Conquering the Worst Case of Infections in Networks

WEN YAN1,(Student Member, IEEE), PO-LING LOH2, CHUNGUO LI1, (Senior Member, IEEE), YONGMING HUANG1, (Senior Member, IEEE), and LUXI YANG1, (Senior Member, IEEE)

1School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: {wyan, chunguoli, huangym, lxyang}@seu.edu.cn)

2Department of Statistics, University of Wisconsin-Madison, WI 53706, USA (e-mail: ploh@stat.wisc.edu)

Corresponding author: Luxi Yang (e-mail: lxyang@seu.edu.cn).

This work was supported in part by the State Key Program of National Natural Science Foundation of China under Grant U1936201, in part by the National Natural Science Foundation of China under Grant 61971128, in part by the National Science Foundation under Grant DMS-1749857, and in part by the China Scholarship Council (CSC) under Grant 201706090055.

ABSTRACT We develop algorithms to control the scope of an infection spread on a network by allocating a fixed immunization budget to edges of the graph. We assume that the infection propagates according to an independent cascade model and interventions operate by reducing the propensities of edges to transmit the infection. We formulate this problem as a constrained min-max optimization problem with respect to the placements of interventions and the location of the worst-case seed nodes. However, the result is a challenging bilevel mixed integer optimization problem. Furthermore, gradients of the objective function with respect to the continuous variables are unavailable in closed form. We employ tools from derivative-free optimization and stochastic optimization to optimize the objective by iterating between the outer minimization and inner maximization problems. In the inner loop, we use a weighted degree discount (WDD) method to select the seed set of the influence maximization problem. In the outer loop, we utilize two methods: a sample-based simultaneous perturbation Nelder-Mead (SBSP-NM) algorithm and a simultaneous perturbation stochastic approximation (SPSA) algorithm. We perform simulations on synthetic graphs and three larger-scale real-world datasets and illustrate the computational feasibility of our algorithms and their efficacy at controlling epidemic spreads.

INDEX TERMS Immunization, bilevel optimization, derivative-free optimization, influence maximization, mixed integer program, networks

I. INTRODUCTION IMMUNIZATION plays an important role in real-world applications, such as rumor control in social networks [1], epidemic control in public health [2], and the control of terrorist attacks for public safety [3]. However, the problem of choosing optimal interventions, subject to a limited budget, is somewhat challenging. First, the interventions must often be selected under uncertainty on the properties of the epidemic spread. For instance, the seed nodes at which the epidemic originates may be unknown, and the epidemic may propagate in a stochastic fashion. Second, the influence of different nodes on the spread of the disease may be highly coupled due to various paths connecting nodes in the network.

In our work, we adopt the approach of finding the optimal allocation of interventions with respect to the worst-case choice of seed nodes for an infection. We assume that the network and the probabilities of transmitting the infection over each edge are known or could be estimated. However, the set of seed nodes is unknown a priori. For instance, to prevent a terrorist attack, the police may already intercept the communication frequencies between terrorists but need to allocate limited police resources to control the order spread and the potential attack as much as possible without knowing which terrorists may issue the attack orders. To go one step further, we formulate the immunization problem as a min-max problem with constraints. The goal is to minimize the maximal expected number of infected nodes in a network, subject to the worst-case seed set, by reducing the propagation probabilities of an optimally chosen subset of edges. Note that unlike minimizing undirect objectives like the leading eigenvalue of the adjacent matrix, optimizing the expected number of infected nodes directly is much more practical but more challenging. What’s more, instead of adopting the node removal or edge removal strategies in complete immunization, we consider a more practically
fractional immunization scenario in this paper.

Nevertheless, the min-max immunization problem brings several difficulties, as well: (i) the inner maximization problem, known as a classical influence maximization problem, is an NP-hard integer optimization problem; (ii) the outer minimization problem is a continuous optimization problem with no convenient closed form for the objective function; and (iii) the two layers of the min-max problem are highly coupled with respect to the discrete solution space of the seed set $A$ and the continuous solution space of the edge intervention vector $\theta$. To deal with these difficulties, we adopt an iterative strategy to alternately optimize over $A$ and $\theta$ separately. This produces an overall upper bound on the worst-case infection spread under the optimal intervention. In order to perform the outer minimization over $\theta$, we utilize two methods from derivative-free optimization and stochastic optimization, respectively: the Nelder-Mead (NM) and simultaneous perturbation stochastic approximation (SPSA) algorithms. Furthermore, to make the procedure computationally feasible for large networks, we modify the traditional NM algorithm by combining sampling and simultaneous perturbation ideas, leading to a sample-based simultaneous-perturbation NM (SBSP-NM) method. For the inner maximization problem, we introduce a weighted degree discount (WDD) algorithm, which greedily selects seed nodes according to their discounted weighted degrees and is inspired by the DegreeDiscountIC algorithm [4]. In summary, the main contributions of this paper are as follows:

- **Problem formulation of finding the optimal allocation of interventions with respect to the worst-case choice of seed nodes for an infection**: A min-max problem with constraints is formulated to minimize the maximal expected number of infected nodes in a network for the first time, subject to the worst-case seed set, by reducing the propagation probabilities of an optimally chosen subset of edges.
- **Optimization to solve the min-max immunization problem**: The min-max immunization problem brings multiple difficulties, thus, an alternating optimization procedure is applied to solve the min-max problem, especially where SBSP-NM and SPSA from derivative-free optimization and stochastic optimization respectively are proposed to solve the outer continuous minimization problem with no convenient closed form for the objective function.
- **Simulations on synthetic graphs and three larger-scale real-world datasets**: Experiments on six synthetic graphs are performed to test the effectiveness of our algorithm on different types of graphs, and further on three larger-scale public real-world social network datasets, NetHEPT, ego-Facebook, and wiki-Vote, to explore the effectiveness and computational feasibility of our immunization algorithms.

**Related work**: Several other authors have proposed methods for developing optimal immunization strategies on networks.

Tong et al. [5] propose an effective algorithm to optimize the leading eigenvalue based on the edge removal strategy to solve the NetMelt Problem. Prakash et al. [6] formulate a minimization problem involving the top eigenvalue of the adjacency matrix, and propose a linear-time algorithm to batch-allocate resources greedily. Chen et al. [7] study immunization strategies that operate on nodes and propose scalable algorithms to find an optimal set of nodes to be immunized based on a novel score definition. Zhang et al. [8] propose algorithms based on greedily reducing the infection count on a graph based on sampled cascades, using additional information from data logs. From the viewpoint of control theory, Nowzari et al. [9] consider methods for controlling epidemics using a generalized susceptible-exposed-infected-vigilant model and propose several cost-optimal strategies based on geometric programming. Preciado et al. [10] study the problem of finding the optimal distribution of vaccines based on an optimization framework under semidefinite programming constraint. Besides, several studies take network measures in complex networks as the objective of immunization, such as both Chan et al. [11] and Gunasekara et al. [12] study the immunization problem from the network robustness viewpoint by edge deletions and additions, and Shen and Cao [13] aim to maintain the network efficiency by reducing the edge weights on weighted networks. However, none of these papers take the approach of controlling the worst-case spread with respect to adversarially chosen seed nodes, which is the method we adopt in this paper.

We also discuss several papers related to the influence maximization problem, which arises as the inner maximization problem in our iterative optimization approach. Influence maximization has long been a popular topic of research in graph mining. In a seminal paper, Kempe et al. [14] formulated the influence maximization problem as a submodular problem and proved that greedy hill-climbing algorithms achieve a $(1 - 1/e)$-approximation of the optimal solution when the infection spreads according to a linear threshold or independent cascade model. Liontis and Pitoura [15] aim to increase the influence of a seed set by boosting the activation probabilities of edges adjacent to an optimally chosen set of nodes. However, unlike our approach, the authors assume that the initial seed set is known. Khim et al. [16] propose algorithms for influence maximization in stochastic and adversarial contagion settings and establish related upper and lower bounds. Chen et al. [17] address the scalability issue in influence maximization by restricting computations on local influence regions of nodes. Other popular state-of-the-art methods on influence maximization include the Cost-Effective Lazy Forward (CELF) algorithm of Leskovec et al. [18], the CELF++ algorithm of Goyal et al. [19], the Reverse Influence Sampling (RIS) algorithm of Borgs et al. [20], etc. However, to the best of our knowledge, we are the first to combine influence maximization and immunization in a min-max bilevel optimization framework.

The remainder of our paper is organized as follows: Section II provides the mathematical formulation of our problem,
as well as related notation. Section III contains our proposed algorithms for solving the immunization problem. Section IV provides experimental assessments of our method on different types of synthetic and real-world datasets. We conclude the paper in Section V.

II. PROBLEM SETUP

Consider an undirected graph $G := (V, E)$, where $V = \{1, \ldots, N\}$ denotes the set of nodes and $E \subseteq V \times V$ denotes the set of edges. If nodes $i$ and $j$ are connected by an edge, we include both $(i,j)$ and $(j,i)$ in $E$. Let $p := \{p_{ij}, (i,j) \in E\} \in \mathbb{R}^{N \times N}$ denote the matrix of edge weights, where $0 \leq p_{ij} = p_{ji} \leq 1$ for each $(i,j) \in E$.

We assume that given a set of seed nodes $A \subseteq V$, an infection spreads in discrete time according to the independent cascade (IC) model: if $A_i$ is the set of newly infected nodes at time $t$, the probability of an uninfected node $i$ becoming infected at time $t+1$ is $1 - \prod_{j: (i,j) \in E,j \in A_t}(1-p_{ij})$. In other words, each newly infected neighbor $j$ has probability $p_{ij}$ of independently infecting node $i$. The influence of a seed set $A$ is the expected number of infected nodes at the conclusion of the epidemic, assuming $A_0 = A$. Let $f_0(G,A,p)$ denote this expected quantity, assuming the epidemic spreads on the graph $G$ with edge weights $p$. Note that since we focus on the operations on the graph based directly on $f_0(G,A,p)$ other than how the infection spreads, our work is very general and not constrained to fit the IC model, which we adopt here for demonstration and simulation.

The influence maximization problem aims to solve the problem $\max_{A \subseteq V, |A| \leq k} f_0(G,A,p)$, where $G$ and $p$ are fixed and assumed to be known as the prior information. In our setting, we assume we have a fixed budget $B$ that can be used to immunize edges in the graph, thus decreasing the corresponding edge weights. Thus, we arrive at the following problem formulation:

$$\min_{\theta} \max_A f_0(G,A,p - \theta)$$

s.t. $\sum_{(i,j) \in E} \theta_{ij} \leq B$ (1a)

$$0 \leq \theta_{ij}, \forall (i,j) \in E$$ (1b)

$|A| \leq k$. (1c)

Here, $\theta := \{\theta_{ij}, (i,j) \in E\}$ denotes the matrix of interventions, so the expected number of infected nodes is $f_0(G,A,p - \theta)$ after the intervention is applied. Moreover, (1a) and (1b) are constraints on $\theta$ to characterize the budget of immunization and the value range, respectively, and (1c) is a widely adopted condition to constrain the number of seed nodes. Note that the interventions here belong to fractional immunizations other than complete immunizations by edge removal or node removal, which are more practical since there is no need to change the topology of the network.

For better demonstration, we take the epidemic control network for example. We assume that we can get the sets of the concerned persons (nodes) and their connections (edges), and estimate the propagation probabilities as the edge weights ($p$). But we don’t know the set of the sudden infected persons ($A$), and they will spread to infect their neighbors. We aim to minimize the number of infected persons (denoted as $f_0$) by reducing the propensities of edges to transmit the infection ($\theta$), namely assigning a constrained number of disinfectants to decrease the infected probabilities. However, the explicit form of $f_0$ is unavailable, let alone the properties like continuity, convexity, and derivability, which makes the problem (1) extremely difficult.

In order to make the operations of our algorithms on each dimension of $\theta$ explicit, we will use the vector notation $\theta = [\theta_i]$, with index $i$ ranging from 1 to $M := |E|$, and the corresponding notation $p = [p_i]$. What’s more, it’s trivial to consider $k$ as an upper bound on the size of the infection seed set.

III. ALGORITHMS

In this section, we describe our proposed algorithms for solving the min-max problem (1). We apply an alternating optimization procedure [21] [22], where we iteratively maximize with respect to the discrete seed set $A$ and minimize with respect to the continuous intervention vector $\theta$. What’s more, for better reproducibility, we state all the algorithms as detailed as possible.

A. SOLVING INNER OPTIMIZATION WITH FIXED $\theta$

To optimize over $A$ with respect to a fixed vector $\theta$, problem (1) simply reduces to the conventional influence maximization problem (2), with fixed edge weights $p - \theta$.

$$\max_A f_0(G,A,p - \theta)$$

s.t. $|A| \leq k$. (2a)

One popular method for solving the influence maximization problem is the general greedy algorithm [14], which greedily adds nodes to the selected seed set to successively maximize the overall influence. However, the simple greedy algorithm can still be computationally expensive on large graphs, especially since the influence function $f_0(G,A,p - \theta)$ can only be approximated for each candidate set $A$ by simulating multiple runs of the epidemic starting from the seed set $A$, and taking an average over the number of infected nodes on each run. For demonstration, we assume all the evaluations of the influence function are estimated by $R_{cas}$ runs of the epidemic according to the independent cascade model in the following paper.

Inspired by the DegreeDiscountIC algorithm [4], we propose a weighted degree discount (WDD) heuristics algorithm. The main differences between WDD and DegreeDiscountIC are that (i) WDD greedily selects nodes according to their discounted weighted degrees, which is much more
suitable for cases where edge weights matter; and (ii) we filter out the nodes that still have strong connections with seeds even after the immunization. The WDD algorithm is stated in Algorithm 1. Here, we set strong_connection_thres to be a predetermined threshold to decide whether the nodes have strong connections to seed nodes after immunization. The value of $\epsilon$ is chosen to be a suitably small number. Since we pay more attention to the outer minimization problem, further improvements could be to test the performance under different combinations including like RIS [20]. It’s also reasonable to adopt WDD to solve the inner maximization problem for demonstration, because it could provide comparisons under different outer solvers and always lower the adversarial influence spread compared to the optimal solutions.

Algorithm 1 Weighted degree discount (WDD) algorithm to optimize over $A$ with fixed $\theta$, WDD($G,k,\theta$)

1. input graph $G$, seed size $k$, edge weight modification $\theta$
2. initialize $A = \emptyset$, strong_connection_thres, $\epsilon$
3. for each node $v \in V$ do
4. compute its weighted degree $wd_v$ based on $p - \theta$
5. initialize its weighted degree discount $wdd_v = wd_v$
6. end for
7. for $i = 1$ to $k$ do
8. select $u = \arg \max_{v \in V/S} \{wdd_v\}$
9. $A = A \cup \{u\}$
10. for each neighbor $v$ of $u$ and $v \in V/S$ do
11. if $p_{uv} - \theta_{uv} > $ strong_connection_thres
12. $wdd_v = \epsilon$
13. else
14. $wdd_v = wdd_v - (p_{uv} - \theta_{uv})$
15. end if
16. end for
17. end for
18. output $A$

B. SOLVING OUTER OPTIMIZATION WITH FIXED $A$

Turning to the problem of maximizing over $\theta$ for a fixed seed set $A$, we obtain a continuous constrained problem (3) as follows:

$$\min_{\theta} f(\theta) := f_0(G,A,p - \theta)$$

$$\text{s.t.} \sum_{i=1}^{M} \theta_i \leq B \quad \text{(3a)}$$

$$0 \leq \theta_i \leq p_i, \forall i \in [1,M]. \quad \text{(3b)}$$

However, computing the influence spread $f$ given a fixed seed set $A$ is #P-hard [17], let alone the explicit form of $f$ as a function of $\theta$, even if $G$ and $p - \theta$ are known explicitly. To overcome the challenge, we propose to use two methods from derivative-free [23] and stochastic optimization [24]. The first method is based on the Nelder-Mead (NM) method, which proceeds by evaluating the value of the objective function on vertices of a successively changing simplex. Hence, we propose an adaptation of the NM algorithm that is more suitable for large-scale problems and can also incorporate constraints on the parameters. The second method is to use the simultaneous perturbation stochastic approximation (SPSA) algorithm, which obtains successive iterates by approximating the gradient of the objective function in a stochastic manner using simultaneous perturbations in all dimensions.

For both methods, note that the objective function $f$ also needs to be approximated for each value of $\theta$ in the solution path. This is conducted by simulating the infection spread with seed set $A$ and edge weights $p - \theta$ and taking an average over the infection count after a suitably large number of simulations. This simulation-based approach has the advantages of model generality and a good theoretical property that it commonly returns a solution with a constant bounded approximation ratio, while suffering significant computational overheads [25].

1) Nelder-Mead method

The NM algorithm was proposed in 1965 [26], and soon became one of the most popular algorithms in derivative-free optimization, due to its easy understanding and implementation. What’s more, its “best-case efficiency” frequently outweighs its lack of convergence theory [27]. The algorithm evaluates the value of the objective function at the vertices of a simplex and performs a direct search over simplex transformations until convergence [27]. For an $M$-dimensional problem, the NM algorithm first generates a simplex with $M + 1$ points and then performs successive simplex transformations using five operations: reflection, expansion, outside contraction, inside contraction, and shrinkage.

However, applying the NM algorithm to large graphs with tens of thousands of edges is not computationally feasible. Furthermore, the function evaluations must all be approximated via simulations of the infection. Instead, we modify the traditional NM method by combining sampling and simultaneous perturbation, leading to a sample-based simultaneous perturbation NM (SBSP-NM) method. The difference between SBSP-NM and the traditional NM method is that our initial simplex lies in a much smaller predetermined $n$-dimension subspace. To form the initial simplex, we first divide the $M$ coordinate dimensions into $m$ blocks and then generate $n$ points by randomly perturbing the corresponding $m$ blocks by a fixed step size around the origin. This forms an initial simplex using the $n$ points and the origin. As a result, our SBSP-NM method can largely reduce the number of function evaluations from $M + 1$ to $n + 1$ on simplex initialization and shrinkage steps, while revealing $m$-dimensional information with each initial vertex of the simplex. In general, we can choose $m$ such that $nm$ is several times larger than $M$.

Moreover, the traditional NM method is only suitable for unconstrained problems. Thus, we adopt a box constraint.
method to trim $\theta$ with its lower and upper bounds as used in Cauchy point computation [28], and use a normalized scaling $\theta = (B/|\theta|)_1 \ast \theta$ when $|\theta|$ exceeds the sum- 

Algorithm 2 SBSP-NM algorithm to optimize over $\theta$ with fixed $A$, algNM(G,A,B)

1: input graph $G$, seed set $A$, modification summation constraint $B$
2: initialize $x^0 = 0$, step, no_improve_thres, no_improve_break,
3: no_improve_num = 0, $\alpha_{nm} = 1$, $\gamma_{nm} = 2$, $\rho = 0.5$, $\sigma_{nm} = 0.5, m, n$
4: compute $f(x^0)$, and define $f_{prev\_best} = f(x^0)$
5: block_size = floor($M/m$)
6: for $j = 1$ to $n$ do
7: $y^j = x^0$
8: for $k = 1$ to $m$ do
9: $kk = \text{int}$(floor($\text{rand} \ast \text{block\_size} + (k - 1) \ast \text{block\_size}$))
10: $y^j_{kk} = y^j_{kk} + \text{step}$
11: end for
12: $y^j = \text{transform}(y^j, 0, p, B)$
13: evaluate $f(y^j)$
14: end for
15: generate the initiate simplex $Y = \{x^0, y^1, ... , y^n\}$
16: while 1:
17: order the vertices of $Y$ and change notations $Y = \{x^1, ... , x^n, x_{n+1}\}$, s.t. $f(x^1) < \cdots < f(x^{n+1})$
18: define $f_{best} := f(x^1)$
19: $x = x^1$
20: if $f_{best} < f_{prev\_best} - \text{no\_improve\_thres}$ then
21: no_improve_num = 0
22: $f_{prev\_best} = f_{best}$
23: else no_improve_num + 1
24: end if
25: if no_improve_num > no_improve_break then
26: break
27: end if
28: compute the centroid of best $n$ points $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x^i$
29: (reflection) compute $x^r = \bar{x} + \alpha_{nm}(\bar{x} - x^{n+1})$, then transform it by $x^r = \text{transform}(x^r, 0, p, B)$ and evaluate $f(x^r)$
30: if $f(x^r) = f(x^{n+1})$ then
31: replace $x^{n+1} \in Y$ by $x^r$, continue
32: else if $f(x^r) < f(x^{n+1})$ then
33: (expansion) compute $x^e = \bar{x} + \gamma_{nm}(\bar{x} - x^{n+1})$, then transform it by $x^e = \text{transform}(x^e, 0, p, B)$ and evaluate $f(x^e)$
34: if $f(x^e) < f(x^r)$ then
35: replace $x^{n+1} \in Y$ by $x^e$, continue
36: else replace $x^{n+1} \in Y$ by $x^r$, continue
37: end if
38: else $f(x^r) \geq f(x^{n+1})$ then
39: (outside contraction) compute $x^{oc} = \bar{x} + \rho_{nm}(\bar{x} - x^{n+1})$, then transform it by $x^{oc} = \text{transform}(x^{oc}, 0, p, B)$ and evaluate $f(x^{oc})$
40: if $f(x^{oc}) \leq f(x^r)$
41: replace $x^{n+1} \in Y$ by $x^{oc}$, continue
42: else go to shrinkage step
43: end if
44: else $f(x^r) \geq f(x^{n+1})$ then
45: (inside contraction) compute $x^{ic} = \bar{x} + \rho_{nm}(\bar{x} - x^{n+1})$, then transform it by $x^{ic} = \text{transform}(x^{ic}, 0, p, B)$ and evaluate $f(x^{ic})$
46: if $f(x^{ic}) < f(x^{n+1})$ then
47: replace $x^{n+1} \in Y$ by $x^{ic}$, continue
48: else go to shrinkage step
49: end if
50: end if
51: (shrinkage) define $n$ new vertices $y^i = x^1 + \sigma_{nm}(x^1 - x^i), i = 1, ... , n$, then transform them by $y^i = \text{transform}(y^i, 0, p, B), i = 1, ... , n$ and new simplex $Y^+ = \{x^1, y^1, ... , y^n\}$, continue
52: end if
53: end if
54: output $\theta$

And in detail, the $\text{transform}$ function of the box constraint method is stated in Algorithm 3.

Algorithm 3 Box constraint method to deal with the constraints (1a) and (1b), $\text{transform}(x, LB, UB, \text{ConstraintSum})$

1: input $x$, LB, UB, ConstraintSum
2: for $i = 1$ to $M$ do
3: if $x_i > UB_i$ then
4: $x_i = UB_i$
5: end if
6: if $x_i < LB_i$ then
7: $x_i = LB_i$
8: end if
9: end for
10: if $\|x\|_1 < \text{ConstraintSum}$ then
11: break
12: else $x = (\text{ConstraintSum}/\|x\|_1) \ast x$
13: end if
14: output $x$

2) SPSA method

We now describe the simultaneous perturbation stochastic approximation (SPSA) algorithm from stochastic optimization [29]. The SPSA algorithm proceeds by approximating the gradient of $f$ based on stochastic function evaluations, and then performs an analog of gradient descent [30] [31].

We also incorporate the constraints on $\theta$ using the box constraint method adopted in the Nelder-Mead method. The SPSA algorithm is described in Algorithm 4, where the parameters $\alpha_{\text{spsa}} = 0.602$ and $\gamma_{\text{spsa}} = 0.101$ are demonstrated to perform well in practical settings [29]. The convergence guarantee of SPSA requires $f$ to be Lipschitz continuous and
differentiable with bounded second-order derivatives [30], which are not satisfied here, while the algorithm could still be workable under its monotone decreasing property with respect to $\theta$.

Algorithm 4 SPSA algorithm to optimize over $\theta$ with fixed $A$, $\text{algSPSA}(G,A,B)$

1: **input** graph $G$, seed set $A$, modification constraint $B$
2: **initialize** $\theta(1) = 0$, $R_{\text{spsa}}$, $\alpha_{\text{spsa}} = 0.602$, $\gamma_{\text{spsa}} = 0.101$, $\alpha$, $c$, $d_{\text{big}}$
3: **for** $n = 1$ to $R_{\text{spsa}}$ **do**
4: **define** $f_{\text{best}} := f(\theta(1))$
5: $\theta = \theta(1)$
6: **for** $i = 1$ to $M$ **do**
7: $\Delta_i(n) \leftarrow$ Bernoulli$(1/2)$
8: **end for**
9: $a(n) = \frac{\alpha}{(\alpha n + n + 1)^\gamma}$
10: $\delta(n) = \frac{\gamma n + 1}{(n + 1)^{1/2}}$
11: $\theta(n)^+ = \text{transform}(\theta(n) + \delta(n)\Delta(n), 0, p, B)$
12: $\theta(n)^- = \text{transform}(\theta(n) - \delta(n)\Delta(n), 0, p, B)$
13: $f(\theta(n)^+) = f_0(G, A, p - \theta(n)^+)$
14: $f(\theta(n)^-) = f_0(G, A, p - \theta(n)^-)$
15: **for** $i = 1$ to $M$ **do**
16: $\theta_i(n + 1) = \theta_i(n) - a(n)\frac{f(\theta(n)^+) - f(\theta(n)^-)}{\Delta_i(n)}$
17: **end for**
18: $\theta(n + 1) = \text{transform}(\theta(n + 1), 0, p, B)$
19: **if** $f(\theta(n + 1)) < f_{\text{best}}$ **then**
20: $f_{\text{best}} = f(\theta(n + 1))$
21: $\theta = \theta(n + 1)$
22: **end if**
23: **end for**
24: **output** $\theta$

C. SOLVING THE MIN-MAX PROBLEM

To solve the original min-max problem (1), we alternately optimize with respect to $A$ and $\theta$, using WDD and SBSP-NM/SPSA, respectively, up to a predetermined iteration number $R_{\text{alt}}$. The procedure is summarized in Algorithm 5, where the hyper-parameter $R_{\text{alt}}$ provides a tradeoff between effectiveness and efficiency. For a better explanation, we also provide a graphical demonstration to show the optimization process of the algorithms more explicitly in Fig. 1.

Next, we do a theoretical analysis on the output of Algorithm 5.

Remark 1. The output $\hat{A}$ serves as a lower bound on the decrement of the expected infection count in the worst-case epidemic spread under the optimal intervention, even if Algorithm 5 does not converge to the optimum of the min-max problem (1).

In detail, if $\hat{A} = \arg \max_{|A| \leq k} f_0(G, A, p - \hat{\theta})$ and $(A^*, \hat{\theta}^*)$ is the solution to the problem (1), we have $f_0(G, A^*, p - \hat{\theta}^*) \leq f_0(G, \hat{A}, p - \hat{\theta})$. Hence, a smaller value of $f_0(G, \hat{A}, p - \hat{\theta})$ in comparison to $\max_{|A| \leq k} f_0(G, A, p)$ ensures that $\hat{\theta}^*$ will also decrease the expected infection count in the worst case by at least the same amount.

Remark 2. It is guaranteed that intervening according to $\theta^*$ will control the worst-case epidemic to the level of $f_0(G, \hat{A}, p - \hat{\theta})$, which is in turn lower bounded by $f_0(G, A, p - \theta)$.

D. COMPLEXITY ANALYSIS OF THE ALGORITHMS

In this subsection, we analyze the computational complexity of the algorithms in sequence.

The computational complexity of Algorithm 1 could be $O(k\log N + M)$ by using the Fibonacci heap, much smaller than the $O(kNR_{\text{cas}}M)$ of the general greedy algorithm.

For Algorithm 2, we analyze each part of Algorithm 2

FIGURE 1: Graphical demonstration of the optimization process of the algorithms.

Algorithm 5 Alternating procedure to optimize problem (1) with respect to $A$ and $\theta$ using WDD and SBSP-NM/SPSA

1: **input** graph $G$, seed size $k$, $B$
2: **initialize** $\theta(0) = 0$, $R_{\text{alt}}$, $f_{\text{best}} = N$
3: **for** $n = 0$ to $R_{\text{alt}} - 1$ **do**
4: $A(n) = \text{WDD}(G, k, \theta(n))$
5: **if** $f_0(G, A(n), p - \theta(n)) < f_{\text{best}}$ **then**
6: $f_{\text{best}} = f_0(G, A(n), p - \theta(n))$
7: $\hat{A} = A(n)$, $\hat{\theta} = \theta(n)$
8: **end if**
9: $\theta(n + 1) = \text{algNM}(G, A(n), B)\theta(n + 1) = \text{algSPSA}(G, A(n), B)$
10: **end for**
11: **output** $\hat{A}$, $\hat{\theta}$
separately and get $O((n + R_{nm})R_{cas}M)$, largely reduced compared to the $O(R_{cas}M^2 + R_{nm}R_{cas}M)$ of the traditional NM method, where $R_{nm}$ denotes the iteration number of the simplex transformations. Moreover, the adoption of the shrinkage operation is considered to be a rarity, which has been verified by Torczon in [32].

The computational complexity of Algorithm 4 is $O(R_{spsa}R_{cas}M)$. It’s noteworthy that unlike solving the integer optimization problem with submodular property in the inner optimization, the outer optimization works on the dynamic graph with varying edge weights $p - \theta$ and solves the continuous optimization problem with no convenient closed form for the objective function. Thus, it’s difficult to avoid the evaluations of influence function, which also are the computational burden both in Algorithm 2 and Algorithm 4.

The computational complexity of Algorithm 5 is $O(R_{alt}(k \log N + (n + R_{nm})R_{cas}M))/O(R_{alt}(k \log N + R_{spsa}R_{cas}M))$ for the adoption of SBSP-NM/SPSA, respectively.

Finally, they could be summarized in Table 1. From Table 1, we could see that Algorithm 2 has comparable computational complexity with Algorithm 4. Note that since our focus is on designing effective algorithms to deal with the min-max problem (1), their efficiencies could further be improved by relieving the heavy computational burden $O(R_{cas}M)$ resulted from each evaluation of influence function, like replacing this simulation-based approach with proxy-based or sketch-based approaches [25].

| Algorithm       | Computational complexity                                      |
|-----------------|--------------------------------------------------------------|
| Algorithm 1     | $O((k \log N + M)$                                          |
| Algorithm 2     | $O((n + R_{nm})R_{cas}M)$                                   |
| Algorithm 4     | $O(R_{spsa}R_{cas}M)$                                       |
| Algorithm 5 (SBSP-NM) | $O(R_{alt}(k \log N + (n + R_{nm})R_{cas}M))$            |
| Algorithm 5 (SBSPA) | $O(R_{alt}(k \log N + R_{spsa}R_{cas}M))$           |

We compare the proposed algorithm with the following four edge removal methods and one fractional immunization method. (a) Degree: selecting the edges with the largest degrees successively (the degree of an edge is the minimum of the two degrees of the endpoints) until the summation constraint of the intervention meets. Here, for all four edge removal baselines, we assume the selected edge could have a 99% edge weight reduction as the intervention, namely $\theta_{ij} = 0.99*p_{ij}$ for each selected edge $(i, j)$, and the degree of a node is computed. (b) Rich-Rich: selecting the edges with the highest scores successively (the score of an edge is the product of the two degrees of the endpoints) until the summation constraint of the intervention meets. (c) Pagerank: selecting the edges with the largest PageRank scores [33] (the PageRank score of an edge is the minimum PageRank score among its two endpoints) until the summation constraint of the intervention meets. (d) NetMelt: selecting the edges with the highest scores that minimize the leading eigenvalue of the network [5]. (e) UNI: uniformly assigning the intervention according to their edge weights, namely $\theta = \frac{B}{|p|}$. Note that the degrees of nodes in Degree and Rich-Rich are computed based on the corresponding converted directed graphs by replacing each edge with two directed edges, likewise for the PageRank scores in Pagerank.

A. DATASET DESCRIPTION

G10-related graphs: We generated random graphs with various topologies on 10 nodes. We studied the following six graphs: (i) the G-uniform graph, with 20 edges and weights uniformly sampled from (0, 1); (ii) the G-sparse graph, with 13 edges and weights uniformly sampled from (0, 1); (iii) the G-dense graph, with 40 edges and weights uniformly sampled from (0, 0.5); (iv) the G-isolated graph, with two smaller isolated graphs to indicate largely changed topology, and 20 edges with weights uniformly sampled from (0, 1); (v) the G-gaussian0301 graph, with 20 edges and weights sampled from the N(0.3, 0.1) distribution; and (vi) the G-gaussian05015 graph, with 20 edges and weights sampled from the N(0.5, 0.15) distribution.

NetHEPT: This dataset has been extensively used in influence maximization studies [14]. It represents the collaboration network from the “High Energy Physics Theory” section of arXiv from 1993–2003, where nodes represent authors, and each edge indicates co-authorship between the pair of authors. The graph contains 12,008 nodes and 237,010 undirected edges (including duplicates). After removing the duplicated edges and cleaning the dataset, we obtained a graph with 118,521 undirected edges. We then randomly generated edge weights uniformly sampled from (0, 1) to obtain a weighted graph G-NetHEPT.

go-Facebook: This dataset describes social circles from Facebook, with 4,039 nodes and 88,234 undirected edges [34]. Similar to the G-NetHEPT graph, we randomly generated uniform edge weights from (0, 0.1) to obtain the weighted graph G-Facebook. Note that we chose smaller...
weights for the G-Facebook graph than the G-NetHEPT graph, due to the relative density of the Facebook network.

wiki-Vote: This dataset contains all the Wikipedia voting data from the inception of Wikipedia till January 2008 [35]. It contains 7,115 nodes and 100,762 undirected edges (converted from 103,689 directed edges), where nodes represent Wikipedia users and each edge indicates a voting action. We randomly generated edge weights uniformly sampled from (0,0.1) to obtain the weighted graph G-wikiVote.

In summary, though the three real-world public graphs relate to different social behaviors like collaboration and voting, the immunization considered here could all be regarded as the control of the influence spread on these networks.

B. RESULTS FOR G10-RELATED GRAPHS

In this subsection, we report the results of our simulations, providing comparisons within two groups: (i) \{NM, SPSA\} for optimization over \(\theta\); and (ii) \{WDD, GG, EX\} for solving \(A\), where GG and EX are the general greedy algorithm [4] and exhaustive search algorithms, respectively. We applied the exhaustive search algorithm only for \(k = 2\). Note that we used the traditional NM algorithm for these experiments, rather than the SBSP-NM algorithm, since the graph size is fairly small.

Varying seed size \(k\): Here, we have four combinations of algorithms (without including EX), and vary the seed size \(k\) from 1 to 5. We set the summation constraint \(B = 0.2\|p\|_1\). For GG, we simulated the infection process 10,000 times to approximate the influence function when selecting successive seed nodes. Furthermore, we used 500 simulations of the infection to approximate the value of \(f\) in our algorithms. We ran the iterative algorithm with 10 alternating steps. From Figure 2, we can see that WDD generally returns smaller values than GG in the influence maximization task (larger is better for the same outer optimization solver), indicating that WDD is slightly inferior to GG under our configurations. However, considering the substantial reduction in computational complexity (as shown in Section III-D), WDD is still acceptable for large graphs, since GG is usually computationally infeasible for them. What’s more, NM is slightly superior to SPSA on these small graphs (smaller is better for the same inner optimization solver).

Varying summation constraint \(B\): We have 6 combinations of possible algorithms. We ran simulations with \(B\) ranging from 0 to \(\|p\|_1\), and set the seed size \(k = 2\). The other parameters were set in the same manner as before. Inspecting Figure 3, we can conclude that both WDD and GG perform worse than EX (the best influence maximization method when the seed size \(k = 2\)), which fits the intuition well. But WDD also returns acceptable results compared to EX and
GG, especially considering the large computational complexity reduction, consistent with the results in Figure 2. Another interesting phenomenon is that SPSA sometimes outperforms NM when $B$ is large (like larger than 50% $|p|$). This suggests that NM is better than SPSA in weak immunization cases, while SPSA may be better in strong immunization cases.

C. RESULTS FOR NETHEPT, EGO-FACEBOOK AND WIKI-VOTE GRAPHS

In this subsection, we make comparisons between the algorithms in the group \{SBSP-NM, SPSA, Degree, Rich-Rich, Pagerank, NetMelt, UNI\} to optimize over $\theta$. In all cases, we use WDD to perform the inner optimization over $A$ for computational complexity concern.

Varying seed size $k$: To use a unified scale for the three
datasets, we varied the seed size $k$ from 6% of the respective number of nodes to 30%. We set $B = 0.1\|p\|_1$. For G-NetHEPT, G-Facebook, and G-wikiVote, we ran 10, 30, and 10 independent cascade model simulations to approximate the infection counts, and ran the alternating algorithm all with 10 iterations, respectively. From Figure 4, we can see that SPSA and SBSP-NM outperform the other five baselines most of the time. Furthermore, SBSP-NM performs much better than SPSA on all three graphs, under the weak immunization cases where fixed summation constraint $B = 0.1\|p\|_1$ is adopted. This may due to that the simplex transformation in SBSP-NM may work more efficiently in lower-dimensional stochastic search, while SPSA shows its advantage in higher-dimensional stochastic search by the usage of the approximation of the gradient.

Varying summation constraint $B$: We varied $B$ from 0.05\|p\|_1 to 0.25\|p\|_1. We set the seed size to be 10% of the respective number of nodes and set the other parameters in the same manner as before. Figure 5 shows that SPSA and SBSP-NM perform better than the other baselines in weak immunization cases, while they may get worse results in strong immunization cases, which need to be improved in the future.

Runtime comparisons: We also compare the runtime of the algorithms {SBSP-NM, SPSA, Degree, Rich-Rich, Pagerank, NetMelt, UNI} with all used WDD to solve inner maximization problem (because the adoption of GG could not make the programs converge in 72 hours). All programs were run on HTCondor [36]. The runtime includes the time required for running the special algorithm and writing output files, and is included as Table 2. We can see that SBSP-NM and SPSA would consume more runtime than the other baselines, which both result from the stochastic search by optimization techniques in the high-dimensional space of $\theta$. Nevertheless, considering the cost and the performance gain, they could still be useful in non-real-time tasks. What’s more, the runtime and the scalability could also further be improved by relieving the heavy burden of the simulation-based function evaluations, like adopting proxy-based or sketch-based approaches [25], which nevertheless goes beyond the scope of our paper. Moreover, SBSP-NM always shows comparable runtime with SPSA, validating the aforementioned analysis of their computational complexities.

V. CONCLUSION

We have studied the problem of controlling the worst-case infection in a network. We first formulated the optimal immunization problem as a constrained min-max problem, and then proposed several effective algorithms based on alternating optimization. To solve the inner layer discrete optimization over $A$ with a fixed value of $\theta$, we proposed a WDD method to select the seed set greedily and efficiently. Next and more importantly, to solve the outer layer continuous optimization over $\theta$ with a fixed value of $A$, we proposed two methods: SBSP-NM and SPSA. Through the simulations on different types of small synthetic graphs and larger-scale real-world datasets, we conclude that the WDD algorithm is computationally feasible for large graphs with acceptable performance. Furthermore, both SBSP-NM and SPSA demonstrate better performance than the other five baselines most of the time, and SBSP-NM usually outperforms SPSA under weak immunization conditions. We hope our work could give some helpful inspirations on the immunization in the worst case by the use of the optimization techniques, especially for decision-makers to deal with the unknown seeds. Note that the initials of $\theta$ in all the simulations of SBSP-NM and SPSA are set as an all-zero vector, which could further improve the performance if replaced by the results of the edge removal baselines, thus to combine the advantages of the stochastic search based on optimization techniques and the edge removal methods based on graph theory. Future research directions include establishing theoretical guarantees based on more tractable formulations and extending the algorithms to other epidemic models and intervention tasks.
ACKNOWLEDGMENT
The authors would like to thank the anonymous reviewers for their valuable and constructive comments that greatly contributed to improving the quality of the paper and the editors for their support during the review process.

REFERENCES
[1] Z. He, Z. Cai, J. Yu, X. Wang, Y. Sun, and Y. Li, “Cost-efficient strategies for restraining rumor spreading in mobile social networks,” IEEE Transactions on Vehicular Technology, vol. 66, no. 3, pp. 2789–2800, 2016.

[2] C. Nowzari, V. M. Preciado, and G. J. Pappas, “Optimal resource allocation for control of networked epidemic models,” IEEE Transactions on Control of Network Systems, vol. 4, no. 2, pp. 159–169, 2015.

[3] Z. Su, K. Ren, R. Zhang, and S.-Y. Tan, “Disrupting terrorist networks based on link prediction: A case study of the 9–11 hijackers network,” IEEE Access, vol. 7, pp. 61 689–61 696, 2019.

[4] W. Chen, Y. Wang, and S. Yang, “Efficient influence maximization in social networks,” in KDD, 2009, pp. 199–208.

[5] H. Tong, B. A. Prakash, T. Eliassi-Rad, M. Faloutsos, and C. Faloutsos, “Gelling, and melting, large graphs by edge manipulation,” in Proceedings of the 21st ACM international conference on Information and knowledge management. ACM, 2012, pp. 245–254.

[6] B. A. Prakash, L. Adamic, T. Iwashyna, H. Tong, and C. Faloutsos, “Fractional immunization in networks,” in ICDM, 2013, pp. 659–667.

[7] C. Chen, H. Tong, B. Prakash, T. Eliassi-Rad, C. Faloutsos, and D. Chau, “Node immunization on large graphs: Theory and algorithms,” IEEE Transactions on Knowledge & Data Engineering, no. 1, pp. 1–1, 2016.

[8] Y. Zhang, A. Ramanathan, A. Vullikanti, L. Pullum, and B. A. Prakash, “Data-driven immunization,” in ICDM, 2017, pp. 615–624.

[9] C. Nowzari, V. M. Preciado, and G. J. Pappas, “Optimal resource allocation for control of networked epidemic models,” IEEE Transactions on Control of Network Systems, vol. 4, no. 2, pp. 159–169, 2017.

[10] Y. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. Pappas, “Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks,” in Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on, 2013, pp. 7486–7491.

[11] H. Chan, L. Akoglu, and H. Tong, “Make it or break it: Manipulating robustness in large networks,” in Proceedings of the 2014 SIAM International Conference on Data Mining. SIAM, 2014, pp. 324–333.

[12] R. C. Gunasekara, C. K. Mohan, and K. Mehlroth, “Multi-objective optimization to improve robustness in networks,” in Multi-Objective Optimization. Springer, 2018, pp. 115–139.

[13] D. Shen and S. Cao, “An efficient immunization strategy based on transmission limit in weighted complex networks,” Chaos, Solitons & Fractals, vol. 114, pp. 1–7, 2018.

[14] D. Kempe, J. Kleinberg, and É. Tardos, “Maximizing the spread of influence through a social network,” in KDD, 2003, pp. 137–146.

[15] K. Lioniis and E. Pitoura, “Boosting nodes for improving the spread of influence,” arXiv preprint arXiv:1609.03478, 2016.

[16] J. T. Kim, V. Jog, and P.-L. Loh, "Computing and maximizing influence in linear threshold and triggering models," in NIPS, 2016, pp. 4538–4546.

[17] W. Chen, C. Wang, and Y. Wang, “Scalable influence maximization for prevalent viral marketing in large-scale social networks,” in KDD, 2010, pp. 1029–1038.

[18] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, C. Faloutsos, J. VanBriesen, and 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.

[19] A. Goyal, W. Lu, and 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.

[20] C. Borgs, M. Brautbar, J. Chayes, and 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.

[21] Y. Hu, Y. Koren, and C. Volinsky, “Collaborative filtering for implicit feedback datasets,” in ICDM, 2008, pp. 263–272.

[22] B. Colson, P. Marcotte, and G. Savard, “An overview of bilevel optimization,” Annals of operations research, vol. 153, no. 1, pp. 235–256, 2007.

[23] O. Kramer, D. E. Ciaurri, and S. Koziel, “Derivative-free optimization,” in Computational optimization, methods and algorithms. Springer, 2011, pp. 61–83.

[24] J. C. Spall, “Stochastic optimization,” in Handbook of computational statistics. Springer, 2012, pp. 173–201.

[25] Y. Li, J. Fan, Y. Wang, and K.-L. Tan, “Influence maximization on social graphs: A survey,” IEEE Transactions on Knowledge and Data Engineering, vol. 30, no. 10, pp. 1852–1872, 2018.

[26] J. A. Nelder and R. Mead, “A simplex method for function minimization,” The computer journal, vol. 7, no. 4, pp. 308–313, 1965.

[27] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, “Convergence properties of the nelder–mead simplex method in low dimensions,” SIAM Journal on optimization, vol. 9, no. 1, pp. 112–147, 1998.

[28] J. Nocedal and S. J. Wright, “Numerical optimization 2nd.” 2006.

[29] J. C. Spall, “Implementation of the simultaneous perturbation algorithm for stochastic optimization,” IEEE Transactions on aerospace and electronic systems, vol. 34, no. 3, pp. 817–823, 1998.

[30] S. Bhatnagar, H. Prasad, and L. Prashanth, Stochastic recursive algorithms for optimization: simultaneous perturbation methods. Springer, 2012, vol. 434.

[31] S. Yin, H. Luo, and S. X. Ding, “Real-time implementation of fault-tolerant control systems with performance optimization,” IEEE Transactions on Industrial Electronics, vol. 61, no. 5, pp. 2402–2411, 2013.

[32] V. J. Torczon, “Multidirectional search: a direct search algorithm for parallel machines,” Ph.D. dissertation, Rice University, 1989.

[33] L. Page, S. Brin, R. Motwani, and T. Winograd, “The pagerank citation scoring system: Bringing order to the web.” Stanford InfoLab, Tech. Rep., 1999.

[34] J. Leskovec and J. M. McAuley, “Learning to discover social circles in ego networks,” in NIPS, 2012, pp. 539–547.

[35] J. Leskovec, D. Huttenlocher, and J. Kleinberg, “Predicting positive and negative links in online social networks,” in Proceedings of the 19th international conference on World wide web. ACM, 2010, pp. 641–650.

[36] “Hicondor,” Accessed: March 11, 2018. [Online]. Available at: http://research.cs.wisc.edu/hicondor/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109ACCESS.2019.2962197, IEEE Access

TABLE 2: Runtime comparisons of the algorithms (seconds)

| algorithm/dataset | NetHEPT-k | ego-Facebook-k | wiki-Vote-k | NetHEPT-B | ego-Facebook-B | wiki-Vote-B |
|-------------------|-----------|----------------|-----------|-----------|----------------|-----------|
| SBSP-NM           | 40,506    | 26,112         | 16,626    | 40,014    | 26,406         | 18,725    |
| SPSPA             | 26,501    | 64,178         | 25,365    | 52,770    | 93,079         | 40,867    |
| Rich-Rich         | 4,256     | 8,224          | 3,151     | 4,051     | 7,963          | 2,611     |
| Paperunk          | 3,095     | 8,464          | 6,111     | 4,481     | 8,276          | 6,338     |
| NetMelt           | 4,501     | 6,673          | 3,443     | 4,334     | 9,162          | 3,379     |
| UNI               | 148       | 358            | 113       | 142       | 346            | 110       |