Designing Effective and Practical Interventions to Contain Epidemics

Prathyush Sambaturu[1], Bijaya Adhikari[2], B. Aditya Prakash[3], Srinivasan Venkatramanan[4], Anil Vullikanti[1 4]
[1] University of Virginia [2] Virginia Tech [3] Georgia Institute of Technology
[4] Biocomplexity Institute at University of Virginia

Outline

1. Introduction
   - Preliminaries
   - Problem Statement
2. Algorithm
   - SAAROUND Algorithm
   - Approximation Guarantees
   - Extensions and Improvements
3. Experiments
   - Performance Guarantees
   - Two stage Intervention
4. Conclusions

Designing Interventions to Contain Epidemics

- Vaccination is a standard public health intervention
- Limited Supply of vaccines, and are produced over time
- We study the problem EpiControl of designing vaccination strategies, within available budget constraints, to minimize the spread of an outbreak
- Challenging stochastic optimization problem
- Bi-criteria approximation algorithm
- Techniques: Linear programming (LP) based rounding and the sample average approximation (SAA)

Epidemic Model

- $G = (V, E)$ denotes a contact graph where V is the set of nodes and $e = (u, v) \in E$ if nodes $u, v \in V$ come into direct contact.
- SIR model of disease spread on networks, in which each node is in one of the following states: susceptible (S), infectious (I) or recovered (R).

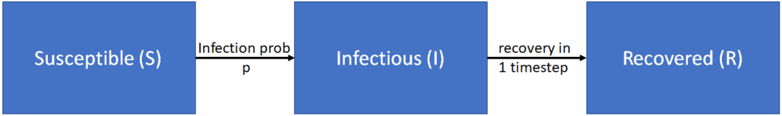

Figure: SIR model

- s_v is the probability that v is initially infected; $\mathbf{s}$ denotes the initial infection vector.

Interventions and Objective

- x_{vt} is an indicator variable, which is 1 if node v gets vaccinated at time t; let $\mathbf{X}_t = \{v : x_{vt} = 1, v \in V\}$, and let B_t denote the number of vaccines available for use at time t.
- **Single stage**: $\texttt{EInf}(G, \mathbf{s}, \mathbf{X}_0)$ denotes the expected number of infections when the intervention is done for set $\mathbf{X}_0$ at time $t = 0$.
- **Two stage**: $\texttt{EInf}(G, \mathbf{s}, \mathbf{X}_0, \mathbf{X}_T)$ denotes the expected number of infections if the interventions are done on sets $\mathbf{X}_0$ and $\mathbf{X}_T$ at times 0 and T, respectively.

Example

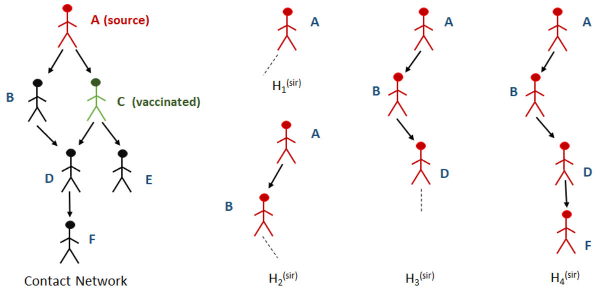

Figure: Contact network $G = (V, E)$ where $V = \{A, B, C, D, E, F\}$. Node A is initially infected, and node C is vaccinated. The subgraphs $H_1^{(sir)}, H_2^{(sir)}, H_3^{(sir)}, H_4^{(sir)}$ are possible stochastic outcomes in the SIR model, which occur with probabilities $1 - p$, $p(1 - p)$, $p^2(1 - p)$, and p^3, respectively.

Suppose, $x_{C0} = 1$, and $\mathbf{X}_0 = \{C\}$. We have

$$\texttt{EInf}(G, \mathbf{s}, \mathbf{X}_0) = (1 - p) + 2p(1 - p) + 3p^2(1 - p) + 4p^3$$

.

Problem Statement

Given a contact network $G = (V, E)$, and an initial infection vector $\mathbf{s}$, we consider the following problems

- **Single stage vaccination problem (1sEpiControl):** given budget B_0, choose $\mathbf{X}_0$ such that $|\mathbf{X}_0| \leq B_0$ and $\texttt{EInf}(G, \mathbf{s}, \mathbf{X}_0)$ is minimized.

- **Two stage vaccination problem (2sEpiControl):** given budget B_0, B_T, choose $\mathbf{X}_0, \mathbf{X}_T$ such that $|\mathbf{X}_0| \leq B_0$, $|\mathbf{X}_T| \leq B_T$, and $\texttt{EInf}(G, \mathbf{s}, \mathbf{X}_0, \mathbf{X}_T)$ is minimized.

Our Contributions

- Design SAAROUND for 1SEPICONTROL and show rigorous worst case guarantees on its performance.
- We show that SAAROUND is a good heuristic for the multi-stage problem.
- Scaling SAAROUND to large networks.
- Evaluate of our algorithm on diverse real and random networks and comparison to baselines.

Outline

1. Introduction
   - Preliminaries
   - Problem Statement
2. Algorithm
   - SAAROUND Algorithm
   - Approximation Guarantees
   - Extensions and Improvements
3. Experiments
   - Performance Guarantees
   - Two stage Intervention
4. Conclusions

Steps in SAARound

Variables y_{vj} are indicators for node v getting infected in sample H_j (i.e., there is a path from $\texttt{src}(H_j)$ to v with no node on it vaccinated).

1. Construct M sampled outcomes: $H_j = (V, E_j)$, for $j = 1, \ldots, M$. The idea is that it suffices to get a solution which minimizes the average number of infections in a set of M sampled outcomes, in order to minimize $\texttt{EInf}(\cdot)$.
2. Solve the following linear program (LP_{saa})

$$(LP_{saa}) \qquad \min \frac{1}{M} \sum_j \sum_v y_{vj} \tag{1}$$

$$\forall j, \forall u \in V: \; y_{uj} \leq 1 - x_{u0} \tag{2}$$

$$\forall j, \forall u \in V, \; (w, u) \in E_j: \; y_{uj} \geq y_{wj} - x_{u0} \tag{3}$$

$$\forall j, \forall s \in src(H_j): \; y_{sj} = 1 - x_{s0} \tag{4}$$

$$\sum_{u \in V} x_{u0} \leq B_0 \tag{5}$$

$$\text{All variables} \in [0, 1] \tag{6}$$

Steps in SAAROUND (Contd.)

3. Let x, y be the optimal fractional solution to (LP). We round it to an integral solution X, Y in the following manner
   - If y_{vj} is binary. Set $Y_{vj} = y_{vj}$. Similarly, for x_{v0}.
   - Round $Y_{vj} = 1$ for each (v, j) if $y_{vj} \geq \frac{1}{2}$, otherwise set $Y_{vj} = 0$.
   - For each v, set $X_{v0} = 1$ with probability $\min\{1, 2x_{v0} \log(4nMN)\}$, where N is the maximum number of paths from $\texttt{src}(H_j)$ to any node v in H_j.
   - $X_0 = \{v : X_{v0} = 1\}$ is the set of nodes vaccinated.
4. **return** X_0.

Approximation Guarantee

Theorem

Let X_0 denote the vaccination set computed by algorithm $\textsc{saaRound}$ *and $M \geq 24n^2 \log n$. Then, with probability at least $1/2$, we have $\mathtt{EInf}(X_0) \leq 6\mathtt{EInf}(X_{opt})$, and $|X_0| \leq 12\log(4nMN)B_0$.*

where,
$n = |V|$,
M is the number of sampled subgraphs considered,
N is the number of paths from sources to any node in H_j

Improvements to speed up saaRound

To improve the scaling of saaRound, we use the following methods.

- **Reduced number of samples**: In practice, we find that fewer samples are needed.
- **Reducing the number of variables** We can set $x_{vt} = 0$ for nodes with vulnerability (probability that it gets infected) at most γ.

Outline

1. Introduction
   - Preliminaries
   - Problem Statement
2. Algorithm
   - SAAROUND Algorithm
   - Approximation Guarantees
   - Extensions and Improvements
3. Experiments
   - Performance Guarantees
   - Two stage Intervention
4. Conclusions

Datasets

Dataset	Nodes	Edges
Montgomery	70729	198138
Portland	1409197	8307767
CA-GrQc	5242	14496
Small World (SW)	2500	14833
Preferential1 (PA1)	1000	1996
Preferential2 (PA2)	100000	199996

Table: Description of datasets

Impact of Pruning

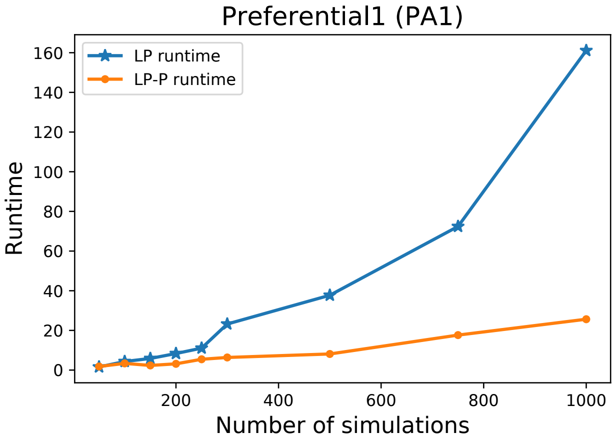

Figure: Comparison of runtimes of Linear Programs with (LP-P) and without pruning (LP).

Comparison to Baselines

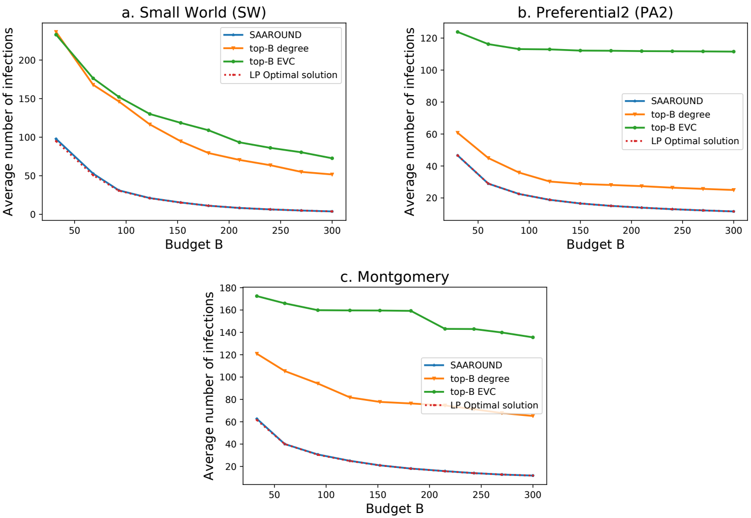

Figure: Objective value (y-axis) vs budget (x-axis) for SAARound, and the degree and eigenscore baselines for four networks.

Budget Violation

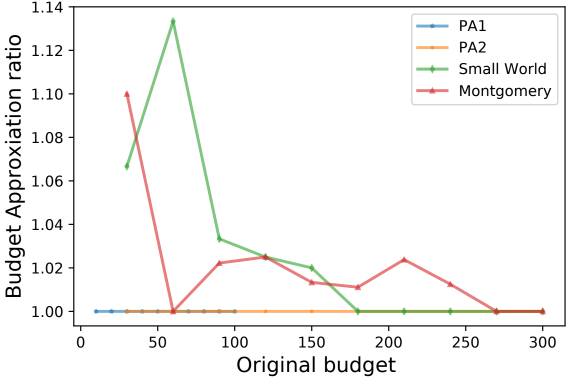

Figure: Budget Violation

Two Stage Intervention

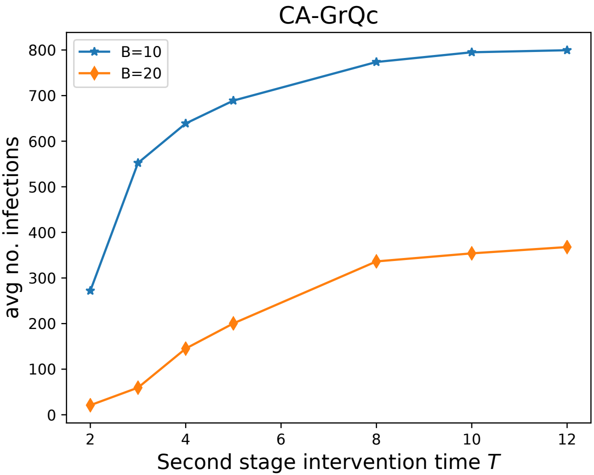

Figure: Two-stage Intervention.

Outline

1 Introduction
- Preliminaries
- Problem Statement

2 Algorithm
- SAAROUND Algorithm
- Approximation Guarantees
- Extensions and Improvements

3 Experiments
- Performance Guarantees
- Two stage Intervention

4 Conclusions

Conclusions

- LP based rounding and the sample average approximation technique provide good approximation guarantees in practice.
- Our results are the first to examine multi-stage interventions, and we find that the temporal dimension leads to significant changes in the solution quality and structure.
- Improving the approximation guarantees by better rounding techniques is an important open problem.
- Our methods can help in public health policy planning and response to large outbreaks.