Impact of Preventive Responses to Epidemics in Rural Regions

P. Schumm[1], W. Schumm[2], and C. Scoglio[1]

[1] K-State Epicenter, Electrical and Computer Engineering Department
[2] Family Studies and Human Services School
Kansas State University, Manhattan, Kansas, USA

Abstract
Various epidemics have arisen in rural locations through human-animal interaction, such as the H1N1 outbreak of 2009. Through collaboration with local government officials, we have surveyed a rural county and its communities and collected a dataset characterizing the rural population. From the respondents' answers, we build a social (face-to-face) contact network. With this network, we explore the potential spread of epidemics through a Susceptible-Latent-Infected-Recovered (SLIR) disease model. We simulate an exact model of a stochastic SLIR Poisson process with disease parameters representing a typical influenza-like illness. We test vaccine distribution strategies under limited resources. We examine global and location-based distribution strategies, as a way to reach critical individuals in the rural setting. To explore an array of potential diseases, we vary the infection rate across the spectrum of outbreaks and quantify the network susceptibility through the whole spectrum. The extent to which social dynamics can control the spreading process is studied. We explore two models of a susceptible individual's dynamics in response to infections observed among the individuals in his neighborhood. Through extensive simulations, our investigation reveals the potentially powerful impacts of social spontaneous responses in rural settings. We compare the strategies over the spectrum and demonstrate that vaccination strategies are the most effective interventions for extremely contagious diseases.

Keywords: human epidemics – data driven modeling – contact network – disease prevention and control – vaccination – social behavior –survey

1. Introduction

In general, the spread of infectious diseases can be contained by human response using different approaches. Susceptible people can acquire immunization through vaccination, or can protect themselves from the diseases using preventive behaviors, such as avoiding close physical contacts with infected individuals or using hygienic habits. Correspondingly, human responses can be modeled using three classes of models distinguished by changes taking place in compartments, parameters, or contact levels to take into account the behavioral changes [1].

A vast literature exists on efficient vaccination strategies, given the need for efficient strategies to distribute vaccines that can often be insufficient for the entire population. Some of these strategies assume that human contact networks are well represented by scale free networks. One popular strategy aims at immunizing those individuals having the highest number of contacts, as the most critical actors for spreading the infection [2]. However, local strategies are more efficient and implementable and often require a lower fraction of the population to be vaccinated than random global immunization to contain epidemics. The strategy of acquaintance

immunization proposes the immunization of random acquaintances of random individuals [3]. Another local strategy proposes to vaccinate highly connected acquaintances of randomly selected people; based on the properties of scale free networks, with this approach the probability of targeting the highly connected individuals in the contact network increases with respect to the simple random selection [4]. In the case of a limited amount of available vaccines, the authors of [5] use stochastic simulations of epidemic and numerical optimization methods to find near-optimal vaccine distributions to minimize the epidemic size. Again in the case of a limited amount of available vaccines, the best strategy suggests to vaccinate schoolchildren, the population group with highest contact in different communities, and the high-risk groups, the population groups that need protection [6]. Since a strong community structure can be detected in social contact networks, the approach in [7] aims at immunizing individuals bridging communities rather than simply targeting highly connected individuals. An extensive set of simulations performed in [8] suggests two strategies based on age classes: In the first strategy vaccinating older children, adolescents, and young adults minimizes the number of infections, while in the second strategy vaccinating either younger children and older adults or young adults minimizes the number of deaths. Using game theory, the authors of [9] show that when vaccination is an individual's choice, a periodic behavior can be seen in simulations. A severe epidemic in one year incentivizes high vaccination rates in the following year, causing a milder epidemic for which individuals have less motivation for vaccination in the subsequent year. In [10], authors develop a vaccination strategy based on optimizing the susceptible size by a partitioning of the contact network through vaccination. Based on the authors' simulations, this strategy is more efficient than those based on vaccinating the highest betweenness or contact individuals. Using a decision-making framework for vaccine distribution policies based on a geographical and demographical data in USA, the authors of [11] find that distributing vaccines first to counties where the latest epidemic waves are expected is the most efficient policy.

When vaccines are not available, it is important to resort to spontaneous changes in individuals' behavior toward contagion-preventive habits. Modeling spontaneous human reactions to the spread of infectious disease is an extremely important topic in current epidemiology [12], and has recently attracted substantial attention. A review of the existing results on the interaction of the epidemic spreading and the human behavior can be found in [1]. In particular, Poletti et. al. developed a population-based model where susceptible individuals could choose between two behaviors in response to presence of infection [13]. Funk et. al. showed that awareness of individuals on the presence of a disease has interconnected dynamics with the disease itself, and awareness can help reducing the epidemic size [14]. Perra et. al. [15] considered the case where individuals go to a "feared" state when they sense infection. Since most of the existing results are for population-based models, they are suitable for a society of well-mixed individuals. Concerning individual-based models, a new model in which an "alert" state is considered has been proposed in [16].

In any case, assessing the effectiveness of mitigation strategies and behavioral responses both from a public health point of view and from individuals' perspectives is a complex and not fully-explored problem. In particular, a thorough evaluation and comparison of feasible mitigation strategies in the specific setting of rural regions is missing. In other words, not only the amount of success a given strategy can provide is not determined, but also its related cost in economical and social terms is unknown.

In this paper, we carry out extensive simulations on a weighted contact network determined by collected data in the City of Chanute and Neosho County in the State of Kansas. In particular we study the impact of limited resource vaccination campaigns, using an exact model of a stochastic SLIR Poisson process. Simulations are run across several scenarios and with stochastic sets of the SLIR model parameters. The evaluation of the vaccination campaigns is performed computing the average number of cases prevented per a single vaccine and the sizes and durations of the outbreaks. In addition we study the impact of contact reductions and social alertness as mitigation strategies over a spectrum of disease strengths.

Our contributions are twofold: we construct and analyze a data-based rural contact network and we provide a thorough analysis and comparison of mitigation strategies in a rural region. Additionally, in the supplementary information document (SI), we derive results on which mitigation strategy has greater potential to be efficient as a function of the epidemic strength [16, 17]. We hope that our results can provide practical guidelines for health officials to contain and suppress epidemics in rural regions.

In the following we describe data collection and analysis in section 2. Section 3 includes the models descriptions for the network, for the epidemic spreading, and also for mitigation strategies and vaccine distributions. In section 4, results are presented and discussed. Finally section 5 concludes our article.

2. Data Collection and Analysis

As of the 2010 U.S. Census, Neosho County was a rural county with 16,512 residents in 571.5 square miles in southeastern Kansas. Most of the population was White (94.1%); a majority were female (50.6%) and many (17.4%) were 65 years of age or older. The median household income was $36,702 with 17.0% living below the poverty level. Between July and October 2010, the towns of Chanute, Thayer, and Galesburg were selected to participate in a survey concerning factors that would predict the spread of epidemics in rural areas. From county public household rosters, households were randomly selected from Chanute (10%, N = 171), Thayer (50%, N = 158), and Galesburg (50%, N = 73) for a total initial $N = 402$. However, there was one duplicate address in both Thayer and Galesburg. Two household heads in Thayer and one in Galesburg had died. One household in Thayer, three in Galesburg, and one in Chanute had moved out of Neosho County. There were 11, 3, and 8 incorrect addresses in Thayer, Galesburg, and Chanute, respectively, based on surveys returned by the post office. Thus, the final number of available and eligible households were 143, 65, and 162 in Thayer, Galesburg, and Chanute, respectively, with total $N = 370$.

The tailored design method was used, with minor modifications, to improve response rates [18-21] with a focus on personalization and multiple follow-up mailings. Households were initially sent a postcard a week in advance of the survey to inform them that a survey would be arriving in a week. The survey was mailed with a cover letter concerning the purpose of the research and informed consent information, as well as a return envelope for the survey and a postcard to be returned separately to allow the household to refuse the survey, to indicate that they had done it, or to ask for results of the research. While the survey was anonymous, the postcard included

household identification. Once that postcard was received, no further follow-ups were sent to the household. The initial survey also included a local news report announcing the impending start of the survey [22]. A week after the survey was mailed, a follow-up postcard was sent thanking the householders for their response if they had returned the survey and reminding them about its importance otherwise. After about three weeks from the first survey, a second survey was mailed, unless a postcard had been received. Again, a follow-up postcard was sent after a week. Finally, a third survey was mailed. All cover letters and follow-up postcards were signed by the senior author in blue ink in order to personalize the survey implementation process. No financial incentives were provided in any of the mailings.

The first survey was received on July 12, 2010 and the last on October 18, 2010. Overall, 242 surveys were for an overall response rate of 65.4%. Over half of the 242 surveys were returned by July 21st (51.2%). By August 18th, 76.4% of the surveys had been returned. Over 90% of the surveys had been returned by September 24th. Only 13 surveys were received in October compared to 27 in September, 51 in August, and 151 in July. The response rate for Chanute was 74.7% (121/162). The response rate for Thayer and Galesburg combined was 55.8% (116/208). Five surveys did not differentiate between Thayer and Galesburg; otherwise, there were 33 (50.8%) from Galesburg and 83 (58.0%) from Thayer. Refusals occurred for 8, 6, and 12 households, respectively, from Thayer, Galesburg, and Chanute. From the postcards that were returned, 54, 24, and 78 households from Thayer, Galesburg, and Chanute, respectively, indicated that they had completed and returned the surveys, representing 64.5% of all surveys returned. The difference in response rates (74.7% vs. 55.8%) between Chanute and Thayer/Galesburg was significant statistically, two-sided Fisher's Exact Test ($p < .001$), odds ratio = 2.34 (95% CI, 1.50 – 3.66, $p < .001$). The difference in response rate for the more urban location was probably related to the content of the survey, which focused on respondents' visiting locations of stores, public sites, and restaurants in Chanute itself. Thus, the survey probably seemed more relevant to Chanute residents, even though we were interested in how often households from outlying towns went to the nearest urban center to visit or shop.

Data Analysis. A majority (56%) of the respondents reported being from Chanute compared to 23% from Thayer and 10% from Galesburg (the remaining percentage did not specify exactly where they were from). Of the 357 participants, the largest number were ages 45 to 64 (47.1%), with 26.1% 65 years of age or older and 18.8% (26-44) and 8.1% (18 to 25) younger than 45. A majority of the participants were females (57.6%). Most of the respondents (75.4%) had lived in their local community for 15 years or more. The vast majority (97.5%) of the respondents lived in a single family home. Very few (6.2%) of the households included a homebound member. Most of the respondents had either the equivalent of a high school degree (22%) or a college (23%) or graduate (12%) degree. Nearly sixty percent had incomes between $25,000 and 100,000 a year with 11% earning more and 30% earning less. Some respondents had type I (1.2%) or type II diabetes (10.4%) or were pre-diabetic (3.2%). Most respondents considered themselves to be slightly (35.7%), somewhat (18.2%), or extremely (8.6%) overweight. Most (56.6%) reported that they ate out one or two times a week with 26% eating out more often and 17% not at all.

In terms of compliance risk, nearly 49% of respondents said they would still visit at least one or two households outside of their home if there was a serious epidemic and radio/TV/internet had

told them to remain at home and not visit with others. Figure 1 presents the distribution of the number of individuals that a respondent expects to still visit against advice. Only half (50.0%) of the respondents had been vaccinated against the flu within the past six months. Nearly 40 percent (38.9%) did not obtain such a vaccination because of concerns about the vaccine's safety or effectiveness. Only about 7% believed they had come down with the flu within the past six months while about 18% thought they might have come down with a cold. About 18% of the respondents reported taking vitamin D supplements; only 6% reported taking zinc supplements. Approximately 80% of the respondents had extensive contact with domestic pets on a daily basis while about 19% of respondents had contact with farm animals or wild animals regularly, shown in figure 2. Contact risk (low, moderate, high) was significantly related statistically with compliance risk (none, low, high) ($p < .001$, ES = 0.50, medium effect size). As contact risk increased from low to high, high compliance risk increased from 4.4% to 21.8%; as contact risk decreased from high to low, the percentage of respondents with no compliance risk increased from 38.6% to 62.3%.

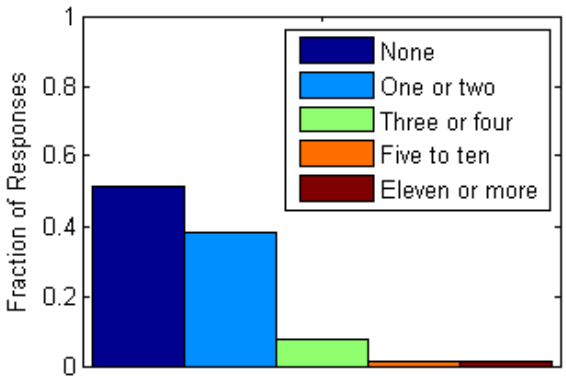


Figure 1. The distribution of the number of households that a respondent expects to still visit in a week against advice during a serious epidemic is shown.

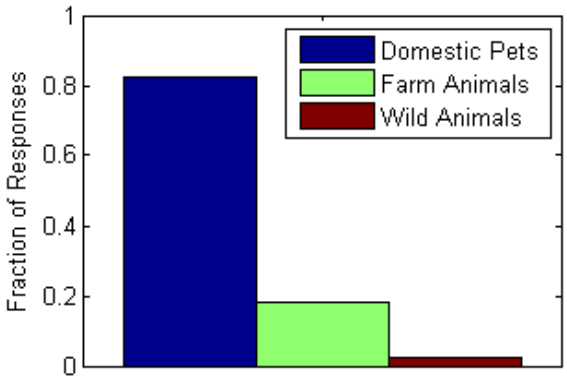


Figure 2. The distribution of types of animals a respondent interacts with in a typical day is shown. Note that the total does not sum to one as respondents can interact with multiple types of animals.

3. Models

In this section the procedure to construct the contact network from survey data is explained. Furthermore, the compartmental model used for simulations, the preemptive vaccination strategies, and the human response models are described. Finally, a cost function, to compare the effectiveness of the different containment strategies is introduced.

Network. From the survey responses, we constructed a rural contact network as an estimation of the social contact structure among the survey respondents. The network is based on two central questions: the number of contacts that a person has, and the locations that a person visits at different times in a typical day. The basis for the interactions between a pair of respondents is the locations that they both visited in common. We considered 4 types of location-based interactions: both visit the same location in the morning, both visit the same location in the afternoon, both visit the same location in the evening, and both visit the same location regardless of time. The fourth category introduces some overlapping in the interactions, but it is added to account for some of the uncertainty in potential pathways of the disease spread. We considered 66 locations in the network construction and therefore 264=66×4 possible interactions between each pair of survey respondents. We compute normalized weights from each respondent i to each other respondent j given by l_{ij}, representing the number of location-based interactions between respondents i and j. For the few respondents who did not complete the section of the survey regarding location visits, we assign them uniform weights of interacting with every other respondent in the network. Letting nodes represent the set of N = 353 respondents and weighted links represent the contact between them; we have a symmetric contact network at this point. Next we uniformly scale the weights on the links directed outward from each respondent i such that the sum of these weights is equal to the number of contacts that respondent i has indicated having with his or her response ($w_{ij} = a_i * l_{ij}$ for every j in 1, 2, … N). (This scaling makes irrelevant the absolute value of the uniform weight of the respondents who lack location data.) This produced a weighted, directed (asymmetrical) contact network of 353 nodes, with each pair of nodes (i and j) connected by two links which are respectively characterized by the weights w_{ij} and w_{ji}.

Six of the vaccination strategies will center on three node metrics: incoming node strength (the sum of the weights incoming to a node), outgoing node strength (the sum of the weights outgoing from a node), and node betweenness (a count of the shortest paths among all pairs that utilize the node) [23, 24]. The incoming node strength of a node is a topology metric that captures the direct impact of the network on the node. The outgoing node strength captures the direct impact that a node can have on the network. The betweenness of a node is a measure which captures the significance of a node in traversing the network. A node with a higher betweenness would be more likely to be traversed (in a shortest-path-type travel across the network between any pair of nodes) than a node with lower betweenness. Although an epidemic is not restricted to following the shortest paths across a network, the betweenness metric still plays an important role in identifying nodes which are likely to catch the disease if it reaches a majority of the nodes in the network. The rural contact network is depicted in figure 3, where the nodes representing individuals are shown in purple in a cloud and they are connected to the locations that they frequent, shown as orange nodes on the map [25].

Figure 3. A depiction of the rural contact network developed from a survey of Neosho County is shown, where the individuals are represented by purple nodes in a "cloud," which is connected by the respondents local travel habits to the set of rural locations shown in orange on the map.

SLIR model. On this weighted network, we model an epidemic outbreak using a Susceptible-Latent-Infected-Recovered compartmental model (SLIR) [23, 26]. In the SLIR model, we assume infections arrive at a susceptible (S) node *j* from an infected (I) node *i* with a rate that is a product of the directed contact w_{ij} and the basic infection rate β. When an infection arrives to a susceptible node, the node takes on a latent infection (transfers from the susceptible compartment to the latent compartment). A node, once latent (L), is considered unable to spread the disease, but is developing to that stage with rate λ. The inverse of the rate λ is the expected time for a node to spend in the latent state. The next stage of the disease, the infected/infectious state, enables the node to spread infections to each of its neighbors at rates proportional to the weights on its outgoing links. Each infected node recovers from the infected state at a rate μ. Once a node is in the recovered state (R), it remains recovered and does not participate in the disease process any further.

We simulate this model exactly using an event-driven simulation of the SLIR process on the weighted rural contact network. We initialize the simulation by assigning a disease state to each node and then drawing exponential waiting times for the next event at each node. Taking the event with the minimum time across all nodes, we advance the event node to its next disease state and re-draw waiting times for all nodes. This step is repeated until all waiting times are infinite, which happens when the disease process is complete. At this point, all nodes will be either susceptible or recovered. In the event-driven simulation, the time periods between successive events will not be regular, but instead they are non-integer stochastic values. This time variability can be observed in figure 4.

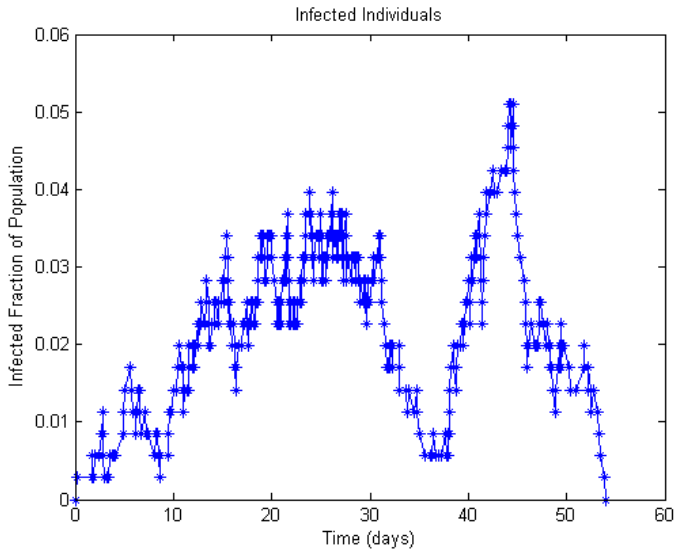


Figure 4. Under no mitigation, with a basic infection rate of $\beta = 0.042$ (days)$^{-1}$, $\varepsilon = (1.1 \text{ days})^{-1}$, and recovery rate of $\mu = (2.5 \text{ days})^{-1}$, a sample outbreak is shown by depicting the total number of infected individuals as a fraction of the population through time.

Social Alertness Responses. Given the estimated survey respondents contact network and the epidemic model, we explore two potential classes of mitigation strategies. In the first class, social responses, we attempt to capture potential human responses to an epidemic through two models, the first being a social alertness. As an extension of the model presented in [16], in this model we take into account the behavior of susceptible individuals. Specifically, as susceptible individuals recognize that an infection exists, they decide to adopt a cautious behavior. We have modeled the cautious behavior as characterized by a lower infection rate. In order to model the behavior of susceptible individuals, we add a new compartment to the considered SLIR model for epidemic spread, to define a new Susceptible-Alert-Latent-Infected-Recovered (SALIR) model. The state transition diagram for the SALIR model is shown in figure 5. Both susceptible and alert individuals can potentially become exposed. However, the infection rate of the alert individuals is lower. Specifically, a susceptible node becomes alert with the alerting rate κ times the number of infected neighbors. An alert individual can get infected in a process similar to a susceptible individual but with a smaller infection rate $0 \leq \beta_a < \beta$. We assume that transition from an alert individual to a susceptible state is much slower than other transitions. In this model, the transition to the alert state is driven by the number of infected neighbors and produces a reduced incoming infection rate.

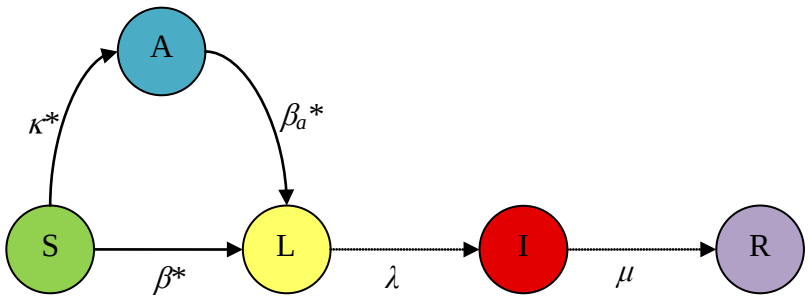
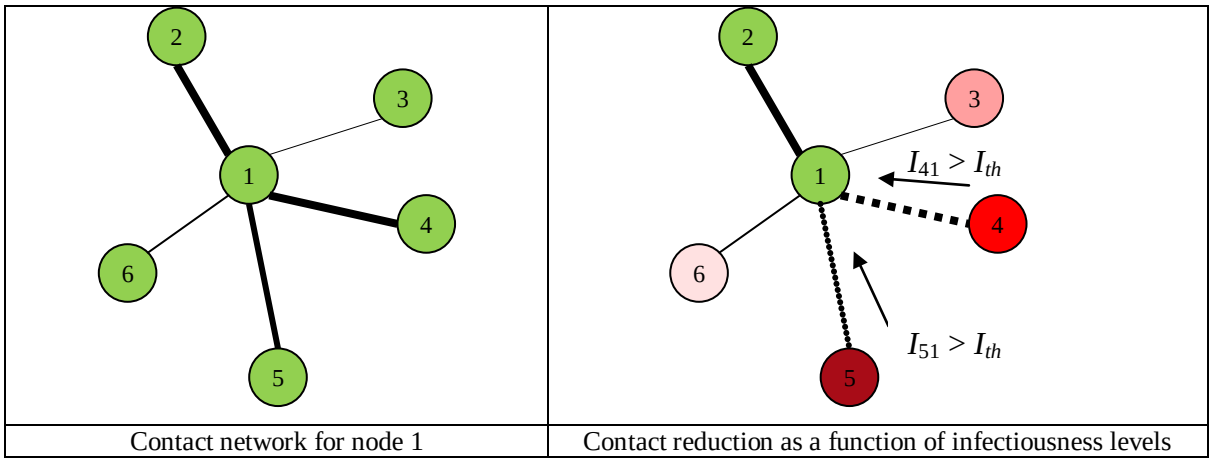


Figure 5. Transition diagram for the Susceptible-Alert-Latent-Infected-Recovered (SALIR) epidemic model with social alertness represented by compartment A. An individual in the rural contact network will transition between

the disease states being driven by the rates along the directions of the arrows. The Exposed-Infected transition and the Infected-Recovered transition are simple rate based processes. We denote the non-basic transition rates with asterisks. The Alerting process (S-A) occurs at a rate κ times the number of infected neighbors. The infection process for the Susceptible (Alert) nodes occurs as a Susceptible (Alert) node receives infections from each infected neighbor at a rate β (β_a). The spontaneous transitions are shown with dashed arrows while the solid arrows depict the interaction based trasitions.

Social Distancing Responses. Extending the approach presented in [17], we consider a scenario where individuals do not change their contact set or infection parameters, but instead reduce their levels of contact. As a social behavioral constraint, we want to guarantee a minimal level of contact all the time. In [17], an optimal control problem is formulated to minimize the total infection cases during the spread of epidemics with minimal social distancing effort. The controller variables are the weights of the contact network. The objective is to find optimal contact weights to minimize a cost function. The cost function is a weighted sum of the new infection cases and the cost associated with the social distancing effort. However, the social distancing patterns are actually decentralized in nature. For example, an individual will adopt a social distancing behavior according to his/her costs and perception of risk. In other words, the social distancing behavior of an individual is not necessarily to benefit the whole network, but rather, to maximize a personal utility function.

In this paper, we define personal social distancing behaviors that are sub-optimal solution of the original social distancing problem. The weight reduction is driven by the infection incoming from each neighbor, using a threshold value for the infection probability, the infection awareness threshold I_{th}, which triggers the reduction. Once a weight on a contact link is reduced, it remains at this reduced level until the infecting node recovers or the (formerly) susceptible node has reached the infected state. An example of contact level reduction is depicted in figure 6. Node 1 is the individual under consideration, and nodes 2-6 are its neighbors. The thicknesses of the links between node 1 and the neighbors represent the levels of contact. Nodes 2, 4, and 5 are in a closer contact with node 1 than nodes 3 and 6, as shown on the left panel of figure 6. Suppose now that an epidemic starts. Some nodes stay susceptible (green), while others become infected (red). Different tones of red represent different level of infections, with higher force of infection being represented by darker red. In the right panel of figure 6, node 1 is healthy or susceptible and needs to protect itself from receiving the infection from the infected nodes 3-6. However, the forces of infection from nodes 3 and 6 are less than the threshold I_{th}, so the contact level is not varied from the original value. The contact level between node 1 and 2 is not reduced since node 2 is healthy; while the contact levels with nodes 4 and 5 is reduced because these neighbors have forces of infection higher than the threshold.



Contact network for node 1 | Contact reduction as a function of infectiousness levels

Figure 6. Example of contact level reduction for node 1 when nodes 5 and 4 reach a value of the infection greater than the threshold.

Vaccination Strategies. Vaccination is carried out by selecting a set of nodes and immunizing them with a certain vaccine efficacy rate. We consider seven different strategies for selecting the set of nodes for vaccination. The first and simplest strategy is a random selection of 10% of the population (35 nodes). The random method represents a blind distribution across the population. The next three strategies consider a targeted selection of nodes (individuals) based respectively on the three node metrics, incoming node strength, outgoing node strength, and node betweenness. These three strategies are idealistically implemented by selecting the 35 nodes with the highest values for the respective metric and administering the vaccine. For less ideal situations, we consider three additional strategies that attempt to represent feasible vaccine distribution strategies for rural populations. Considering again the three above mentioned network metrics, we determine the location which has the highest average value (on the set of nodes that visit the location) of each metric. These locations are a restaurant (outgoing node strength), a pharmacy (node betweenness), and a location used for public events (incoming node strength). After selecting the locations that represent on average the best places to find nodes with higher values of each metric, we consider a random selection within a location of 10% of the entire population for vaccination. This location-based targeting has been proposed in [26]. It allows an indirect (and thus more feasible) targeting of critical populations that ensures a more effective use of resources than widely distributing resources in a global manner. Note that there is an implicit assumption that the entire population is susceptible previous to the distribution of the vaccine. Although this is not a realistic assumption for a commonly occurring strain of influenza, it would likely be the case for any new disease threat. In figure 7, a simple exemplification of these strategies is described.

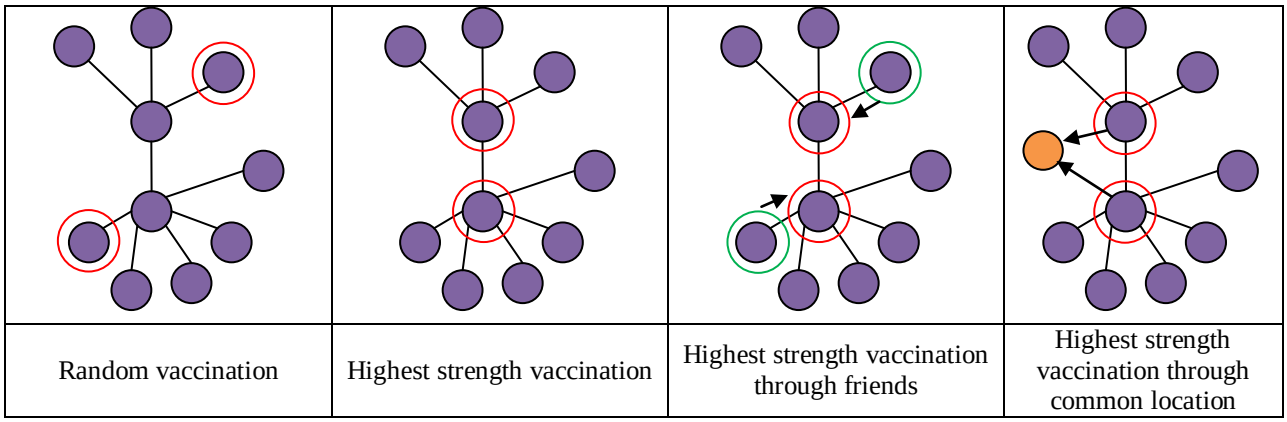


Figure 7. Example of vaccination strategies for individuals (purple nodes) in a contact network. Orange nodes represent locations, red circles represent the selected nodes for vaccination, and the green nodes represent random selected individuals, whose friends will be candidate for vaccination.

Evaluation Function for Comparison. We develop a cost function that enables the comparison of the various mitigation strategies from a pseudo-resource perspective. The function is designed to capture the change in the weights representing social contact similar to [17], but it also considers an "effective" change to the social contact by a reduction in the total infection rate $w_{ij}\beta$. Including the basic infection rate β allows it to also compute a cost for the social alertness response. We express the cost as

$$C = \sum_{i=1}^{N} \sum_{j=1}^{N} \int_{t_0}^{t_{final}} (\beta' w_{ij}' - \beta w_{ij}) dt, \qquad (1)$$

where w_{ij}' and β' are the (constant) weight and (constant) basic infection rate without mitigation. For the vaccination strategies, the basic infection rate and the weights are constant for the duration of the epidemic. The cost is non-zero for vaccination because every vaccinated node will effectively set its incoming weights to zero. This leaves only the duration of the epidemic and the total incoming strengths of the successfully vaccinated nodes to differentiate among the different vaccination strategies. For the social alertness responses, the weights are kept constant, leaving an integration of the change in basic infection rates over time. Similarly, for the dynamic contact responses, the basic infection rate is kept constant, leaving an integration of the changes in the weights over time. The cost function, when applied to the dynamic contact responses, is identical to the cost function used in [17] if it were scaled by the basic infection rate β.

4. Results

Network Analysis. We measured on the network the metrics of interest for the vaccination targeting strategies. Figure 8 shows the diversity found in the weights that measure the levels of contact between each neighboring pair of nodes. Roughly 30 percent of the links carry very small weights, and there are very few links representing the highest weighted contacts. In figure 6 we display two views of the network topology to visualize the estimated rural community contact structure. Since the network is rather dense, we remove the links with lower weights in two different patterns. On the left side of figure 6, we colored the nodes and the links having weights between 0.2 and 1.0, where the weights of the green links are between 0.2 and 0.3 and those of

the purple links are between 0.3 and 1.0. In this depiction, a minority but significant set of individuals (roughly 50 nodes) can been noticed for their state of isolation. These nodes are not strongly connected to the core of the network, but are connected when the links with the lowest weights are considered. This loosely connected "fringe" of the rural community is rarely reached by epidemics until a very strong epidemic comes (notice the strange inflection in figure 10 of the SI in the upper end of the 95 percent confidence interval on the end of the spectrum having the highest disease strengths). On the right side of figure 9, we colored the nodes and the links having weights between 0.4 and 1.0 as well as what we call the "best-friends" links. For each node, we select the link having the highest out-going weight and define this link as the "best-friend" link of the node. This depiction of the network captures the most likely paths (it is composed of the highest weighted links) that an epidemic might take from anywhere in the network towards the center of the network. Although this pattern of visualization may give the false impression that the network is tree-like or scale-free, an epidemic would leave a tree-like pattern as it traces its way through the rural community. Note that figure 8 proves that both of these network visualizations in figure 9 are missing majorities of the links in the complete network.

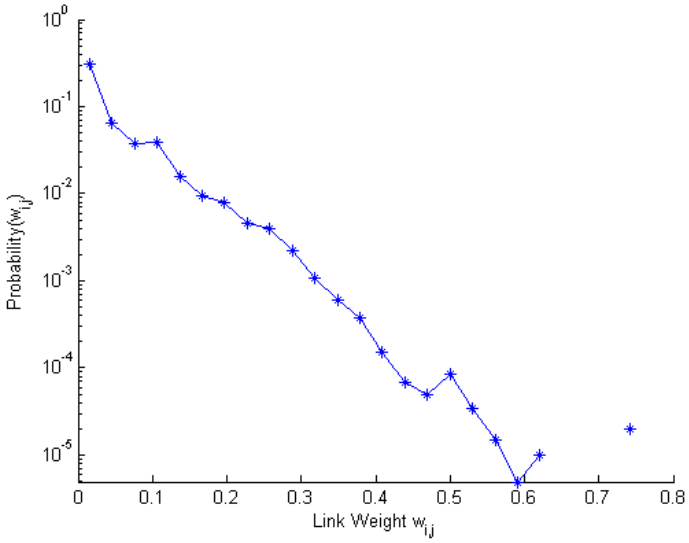


Figure 8. The distribution of the weights representing social contact on the links for the rural community contact network. Note that the vertical axis has a log scale.

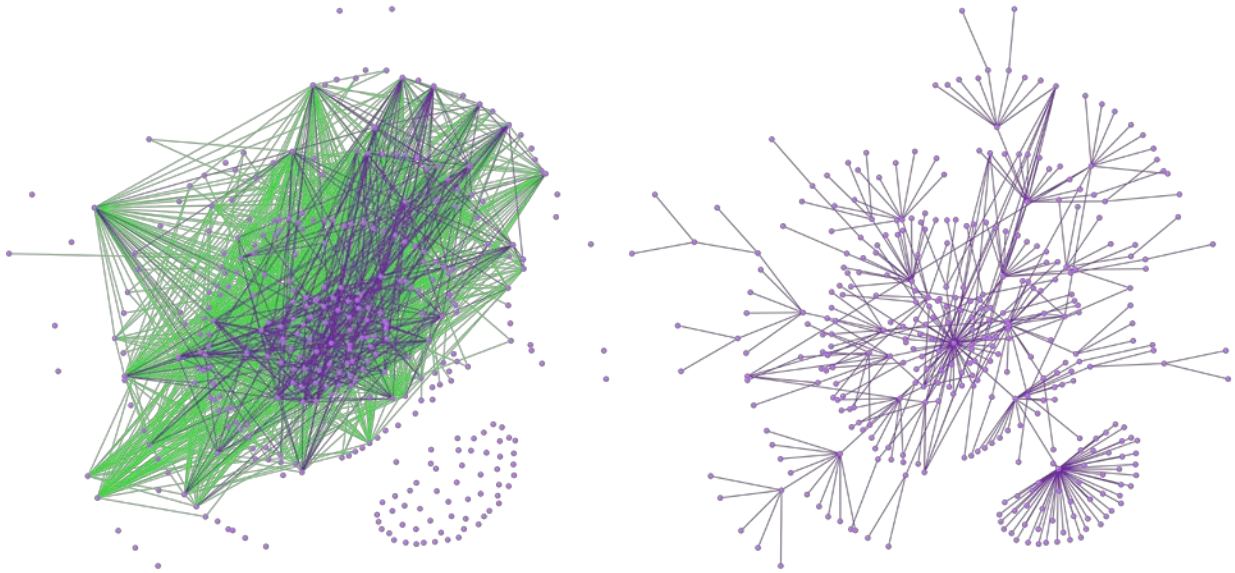


Figure 9. (Left) A visualization of the rural community contact network showing the nodes and the links having weights between 0.2 and 1.0, where the weights of the green links are between 0.2 and 0.3 and those of the purple links are between 0.3 and 1.0. (Right) A visualization of the rural community contact network showing the nodes and the links having weights between 0.4 and 1.0 as well as the "best-friends" links, where the best friend link of a node is defined as the link having the highest out-going weight.

Figure 10 shows the distribution of the node betweenness metric for the network. More than 80 percent of the nodes have very small values of node betweenness, leaving a select group of nodes that are critical connections in the system of shortest paths through the community.

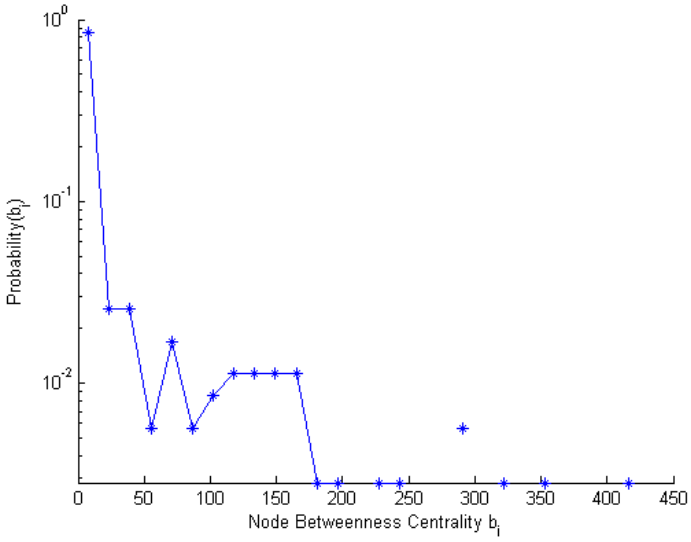


Figure 10. The distribution of the node betweenness values of the individuals in the rural community contact network. Note that the vertical axis has a log scale.

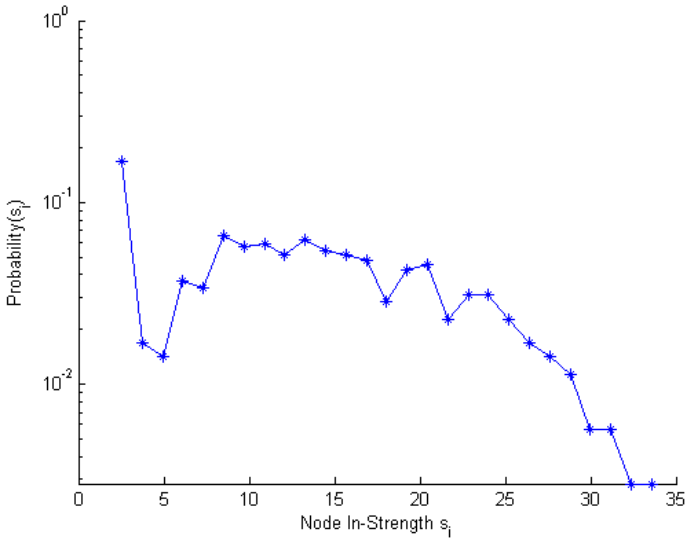


Figure 11. The distribution of the node in-strength values of the individuals in the rural community contact network. Note that the vertical axis has a log scale.

Figure 11 depicts the distribution of the node in-strength metric for the network. It is much less heterogeneous than the node betweenness and link weight distributions as the in-strengths are found rather homogeneous across the nodes. We explored the correlations between the network metrics and various survey responses and found that node betweenness was significantly correlated with age (r = -.15, p < .01), travel time to work (r = .20, p < .001), distance to work (r = .19, p < .001), level of education (r = .12, p < .05), number of non-family friends contacted weekly (r = .51, p < .05), and hours away from home each day (r = .22, p < .001). The outgoing node strength was significantly correlated with age (r = -.20, p < .001), visiting with more family members outside one's residence (r = .18, p < .01), household size (r = .12, p < .05), travel time to work (r = .44, p < .001), distance to work (r = .45, p < .001), compliance risk (r = .13, p < .05), level of education (r = .23, p < .001), income (r = .17, p < .01), having diabetes (r = -.12, p < .05), how often one eats out (r = .16, p < .01), and hours away from home each day (r = .50, p < .001). Node in-strength was correlated with level of education (r = .12, p < .05) and having had the flu in the past six months (r = -.12, p < .05).

In general, while many of these relationships are not especially strong in terms of effect sizes, it appears that residents with higher levels of education, who have longer commutes, who are younger, with more income, those without diabetes or recent flu-like illnesses, who are away from home more hours each day, and who eat out more often are more likely to be important agents in the network measures that influence the potential spread of epidemics. It is interesting to observe that the younger rural residents are likely the most important agents when considering that rural regions are typically characterized by aging populations. This importance appears to be due to them, the younger persons, spending more time away from home, driving longer to work, visiting more businesses, and in all this, having and visiting more persons outside of their homes. Perhaps, the traditional farmer who rarely visits town and is mostly self-sufficient within his home and immediate neighbors is giving way to a younger generation and changing economy where increased travel and social interaction are increasingly required.

Disease Simulations. We performed extensive simulations to investigate potential epidemics and the proposed vaccination strategies for the rural contact network representing a sample population from Neosho County. To mimic a realistic epidemic with the stochastic SLIR model, we utilize average values of λ^{-1} = 0.764 days, μ^{-1} = 1.736 days, and $R_0 = \beta/\mu$ = 1.75 with respective standard deviations of 0.100 days, 0.100 days, and 0.065 [27-33]. We explore the hypothetical outbreaks first by simulating 1,000,000 trials of the considered situation (such as without mitigation or with a specific mitigation strategy). For each trial, a triple of (λ, μ, R_0) is drawn from the three Gaussian distributions with the respective parameters and the outbreak is simulated until it dies out, leaving only susceptible and recovered individuals behind. The second type of experiment we ran was the simulation of sets of 10,000 trials that scan over values R_0 to quantify the range of potential outbreaks. In this second type of experiment, we deterministically vary R_0, while λ and μ are still drawn from their distributions [34, 35].

For each simulation, we track the numbers of nodes in each disease state through time as well as the timings of all event occurrences. We capture the total cases and duration of each outbreak and this is the majority of the data presented. We simplify the presentation of the results of the second type of experiment by computing and plotting the average and 95% range of the resulting total cases for each group of 10,000 simulations on a single R_0 [27-29, 31-33]. Figure 12 summarizes the distributions of the total cases as a fraction of the population in the manner described above. As R_0 increases, the epidemic size increases in a near-linear manner. It can be seen that distributions are broad but have low average values. This figure suggests that around 5 percent of the population might typically fall sick during an influenza season, but severe outbreaks might touch 30-40 percent of the community.

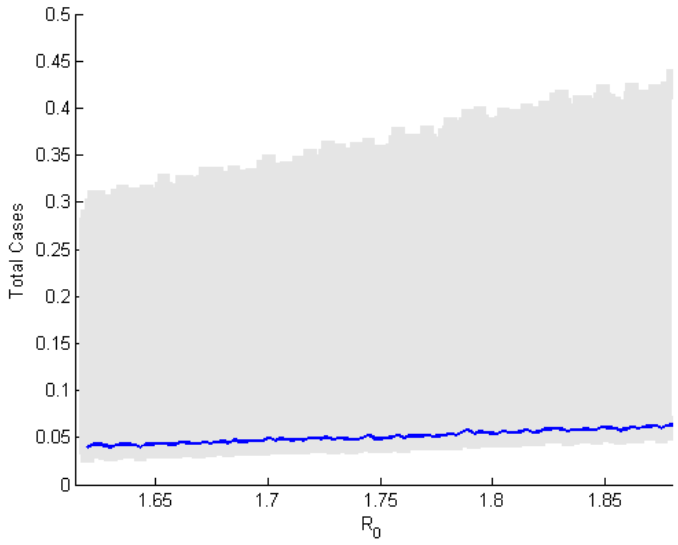


Figure 12. Under no mitigation, the distributions of the total cases as a fraction of the considered population over the estimated range of R_0 are represented by the averages (blue line) and 95% confidence interval (grey shaded region). As the infection rate increases, the epidemic size increases in a near-linear manner.

Vaccination Strategies. We ran seven sets of simulations to consider the seven vaccination strategies described in Section 3 and for each set we ran both types of experiments as described previously. In each trial, we draw a value for vaccine efficacy from a Gaussian distribution with mean of 72.0% and standard deviation of 6.0% to approximate realistic efficacy values [39-41]. The first vaccination strategy, the random distribution over the entire population, is the selection of a group of individuals representing 10 percent of the population and administering vaccines prior to the start of an outbreak with the given efficacy. Figure 13 demonstrates the potential reduction in the distributions of outbreaks by random vaccination.

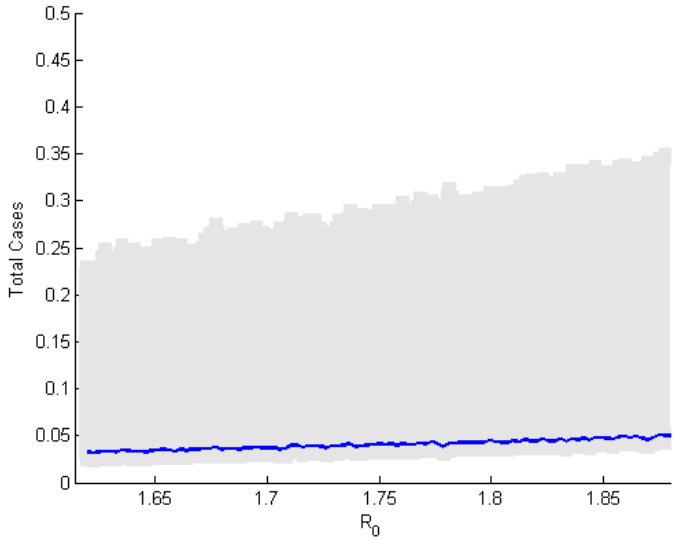


Figure 13. Under a random vaccination of 10 percent of the population, the distributions of the total cases as a fraction of the considered population over the estimated range of R_0 are represented by the averages (blue line) and 95% confidence interval (grey shaded region). As the infection rate increases, the epidemic size increases in a near-linear manner.

The three idealistic vaccination strategies select their targets and vaccinate them by rankings determined by the node metrics. The left side of figure 14 captures the reduced epidemic sizes under an individual targeting strategy which uses node betweenness to select the individuals. For a realistic targeting of a distribution location, the right side of figure 14 captures the potential reductions in the epidemic sizes under the node-betweenness-based location targeting strategy. The location-based strategies are intuitively less successful than the individual targeting methods, but they represent much more feasible options for an administrative intervention.

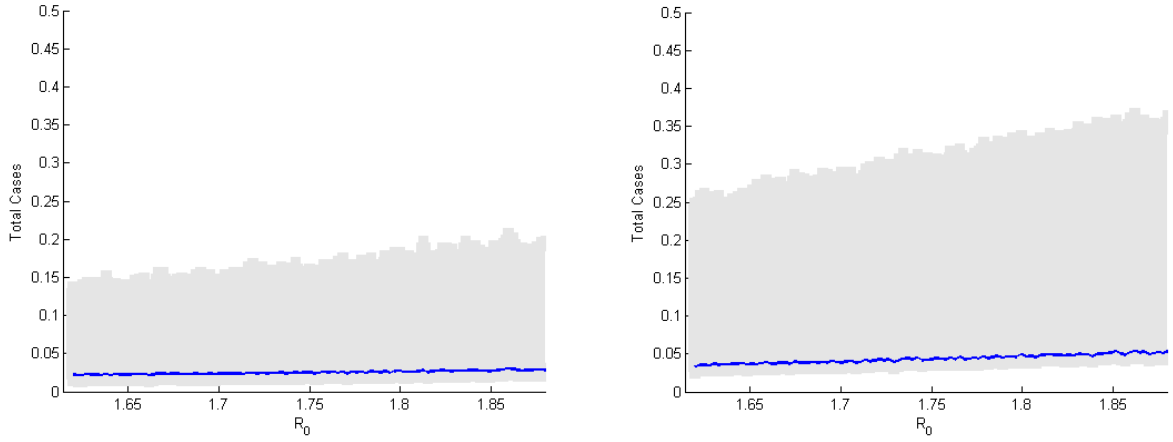


Figure 14. (Left) Under a node-betweenness-based *individual* targeted vaccination of 10 percent of the population, the distributions of the total cases as a fraction of the considered population over the estimated range of R_0 are represented by the averages (blue line) and 95% confidence interval (grey shaded region). As the infection rate increases, the epidemic size increases in a near-linear manner. (Right) Under a node-betweenness-based *location* targeted vaccination of 10% of the population, the distributions of the total cases as a fraction of the considered population over the estimated range of R_0 are represented by the averages (blue line) and 95% confidence interval (grey shaded region).

We summarize the comparison of the different vaccination strategies in table 1, which lists epidemic sizes as fractions of the total population, the epidemic durations in days, and the average cases prevented per vaccine (CPPV) for the vaccination strategies. It is immediately interesting to notice that each of the vaccination strategies reduces the average epidemic duration, as much as 20 percent on the average value. In table 1, it can be seen that the individual targeting methods have the highest average results, but among the feasible methods, the location-based targeting using the node betweenness metric is the most successful at reducing the total cases on average. The node betweenness also provides the best metric for the individual targeting strategies.

The last column of table 1 displays the cases prevented per vaccine distributed. The value cases prevented per vaccine has an intuitive benchmark of the average vaccine efficacy at 0.72. If a vaccination strategy is very efficient at stopping an outbreak, then we can expect the average number of cases prevented per vaccine to be higher than the typical efficacy of the vaccine. On the other hand, if a vaccinated trial is resulting in an average number of cases prevented per vaccine that is less than the typical vaccine efficacy, it doesn't necessitate that the vaccination strategy will perform poorly in all situations. In general this situation implies that the vaccines are being given to individuals who usually aren't being infected and therefore they made little use of the vaccine in that set of trials. This could arise from either a poor vaccine distribution strategy or from a distribution of vaccines that is larger in size than a typical outbreak. When we have a strong outbreak, the vaccines are almost surely going to be a necessary measure, whereas in a weaker outbreak, most of the population will not be infected and extra vaccines will be "unused" with respect to preventing new cases. Notice that for the first type of experiment when we are not considering any vaccination the epidemic impacts roughly 5.1 percent of the population while the number of vaccines distributed is sufficient for 10 percent of the population.

Table 1. Total cases as a fraction of the population, cases prevented per vaccine (CPPV), and the duration of the epidemic outbreaks in days (*in italics*).

	Average	95% CI	CPPV
No Vaccination	0.0512	(0.0057, 0.3088)	–
	10.7200	*(1.6101, 35.6527)*	
Random Vaccination	0.0407	(0.0057, 0.2493)	0.1065
	9.9031	*(1.6190, 32.4859)*	
Targeted Betweenness (TB)	0.0251	(0.0057, 0.1388)	0.2638
	8.3638	*(1.5960, 25.2766)*	
Targeted In-strength (TIS)	0.0324	(0.0057, 0.1955)	0.1898
	9.1509	*(1.6046, 29.4892)*	
Targeted Out-strength (TOS)	0.0261	(0.0057, 0.1445)	0.2537
	8.4930	*(1.6046, 25.8235)*	
Location Targeted Betweenness (LTB)	0.0433	(0.0057, 0.2635)	0.0797
	10.1597	*(1.6073, 33.5715)*	
Location Targeted In-strength (LTIS)	0.0433	(0.0057, 0.2635)	0.0795
	10.1652	*(1.6231, 33.4379)*	
Location Targeted Out-strength (LTOS)	0.0434	(0.0057, 0.2635)	0.0790
	10.1558	*(1.6153, 33.5992)*	

Disease Spectrum Simulations. We performed 16,000,000 simulations to investigate potential epidemics and the proposed vaccination and spontaneous response strategies for the rural contact network across a wide spectrum. To mimic a realistic epidemic with the stochastic SLIR model, we set λ = (1.1 days) $^{-1}$ and μ = (2.5 days) $^{-1}$ [26, 32, 33]. We explored a wide range of values for the basic infection rate: 500 values varying from β = (100 days) $^{-1}$ to β = (0.2 days) $^{-1}$. Previous work had explored only a single value of β = (2.5 days)$^{-1}$ [26]. For each strategy and each value of β, we ran 1000 simulations to sample the stochastic space of the epidemic model. Therefore one "set" of simulations contains 0.5 million simulations.

For each simulation, we track the numbers of nodes in each disease state through time as well as the timings of all event occurrences. Additionally we capture the total cases of each outbreak and the cost of the associated mitigation strategy, and this is the majority of the data we present. We simplify the presentation of the data by computing and plotting the average and 95% range of the resulting total cases and costs for each group of 1000 simulations on a single β. For displaying the range of β values, we adopt an inverted infection rate axis that has been used for a network metric known as Viral Conductance, which explores a spectrum of epidemic strengths on a network [29, 30]. We summarize in figure 15 the estimated ranges of disease strengths of various known threats with respect to the spectrum of disease strengths which we explore through simulation ($0 < \mu/\beta \leq 40$) [32, 36-38]. We have scaled the estimated ranges by the average weight on the links of the contact network to separate the contact from the basic infection rate.

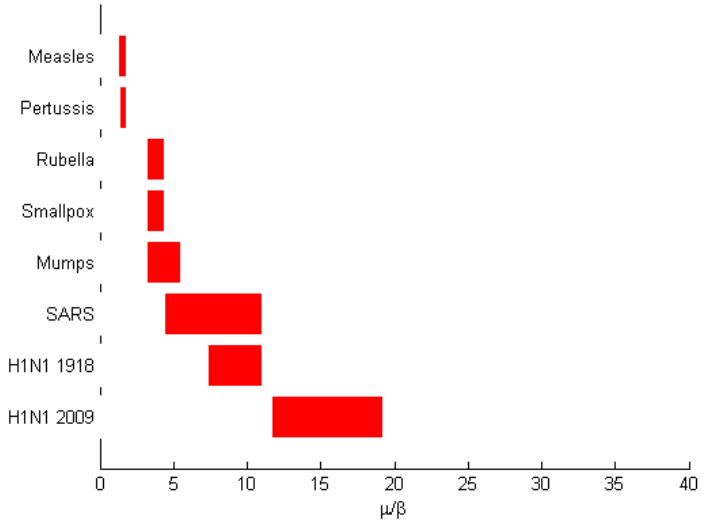


Figure 15. The estimated ranges of disease strengths of current and historical epidemic threats are stacked in increasing level of infection rates over the spectrum of disease strengths ($0 < \mu/\beta \leq 40$) explored in each set of simulation for various mitigation strategies.

The first case to consider is the set of potential epidemics when there is no mitigation response. From the "set" of simulations without any mitigation, we show a sample of an outbreak for a basic infection rate of $\beta = 0.042$ (days) $^{-1}$ in figures 4 and 16. Figure 17 summarizes the distributions of the total cases and thus the total sizes of epidemics within the population in the manner described above. As the infection rate decreases, the epidemic size decreases in a non-linear manner.

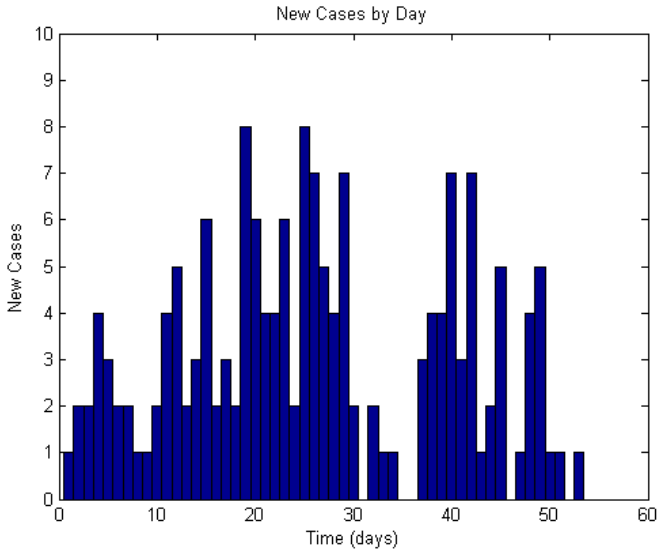


Figure 16. Under no mitigation, with a basic infection rate of $\beta = 0.042$ (days)$^{-1}$, $\lambda = (1.1$ days$)^{-1}$, and recovery rate of $\mu = (2.5$ days$)^{-1}$, a sample outbreak is shown by depicting the new cases of infection per day.

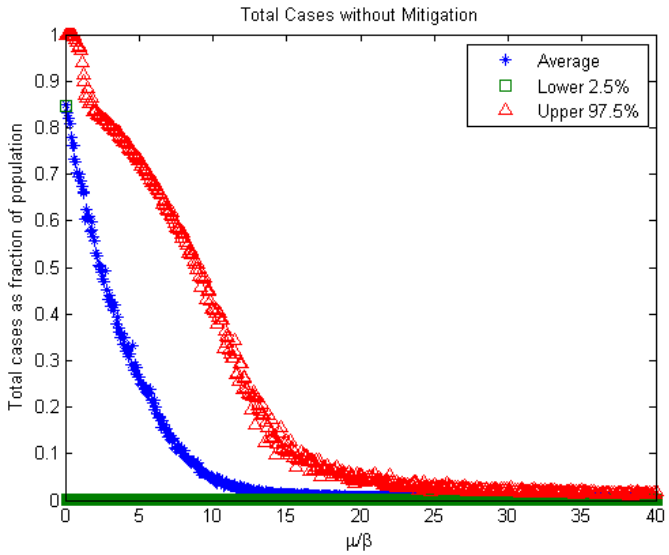


Figure 17. Under no mitigation, the distributions of the total cases as a fraction of the considered population over several values of the basic infection rate are represented by their averages (blue stars), lower 2.5% sample (green squares), and upper 97.5% sample (red triangles). As the infection rate decreases, the epidemic size decreases in a non-linear manner.

Social Alertness Responses. We ran twelve sets of simulations to consider the social alertness strategy described in Section 3 under various values of the alerting rate κ. We set the alerted basic infection rate to $\beta_a = 0.1\beta$ and tested values of κ from 10^{-6} to 0.5 [16]. A higher κ value represents a stronger mitigation strategy, but it will also come at a higher cost as it causes more nodes to reduce their incoming infection rates. Figures 18 and 19 present the prevented cases and cost distributions for κ=0.005 and a stronger mitigation with κ=0.1.

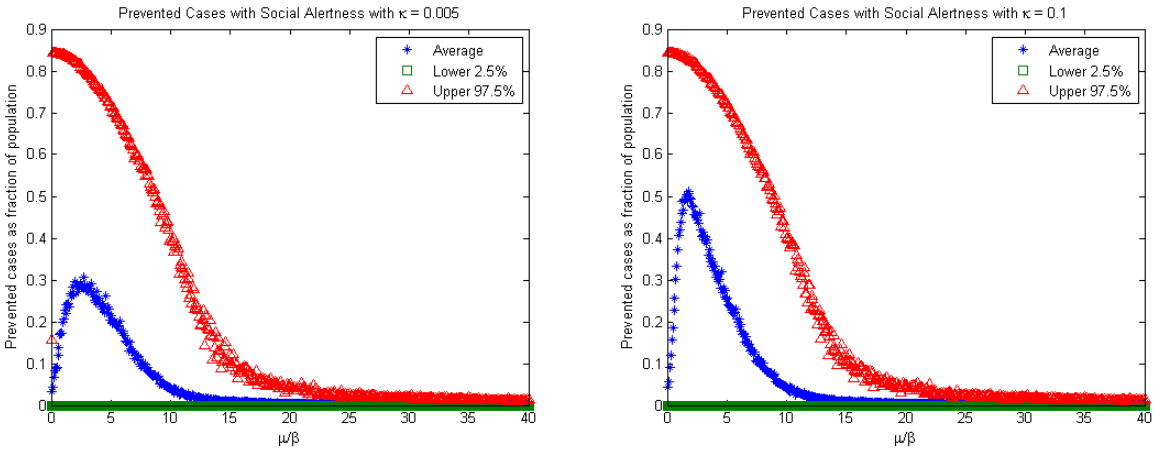


Figure 18. Under spontaneous social alertness responses to the epidemics with alerting rates of κ=0.005 (left) and κ=0.1 (right), the distributions of the prevented cases (from the social response intervention) as a fraction of the considered population over several values of the basic infection rate are represented by their averages (blue stars), lower 2.5% sample (green squares), and upper 97.5% sample (red triangles).

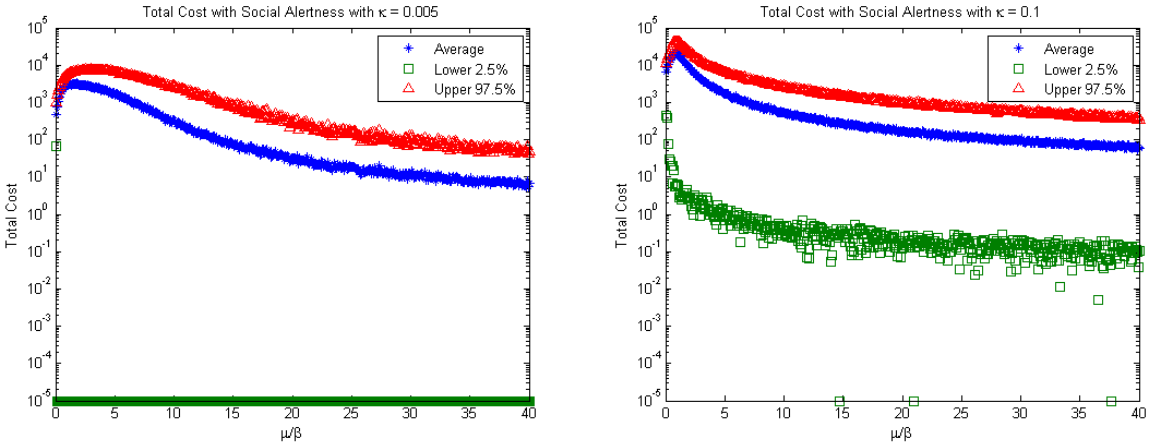


Figure 19. Under spontaneous social alertness responses to the epidemics with alerting rates of κ=0.005 (left) and κ=0.1 (right), the distributions of the total mitigation cost values (From Equation 1) over several values of the basic infection rate are represented by their averages (blue stars), lower 2.5% sample (green squares), and upper 97.5% sample (red triangles) on a vertical log scale.

Social Distancing Responses. We ran twelve sets of simulations to consider the dynamic social distancing response described in Section 3, under various values of the infection awareness threshold I_{th}. We set the minimal weight parameter to α = 0.1 and tested values of I_{th} from 10^{-6} to 0.5 [17]. A lower I_{th} value represents a stronger mitigation strategy, but it will also come at a higher cost as it causes more nodes to reduce their incoming contact levels (the weights on the incoming weights). Figures 20 and 21 present the prevented cases and cost distributions for I_{th}= 0.005 and for a stronger mitigation with I_{th}= 10^{-6}.

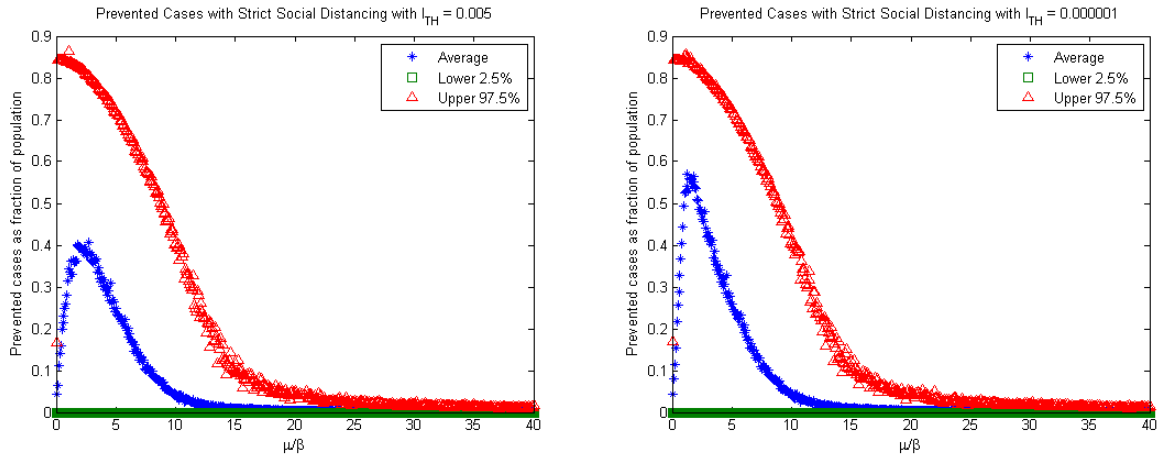


Figure 20. Under spontaneous social distancing responses to the epidemics with infection thresholds of I_{th}= 0.005 and I_{th}= 0.000001, the distributions of the prevented cases as a fraction of the considered population over several values of the basic infection rate are represented by their averages (blue stars), lower 2.5% sample (green squares), and upper 97.5% sample (red triangles).

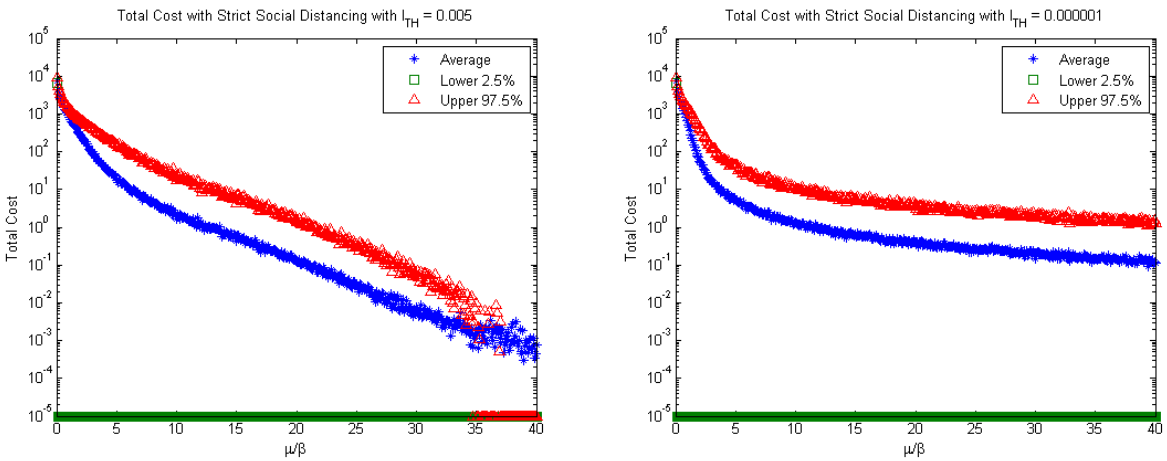


Figure 21. Under spontaneous social distancing responses to the epidemics with infection thresholds of I_{th}= 0.005 and I_{th}= 0.000001, the distributions of the total mitigation cost values (From Equation 1) over several values of the basic infection rate are represented by their averages (blue stars), lower 2.5% sample (green squares), and upper 97.5% sample (red triangles) on a vertical log scale.

Vaccination Strategies. We ran seven sets of simulations to consider the seven vaccination strategies described in Section 3. We use a vaccine efficacy of 70% to represent a typical value [41]. The first vaccination strategy, the random distribution over the entire population, is summarized over the same spectrum in figure 22. The cost distributions are shown with a vertical log scale in figure 23.

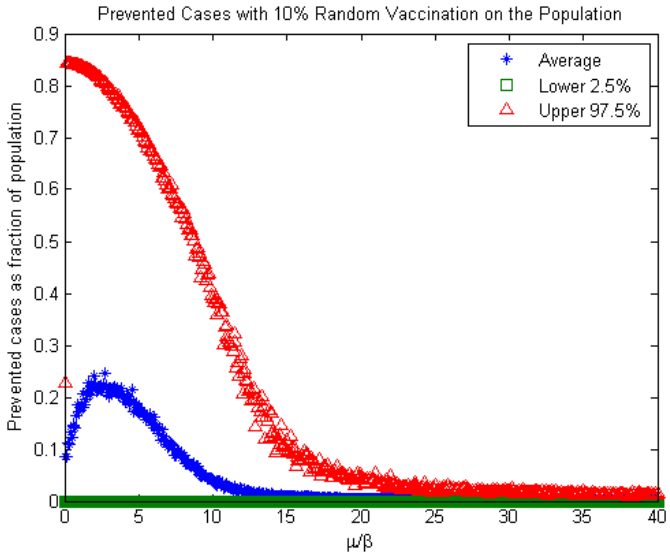


Figure 22. Under a random vaccination of 10% of the population, the distributions of the prevented cases as a fraction of the considered population over several values of the basic infection rate are represented by their averages (blue stars), lower 2.5% sample (green squares), and upper 97.5% sample (red triangles).

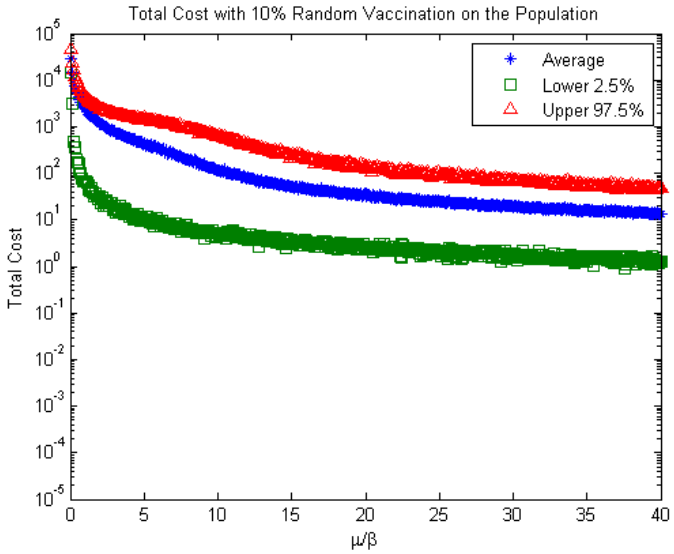


Figure 23. Under a random vaccination of 10% of the population, the distributions of the total mitigation cost values (From Equation 1) over several values of the basic infection rate are represented by their averages (blue stars), lower 2.5% sample (green squares), and upper 97.5% sample (red triangles) on a vertical log scale.

The three idealistic vaccination strategies select their targets and vaccinate them by rankings determined by the node metrics. Figure 24 captures the potential reductions in the epidemic sizes under an individual targeting strategy. For a realistic targeting of a distribution location, figure 25 captures the potential reductions in the epidemic sizes. The location-based strategies are intuitively less successful than the individual targeting methods, but they represent much more feasible options for an administrative intervention.

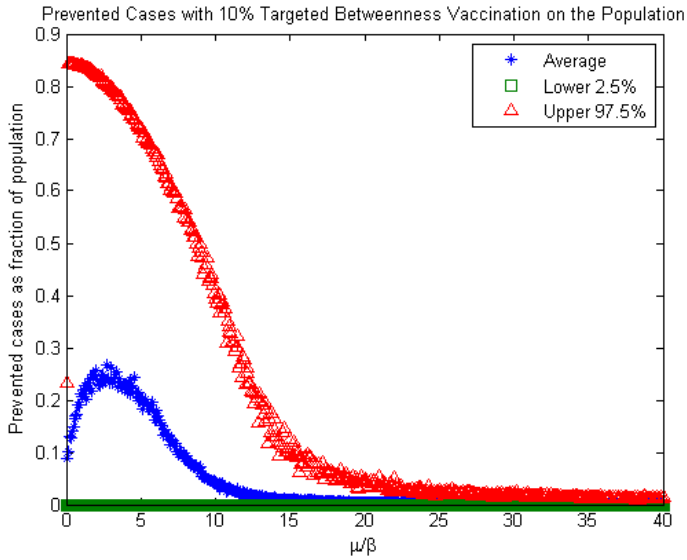


Figure 24. Under a node-betweenness-based *individual* targeted vaccination of 10% of the population, the distributions of the prevented cases as a fraction of the considered population over several values of the basic infection rate are represented by their averages (blue stars), lower 2.5% sample (green squares), and upper 97.5% sample (red triangles).

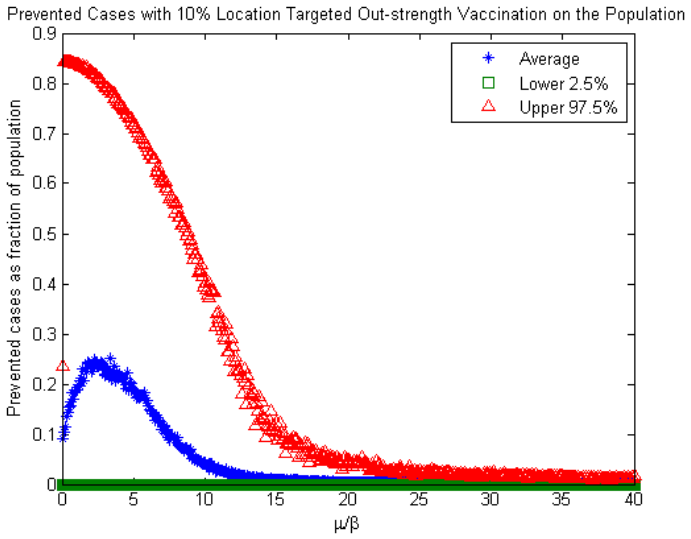


Figure 25. Under a node-out-strength-based *location* targeted vaccination of 10% of the population, the distributions of the prevented cases as a fraction of the considered population over several values of the basic infection rate are represented by their averages (blue stars), lower 2.5% sample (green squares), and upper 97.5% sample (red triangles).

We summarize the comparison of the different vaccination strategies in table 2, which lists the average cases prevented per vaccine distributed in four perspectives. The first perspective is the average cases saved per vaccine across all tested values of beta. This column is biased significantly by the epidemics having small values of beta, those in the "tail." The vaccines have little impact on the weaker outbreaks and thus averaging over all beta values has less significance except for comparisons between vaccination strategies. More realistic results are observed in the next three columns, which list the average cases prevented per vaccine distributed for three potential strong epidemics.

Table 2. Average number of cases prevented per vaccine.

	Average	$\beta = 5.00$	$\beta = 0.50$	$\beta = 0.05$
Random	0.0112	0.6237	0.5696	0.2243
Targeted Out-strength (TOS)	0.0218	0.7200	0.6225	0.5878
Targeted In-strength (TIS)	0.0181	0.6574	0.7029	0.4880
Targeted Betweenness (TB)	0.0228	0.6290	0.8182	0.6121
Location Targeted Out-strength (LTOS)	0.0171	0.6698	0.6224	0.4746
Location Targeted In-strength (LTIS)	0.0165	0.6422	0.6654	0.3829
Location Targeted Betweenness (LTB)	0.0165	0.7619	0.7339	0.4522

In table 2, it can be seen that the individual targeting methods have the highest average results, but among the feasible methods, the location-based targeting using the node out-strength metric is the most successful on average. Given that we have simulated each vaccine having an efficacy of 70%, we can see that some combinations of strategy and disease strength improve this base level of efficiency, while others fail to do so. To summarize the comparison among the three location-based strategies, figure 26 plots the efficiencies of the three methods against all tested disease strengths. The ranges in which each strategy is most efficient are signaled by the colored bar across the top of figure 26.

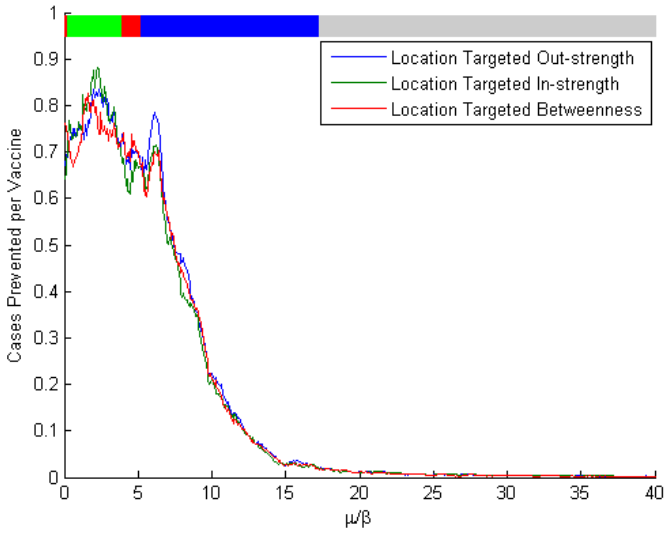


Figure 26. Three location-based vaccination distribution strategies are compared by vaccine efficiency across a spectrum of disease strengths. Each strategy distributes vaccines randomly among visitors respectively at a specific restaurant (blue), a specific pharmacy (red), and a location used for public events (green). The locations are chosen based on having the highest average of the node metrics outgoing node strength (blue), node betweenness (red), and incoming node strength (green). The bar across the top of the figure uses color to signify the ranges in which each strategy is the most successful one, even if only marginally. There is no strategy which distinguishes itself above the others in the grey region, and all are inefficient due to the small size of the outbreaks.

Comparison. In each of the social and vaccination based mitigation strategies, a peak can be observed in the average number of cases prevented. This peak is due to each strategy working best in some intermediate range. For extremely strong epidemics, the mitigation strategies are insufficient and they disease will overcome them as much as possible. As the diseases are weaker at the higher values of μ/β, the mitigation strategies are exerting more effort than necessary to impede the disease and are thus less efficient. Although interesting trends can be observed in the above curves, table 3 simplifies the comparison among the different mitigation methods through aggregation. Table 3 integrates the total cases curves in a manner similar to Viral Conductance to give a measure of the overall epidemic potential under each method [34, 35]. The first column represents the integration using the average total cases curve, while the second column uses the 97.5% samples. A *lower* value in table 3 represents a more successful mitigation (on average across the infection rate range). A comparison of the results for the vaccination strategies gives results similar to those in table 2, with the outgoing node strength again being the best among the location-based strategies. The success of the two social response approaches depends significantly on how strong of a response is carried out, as given respectively by κ and I_{th}. While the social alertness exhibits a smooth transition in the aggregated epidemic strengths as κ varies, the social distancing has a relatively sharp transition as I_{th} varies.

Table 3. Aggregated Epidemic Strengths for Mitigation Strategies and Responses. The strategies listed under Vaccination are abbreviations of the strategies listed in table 2.

Without Mitigation			Social Alertness			Social Distancing		
	Average	97.50%	κ	Average	97.50%	I_{TH}	Average	97.50%
	8.7598	23.0779	10^{-6}	8.7549	23.0994	10^{-6}	2.1146	4.4173
			10^{-5}	8.7118	23.1187	10^{-5}	2.1090	4.4099
Vaccination			10^{-4}	8.6297	22.8042	10^{-4}	2.1026	4.3958
Strategy	Average	97.50%	0.001	7.8554	20.3000	0.001	2.2156	4.9116
Random	7.6480	20.3901	0.002	7.2693	18.4994	0.002	2.4969	6.1980
TOS	6.5947	17.1671	0.005	6.1795	15.2139	0.005	4.0062	12.0555
TIS	6.9671	18.8768	0.01	5.2181	12.5227	0.01	6.6956	20.2309
TB	6.5032	17.0014	0.02	4.3147	10.0819	0.02	8.3960	22.5924
LTOS	7.0676	18.8932	0.05	3.3443	7.4649	0.05	8.7071	22.9317
LTIS	7.1273	19.2669	0.1	2.8098	6.0598	0.1	8.6845	23.0666
LTB	7.1282	19.1612	0.2	2.4143	5.0334	0.2	8.7322	23.2048
			0.5	2.0291	4.1300	0.5	8.7294	23.1232

When considering the social contact cost function, we can see the efficiencies of the various strategies across the tested disease strengths by computing the average cases prevented per unit of cost. Figure 27 compares the social cost efficiencies of the three location-based strategies, with a color bar similar to figure 26 representing the most efficient of the strategies by region. Figures 28 and 29 present the efficiencies of the social alertness and social distancing strategies. The strongest mitigations of both strategies also have the highest efficiencies, but the weaker responses tend to have more sensitive efficiencies. Due to this sensitivity caused by the very low costs associated with the weaker responses, we do not present the efficiencies of the weakest responses for each strategy. The social distancing, having cheaper costs as it is a link-based method as opposed to node-based, has greater sensitivity, but also higher efficiencies. Similar to the results of table 3, we observe a much sharper transition in the social distancing results (figure 28) than seen in the social alertness (figure 29).

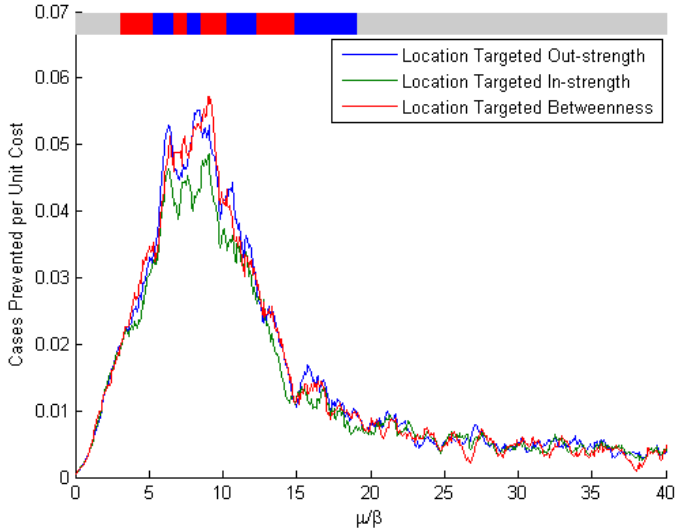


Figure 27. Three location-based vaccination distribution strategies are compared by social cost efficiency across a spectrum of disease strengths. Each strategy distributes vaccines randomly among visitors respectively at a specific restaurant (blue), a specific pharmacy (red), and a location used for public events (green). The locations are chosen based on having the highest average of the node metrics outgoing node strength (blue), node betweenness (red), and incoming node strength (green). The bar across the top of the figure uses color to signify the ranges in which each strategy is the most successful one, even if only marginally. There is no strategy which distinguishes itself above the others in the grey regions.

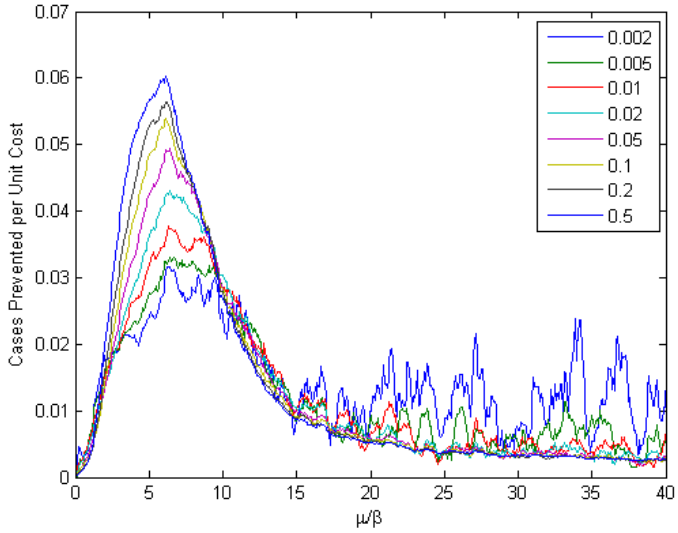


Figure 28. Eight social alertness strategies (κ = 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5) are compared by social cost efficiency across a spectrum of disease strengths. The remaining four responses were too sensitive to display. The strongest responses are also the most efficient according the costs described by Equation 1.

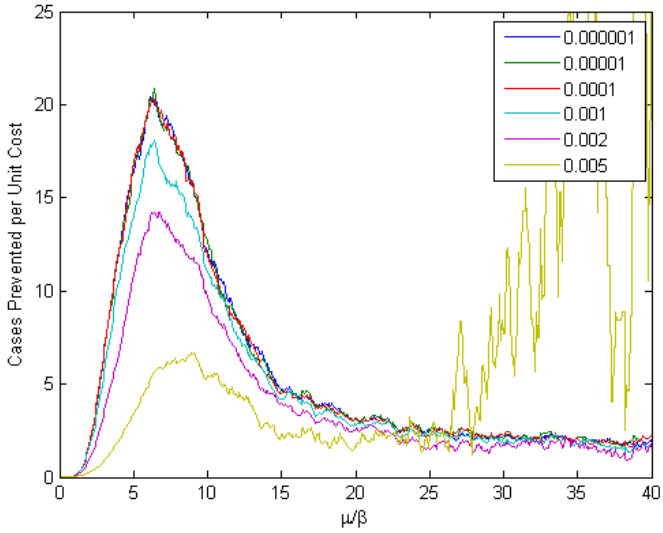


Figure 29. Six social distancing strategies (I_{th} = 10^{-6}, 10^{-5}, 10^{-4}, 0.001, 0.002, 0.005) are compared by social cost efficiency across a spectrum of disease strengths. The remaining six strategies were too sensitive to display. The strongest responses are also the most efficient according the costs described by Equation 1.

5. Discussion and Conclusions

Rural Community. From the network analysis, we observed that the rural contact structure displays a significant amount of heterogeneity in the considered metrics. This heterogeneity suggests that the small number of nodes having the highest values of each metric might present strategic sub-populations for mitigation objectives. The rural contact network also contained a relatively disease-resistant sub-population due to their poor level of connectivity and location on the "fringes" of the rural community network. From statistical correlations, it appears that residents with higher levels of education, who have longer commutes, who are younger, with more income, those without diabetes or recent flu-like illnesses, who are away from home more hours each day, and who eat out more often are more likely to be important players in the according to the network metrics that influence the potential spread of infectious diseases.

Social Behavioral Responses. The greatest challenge to implementing or encouraging a social response could reside in the decentralized and spontaneous nature of the process. As seen in figure 1, compliance is definitely an issue as captured in the respondents' responses to a hypothetical epidemic. Adoption of either behavioral response to an epidemic would come with a reduction in social relationships. Our cost function attempts to quantify the social cost to individuals through the change in effective contact levels. It can be seen (in figures 18-21) that for all of the cost associated with strongest epidemics, there are relatively few cases prevented because the disease still survives and touches most of the population. There is also a lack of any significant outbreak for the weaker diseases, which leaves few cases to be prevented by a social response in this range. Both the responses are most efficient for an intermediate range of disease strengths (figures 28, 29).

The social responses we model do not necessarily reflect realistic responses, but rather represent possible behavioral changes. When considering an entire county, the population would be expected to exhibit a heterogeneous set of responses to an epidemic. The social alertness, based on social and/or habits, and the social distancing, based on individual relationships, are both capable of being very successful in stopping outbreaks of any realistic disease strength. The most interesting difference noticed between the two methods is the relatively sharp transition in the effectiveness of the social distancing as the threshold parameter is varied. This is likely due to the difference in functionality between the threshold method and the rate-based method. The threshold should be ineffective when it is not low enough and possibly useful when it is. Within a population, we would also expect heterogeneity in whether individuals respond to epidemics based on their observances being more than a threshold or at a rate proportional to their observances.

Vaccination Strategies. For vaccine distribution we considered seven strategies, but only four are reasonably feasible for local administrators to implement, those being the random distribution across the population and the three location-based distributions. The traditional targeted groups for distribution such as the health-care personnel, the very young (6-59 months), the elderly (50 years or older), pregnant women, those with chronic health issues, and American Indians are not completely identifiable from our survey results [42]. We could identify respondents by age range, but occupation and maternity status are transitory positions and were not explored by the survey. The global random distribution of vaccines gives a simplest method to compare the other vaccination methods to. The location-based methods are indicative for anonymously targeting subpopulations, not only for vaccination campaigns, but also for educational outreach to encourage social responses such as adoption of preventative health practices. From figures 18, 20, 22, 24, and 25, it is possible to observe that vaccinations are the most effective interventions when considering extremely infectious diseases. When individuals are successfully vaccinated, they have full protection from the pathogen, full protection that cannot be guaranteed by the social responses considered here.

Interestingly, using the network metrics to select locations does not necessarily produce intuitive results. The restaurant chosen to represent locations that are frequented by nodes with high node outgoing node strength (as it had the highest average value) had less than one-third of the survey respondents frequenting it than some of the more popular restaurants in the region. Although diseases are partially mitigated, there is a limit to the reduction that can be observed in the total cases for the strongest diseases due to the resource limitation. Therefore when considering limited-resource vaccine distribution, local administrators should probably follow the traditional priority schedule. However, the identification of the critical locations would be useful for preventative education efforts, real-time epidemic alerts, and emergency resource distribution.

Administrative Responses. The results of this analysis are intended to help guide responses to a rural epidemic threat. The first step in connecting an outbreak to these results would be to identify the estimated range of the disease strength. This information would identify the portion of the studied disease strength spectrum to consider. Once knowing this, responders can see the theoretical impacts that might be had from a limited-resource vaccine distribution and the social responses. Seeing the strength of the outbreak, administrators can consider how strong of a social response or vaccination campaign is necessary to acquire a successful mitigation of the epidemic.

For a mild influenza, it might not require much of a social response to keep it from spreading. Yet, the personal impacts are often not significant, which may leave little motivation for an individual to significantly change his behavior. For more dangerous diseases such as smallpox, there is more public concern; still this disease would also require a much stronger vaccination campaign and/or social responses than influenza to stop it. Social behavior and human interaction (contact) are not exact sciences, so the theoretical mitigation results should be considered possibilities and aspirations rather than deterministic outcomes for any rural county or town.

Summary. Starting with a survey of a rural community, demographics were analyzed and an estimation of the social contact structure was built. This network was measured and the metrics were correlated with various demographics from the survey. Through the use of an exact model of a stochastic SLIR Poisson process, we have characterized a typical influenza-like outbreak in the community and investigated vaccination strategies and social behavioral responses. When considering resource-limited vaccine distribution strategies, we identified critical locations for ethical targeting subpopulations for effective disease prevention. Our aspiration is that this analysis will be a valuable resource for both the rural community on which this study focused, and also for several similar communities in the region.

References

1. Funk S, Salathé M, Jansen VAA. Modelling the Influence of Human Behaviour on the Spread of Infectious Diseases: A Review. *Journal of the Royal Society Interface*. 2010 May 26; 1742–5662.

2. R Pastor-Satorras, A Vespignani. Immunization of complex networks. *Physical Review E*. 2002; 65(3):036104.

3. Cohen R, Havlin S, ben-Avraham D. Efficient Immunization Strategies for Computer Networks and Populations. *Physical Review Letters*. 2003 Dec; 91:247901.

4. Gallos LK, Liljeros F, Agryrakis P, Bunde A, Havlin S. Improving immunization strategies. *Physical Review E*. 2007 Apr; 75:045104R.

5. Patel R, Longini IJ, Halloran M. Finding optimal vaccination strategies for pandemic influenza using genetic algorithms. *Journal of Theoretical Biology*. 2005 May 21; 234(2):201–12.

6. Longini IMJ, Halloran ME. Strategy for Distribution of Influenza Vaccine to High-Risk Groups and Children. *American Journal of Epidemiology*. 2005; 161(4):303–6.

7. Salathé M, Jones J. Dynamics and control of diseases in networks with community structure. *PLoS Computational Biology*. 2010 Apr 8; 6(4):e1000736.

8. Glasser J, Taneri D, Feng Z, Chuang J-H, Tüll P, Thompson W, et al. Evaluation of Targeted Influenza Vaccination Strategies via Population Modeling. *PLoS ONE*. 2010; 5(9):e12777.

9. Cornforth D, Reluga T, Shim E, Bauch C, Galvani A, al. et. Erratic Flu Vaccination Emerges from Short-Sighted Behavior in Contact Networks. *PLoS Computational Biology*. 2011 Jan; 7(1):e1001062.

10. Schneider C, Mihaljev T, Havlin S, Herrmann H. Suppressing epidemics with a limited amount of immunization units. *Physical Review E*. 2011; 84:061911.

11. Araz OM, Galvani A, Meyers LA. Geographic prioritization of distributing pandemic influenza vaccines. *Health Care Management Science*. 2012 May 18; *In Press.*

12. Ferguson N. Capturing human behaviour. *Nature*. 2007 Apr 11; 446(7137):733–733.

13. Poletti P, Caprile B, Ajelli M, Pugliese A, Merler S. Spontaneous behavioural changes in response to epidemics. *Journal of Theoretical Biology*. 2009 Sep 7; 260(1):31–40.

14. Funk S, Gilad E, Watkins C, Jansen VAA. The Spread of Awareness and Its Impact on Epidemic Outbreaks. *PNAS*. 2009 Mar 30; 106(16):6425–6.

15. Perra N, Balcan D, Gonçalves B, Vespignani A. Towards a Characterization of Behavior-Disease Models. *PLoS ONE*. 2011; 6(8):e23084.

16. Sahneh FD, Scoglio C. Epidemic spread in human networks. *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), 2011*. Orlando; 2011. p. 3008 –3013.

17. Youssef M, Scoglio C. Optimal Mitigation Strategies for Epidemics in Social Networks. *Under Review*. 2012;

18. Dillman DA. Mail and Telephone Surveys: The Total Design Method. 1st ed. John Wiley & Sons; 1978.

19. Dillman DA. Mail and Internet Surveys: The Tailored Design Method. 2nd ed. Wiley; 1999.

20. Dillman DA, Smyth JD, Christian LM. Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method. 3rd ed. Wiley; 2008.

21. Smyth JD, Dillman DA, Christian LM, O'Neill AC. Using the Internet to Survey Small Towns and Communities: Limitations and Possibilities in the Early 21st Century. *American Behavioral Scientist*. 2010 May 1; 53(9):1423–48.

22. Campbell H. Surveys to study epidemics. Chanute Tribune. Chanute, Kansas; 2010 Jul 8;

23. Barrat A, Barthélemy M, Vespignani A. Dynamical Processes on Complex Networks. 1st ed. Cambridge University Press; 2008.

24. Newman M. Networks: An Introduction. 1st ed. USA: Oxford University Press; 2010.

25. Neosho County, KS - Google Maps [Internet]. [cited 2012 Jun]. Available from: http://maps.google.com/maps?hl=en&tab=wl

26. Scoglio C, Schumm W, Schumm P, Easton T, Roy Chowdhury S, Sydney A, et al. Efficient Mitigation Strategies for Epidemics in Rural Regions. *PLoS ONE*. 2010 Jul 13; 5(7):e11569.

27. Boëlle P-Y, Ansart S, Cori A, Valleron A-J. Transmission parameters of the A/H1N1 (2009) influenza virus pandemic: a review. *Influenza Other Respi Viruses*. 2011 Sep; 5(5):306–16.

28. White LF, Wallinga J, Finelli L, Reed C, Riley S, Lipsitch M, et al. Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the USA. *Influenza Other Respi Viruses*. 2009 Nov; 3(6):267–76.

29. Ghani A, Baguelin M, Griffin J, Flasche S, van Hoek AJ, Cauchemez S, et al. The Early Transmission Dynamics of H1N1pdm Influenza in the United Kingdom. *PLoS Curr* 2010 Jun 13; 1.

30. Cauchemez S, Donnelly CA, Reed C, Ghani AC, Fraser C, Kent CK, et al. Household Transmission of 2009 Pandemic Influenza A (H1N1) Virus in the United States. *New England Journal of Medicine*. 2009; 361(27):2619–27.

31. Balcan D, Hu H, Goncalves B, Bajardi P, Poletto C, Ramasco J, et al. Seasonal transmission potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility. *BMC Medicine*. 2009 Sep 10; 7(1):45.

32. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Kerkhove MDV, Hollingsworth TD, et al. Pandemic Potential of a Strain of Influenza A (H1N1): Early Findings. *Science*. 2009 Jun 19; 324(5934):1557–61.

33. Balcan D, Gonçalves B, Hu H, Ramasco JJ, Colizza V, Vespignani A. Modeling the spatial spread of infectious diseases: the GLobal Epidemic and Mobility computational model. *Journal of Computational Science*. 2010 Aug 1; 1(3):132–45.

34. Kooij RE, Schumm P, Scoglio C, Youssef M. A New Metric for Robustness with Respect to Virus Spread. *Proceedings of the 8th International IFIP-TC 6 Networking Conference*. Berlin, Heidelberg: Springer-Verlag; 2009; p. 562–72.

35. Youssef M, Kooij R, Scoglio C. Viral conductance: Quantifying the robustness of networks with respect to spread of epidemics. *Journal of Computational Science*. 2011 Aug; 2(3):286–98.

36. CDC Smallpox | Course: "Smallpox: Disease, Prevention, and Intervention" [Internet]. [cited 2012 Jun]. Available from: http://www.bt.cdc.gov/agent/smallpox/training/overview/

37. Wallinga J, Teunis P. Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures. *American Journal of Epidemiology*. 2004 Sep 15; 160(6):509–16.

38. Mills CE, Robins JM, Lipsitch M. Transmissibility of 1918 pandemic influenza. *Nature*. 2004 Dec 16; 432(7019):904–6.

39. Monto AS, Ohmit SE, Petrie JG, Johnson E, Truscon R, Teich E, et al. Comparative Efficacy of Inactivated and Live Attenuated Influenza Vaccines. *New England Journal of Medicine*. 2009; 361(13):1260–7.

40. Fielding J, Grant K, Garcia K, Kelly H. Effectiveness of Seasonal Influenza Vaccine against Pandemic (H1N1) 2009 Virus, Australia, 2010. *Emerging Infectious Diseases*. 2011 Jul; 17(7):1181–7.

41. CDC - Seasonal Influenza (Flu) - Q & A: How Well Does the Seasonal Flu Vaccine Work? [Internet]. [cited 2012 Jun]. Available from: http://www.cdc.gov/flu/about/qa/vaccineeffect.htm

42. CDC - Seasonal Influenza (Flu) - Who Should Get Vaccinated Against Influenza [Internet]. [cited 2012 Jun]. Available from: http://www.cdc.gov/flu/protect/whoshouldvax.htm