

Effective Social Network-Based Allocation of COVID-19 Vaccines

Jiangzhuo Chen
University of Virginia
Charlottesville, VA, USA
chenj@virginia.edu

Stefan Hoops
Achla Marathe
University of Virginia
Charlottesville, VA, USA

Henning Mortveit
Bryan Lewis
University of Virginia
Charlottesville, VA, USA

Srinivasan Venkatramanan
University of Virginia
Charlottesville, VA, USA

Arash Haddadan
Amazon Inc.
Bellevue, WA, USA

Parantapa Bhattacharya
University of Virginia
Charlottesville, VA, USA

Abhijin Adiga
Anil Vullikanti
University of Virginia
Charlottesville, VA, USA

Aravind Srinivasan
University of Maryland
College Park, MD, USA

Mandy L Wilson
University of Virginia
Charlottesville, VA, USA

Gal Ehrlich
Maier Fenster
Ehrlich & Fenster of the Ehrlich
Group
Ramat-Gan, Israel

Stephen Eubank
Christopher Barrett
University of Virginia
Charlottesville, VA, USA

Madhav Marathe
University of Virginia
Charlottesville, VA, USA
marathe@virginia.edu

ABSTRACT

We study allocation of COVID-19 vaccines to individuals based on the structural properties of their underlying social contact network. Using a realistic representation of a social contact network for the Commonwealth of Virginia, we study how a limited number of vaccine doses can be strategically distributed to individuals to reduce the overall burden of the pandemic. We show that allocation of vaccines based on individuals' degree (number of social contacts) and total social proximity time is *significantly more effective* than the usually used age-based allocation strategy in reducing the number of infections, hospitalizations and deaths. The overall strategy is robust even: (i) if the social contacts are not estimated correctly; (ii) if the vaccine efficacy is lower than expected or only a single dose is given; (iii) if there is a delay in vaccine production and deployment; and (iv) whether or not non-pharmaceutical interventions continue as vaccines are deployed. For reasons of implementability, we have used degree, which is a simple structural measure and can be easily estimated using several methods, including the digital technology available today. These results are significant, especially for resource-poor countries, where vaccines are less available, have lower efficacy, and are more slowly distributed.

CCS CONCEPTS

• **Computing methodologies** → **Modeling and simulation**; *Agent / discrete models*; • **Applied computing** → **Health informatics**.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
KDD '22, August 14–18, 2022, Washington, DC, USA
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9385-0/22/08.
<https://doi.org/10.1145/3534678.3542673>

KEYWORDS

COVID-19, vaccine prioritization, network degree, agent-based model, scenario modeling

ACM Reference Format:

Jiangzhuo Chen, Stefan Hoops, Achla Marathe, Henning Mortveit, Bryan Lewis, Srinivasan Venkatramanan, Arash Haddadan, Parantapa Bhattacharya, Abhijin Adiga, Anil Vullikanti, Aravind Srinivasan, Mandy L Wilson, Gal Ehrlich, Maier Fenster, Stephen Eubank, Christopher Barrett, and Madhav Marathe. 2022. Effective Social Network-Based Allocation of COVID-19 Vaccines. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22)*, August 14–18, 2022, Washington, DC, USA. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3534678.3542673>

1 INTRODUCTION

New vaccines typically take a decade to develop and distribute, but vaccines for COVID-19, the disease caused by the novel coronavirus SARS-CoV-2, have been developed in record time to help mitigate the raging pandemic. As of May 23, 2022, the reported numbers of confirmed cases and deaths are over 83M and 1M in the US, and over 525M and 6M worldwide, respectively. Vaccines offer a safe and effective way to contain the pandemic quickly. However, the supply of COVID-19 vaccines is limited at the beginning of the pandemic, and in underdeveloped countries. It is a challenge how to distribute vaccines in a timely manner to bring a pandemic under control before most of the population are infected.

Vaccination priority is complex and intertwined with age, race, occupation, health equity, geography, and politics. Data shows that COVID-19 disproportionately affects individuals with comorbidities, as well as people of low socio-economic status and high social vulnerability. There can be many criteria for prioritization, for example: (i) risk of infection; (ii) risk of death; (iii) risk of transmission if infected; and (iv) occupation, such as healthcare workers, teachers, cashiers, etc. The public seemed to agree with prioritizing people

with serious comorbidities, certain occupations, and those who had been disproportionately affected by COVID-19 [40]. Estimating the consequences of different prioritization strategies is complicated by production limitations, however. Additionally, vaccine distribution requires complex logistical support, such as cold-chain storage, transportation, qualified personnel, and scheduling etc. for any prioritization scheme to achieve its results in an effective and equitable manner. See [4] for a comprehensive discussion on this topic.

The prioritization order used by the US Centers for Disease Control and Prevention (CDC) recommended healthcare personnel and long-term facility care residents be vaccinated first; followed by frontline essential workers, and those aged 75 years and older because they are at a higher risk of hospitalization, illness and death; followed by those aged 65–74 years; followed by those aged 16–64 years with underlying medical conditions and other essential workers. **In this paper we retrospectively study the vaccine allocation problem assuming a situation in January 2021, when COVID-19 vaccines became available but were still limited in the US.** We are interested in a better allocation strategy, that could have had a larger reduction of infections, hospitalizations, and deaths, than the age-based one.

Our contributions. When vaccine supply was limited and emerging variants were accelerating the pandemic in different parts of the world, as was the situation in early 2021, a natural question is: *can we prioritize vaccine distribution so as to significantly reduce the overall burden of COVID-19 quickly?*

We propose prioritization schemes based on properties of individuals within social contact networks with the goal of bending the pandemic curve and improving overall pandemic outcome. We synthesize a digital twin of Virginia, which is a detailed social contact network model for the Commonwealth of Virginia (8 million individuals), and use an agent-based model (ABM) to study the effectiveness of various prioritization schemes. In contrast to other such networks, our networks incorporate detailed information about the population, their activities and the built infrastructure. Further information on how such a digital twin is constructed and its structural properties can be found in [13]. Our ABM simulates disease propagation and a complex set of interventions, including various non-pharmaceutical interventions (NPIs) and vaccine allocation schemes.

Our prioritization schemes based on simple, individual-based yet computable, structural properties of the underlying social contact network are motivated by: (i) recent advances in network science that have studied such schemes in more abstract settings; (ii) our ability to construct detailed, realistic social contact networks at scale; (iii) our ability to simulate and assess such strategies even for complex disease transmission models and public health control measures; and (iv) recent progress in development of digital apps that can be used for measuring structural properties in large populations relatively accurately, rendering such schemes potentially operationalizable.

Our prioritization schemes can be stated simply as follows: *vaccinate individuals who typically exhibit high social contact (degree or total contact time in the social contact network).* Some key points to note: (i) we focus on simple network structural properties that can

be estimated in a privacy-preserving way, (ii) we do not insist on strict ordering of individuals nor an exact estimation of their social contacts, and (iii) while our *analysis* uses a realistic representation of the social contact networks, implementation of the policy does not require one to *synthesize* the social network.

There is folklore that degree-based heuristics to allocate vaccines often work well. The folklore is based on mathematical results for highly structured random networks or on computational experiments based on relatively simple classes of social contact networks [8, 37, 46]. But the folklore has never been tested in time-varying realistic social contact networks such as the one constructed here and intended to capture the network evolution due to adaptive NPIs and vaccine allocation that is undertaken in a time varying manner. Our results show for the first time that degree-based heuristics are likely to work even for such time-varying social contact networks; see Sections 5 and [13] for further discussions on this topic.

Our results suggest that in just two months (i.e. by the end of March 2021), compared to age-based allocation, the proposed degree-based strategy can result in averting an *additional 56–110k infections (8–16%), 3.2–5.4k hospitalizations (8–13%), and 700–900 deaths (6–8%) just in the state of Virginia. Extrapolating these results per capita for the entire US, we estimate this strategy will lead to 3–6 million fewer infections, 181–306k fewer hospitalizations, and 51–62k fewer deaths compared to the age-based allocation.* The results continue to hold qualitatively and show that we can avert many more infections, hospitalizations, and deaths even if the current social distancing measures are relaxed. Furthermore, similar results hold even for vaccines with 50% efficacy; this is important, as most resource-poor countries did not have access to high efficacy vaccines at this point in time. The basic intuition behind our results is that vaccinating individuals with high degree not only protects them but also confers significant protection to individuals who come in close proximity in their contact network.

A natural question is: *how can such individuals be identified?* A person could be designated as “high degree” through identifying data or proxy characteristics to necessary statistical precision that show the individual belongs to such a critical group identified by the model. We discuss how currently deployed digital contact tracing apps can be modified in a very simple manner to achieve the goal of identifying high degree individuals in [13]. Such individuals can also be identified by observing that certain occupations naturally lead to a high level of social interactions. Our methods are robust to partial mis-estimation of these social contacts and their implementation does *not* require access to the social contact network. **It’s worth pointing out that our results are consistent with high degree-based strategies but implementation of such policies is the key.**

2 EXPERIMENT SETTINGS AND DESIGN

For the experiments, we use an agent-based simulation model, Epi-Hiper. The simulation’s input parameters specify the population demographics and contact network, COVID-19 disease model, initial configuration S_0 , NPIs, and vaccination schedule. The simulation output is a dendrogram: a directed graph that tells us who infects whom and on what day. From the output data, we can compute epidemiological measures such as daily new infections, cumulative

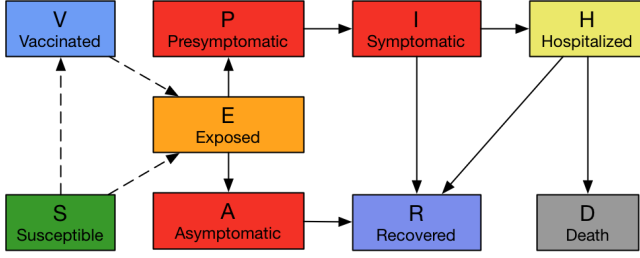


Figure 1: The COVID-19 disease model is represented as a probabilistic timed transition system (PTTS): the state transitions are probabilistic, and, in many cases, are timed, i.e., transitions after a given time period. An individual starts from the Susceptible state. The dashed lines represent state transitions triggered by either interactions with infectious individuals or vaccination. The solid lines represent probabilistic timed state transitions.

infections, prevalence in each age group, total hospitalizations, and deaths, as well as many other measures.

2.1 Simulation parameterization

The initial settings are calibrated to the conditions in Virginia as of January 1, 2021. Every simulation is run for 90 days, until March 31, 2021. Since the simulations are stochastic, each simulation is repeated for 30 replicates, and distributions of the measures are computed. The boxplots and curves in figures presented in Section 3 (Figure 2 through 8) are all based on data from 30 replicates. The curves show an uncertainty of one standard deviation above and below the mean.

Disease model. The disease model is the *best guess* version of “COVID-19 Pandemic Planning Scenarios” prepared by the US Centers for Disease Control and Prevention (CDC) SARS-CoV-2 Modeling Team [11]. It is an SEIR model where states and transitions are shown in Figure 1. Individuals of different age groups have different infectivity and susceptibility; dwell time distributions and state transition probability distributions are stratified by the following age groups: preschool (0–4 years), students (5–17), adults (18–49), older adults (50–64) and seniors (65+). Furthermore, individuals that are vaccinated have different disease parameter values than those that are not vaccinated. Detailed parameterization for unvaccinated individuals is summarized in [13].

Initializations. The simulations are initialized at the county level by age group using the detailed data of confirmed cases from [49]. The initialization specifies the health state of each individual. Based on county-level cumulative confirmed cases through December 19, 2020, we derive the number of prior infections in each county by scaling the cumulative number by a case ascertainment ratio of 3 (i.e., only one third of all infections are reported), then computing the number of prior infections in each age group of this county using the age distribution in cases. We randomly choose individuals in each age group in each county and set their health states to recovered to reflect that they have already been infected. Based on county-level daily confirmed cases from December 20, 2020 to January 5, 2021, we derive the number of individuals that are infected each day by the same scaling, and seed the simulation by

setting randomly chosen individuals to *exposed* by day in each age group of each county.

Non-pharmaceutical interventions. We consider four NPIs: (i) *Infectivity reduction (IR)*. Infectivity is universally reduced (by 60%) through preventive behavior, e.g., mask wearing and hand washing. (ii) *Generic social distancing (GSD)*. A fraction (25%) of the population chooses to reduce non-essential (shopping, religion, and other) activities. (iii) *Virtual learning (VL)*. A fraction (50%) of K-12 students choose virtual learning. (iv) *Voluntary home isolation of symptomatic cases (VHI)*. With probability 75%, a symptomatic person chooses to stay home for 14 days, reducing the weights on household contacts by 50%. For this person, all outside contacts are disabled and at-home contacts are reduced by 50% temporarily during these 14 days.

Scenarios based on relaxing social distancing measures. We assume that these NPIs are in place when a simulation starts, but adherence may change during the simulation. We consider three scenarios for adherence to the NPIs:

- *As-is*. NPI parameters remain the same for the duration of the simulation.
- *Slow relaxation*. NPI parameters change every 30 days from January 30, 2021, so that in 7 months, infectivity reduction decreases from 60% to 10%, generic social distancing decreases from 25% to 10%, and virtual learning decreases from 50% to 25%. Note that this is used to specify the speed of relaxation. Nevertheless the results are only reported for the period until end of March.
- *Fast relaxation*. NPI parameters change every 30 days from January 30, 2021, so that in 5 months, they reach the same levels as in the *slow relaxation* scenario.

2.2 Vaccination: supply, schedule and priority groups

Vaccine schedule. In the experiment, we consider a vaccination schedule of 25 million people being vaccinated per month in the US, starting from late December 2020. Assuming that vaccines are allocated to all states proportional to population size, Virginia has 650K people being vaccinated per month. We also consider a schedule where vaccination in Virginia occurs at half this rate. Therefore we consider three vaccination schedules: *none* (no vaccination), *fast* (vaccinating 650K people per month), and *slow* (vaccinating 325K people per month). The later schedule is intended to capture the current challenges faced in distributing the vaccines to individuals. For simplicity, all individuals vaccinated during each month are assumed to be vaccinated on the first day of that month; spreading the vaccines over the month does not change the overall results by much.

Vaccine efficacy. Overall vaccine efficacy is characterized by three numbers: (i) e_I , efficacy against infection; (ii) e_D , efficacy against severe illness (requiring hospitalization or leading to death) given infection; and (iii) e_T , efficacy against onward transmission given infection. We assume that $e_I = 90\%$ and $e_D = 50\%$ starting only 21 days after vaccination. In our sensitivity analyses, we also consider $e_I = 50\%$. In all cases, we ignore e_T .

Vaccination prioritization. The Pfizer-BioNTech vaccine and the Moderna vaccine were not approved for people younger than 16 years and 18 years, respectively, until May 2021. Therefore, we only allocate vaccines to people who are at least 18 years old. Among those people, we consider the following prioritization strategies.

- *No priority.* Everyone 18+ years old is vaccinated with the same probability. This is our baseline strategy.
- *Essential workers.* This strategy targets those who work for medical, care facilitation, retail, education, military, and government.
- *Older people.* This strategy prioritizes those who are at least 50 years old.
- *High degree.* Degree of an individual is the number of contacts per day. This strategy targets those in the top quartile among all 18+ years old in terms of degree.
- *Long total contact (also denoted as weighted degree).* Weighted degree of an individual is the total contact time this individual has with other people in a day. This strategy targets those in the top quartile among all 18+ years old in terms of weighted degree.

Most vaccines are allocated to the targeted groups, but we allow some to be given to other groups. This accounts for potential inaccuracy and precision in identifying and locating the targeted people. For example, since we do not know people’s daily number of contacts, which may vary, we can only estimate it using proxy attributes, such as age, household size and occupation, or from data collected through digital devices. We consider the following rates of enforcement: 100%, 80%, and 60%.

2.3 Experimental design

The design consists of 4 factors: (i) 3 adherence scenarios (as-is, slow relaxation, fast relaxation); (ii) 3 vaccination schedules (none, fast, slow); (iii) 5 prioritization targets (no priority, essential workers, older people, high degree, high weighted degree); and (iv) 3 levels of priority enforcement (100%, 80%, 60%). Combining (iii) and (iv) we have the baseline (no-priority) plus 12 prioritized strategies named according to the target group (essential, old-age, high degree, high weighted degree) and the fraction of vaccine given to the target group (100%, 80%, 60%), e.g., “essential 100%” or “high degree 60%”. We also consider vaccines with a 50% efficacy against infection (e_I) and compare the effectiveness of degree-based vaccinations under this assumption against that under 90% efficacy.

3 RESULTS AND ANALYSIS

In Figure 2, we show daily new infections under three scenarios (as-is, slow relaxation and fast relaxation) *without vaccination*. If NPI adherence can be maintained, then we expect infections to decrease after January. With slow relaxation, the infections will fluctuate around a level that will be a little higher than the January-end level. With fast relaxation, the infections show a steady increase in the next three months. The sharp increase every 30 days is caused by the implementation of the relaxation of NPIs and does not have any influence on the results presented. Next we show the effectiveness of vaccination strategies under various assumptions. Due to the page limit, many figures were omitted here but are available in Section 3 of Supplementary Information [13].

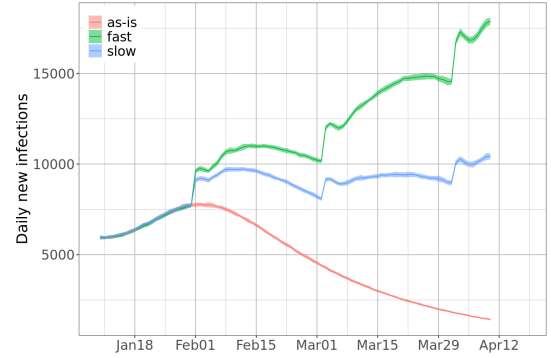


Figure 2: Daily number of new infections without vaccination under different NPI relaxation scenarios. The sudden surge in the beginning of each month is caused by NPI relaxation. Without NPI relaxation, the incidence peaks in early February, then starts decreasing. With slow relaxation, daily incidence increases then fluctuates. With fast relaxation, the incidence keeps rising.

3.1 Effectiveness of degree- and weighted degree-based strategies

Prioritizing vaccinations based on individual degree and weighted degree are extremely effective in controlling the pandemic. In particular, depending on the scenario, the reductions in the number of infections and hospitalizations by these schemes are over 50% more than the reductions from the age-based prioritization schemes. For example, assuming that the current non-pharmaceutical interventions remain at the same level over the next few months, our experiment shows that by the end of March 2021, degree-based schemes can result in 56–110k fewer infections, 3.2–5.4k fewer hospitalizations, and 700–900 fewer deaths in the state of Virginia, compared to age-based schemes. Note that the ranges come from different levels of priority enforcement (three levels for both age-based and degree-based schemes). Figure 3 shows the estimated reductions by one of the age-based schemes and the further reductions by one of the degree-based schemes. Extrapolating these results for the entire US, we estimate that degree-based schemes will lead to 3–6 million fewer infections, 181–306k fewer hospitalizations, and 51–62k fewer deaths by the end of March, compared to age-based schemes. If the NPIs are relaxed, the reductions in infections, hospitalizations, and mortality are even more substantial. This implies that when conditions worsen, the marginal gains from a more effective strategy are even higher.

Figure 4 compares incidence reduction up to March 31, 2021, under different prioritization strategies for the *fast* vaccine distribution schedule. We find that all strategies targeting either essential workers or high degree people outperform the no-priority distribution. The degree-based strategies reduce incidence more than any other strategy. For example, with no NPI relaxation (as-is), all degree-based strategies can reduce infections by over 20% while all other strategies can reduce infections by at most 20%. Strategies targeting older people perform worse than the no-priority distribution in terms of reducing incidence. Similar results are obtained for the *slow* vaccine distribution schedule, as shown in Figure 5. One reason to consider weighted degree-based heuristics is that they are

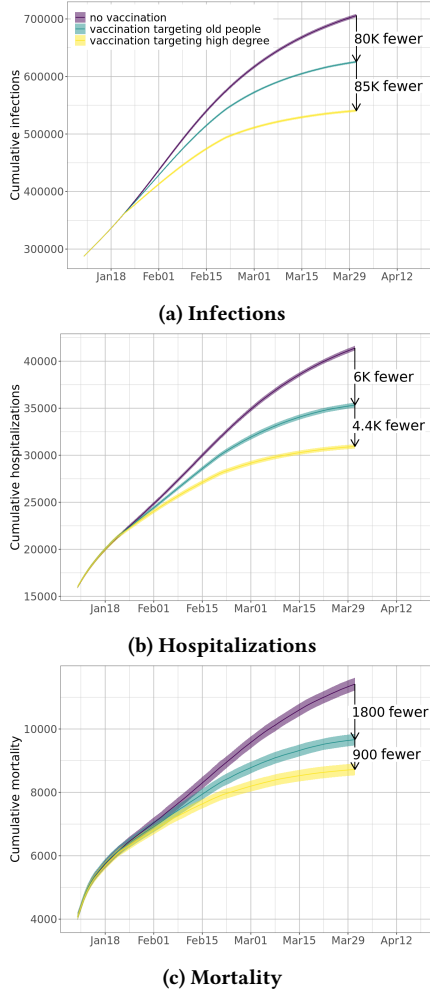


Figure 3: Vaccination targeting old people can reduce (a) total infections, (b) total hospitalizations, and (c) total mortality significantly, assuming current non-pharmaceutical interventions remain at the same level. Vaccination targeting high degree people can further reduce total infections, hospitalizations, and mortality. Numbers in the plots show total reductions up to the end of March 2021.

potentially easier to implement in the current digital apps; we will discuss this further in later sections. **All degree-based strategies outperform the other strategies.**

Targeting high degree people is also the most effective strategy for reducing mortality. Prioritization of older people is effective in reducing mortality compared to other strategies, but not when compared to a high degree strategy. See additional figures in [13]. Figure 6 shows that prioritizing people with high weighted degree (total contact durations) is even more effective than prioritizing those with high degree. For example, with no relaxation of NPIs, targeting people of high weighted degree can reduce infections by about 23-30%, compared to targeting high degree people, which can reduce infections by about 21-26%. In the case where NPIs are relaxed, the strategy prioritizing high weighted degree can cause

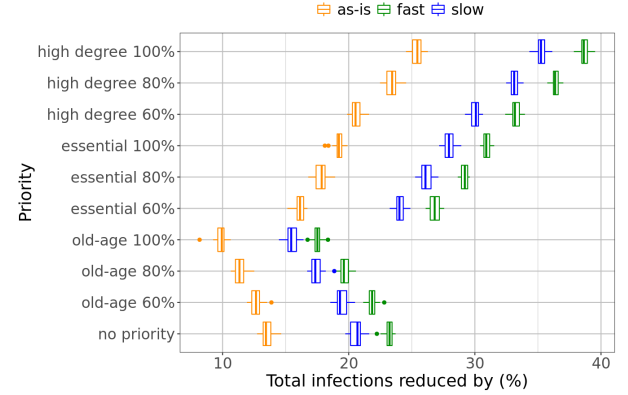


Figure 4: Total reduction in incidence under the fast vaccine distribution schedule. Degree-based strategies outperform all other ones, while age-based strategies are outperformed by all other ones.

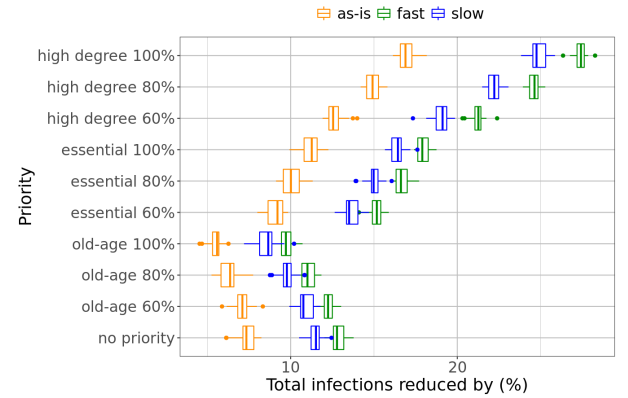


Figure 5: Total reduction in incidence under the slow vaccine distribution schedule. Degree-based strategies still reduce more infections than other strategies and are more effective if accuracy is higher.

over 40% reduction in infections if it can be implemented with high precision.

3.2 The high degree prioritization schemes are effective even when we cannot accurately estimate the degree of a node

Our results show that prioritization schemes based on degree and weighted degree (total contact time) work even when they are not accurately estimated. Specifically, even when we can only estimate the degree for 60% of the nodes (as being in the first quartile or not), we notice significant improvement in the overall control of the pandemic. This is highlighted in Figure 7, where we compare degree-based schemes of various accuracies with the age-based scheme and show improvement even at lower levels of accuracy. Targeting high degree people with only 60% accuracy improves the reduction from 10% by the age-based strategy to 20% (with no relaxation), from 15% to 30% (with slow relaxation), or from

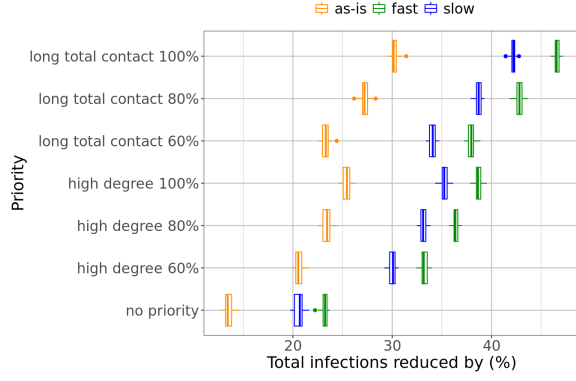


Figure 6: Comparison of degree and weighted degree-based strategies under fast vaccine distribution schedule. Both can reduce infections much more than the baseline strategy. The weighted degree-based strategy outperforms the degree-based one at any prioritization level.

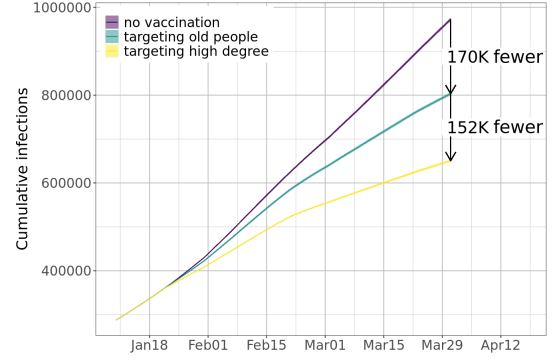


Figure 7: Even with lower (80% or 60%) accuracy in identifying and vaccinating high degree people, this strategy is still much more effective than the age-based strategy in reducing infections.

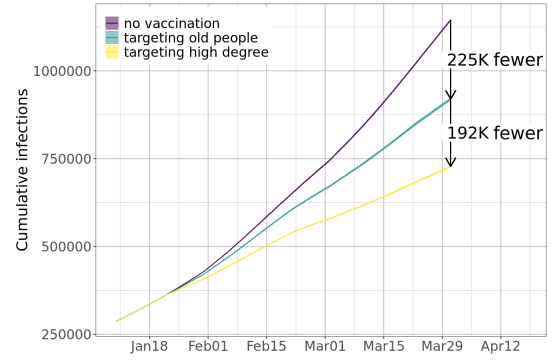
17.5% to 33% (with fast relaxation). In fact, these strategies require neither knowledge of the exact degree of each person, nor that of the complete ranking of people by degree. They only depend on knowing which nodes have high degrees (are in the top quartile); they are tolerant to a certain amount of inaccuracy.

3.3 Effectiveness when social distancing measures are relaxed

The effectiveness of degree-based strategies holds in three hypothetical scenarios for social distancing: one in which there is no relaxation, and the other two wherein social distancing is progressively relaxed 5 or 7 months from January 2021. Our results show that the value of these prioritization schemes is even higher when social distancing measures are relaxed quickly. Recall in Figure 3 we observe that, with no relaxation, the degree-based strategy results in another reduction of 85K infections and an additional reduction of 900 mortality, compared to the age-based strategy. In Figure 8, we find that with relaxation of NPIs, the degree-based strategy can



(a) Reduction in infections with slow relaxation of NPIs



(b) Reduction in infections with fast relaxation of NPIs

Figure 8: Reductions in infections from degree-based allocation strategies are even larger when NPIs are relaxed when compared to the age-based schemes. This figure shows reductions in infections with: (a) slow relaxation of NPIs, and (b) fast relaxation of NPIs.

reduce even more infections (152K with slow relaxation and 192K with fast relaxation) and more mortality (1.3K with slow relaxation and 1.5K with fast relaxation; mortality figures available in [13]). These observations highlight the importance of vaccination prioritization if the current NPIs are relaxed, which will likely happen as vaccines get distributed.

3.4 Effectiveness with low efficacy vaccines

We have assumed that vaccines have 90% efficacy regarding protection against infection (e_I). Our results also hold when the vaccine efficacy is lower than that of the Pfizer and Moderna vaccines. We study this for two reasons: (i) there was a discussion about giving just one dose of these vaccines which may result in lower efficacy (about 50%) or approving a low efficacy vaccine¹, and (ii) most other vaccines are traditional vaccines with a lower efficacy.

To this end we study the degree-based strategies assuming 50% vaccine efficacy. We find that while the reduction in infections decreases with low efficacy vaccines, it is still significant. For example, regardless of NPI relaxation, a degree-based strategy with even 60% accuracy can reduce infections by over 10% with fast

¹<https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-takes-action-help-facilitate-timely-development-safe-effective-covid>

vaccine distribution or over 5% with slow vaccine distribution, assuming $e_I = 50\%$. The reductions on hospitalizations and mortality are also significant. The figures are omitted due to the page limit, but interested readers can find more detailed results in Section 3.3 of [13].

4 DISCUSSION

These results are obtained using a realistic, data-driven and highly resolved agent-based model and individual-based social contact network of the Commonwealth of Virginia. The agent-based model represents individual-level activities that are spatially explicit. It allows us to: (i) capture details of within-host disease progression, as well as between-host transmission, including the impact of vaccines, (ii) model the complicated set of interventions that are currently in play, (iii) represent network-based vaccine prioritization schemes, (iv) represent the expected vaccine deployment schedule, including the expected mix of vaccine efficacy against infection, severe illness, and onward transmission estimates, (v) incorporate current surveillance data, and (vi) study counter-factual and hypothetical scenarios, such as a steady relaxation of social distancing measures. *This is the first study we know of that accounts for all of these components, not just for COVID-19, but for any infectious disease outbreak.*

The efficacy of the proposed policy is based on the assumption that the synthetic contact network is a realistic representative of the real-world social contact world. While the structural metrics may vary over time, we show the results are fairly robust to misidentification of high degree individuals. We believe both these assumptions hold and discuss this in more detail below. Further discussion on this topic can be found in the Supplementary Information [13], where we describe how our networks are synthesized, their structural properties, and the way the pandemic is simulated.

The potential efficacy of degree-based heuristics has been discussed in several earlier papers—this includes both provable analyses on different random graph models (under mean field assumptions in some cases), e.g., [3, 8, 39], and empirical analysis in various real world networks, e.g., [3, 15, 54]; a notion of weighted degree is also considered in [15]. However, it is important to note that these results are *not directly applicable* in our context for the following reasons. First, many of the theoretical results show the efficacy of these methods for simple power law-type models – the networks we generate are similar to power law networks, but with a very different exponent. Additionally, the network exhibits other features of social networks (local clustering, low diameter, and relatively high expansion). Second, many of the results are shown when vaccines are applied at the start of the epidemic process, and the results do not say anything of what happens when the vaccine is applied temporally – this is important, because the temporal epidemic process infects individuals, thereby changing the network structure substantially, including the application of NPIs.

Nevertheless, the intuition behind the efficacy of such methods is simply stated as follows: vaccinating high degree nodes not only protects them, but also confers a higher level of indirect protection on their neighbors as they interact with many individuals who might themselves be conferred similar protection. Our data-driven approach shows, in fact, that real-world social networks

have sufficient nodes of high degree to ensure that such heuristics are effective. Note that, by virtue of degree bias in social networks, even traditional approaches such as contact tracing will lead us to high degree individuals. The proposed approach makes identifying these individuals as a proactive, rather than reactive, step in infection control. It is important to note, however, that just the presence of high degree nodes does not guarantee that degree-based heuristics would work. See [13] for further discussion.

Identification of nodes with high degrees can be done in multiple ways, including using digital apps that have been deployed for contact tracing, interviewing individuals, and identifying typical job categories or other demographic attributes that entail higher social interactions. Further, even when other prioritization schemes are considered, one can use high social contact to further prioritize the distribution. For example, when distributing vaccines based on age, one can further subselect individuals with higher social contact in the case of limited supply.

Our results suggest that degree-based prioritization should be considered by larger and resource-poor countries to quickly bend the epidemic curve. The benefits of the proposed degree-based prioritization are so significant that even a partially successful campaign will likely have a large impact.

5 RELATED WORK

There has been a lot of work on analyzing interventions to control epidemics, and this falls into two broad categories. The first involves using a system of coupled differential equations to represent the dynamics, e.g., [34, 35, 45, 52, 53]. Even though closed-form solutions are not available even for simple models, when the system is not very large, it can be solved by brute-force local search methods, e.g., [35]. For some types of models, greedy strategies have been used [51, 53].

The second class of methods is network or agent based, of the form we study here, e.g., [16, 21, 23, 31, 32]. Analyzing interventions to minimize the expected outbreak size (or to optimize other epidemic outcomes) in network models is much harder. Prior work has generally attempted to solve these problems by either simplifying the network (e.g., assuming random graph models), or simplifying the disease model. The simplest setting is that of transmission probability of 1 (modeling a highly contagious disease), with a fixed source. Even this setting is challenging, and work by [17, 24] designs bicriteria approximation algorithms for this problem.

A variation of this setting is when the source is chosen randomly, and, in this case, the problem of minimizing the number of infections corresponds to deleting a subset of nodes such that the sum of squares of the component sizes in the residual network is minimized. A minor modification of the results of [6, 29] gives approximation algorithms for this objective. We note that [47] uses a stochastic optimization approach for minimizing the expected number of infections. While their worst case approximation factor can be quite large, their empirical performance is quite good. The work of [3, 8] on the robustness of networks can be viewed as interventions to reduce the spread of an outbreak.

It is well understood that the network structure has a significant impact on the dynamics of epidemic spread. This has motivated

a lot of research on modifying network properties to control epidemic spread. One of the most studied properties is degree, and in many network models, as well as in a broad class of real world networks, it has been found that removing the highest degree nodes (equivalently, vaccinating high degree nodes) turns out to be very effective [3, 8, 14, 15, 39, 54]. Cohen et al [14] show that a simple decentralized strategy of “acquaintance immunization” has the effect of selecting high degree nodes. Another set of properties that has been studied extensively are spectral properties, namely the eigenvalues and eigenvectors associated with the adjacency matrix of the graph and its Laplacian. It has been shown using multiple approaches [20, 36, 41] that epidemic spread exhibits a threshold behavior—if the spectral radius (the largest absolute value of an eigenvalue) is below a certain threshold, the disease dies out. This has motivated a considerable amount of work on reducing the spectral radius to control the outbreak [37, 42, 43, 46, 50]. Interventions based on other structural measures, such as betweenness centrality [26], coreness [27], and complex centrality [22], are also proven to be effective in epidemic containment. But these measures are difficult to estimate in real world contact networks. *In general, the theoretical studies do not apply to temporal vaccine allocation problems – in such cases the network is constantly changing as the epidemic spreads and vaccines are distributed in time.*

In the context of COVID-19, where we have multiple approved vaccine candidates, the role of vaccine efficacy, especially whether it reduces susceptibility to disease or transmission becomes important [30]. A study by Bubar et al. [9] identified that under different underlying assumptions, vaccine prioritization policies vary from 20-49 years to adults over 60 years old. They also note that prioritizing seronegative individuals could improve the marginal impact of a given policy. A similar study at a global scale using different supply assumptions was reported in [25]. See [1, 7, 19, 25, 33, 45, 48] for other recent papers on this topic. Multiple studies have also identified the tradeoffs based on the underlying policy objectives [10, 33] using compartmental models. In [44] a vaccination plan starting with superspreaders followed by descending age groups is shown to be more effective than an age-based plan, using an age-stratified compartmental model where a *superspreader* group is defined in addition to the age groups. The current allocation policy in the US at the federal level is centered around the framework developed by the National Academies of Sciences, Engineering, and Medicine (NASEM) [38].

Very few papers have studied vaccine allocation problems when there is a vaccine schedule (temporal vaccine allocation). Furthermore, they do not study how robust such methods are against uncertainty in estimating the structural properties, which is a crucial contribution of this paper. Nevertheless, these results do suggest the potential value of such methods.

Digital apps to estimate network properties. Digital contact tracing apps have been deployed in several countries [2, 5, 12, 18, 28] with mixed success. Challenges include low penetration levels, compliance, and accuracy of the apps in discovering neighbors accurately. Our allocation method is based on exploiting simple network properties that can be estimated using digital devices. Digital contact tracing apps can potentially measure both degree and weighted degree measures that we use here quite accurately.

6 CONCLUSIONS AND LIMITATIONS

We present an analysis of various vaccine prioritization strategies based on demographic attributes, occupation, and structural attributes of social contact networks. Our results show that vaccine prioritization schemes based on network degrees and total contact time can provide significant reductions in incidence, mortality, and hospitalizations. The results hold even for low efficacy vaccines and even when degrees and contact networks are estimated only approximately. Network-based prioritization is often more than twice as effective as other strategies. The results suggest that such methods should be considered when vaccines are available in limited supply; the benefits are likely to be greater in resource-poor and highly populated regions of the world. The advantage of our approach is in leveraging the mechanistic and network-based understanding of disease spread, and creating priority categories that cut across age, risk, and other demographic characteristics. Similarly, other high degree-based interventions, e.g., mask mandate on high degree people or closure of places where many people mix, may be effective too.

The study has a number of limitations: (i) our assumptions regarding the background interventions are our best estimates; (ii) in the case that a vaccinated node gets infected, we assume that they can transmit like any other node; (iii) our results depend on estimating the degrees and weighted degrees of nodes. While we have shown that the results are robust to mis-estimation, the overall efficacy of the scheme does depend on the ability to infer these degrees. Finally, the effectiveness results need to be combined with ethicalness while being considered for vaccine prioritization policies.

ACKNOWLEDGMENTS

The authors would like to thank members of the Biocomplexity COVID-19 Response Team and the Network Systems Science and Advanced Computing (NSSAC) Division for their thoughtful comments and suggestions related to epidemic modeling and response support. We thank members of the Biocomplexity Institute and Initiative, University of Virginia, Paul Eastham, Jeff Shaman, Chadi Saad-Roy, Theo Gibbs and Adam Sadilek for useful discussion and suggestions. This work was partially supported by NIH Grant 1R01GM109718, NSF BIG DATA Grant IIS-1633028, NSF Grant OAC-1916805, NSF Expeditions in Computing Grant CCF-1918749, CCF-1918656, CCF-1917819, NSF RAPID CNS-2028004, NSF RAPID OAC-2027541, US CDC 75D30119C05935, University of Virginia Strategic Investment Fund award number SIF160, Google Grant, and Defense Threat Reduction Agency (DTRA) under Contract No. HDTRA1-19-D-0007. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies.

REFERENCES

- [1] Babak Abbasi, Masih Fadaki, Olga Kokshagina, Naima Saeed, and Prem Chhetri. 2020. Modeling Vaccine Allocations in the COVID-19 Pandemic: A Case Study in Australia. Available at SSRN 3744520 (2020).
- [2] N. Ahmed, R. A. Michelin, W. Xue, S. Ruj, R. Malaney, S. S. Kanhere, et al. 2020. A Survey of COVID-19 Contact Tracing Apps. *IEEE Access* 8 (2020), 134577–134601.
- [3] R. Albert and A.-L. Barabasi. 1999. Emergence of scaling in random networks. *Science* 286 (1999), 509–512.

- [4] Roy M Anderson, Carolin Vegvari, James Truscott, and Benjamin S Collyer. 2020. Challenges in creating herd immunity to SARS-CoV-2 infection by mass vaccination. *The Lancet* 396, 10263 (2020), 1614–1616.
- [5] Andrew Anglemeyer, Theresa HM Moore, Lisa Parker, Timothy Chambers, Alice Grady, Kellia Chiu, et al. 2020. Digital contact tracing technologies in epidemics: a rapid review. *Cochrane Database of Systematic Reviews* 8 (2020).
- [6] J. Aspnes, K.L. Chang, and A. Yampolskiy. 2006. Inoculation strategies for victims of viruses and the sum-of-squares partition problem. *J. Comput. Syst. Sci.* (2006).
- [7] Dimitris Bertsimas, Joshua Kiefer Ivanhoe, Alexandre Jacquillat, Michael Lingzhi Li, Alessandro Previero, Omar Skali Lami, and Hamza Tazi Bouardi. 2020. Optimizing Vaccine Allocation to Combat the COVID-19 Pandemic. *medRxiv* (2020).
- [8] B. Bollobás and O. Riordan. 2004. Robustness and vulnerability of scale-free random graphs. *Internet Mathematics* (2004).
- [9] Kate M Bubar, Kyle Reinholt, Stephen M Kissler, Marc Lipsitch, Sarah Cobey, Yonatan H Grad, and Daniel B Larremore. 2021. Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science* (2021).
- [10] Jack H Buckner, Gerardo Chowell, and Michael R Springborn. 2020. Optimal dynamic prioritization of scarce COVID-19 vaccines. *medRxiv* (2020).
- [11] Centers for Disease Control and Prevention. 2020. COVID-19 Pandemic Planning Scenarios. <https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios-h.pdf>. [Online, accessed January 15, 2021].
- [12] Justin Chan, Shyam Gollakota, Eric Horvitz, Joseph Jaeger, Sham Kakade, Tadayoshi Kohno, et al. 2020. PACT: Privacy Sensitive Protocols and Mechanisms for Mobile Contact Tracing. *arXiv preprint arXiv:2004.03544* (2020).
- [13] Jiangzhuo Chen, Stefan Hoops, Achla Marathe, Henning Mortveit, Bryan Lewis, Srinivasan Venkatramanan, Arash Haddadan, Parantapa Bhattacharya, Abhijin Adiga, Anil Vullikanti, Aravind Srinivasan, Mandy L Wilson, Gal Ehrlich, Maier Fenster, Stephen Eubank, Christopher Barrett, and Madhav Marathe. 2022. Effective Social Network-Based Allocation of COVID-19 Vaccines - Supplementary Information. http://nssac.bii.virginia.edu/~chenj/paper/vax_allocation-si.pdf.
- [14] Reuven Cohen, Shlomo Havlin, and Daniel Ben-Avraham. 2003. Efficient immunization strategies for computer networks and populations. *Physical review letters* 91, 24 (2003), 247901.
- [15] Ken T.D. Eames, Jonathan M. Read, and W. John Edmunds. 2009. Epidemic prediction and control in weighted networks. *Epidemics* 1, 1 (2009), 70 – 76.
- [16] S. Eubank, H. Guclu, V. S. Anil Kumar, M. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* 429 (2004), 180–184. Issue 6988.
- [17] S. Eubank, V. S. Anil Kumar, M. V. Marathe, A. Srinivasan, and N. Wang. 2006. Structure of Social Contact Networks and Their Impact on Epidemics. In *Discrete Methods in Epidemiology*. Vol. 70. American Math. Soc., Providence, RI, 179–200.
- [18] Luca Ferretti, Chris Wymant, Michelle Kendall, Lele Zhao, Anel Nurtay, Lucie Abeler-Dörner, Michael Parker, David Bonsall, and Christophe Fraser. 2020. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* (2020).
- [19] Brody H Foy, Brian Wahl, Kayur Mehta, Anita Shet, Gautam I Menon, and Carl Britto. 2020. Comparing COVID-19 vaccine allocation strategies in India: a mathematical modelling study. *International Journal of Infectious Diseases* (2020).
- [20] Ayalvadi Ganesh, Laurent Massoulié, and Don Towsley. 2005. The effect of network topology on the spread of epidemics. In *Proceedings of 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, Vol. 2. 1455–1466.
- [21] T.C. Germann, K. Kadau, I.M. Longini Jr, and C.A. Macken. 2006. Mitigation strategies for pandemic influenza in the United States. *Proceedings of the National Academy of Sciences* 103, 15 (2006), 5935–5940.
- [22] Douglas Guilbeault and Damon Centola. 2021. Topological measures for identifying and predicting the spread of complex contagions. *Nature communications* 12, 1 (2021), 1–9.
- [23] M. Elizabeth Halloran, Neil M. Ferguson, Stephen Eubank, Ira M. Longini, Derek A. T. Cummings, Bryan Lewis, et al. 2008. Modeling targeted layered containment of an influenza pandemic in the United States. In *Proceedings of the National Academy of Sciences (PNAS)*. 4639–4644.
- [24] A. Hayrapetyan, D. Kempe, M. Pál, and Z. Svitkina. 2005. Unbalanced Graph Cuts. In *ESA*. 191–202.
- [25] A Hogan, P Winskill, O Watson, P Walker, C Whittaker, M Baguelin, D Haw, A Lochen, K Gaythorpe, K Ainslie, et al. 2020. Report 33: Modelling the allocation and impact of a COVID-19 vaccine. (2020).
- [26] Petter Holme, Beom Jun Kim, Chang No Yoon, and Seung Kee Han. 2002. Attack vulnerability of complex networks. *Physical review E* 65, 5 (2002), 056109.
- [27] Maksim Kitsak, Lazaros K Gallos, Shlomo Havlin, Fredrik Liljeros, Lev Muchnik, H Eugene Stanley, and Hernán A Makse. 2010. Identification of influential spreaders in complex networks. *Nature physics* 6, 11 (2010), 888–893.
- [28] Mirjam E Kretzschmar, Ganna Rozhnova, Martin CJ Bootsma, Michiel van Boven, Janneke HHM van de Wijger, and Marc JM Bonten. 2020. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. *The Lancet Public Health* 5, 8 (2020), e452 – e459.
- [29] V S Anil Kumar, Rajmohan Rajaraman, Zhifeng Sun, and Ravi Sundaram. 2010. Existence theorems and approximation algorithms for generalized network security games. In *Distributed Computing Systems (ICDCS), 2010 IEEE 30th International Conference on*. IEEE, 348–357.
- [30] Marc Lipsitch and Natalie E Dean. 2020. Understanding COVID-19 vaccine efficacy. *Science* 370, 6518 (2020), 763–765.
- [31] Eric Lofgren, M. Elizabeth Halloran, C. M. Rivers, J. M. Drake, Travis C. Porco, et al. 2014. Opinion: Mathematical models: A key tool for outbreak response. *PNAS* (2014), 18095–18096.
- [32] M. Marathe and A. Vullikanti. 2013. Computational Epidemiology. *Commun. ACM* 56, 7 (2013), 88–96.
- [33] Laura Matrajt, Julia Eaton, Tiffany Leung, and Elizabeth R Brown. 2020. Vaccine optimization for COVID-19: who to vaccinate first? *medRxiv* (2020).
- [34] Robert M May and Roy M Anderson. 1984. Spatial heterogeneity and the design of immunization programs. *Mathematical Biosciences* 72, 1 (1984), 83–111.
- [35] J. Medlock and A. P. Galvani. 2009. Optimizing influenza vaccine distribution. *Science* 325, 5948 (2009), 1705–1708.
- [36] P. Van Mieghem, J. S. Omic, and R. E. Kooij. 2009. Virus Spread in Networks. *IEEE/ACM Transaction on Networking* (2009).
- [37] P. Van Mieghem, D. Stevanovic, F. Fernando Kuipers, Cong Li, Ruud van de Bovenkamp, Daijie Liu, and Huijuan Wang. 2011. Decreasing the spectral radius of a graph by link removals. *IEEE Transactions on Networking* (2011).
- [38] Medicine National Academies of Sciences, Engineering et al. 2020. Framework for equitable allocation of COVID-19 vaccine. (2020).
- [39] R. Pastor-Satorras and A. Vespignani. 2002. Immunization of complex networks. *Physical Review E* 65, 036104 (2002).
- [40] Govind Persad, Ezekiel J Emanuel, Samantha Sengenito, Aaron Glickman, Steven Phillips, and Emily A Largent. 2021. Public perspectives on COVID-19 vaccine prioritization. *JAMA Network Open* 4, 4 (2021), e217943–e217943.
- [41] 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.
- [42] Victor M. Preciado, Michael Zargham, Chinwendu Enyioha, Ali Jadbabaie, and George J. Pappas. 2013. Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks. In *IEEE Conference on Decision and Control*. IEEE.
- [43] Victor M Preciado, Michael Zargham, Chinwendu Enyioha, Ali Jadbabaie, and George J Pappas. 2014. Optimal resource allocation for network protection against spreading processes. *IEEE Transactions on Control of Network Systems* 1, 1 (2014), 99–108.
- [44] Tarcísio M Rocha Filho, José FF Mendes, Thiago B Murari, Aloísio S Nascimento Filho, Antônio JA Cordeiro, Walter M Ramalho, et al. 2022. Optimization of COVID-19 vaccination and the role of individuals with a high number of contacts: A model based approach. *PLoS ONE* 17, 3 (2022), e0262433.
- [45] Chadi M Saad-Roy, Caroline E Wagner, Rachel E Baker, Sinead E Morris, Jeremy Farrar, Andrea L Graham, et al. 2020. Immune life history, vaccination, and the dynamics of SARS-CoV-2 over the next 5 years. *Science* 370, 6518 (2020), 811–818.
- [46] Sudip Saha, Abhijin Adiga, B. Aditya Prakash, and Anil Kumar S. Vullikanti. 2015. Approximation Algorithms for Reducing the Spectral Radius to Control Epidemic Spread. In *Siam Data Mining (SDM)*.
- [47] 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*. 1187–1195.
- [48] Guy Tennenholtz, Constantine Caramanis, and Shie Mannor. 2020. Sequential Vaccination for Containing Epidemics. *medRxiv* (2020).
- [49] The New York Times. 2020. Coronavirus (Covid-19) Data in the United States. <https://github.com/nytimes/covid-19-data>, last accessed on January 7, 2021.
- [50] H. Tong, B. Aditya Prakash, T. Eliassi-Rad, M. Faloutsos, and C. Faloutsos. 2012. Gelling, and Melting, Large Graphs by Edge Manipulation. In *Proc. of CIKM*.
- [51] Srinivasan Venkatramanan, Jiangzhuo Chen, Arindam Fadikar, Sandeep Gupta, Dave Higdon, Bryan Lewis, et al. 2019. Optimizing spatial allocation of seasonal influenza vaccine under temporal constraints. *PLoS computational biology* 15, 9 (2019), e1007111.
- [52] S. Venkatramanan, J. Chen, S. Gupta, B. Lewis, M. Marathe, H. Mortveit, and A. Vullikanti. 2017. Spatio-Temporal Optimization of Seasonal Vaccination Using a Metapopulation Model of Influenza. In *2017 IEEE International Conference on Healthcare Informatics (ICHI)*. 134–143.
- [53] Bryan Wilder, Sze-Chuan Suen, and Milind Tambe. 2018. Preventing Infectious Disease in Dynamic Populations Under Uncertainty. In *AAAI Conference on Artificial Intelligence*.
- [54] Yingrui Yang, Ashley McKhann, Sixing Chen, Guy Harling, and Jukka-Pekka Onnela. 2019. Efficient vaccination strategies for epidemic control using network information. *Epidemics* 27 (2019), 115 – 122.