6. **MDS** (Yan et al. 2019) greedily removes b edges w.r.t. the importance scores estimated by influenced probability and rumor-spread ability of their endpoints.
7. **BPM** (Kimura, Saito, and Motoda 2009) removes the top- b edges w.r.t. importance scores estimated using the bond percolation method (BPM). BPM does not consider specific seed nodes.
8. **Modified BPM (MBPM)** is a modified version of BPM that considers specific seed nodes.
9. **GREEDY** removes b edges greedily with influence estimated by Monte Carlo (MC) simulation.
10. **RIS** (Yi et al. 2022) estimates the importance of each edge for propagating to other nodes using sampling, and removes the b most important edges.

²The out-degree of each node $v \in V$ is $|\{u \in V : (v, u) \in E\}|$.

For MBPM and GREEDY, a number after their names denotes the number of samplings (e.g., GREEDY-100 denotes GREEDY with 100 samplings). For RIS, a small ϵ increases sampling numbers and improves edge importance estimation accuracy. Note that all these baselines approach IMIN as a discrete combinatorial optimization problem. See Appendix E for more details on the baselines.

DIFFIM. For each training graph, we generated 1,000 random seed sets, using 800 for training and 200 for validation. For each test graph, we generated 50 random seed sets and reported average performance, where we used 10,000 Monte Carlo simulations as the “ground-truth” influence, following the settings by Kempe, Kleinberg, and Tardos (2003). We consistently used a graph convolutional network (GCN) with six layers and a final fully connected layer. For DIFFIM+, the probabilistic decisions \tilde{r} were optimized (see Alg. 1) in $n_{ep} = 100$ epochs for each removal with $\alpha = 0.1$ and $\beta = 1$.

Seed set generations. For each seed set, the size was sampled uniformly between 10 and $\lfloor 0.01|V| \rfloor$ (inclusive), and then the nodes were sampled uniformly.

Probabilistic decisions. The probabilistic decisions \tilde{r} were normalized by a sigmoid function $\sigma_{\text{sigmoid}}(x) = \frac{1}{1+e^{-x}} \in [0, 1]$. We used initial probabilistic decisions $\tilde{r}(e) = \sigma_{\text{sigmoid}}(x) = 1 - \frac{b}{|E|}, \forall e \in E$, which makes the loss $\mathcal{L}_{\text{budget}}$ regarding the budget constraint to be 0 (see Sec. 5.3). See Appendix F for more details of the experimental settings, e.g., hardware information.

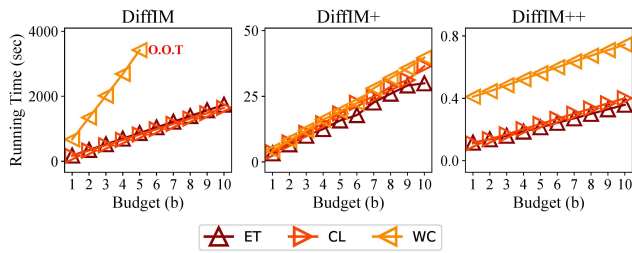


Figure 2: The running time of each DIFFIM version when budget b increases from 1 to 10. The running time of each version grew linearly with b , showing good scalability.

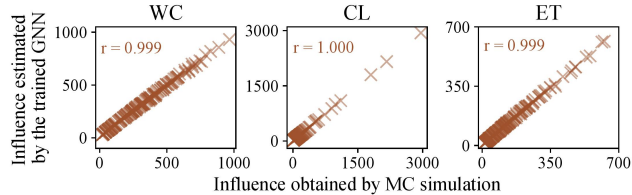


Figure 3: The Pearson correlation coefficients between the ground-truth influence of the validation sets and that estimated by MC simulation of a trained GNN. In all the cases, trained GNNs estimated influences near-perfectly.

6.2 Q1. Performance

We shall show that DIFFIM showed good performance in influence minimizing (IMIN; Problem 1) in terms of effectiveness (the reduced ratio of influence) and efficiency (the running time). Formally, the reduced ratio R_r is defined as

$$R_r(\mathcal{E}; G, p, S) := \frac{\sigma(S; G, p) - \sigma(S; G \setminus \mathcal{E}, p \setminus \mathcal{E})}{\sigma(S; G, p) - |S|}. \quad (2)$$

A higher R_r implies higher effectiveness in IMIN.

In Fig. 1, we reported the average reduced ratio of influence and running time across all the test seed sets for each dataset and method, using budgets $b \in \{5, 10\}$. Methods taking more than one hour on a single seed set were considered “out of time” and excluded. Several baselines (RANDOM, ODC, BC, PR, BPM, and KED) do not depend on the seed set, and can be executed once for all seed sets. Hence, we represented the performance of each of those methods as a horizontal line. All the versions of DIFFIM were Pareto-optimal, i.e., no baseline was faster and more effective than any version, with at least one version outperforming all the baselines w.r.t. effectiveness in each case. In most cases, all the versions outperformed all the baselines w.r.t. effectiveness. Specifically, DIFFIM achieves $30 \times$ to $15,160 \times$ speed-up with similar effectiveness. Notably, for WC, RIS- ϵ ran out of time for each $\epsilon \in \{0.2, 0.4, 0.6\}$. As intended, the more proposed schemes we use, the higher speed we have (w.r.t. speed, DIFFIM < DIFFIM+ < DIFFIM++). The effectiveness of different versions did not vary much, yet we observed an overall trend: DIFFIM > DIFFIM+ > DIFFIM++. See Fig. 5 and Table 7 in Appendix G.1 for the full results with standard deviations with different budgets $b \in \{3, 5, 7, 10\}$.

datasets	WC	CL	ET
time (in seconds)	17,227	7,390	6,818

Table 3: Average training time (in seconds) for GNNs.

estimation method	WC	CL	ET
GNN	0.0082	0.0056	0.0056
MC simulation	3.7757	0.8276	0.7817

Table 4: Average time (in seconds) to estimate influence for each seed set using GNNs and MC simulations. MC simulations were repeated 10,000 times for estimation.

6.3 Q2. Scalability

In Fig. 2, we reported the running time of all DIFFIM versions on each dataset with different budgets. As the budget b increased from 1 to 10, the running time increased almost linearly for all versions, which validated our analysis on the time complexity in Sec. 5.5. Our analysis in Sec. 5.5 also implies that DIFFIM+ and DIFFIM++ have better scalability w.r.t. the input graph size than DIFFIM. Specifically, the time complexity of DIFFIM is quadratic in $|E|$ while that of DIFFIM+ and DIFFIM++ is linear, which was also validated in Fig. 2, where the running time gap between DIFFIM and the other two versions increased on the larger dataset WC. In Table 9 in Appendix G.2, for each method and each dataset, we provided the minimum budget b for the method to run out of time (i.e., take more than one hour on a single seed set).

6.4 Q3. Influence estimation quality

We shall show that our surrogate GNNs were trained well for influence estimation. In Fig. 3, for the validation seed sets (recall that for each training graph, we generated 1,000 random seed sets, using 800 training and 200 for validation; see Sec. 6.1), we reported the Pearson correlation coefficients (Pearson’s r) between the influence obtained by MC simulation (seen as the ground truth) and that estimated by the trained GNNs. The estimation by the GNNs was highly correlated with the ground truth. Specifically, the trained GNNs achieved a Pearson’s r of 0.999 or higher on each dataset. See Appendix G.4 for more details, e.g., how the estimation errors decreased along training.

We reported the training times for GNNs in Table 3 and compared the influence estimation times of GNNs with MC simulations in Table 4. Overall, the estimation times of GNNs were at least $100\times$ faster than those of the MC simulations.

6.5 Q4. Inductivity

In Table 5, we compared the effectiveness (the reduced ratio of influence; see Eq. (2)) of all DIFFIM versions in the transductive setting (trained and tested on the same dataset) and the inductive setting (trained and tested on different datasets) when budget $b = 5$. Note that even for the transductive setting, the training and test graphs were different (spec., different timestamps), and the seed sets were different (see Sec. 6.1). In the inductive setting, on each dataset, we tested

method	DIFFIM			DIFFIM+			DIFFIM++		
	WC	CL	ET	WC	CL	ET	WC	CL	ET
transductive	0.4311	0.6547	0.5613	0.3914	0.6614	0.5332	0.3876	0.6583	0.4718
inductive	0.4256	0.6394	0.5429	0.3910	0.6534	0.5045	0.3692	0.6230	0.5050
(difference)	(-1.3%)	(-2.3%)	(-3.3%)	(-0.1%)	(-1.2%)	(-5.4%)	(-4.7%)	(-5.4%)	(+7.0%)
strongest baseline (transductive)	0.3160	0.5591	0.5284	0.3160	0.5591	0.5284	0.3160	0.5591	0.5284

Table 5: The effectiveness (the reduced ratio of influence) of each DIFFIM version when budget $b = 5$, and the comparison between the performance in transductive and inductive settings, with each percentage being the difference ratio (negative means the effectiveness in inductive settings was lower). In all the cases, the effectiveness of each DIFFIM version was only slightly lower (or even higher) in the inductive setting. Importantly, DIFFIM outperformed all the baselines in most cases in both transductive and inductive settings.

loss	WC	CL	ET
original	0.3914	0.6614	0.5332
$-\mathcal{L}_{\text{budget}}$	0.0052	0.0112	0.0170
$-\mathcal{L}_{\text{certainty}}$	0.3891	0.6479	0.5282

Table 6: The effectiveness (the reduced ratio of influence) of DIFFIM+ on budget $b = 5$ when removing different parts of the loss function (Eq. (1)). Each part was helpful, and $\mathcal{L}_{\text{budget}}$ was much more helpful than that of $\mathcal{L}_{\text{certainty}}$.

the two models that were trained on the other two datasets. The reported performance was the average of the results from these two models. In most cases, the effectiveness of our methods was higher in the transductive setting than in the inductive setting. From transductive to inductive settings, their effectiveness dropped by at most 5.4%. In most cases and settings, the effectiveness of our methods is higher than the most effective baseline (which was transductive and much slower than DIFFIM+ and DIFFIM++), showing the good inductivity of our methods.

6.6 Q5. Ablation studies

We evaluated the importance of $\mathcal{L}_{\text{budget}}$ and $\mathcal{L}_{\text{certainty}}$ in the loss function of DIFFIM+ (see Eq. (1)). In Table 6, we compared the effectiveness (the reduced ratio of influence) of DIFFIM+ on budget $b = 5$ when using the whole loss function and when removing $\mathcal{L}_{\text{budget}}$ or $\mathcal{L}_{\text{certainty}}$. We observed the effectiveness of DIFFIM+ dropped significantly without $\mathcal{L}_{\text{budget}}$, showing the significance of $\mathcal{L}_{\text{budget}}$. The effectiveness of $\mathcal{L}_{\text{certainty}}$ was marginal, but it was necessary for theoretical guarantees (see Lem. 2 in Appendix A.3).

6.7 Additional experiments

We also conducted experiments on **two more influence diffusion models**, specifically, the linear threshold (LT) model (Kempe, Kleinberg, and Tardos 2003) and the general Markov chain susceptible-infected-recovered (G-SIR) model (Yi et al. 2022). On these models, DIFFIM still showed empirical superiority, being significantly faster than the most effective baseline, while achieving a similar reduced ratio of influence. See Appendix B for more details.

In addition, we conducted experiments on large-scale datasets. Due to the absence of realistic activation probabilities, for these datasets, we used the weighted cascade model (Kempe, Kleinberg, and Tardos 2003), a special case of the IC model where the activation probability of each edge from node u to v is 1 divided by the in-degree of v . On these datasets, many strong baselines ran out of time or memory, and DIFFIM consistently outperformed those that completed within the given limits. See Appendix G.5 for details.

Furthermore, in Appendix G, we presented additional experimental results on (1) the detailed trade-off between time and reduction ratio, (2) scalability w.r.t. budgets, and (3) comparisons with variants of DIFFIM missing certain components. Overall, we demonstrate the superiority of DIFFIM over baseline methods in terms of trade-offs and scalability, and the importance of each component in DIFFIM.

7 Conclusions

In this work, we studied influence minimization (IMIN) with edge removal under the independent cascade (IC) model, with more models discussed in Appendix B. We proposed DIFFIM (Sec. 5), which incorporates two key schemes: *surrogate modeling* for efficient influence estimation (Sec. 5.2) and *continuous relaxation* of edge removal (Sec. 5.3). Additionally, we proposed *gradient-driven edge selection* for instant edge selection without test-time gradient descent iterations (Sec. 5.4). Our extensive experiments demonstrated that all three schemes improved the speed of DIFFIM with little (or even no) IMIN performance degradation, in addition to its superior speed and effectiveness over baselines (Sec. 6.2). We also showed its scalability (Sec. 6.3) and ability to perform well when trained and tested on different graphs (Sec. 6.5).

Our future work will extend our approach to other influence-related graph problems. For example, our method can be adapted to the influence maximization problem (Kempe, Kleinberg, and Tardos 2003), whose objective is to identify the most influential seed set, by (1) introducing a global seed node linked to all existing nodes with 100% activation probabilities and (2) selectively removing some of these new edges to maximize the influence of the seed node.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00406985, 50%). This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2022-0-00871 / RS-2022-II220871, Development of AI Autonomy and Knowledge Enhancement for AI Agent Collaboration, 40%) (No. RS-2019-II190075, Artificial Intelligence Graduate School Program (KAIST), 10%).

References

- Blakely, D.; Lanchantin, J.; and Qi, Y. 2021. Time and space complexity of graph convolutional networks. Technical report, University of Virginia.
- Borgs, C.; Brautbar, M.; Chayes, J.; and Lucier, B. 2014. Maximizing social influence in nearly optimal time. In *SODA*.
- Budak, C.; Agrawal, D.; and El Abbadi, A. 2011. Limiting the spread of misinformation in social networks. In *WWW*.
- Chen, W.; Wang, C.; and Wang, Y. 2010. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *KDD*.
- Chiang, W.-L.; Liu, X.; Si, S.; Li, Y.; Bengio, S.; and Hsieh, C.-J. 2019. Cluster-gcn: An efficient algorithm for training deep and large graph convolutional networks. In *KDD*.
- Dai, C.; Chen, L.; Hu, K.; and Ding, Y. 2022. Minimizing the spread of negative influence in SNIR model by contact blocking. *Entropy*, 24(11): 1623.
- Jiang, Z.; Chen, X.; Ma, J.; and Philip, S. Y. 2022. RumorDecay: rumor dissemination interruption for target recipients in social networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(10): 6383–6395.
- Jin, F.; Dougherty, E.; Saraf, P.; Cao, Y.; and Ramakrishnan, N. 2013. Epidemiological modeling of news and rumors on twitter. In *SNA-KDD*.
- Kempe, D.; Kleinberg, J.; and Tardos, É. 2003. Maximizing the spread of influence through a social network. In *KDD*.
- Kempe, D.; Kleinberg, J.; and Tardos, É. 2005. Influential nodes in a diffusion model for social networks. In *ICALP*.
- Kimura, M.; Saito, K.; and Motoda, H. 2009. Blocking links to minimize contamination spread in a social network. *ACM Transactions on Knowledge Discovery from Data*, 3(2): 1–23.
- Kipf, T. N.; and Welling, M. 2017. Semi-supervised classification with graph convolutional networks. In *ICLR*.
- Ko, J.; Lee, K.; Shin, K.; and Park, N. 2020. Monstor: an inductive approach for estimating and maximizing influence over unseen networks. In *ASONAM*.
- Luo, C.; Cui, K.; Zheng, X.; and Zeng, D. 2014. Time critical disinformation influence minimization in online social networks. In *JISIC*.
- Manouchehri, M. A.; Helfroush, M. S.; and Danyali, H. 2021. Temporal rumor blocking in online social networks: A sampling-based approach. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(7): 4578–4588.
- Ni, P.; Zhu, J.; and Wang, G. 2023. Misinformation influence minimization by entity protection on multi-social networks. *Applied Intelligence*, 53(6): 6401–6420.
- Page, L.; Brin, S.; Motwani, R.; and Winograd, T. 1998. The pagerank citation ranking: Bring order to the web. Technical report, Stanford University.
- Platto, S.; Wang, Y.; Zhou, J.; and Carafoli, E. 2021. History of the COVID-19 pandemic: Origin, explosion, worldwide spreading. *Biochemical and biophysical research communications*, 538: 14–23.
- Sabottke, C.; Suci, O.; and Dumitras, T. 2015. Vulnerability disclosure in the age of social media: Exploiting twitter for predicting Real-World exploits. In *USENIX Security*.
- Schneider, C. M.; Mihaljev, T.; Havlin, S.; and Herrmann, H. J. 2011. Suppressing epidemics with a limited amount of immunization units. *Physical Review E*, 84(6): 061911.
- Shelke, S.; and Attar, V. 2019. Source detection of rumor in social network—a review. *Online Social Networks and Media*, 9: 30–42.
- Tong, H.; Prakash, B. A.; Eliassi-Rad, T.; Faloutsos, M.; and Faloutsos, C. 2012. Gelling, and melting, large graphs by edge manipulation. In *CIKM*.
- Tripathy, R. M.; Bagchi, A.; and Mehta, S. 2010. A study of rumor control strategies on social networks. In *CIKM*.
- Wang, X.; Deng, K.; Li, J.; Yu, J. X.; Jensen, C. S.; and Yang, X. 2020. Efficient targeted influence minimization in big social networks. *World Wide Web*, 23(4): 2323–2340.
- Xu, W.; and Chen, H. 2015. Scalable rumor source detection under independent cascade model in online social networks. In *MSN*.
- Yan, R.; Li, Y.; Wu, W.; Li, D.; and Wang, Y. 2019. Rumor blocking through online link deletion on social networks. *ACM Transactions on Knowledge Discovery from Data*, 13(2): 1–26.
- Yang, W.; Brenner, L.; and Giua, A. 2019. Influence maximization in independent cascade networks based on activation probability computation. *IEEE Access*, 7: 13745–13757.
- Yi, Y.; Shan, L.; Paré, P. E.; and Johansson, K. H. 2022. Edge deletion algorithms for minimizing spread in sir epidemic models. *SIAM Journal on Control and Optimization*, 60(2): S246–S273.
- Zareie, A.; and Sakellariou, R. 2022. Rumour spread minimization in social networks: A source-ignorant approach. *Online Social Networks and Media*, 29: 100206.
- Zhou, C.; Zhang, P.; Guo, J.; Zhu, X.; and Guo, L. 2013. Ublf: An upper bound based approach to discover influential nodes in social networks. In *ICDM*.
- Zhu, J.; Ni, P.; Wang, G.; and Li, Y. 2021a. Misinformation influence minimization problem based on group disbanded in social networks. *Information Sciences*, 572: 1–15.
- Zhu, M.; Wang, X.; Shi, C.; Ji, H.; and Cui, P. 2021b. Interpreting and unifying graph neural networks with an optimization framework. In *WWW*.