Optimizing Algorithms for Immune Shielding on Networks

Andreea Măgălie
amagalie@gatech.edu

Jiajia Xie
jxie@gatech.edu

Haotian Yin
hyin79@gatech.edu

ABSTRACT

One proposed way to manage epidemics in the absence of a vaccine aside from social distancing and isolation of infected individuals has been immune shielding. Immune shielding focuses on recovered individuals and places them in highly connected spots in a community in order to decrease overall disease transmission. The recovered individuals act as shields since they are protecting the rest of the population by exposing themselves to infected individuals in order to minimize susceptible-infected interactions. Immune shielding works in theoretical frameworks, but there is a large gap between theoretical work and implementation. In this paper we test the idea of immune shielding on small, sparse networks and aim to assess how much can this preventive strategy change the course of an outbreak in a hospital or care home. We will use an optimization algorithm and risk scores in order to find the optimal way in which we can apply immune shielding and determine the limits by how much can this method reduce the size of an outbreak.

KEYWORDS

immune shielding, network rewiring, optimization algorithms

1 INTRODUCTION

In light of the recent pandemic it became clear that the present highly connected world is at great risk when a novel air-born pathogen appears in the human population. With vaccine development operating at a longer timescale than disease spread, other measures to contain a pathogen need to be put in place. The main containment measures used have been social distancing, contact tracing and isolation of infected patients. However we have seen that short and long term compliance with social distancing norms is a major issue and even with all the present measures, hospitals and health care facilities are hardest hit by the pandemic since their populations are more at risk. The question thus arises if we can use other measures to protect these high-risk networks.

One proposed measure to contain epidemics has been immune shielding. The idea behind immune shielding is that we can use individuals who have already acquired antibodies (recovered individuals) and strategically place them in the high-risk community such that their placement decreases the overall spread of the disease. In practice this could mean reassigning patients to health care workers (HCWs) based on their infectiousness. Given that some HCWs might be recovered we propose that changing patients lists might decrease the risk of disease transmission and eventually decrease the size of the outbreak within the health facility. The motivation behind reassigning patients is to replacing risky connections (susceptible - infected) with safe connections (recovered - infected) which do not further transmit the disease.

The idea of immune shielding has been proposed before in the context of large, well-mixed networks but has not been thoroughly studied on smaller, sparse networks which is where immune shielding is most likely to be implemented. While there are many approaches on implementing this strategy, we are interested in what the optimal one is and how much can it improve the outcome of an outbreak. Finding the implicit limits of this strategy will allow for a more realistic assessment of immune shielding performance in small communities and open new possibilities for further implementations.

2 PROBLEM DEFINITION

We start with a stochastic SIR model on a small network to represent a potential COVID outbreak in a health care facility. Following the assumption that we can use the immune shielding intervention every day, we ask what is the best way to rewire the network at every time step and keep node degrees constant such that the magnitude of an outbreak is minimized? By magnitude of an outbreak or epidemic size, we mean the final number of recovered individuals.

With the purpose defined, this project focuses on how to rewire a network while keeping every individual's degree constant. The purpose of maintaining the node degrees constant is to keep the workload or necessary care of individuals and maintain the overall social structure the same. The standard idea behind rewiring is that we assume that the infected individuals' connection with their susceptible neighbors would be removed, a new connection to a recovered individual will be established, and their susceptible neighbors will be connected to a recovered individual. This is called a swap and it replaces the edges $(S-I, R-R)$ with $(S-R, I-R)$, thus maintaining all node degrees the same. A schematic of this swap is shown below in Fig. 1.

2.1 Formulation

We can formulate the problem of finding the best rewired network at every time step as an integer programming problem. Note that while rewiring is deterministic (based on fixed degree distribution and node states S, I or R), our overall simulation is still stochastic because of the stochastic epidemiological model.

We can associate a cost with every adjacency matrix A, where higher cost means higher probability of spreading the disease to

Figure 1: Swap Schematic

Table 1: Summary of data set

Name	n	m
ia-enron-only	143	623
regular	100	600

agencies or papers of related works. For example, there is an administrative data set that has been transformed into a people-to-people network between physicians and patients, [4]. In addition, it will be useful for us to try out different theoretical networks for controls and ground check such as regular or complete graphs.

3 SURVEY ON RELEVANT PRIOR WORK

We review related works here, which can be categorized into two parts: structural-based approaches and resources-based approaches. None of them is able to simultaneously maintain the social connections within an environment and to prevent an outbreak of pandemics.

In general, approaches for pandemic intervention heavily relies on attacking the network structure, by either quarantining individuals or shutting down locations. For instance, [6] proposes an idea of threshold control of agent-based models, which is to minimize the first eigenvalue of the corresponding adjacency matrix by removing nodes from a graph. Other measure of significance like Pagerank, connectivity of neighbors etc. all work in the same way which eventually will produce a different network. There are two shortcomings need to be addresses. The first is the feasibility of applications. The approaches do not take into account a cost measure of breaking down connections. In real life, we have to be careful to consider shutting down a hospital, train station, or restricting students on attending schools because the social cost is tremendous. Second, they do not capture the alternative behaviors of humans. For example, in real life, a location is not opened does not means some people who have been to the place will disappear from the graph.

The focus of this work, immune shielding, has been introduced in the context of COVID-19 in early May [8]. In that work, immune shielding is framed in a theoretical framework which implies infinitely large populations where everyone is connected to everyone else. While [8] provides a good starting point, it fails to address how immune shielding works on smaller and sparse networks. The original work also assumes that the shields (recovered individuals) undertake a higher workload. This might not be possible in a hospital scenario, where most of the medical personnel is already working at maximum capacity. Here we aim to solve these shortcomings by applying the ideas in [8] on small networks and maintaining node degrees constant.

There are works particularly model constraints on resource as well as saving the social connections among people such as [7] and [1]. The prior aims to optimally allocates a vaccines, a typically limited resource, to reduce the risk of an outbreak. The second works in a smaller health-care setting that optimally surveils facilities and employees to reduce the probability of undetected infections. We see also two challenges for implementing these ideas.
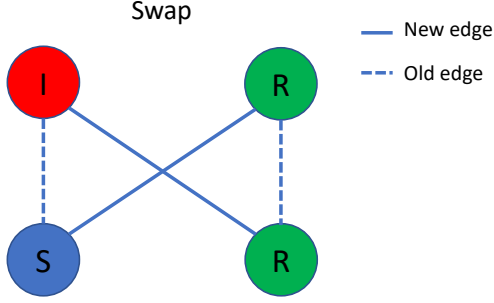
more nodes. Hence, we would change $A(t)$ in order to minimize the cost at every step. The easiest cost function would be the number of edges of the type SI. A more general cost function would use a cost matrix for all possible edge types. Hence we define a cost function which depends on the adjacency matrix and node types

$$f(A, v) = \sum_{\forall i,j} A_{ij} \cdot C(v_i, v_j), \qquad (1)$$

where A_{ij} indicates whether there is an edge between i and j and $C(v_i, v_j)$ represents the cost of the potential edge. In order to do a step-wise optimization we can formulate this as an integer programming problem.

Let $x_{i,j}$ for $i, j \in N$ be the decision variables, where $N = \{1, 2, 3...n\}$. let k be a budget of edges rewiring.
We formulate the rewiring at every time step in our epidemic as a deterministic integer programming problem:

$$\min_x \quad \sum_{\forall i,j} x_{ij} \cdot C(v_i, v_j) \cdot W(v_i, v_j)$$

$$\text{s.t.} \quad \forall i \in N, \qquad \sum_{j \in N} x_{ij} = \sum_{j \in N} A_{ij}$$

$$\forall i \in N, \qquad \sum_{j \in N} x_{ji} = \sum_{j \in N} A_{ji} \quad (2)$$

$$\forall i \in N, \forall j \in N, \qquad x_{i,j} \in \{0, 1\}$$

$$\text{Given } k, \qquad \sum_{i,j \in N} x_{ij} A_{ij} \geq |E| - k$$

where $W(v_i, v_j)$ denotes the weight of each edge $(v_i, v_j) \in E$ and $C(v_i, v_j)$ is the cost associated with every edge type. This cost is predetermined by us based on how 'risky' an interaction between v_i and v_j is. The last constraint is a budget constraint which controls the number of edges rewiring to be less than or equal to k.

2.2 Data Description

Given the scope of the problem, the ideal data used for this project are hospital networks with population in the order of hundreds. In particular, we are going to use existing data sets from medical

One is indeed a scarce resource with a large problem. Even, there is no vaccination of COVID19 available to us at the moment when this paper is written. Another issue is the problem formulated are computationally hard to solve in general.

[2] proposed a preventive rewiring method on a large epidemic network model, which allows the susceptible node to reconnect itself with a random susceptible node or recovered node with certain successful rate. The paper shows that if the successful rate of re-connection increases, the number of infected individuals at each time step decreases which leads to a smaller epidemic size. The result shows that the rewiring could be an effective method to control the epidemic size in a large network. The model used in [2] can arbitrarily reassign the connection and assumes the degree distribution does not change significantly for initial phases. However, not every network in real life is large and dense. For smaller networks or sparser networks, the effect of change of degree distributions should be considered.

[5],further explore the effect of preventive measures of [2], it applies the rewiring algorithm on configuration network, clique network, and Real-world networks. For three networks, certain level of rewiring actually worsen the epidemic size at different stage of the epidemic. The result indicates that with higher rate of preventive measures, these measures will effectively prevent the epidemic to occur since it will decrease R_0 at initial phase of a outbreak. However, if the preventive measures are applied with a moderate rate, they might have negative effects on the epidemic. That is to say, when designing the rewiring algorithm of the project, the frequency of rewiring should be considered.

4 PROPOSED METHODS

4.1 Intuition

We are running a stochastic SIR algorithm on a network. Let $A(t)$ be the adjacency matrix at time t. We aim to change the adjacency matrix at every step such that we minimize the epidemic size, i.e. the final number of infected individuals R_∞. When we rewrite the adjacency matrix we aim to keep node degrees constant, i.e. the sum on all columns and rows the same. To do so, our algorithms rely on minimizing the number of SI links in order to minimize the the number of risky connection at a certain time. The mathematical motivation behind this is that given R_∞ can also be written as the sum of the new number of infections every day:

$$R_\infty = \sum_{t=0}^{\infty} I_{new}(t). \tag{3}$$

At a given time t we can write

$$\mathbb{E}\big(I_{new}(t+1)\big) = \beta \sum_{v \text{ sus}} \frac{\#S_v I}{deg(v)}, \tag{4}$$

since every S node v will become infected with probability

$$\beta \cdot \text{(number of SI links from v)}/\text{(degree of v)},$$

because we are using a frequency dependent model. We have made this choice because it approximates transmission in a health facility better given that in a small dt, one individual is likely to interact only with one of it's neighbors, not all of them.

Hence we see that minimizing the overall number of SI connections also decreases $I_{new}(t)$ from (4) which in turn decreases R_∞ based on (3).

We make one important note that when the networks is regular and every node has degree k, the value in (4) becomes:

$$\mathbb{E}\big(I_{new}(t+1)\big) = \beta \cdot \frac{\text{Total number of SI links}}{k}, \tag{5}$$

meaning that decreasing the number of SI links is enough to guarantee that we obtain an optimal solution in this simplified case. This is also true because in a regular network the nodes become indistinguishable and the local solution of minimizing the number of SI links from day to day becomes the global optimum as well.

4.2 Model Description

The model being used throughout this study is a simple agent-based SIR model with two parameters $\beta = .2/day$ and $\gamma = .1/day$ and $dt = 1\ hour$, where β is the probability of getting infected and γ is the probability of recovery. This provides a baseline shown in Fig. 2 on which we can further improve with our interventions.

We have developed several algorithms which aim to minimize the number of SI connections at every time step: minAreg, GreedySI and GreedySI+.
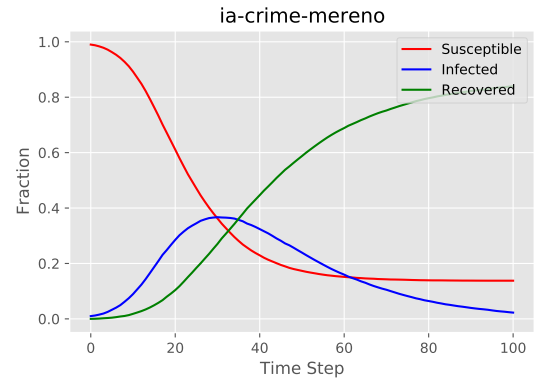

Figure 2: SIR model: $\beta = 0.1$, $\gamma = 0.05$

4.2.1 ***minAreg***. This algorithm takes in node types and a degree distribution for a regular graph and finds an adjacency matrix with the minimum number of SI links. Note that there are multiple such matrices, but because of (5) we can choose either of them and the expected number of new infected nodes will remain the same. We want to focus on this case because it is a particular case which is not too far from real distributions and we can find the optimal solution for.

Writing algorithms that construct an adjacency matrix step-wise is a tricky process. The main problem is that once we decide to connect two nodes, the remaining degree distribution might not be feasible. To complicate matters, knowing if a degree distribution is feasible is a costly and tricky process as well. For example the degree list $[1, 1, 2]$ forms a 'V', $[1, 1, 1, 2]$ cannot be reconstructed into a network without self loops.

The algorithms **minAreg** works by exhaustively connecting I nodes to other I and R nodes and then connecting the remaining edges. For example if there are more I nodes than degree k, we can connect them with each other and we will have no SI edges without needing to make an IR edge. The problem of a degree distribution becoming unfeasible is too hard because of the regularity in the network.

4.2.2 **GreedySI.** The first proposed algorithm for solving the problem formulated with $W(v_i, v_j) = 1$ for every $(v_i, v_j \in E)$ is GreedySI 1. In GreedySI, steps 5-12 divide edges into S and pre-sorted R set in ascending order based on cost of edges based on their category. Then, the algorithm loops over every SI edge from S to find a candidate SR or RR edge in R to rewire. The time complexity of GreedySI is $O(kn^2)$.

Algorithm 1 GreedySI

1: **procedure** GREEDYSI($G(V, E), k$)   ▷ k is a budget constraint
2:     $S \leftarrow [\ ]$
3:     $R \leftarrow [\ ]$
4:     $C = 0$
5:     **for** $(v_i, v_j) \in E$ **do**
6:         **if** $(v_i, v_j) == SI$ **then**
7:             $S \leftarrow (v_i, v_j)$ ▷ Append the edge to the list
8:         **else if** $(v_i, v_j) == RR$ **then**
9:             $R \leftarrow (v_i, v_j)$
10:     **for** $(v_i, v_j) \in E$ **do**
11:         **if** $(v_i, v_j) == SR$ **then**
12:             $R \leftarrow (v_i, v_j)$
13:     **while** $R \neq \phi$ **do**   ▷ Stop when R is empty
14:         $(v_i, v_j) = S.pop()$ ▷ .pop() will return the last element of the list
15:         $(u_i, u_j) = R.pop()$
16:         **if** $\{v_i, v_j\} \cap \{u_i, u_j\} = \phi$ **then**   ▷ No overlap between two edges
17:             **if** $(v_i, u_i) \notin E$ **then**   ▷ new edge cannot accept if it already exist
18:                 **if** $(v_j, u_j) \notin E$ **then**
19:                     **if** $u_j == R$ **then**
20:                         $E \setminus \{(u_i, u_j), (v_i, v_j)\}$
21:                         $E \cup \{(v_i, u_i), (v_j, u_j)\}$
22:                         $C = C + 1$
23:             **else if** $(v_i, u_j) \notin E$ **then**
24:                 **if** $(v_j, u_i) \notin E$ **then**
25:                     $E \setminus \{(u_i, u_j), (v_i, v_j)\}$
26:                     $E \cup \{(v_i, u_j), (v_j, u_i)\}$
27:                     $C = C + 1$
28:         **if** $C \geq k$ **then**
29:             $R = \phi$
30:     **return** $G(V, E)$

Since all weights are the same in this problem there is no distinction among SI, RR, and SR itself which implies the algorithm treats each edge in the same priority as others. Note that this algorithm is considered to be the baseline method for the study. In

the next section, weight will be taken in account as a part of the objective function. We will define the weight of each edge first and design a similar algorithm as GreedySI to approximate the minimum weighted sum of risk.

4.2.3 **GreedySI+.**

The Greatest Eigenvalue Method. Instead of defining a fix cost-matrix, we consider using a different measure against the likelihood of spreading. We are going to implement the idea of "vulnerability" score measured by the greatest eigenvalue of the adjacency matrix [6]. Let $C(v_i, v_j)$ be the score function calculated by taking the product of "Shield-value" between v_i and v_j. The rest of the optimization objective and constraints are the same as above.

We propose an improved version of GreedySI, GreedySI+, as the approach to solve for the weighted sum of risk problem formulated in 4.3. Given a contact network $G(V, E)$, the weight for each edge $(v_i, v_j) \in E$ is defined as:

$$W(v_i, v_j) = p_{v_i} * q_{v_j} \qquad (6)$$

where p, q are the left and right eigenvectors of the adjacency matrix A of $G(V, E)$ respectively corresponding to the greatest eigenvalues. The above defined weight is a good approximation of the number of decrease in the leading eigenvalue of the adjacency matrix if removing edge (v_i, v_j) from the graph $G(V, E)$ based on [3].

In GreedySI+ SI edges and (RR, SR) edges are sorted in ascending order based on the product of two leading eigenvalues. By the property of this greedy algorithm, an SI edge with high importance with respect to the leading eigenvalue will be swapped with a RR or SR edge with high importance as well. Since the leading eigenvalue contributes to the bound of the reproduction number, R0, of agent-based models, we expect the SIR model has a lower R0 by running GreedySI+.

5 EXPERIMENTAL EVALUATIONS

In this section, we propose three different experiments to empirically evaluate the algorithms minAreg, GreedySI and GreedySI+. We will first compare them without any restrictions on a regular network and social network and then we will compare the effectiveness of the last two approximating algorithms with 1) changing in budget 2) changing the time when we start the intervention. We will conduct the experiment by running these algorithms over the data-set defined in Table 1.

To evaluate the algorithms, we will run the simulation a large number of times to obtain a mean outbreak size and standard deviation. We will then compare those numbers to the control when no rewiring is done and see how much the rewiring improves the outcome.

5.1 No restrictions

First we run our intervention algorithms on a regular network with 100 nodes and average degree 6. We start with this particular case because **minAreg** gives us the optimal solution and we can compare how our approximations do compared to the optimal one.

We then run our simulations again with no restrictions and daily
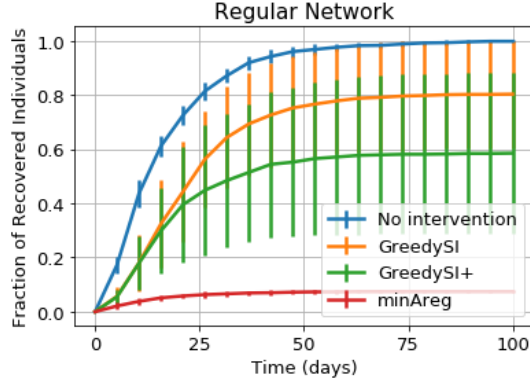
Figure 3: Interventions on regular network with 100 nodes and degree 6. $\beta = 0.2/24$, $\gamma = 0.07/24$. We start the simulation with 5% infected individuals and run the simulation 10 times.

interventions for a social network with 143 nodes and average degree 8. Results are shown in Fig. 4.
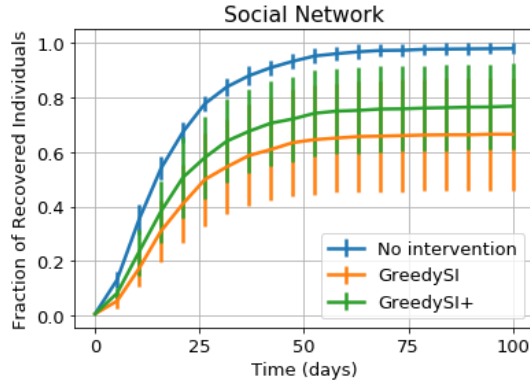

Figure 4: Interventions on social network with 143 nodes and degree 8. $\beta = 0.2/24$, $\gamma = 0.07/24$. We start the simulation with 5% infected individuals and run the simulation 10 times.

5.2 Budgets

In this part, we compare the performance of **GreedySI** and **GreedySI+** with controlled in variant intervention budget. The experimental result will answer question like: what is the minimum of intervention budgets needed to quickly bring down the size of an outbrak?

We present the experimental result on ia-enron-only for with various intervention budgets in Fig. 5. The results implies that for both **GreedySI** and **GreedySI+**, as budgets get larger, the size of an outbreak could be further reduced. However, after the budget is larger than a certain threshold, the size of an outbreak will remain nearly unchanged.
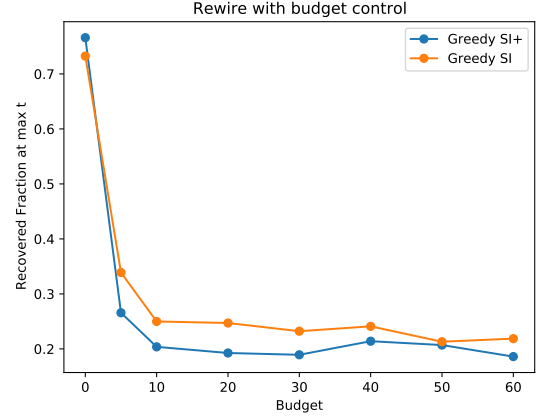
Figure 5: Final Recovered Fraction on ia-enron-only network with different intervention budget. $\beta = 0.2/24$, $\gamma = 0.1/24$. We start the simulation with 1 % infected individuals and run the simulation 15 times. Index 0 means no intervention. Index 60 means infinite budgets. Budgets in [0,5,10,20,30,40,50, infinite]

5.3 Initial Valid Time

In this part, we compare the performance of **GreedySI** and **GreedySI+** with controlled in variant start time given a low or high budget. The experimental result will answer question like: when should we apply these algorithms in practice to minimize the size of an outbreak if given a low budget or high budget?

We present the experimental result on ia-enron-only for both low and high budget in Fig. 6. The result implies an increasing budget for **GreedySI** requires an delayed intervention in order to minimize the size of an outbreak. On the other hand, GreedySI+ requires an earlier intervention. This explains the out-performance of GreedySI+ since it takes advantage of minimizing the leading eigenvalue of the adjacency matrix, which is a bound for an outbreak threshold. Logically, to prevent an outbreak, it must be done in an earlier stage.

6 CONCLUSION AND DISCUSSION

Our first observation across experiments is that all interventions significantly reduce the size of the outbreak. When we first look at Fig.3 we do notice that the optimal solution performs much better than any of the other algorithms. One of the reasons we suspect this is the case is because **minAreg** relies heavily on connecting and separating infected and recovered nodes. In **minAreg** we completely reconstruct the adjacency matrix every day and connect infected nodes between each other. This is different than any of the other interventions where we only connect infected individuals with recovered ones.

Comparing figures 3 with 4 we can say that there is not a consistent difference between GreedySI and GreedySI+ and that their performance might depend on the type of network we are using.

Andreea Măgălie, Jiajia Xie, and Haotian Yin
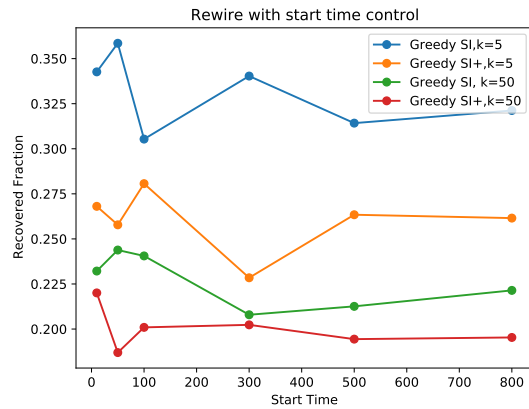
Figure 6: Final Recovered Fraction on ia-enron-only network with different intervention start time and budget. $\beta = 0.2/24$**,** $\gamma = 0.1/24$**. We start the simulation with 1 % infected individuals and run the simulation 15 times.**

When we look at how the budget affects the outcome in Fig. 5 we see that it acts as a threshold. As long as the budget is larger than ≈ 10, the final recovered fraction remains unchanged. This result is in line with what we expected since GreedySI and GreedySI+ don't perform too many switches. The effect of the budget important to consider when thinking about limitations about how often an how much can hospitals reassign their patients.

There is still plenty of work to be done on this topic and these results highlight that while these interventions can be very powerful, we still need to perform further work to understand how they work on different networks. Another important aspect we have not considered is that hospital networks can usually be divided into HCWs and patients, thus imposing further restrictions on the type of networks and rewiring we can perform. While we tried to improve from the baseline GreedySI algorithm, the optimal solution yields probably a much better result based on our special case in Fig. 3. The key change seems to be cohorting infected and recovered individuals as much as possible, i.e. not only swapping SI with RR links, but more broadly reducing the number of SI links and replacing them with both IR and II.

ACKNOWLEDGMENTS

This work is motivated by an ongoing project with the Weitz group at Georgia Tech. In that project we try to apply immune shielding in the context of hospital or care home facilities and focus on practical applications such as patient cohorting and approachable algorithms which can be passed further down to health care providers. A shortcoming of those projects is that we are not trying to answer any theoretical questions, but rather have feasible strategies which work 'well enough'. This work is fundamentally different since we are focusing on algorithm design and testing the limits of immune shielding. While the project in Weitz group is concerned with how can we implement immune shielding in hospitals, we are concerned with how well could immune shielding perform at it's peak.

REFERENCES

[1] Bijaya Adhikari, Bryan Lewis, Anil Vullikanti, José Mauricio Jiménez, and B. Aditya Prakash. 2019. Fast and near-optimal monitoring for healthcare acquired infection outbreaks. *PLOS Computational Biology* 15, 9 (09 2019), 1–22. https://doi.org/10.1371/journal.pcbi.1007284

[2] Tom Britton, David Juher, and Joan Saldaña. 2016. A Network Epidemic Model with Preventive Rewiring: Comparative Analysis of the Initial Phase. *Bulletin of Mathematical Biology* 78, 12 (2016), 2427–2454. https://doi.org/10.1007/s11538-016-0227-4

[3] Chen Chen, Hanghang Tong, B. Aditya Prakash, Tina Eliassi-Rad, Michalis Faloutsos, and Christos Faloutsos. 2016. Eigen-Optimization on Large Graphs by Edge Manipulation. *ACM Trans. Knowl. Discov. Data* 10, 4, Article 49 (June 2016), 30 pages. https://doi.org/10.1145/2903148

[4] Bruce E Landon, Jukka-Pekka Onnela, Nancy L Keating, Michael L Barnett, Sudeshna Paul, A James O'Malley, Thomas Keegan, and Nicholas A Christakis. 2013. Using administrative data to identify naturally occurring networks of physicians. *Medical care* 51, 8 (2013), 715.

[5] Ka Yin Leung, Frank Ball, David Sirl, and Tom Britton. 2018. Individual preventive social distancing during an epidemic may have negative population-level outcomes. *Journal of The Royal Society Interface* 15, 145 (2018), 20180296. https://doi.org/10.1098/rsif.2018.0296 arXiv:https://royalsocietypublishing.org/doi/pdf/10.1098/rsif.2018.0296

[6] B. Aditya Prakash, Deepayan Chakrabarti, Nicholas C. Valler, Michalis Faloutsos, and Christos Faloutsos. 2012. Threshold conditions for arbitrary cascade models on arbitrary networks. *Knowledge and Information Systems* 33, 3 (2012), 549–575. https://doi.org/10.1007/s10115-012-0520-y

[7] Prathyush Sambaturu, Bijaya Adhikari, B. Aditya Prakash, Srinivasan Venkatramanan, and Anil Vullikanti. 2020. Designing Effective and Practical Interventions to Contain Epidemics. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '20)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1187–1195.

[8] Joshua S. Weitz, Stephen J. Beckett, Ashley R. Coenen, David Demory, Marian Dominguez-Mirazo, Jonathan Dushoff, Chung-Yin Leung, Guanlin Li, Andreea Măgălie, Sang Woo Park, Rogelio Rodriguez-Gonzalez, Shashwat Shivam, and Conan Y. Zhao. 2020. Modeling shield immunity to reduce COVID-19 epidemic spread. *Nature Medicine* 26, 6 (2020), 849–854. https://doi.org/10.1038/s41591-020-0895-3