Title

A large-scale location-based social network to understanding the impact of human geo-social interaction patterns on vaccination strategies in an urbanized area

Permalink

https://escholarship.org/uc/item/11t1k5pv

Authors

Luo, Wei
Gao, Peng
Cassels, Susan

Peer reviewed

A large-scale location-based social network to understanding the impact of human geo-social interaction patterns on vaccination strategies in an urbanized area

Wei Luo[a], Peng Gao[b,*], Susan Cassels[c]

[a] School of Geographical Sciences & Urban Planning, Arizona State University, AZ, United States
[b] Department of Geography, University at Buffalo, Buffalo, NY, United States
[c] Department of Geography, University of California, Santa Barbara, CA, United States

ARTICLE INFO

ABSTRACT

Cities play an important role in fostering and amplifying the transmission of airborne diseases (e.g., influenza) because of dense human contacts. Before an outbreak of airborne diseases within a city, how to determine an appropriate containment area for effective vaccination strategies is unknown. This research treats airborne disease spreads as geo-social interaction patterns, because viruses transmit among different groups of people over geographical locations through human interactions and population movement. Previous research argued that an appropriate scale identified through human geo-social interaction patterns can provide great potential for effective vaccination. However, little work has been done to examine the effectiveness of such vaccination at large scales (e.g., city) that are characterized by spatially heterogeneous population distribution and movement. This article therefore aims to understand the impact of geo-social interaction patterns on effective vaccination in the urbanized area of Portland, Oregon. To achieve this goal, we simulate influenza transmission on a large-scale location-based social network to 1) identify human geo-social interaction patterns for designing effective vaccination strategies, and 2) and evaluate the efficacy of different vaccination strategies according to the identified geo-social patterns. The simulation results illustrate the effectiveness of vaccination strategies based on geo-social interaction patterns in containing the epidemic outbreak at the source. This research can provide evidence to inform public health approaches to determine effective scales in the design of disease control strategies.

1. Introduction

Infectious diseases (e.g., influenza) have posed great challenges to society. Dangerous infectious diseases, such as severe acute respiratory syndrome (SARS), anthrax, H1N1 flu, and Ebola pandemics, have been exacerbated by increasing human density and population mobility seen in contemporary society (Anderson et al. 2004; Fraser et al. 2009; Leung et al. 2004). Infectious disease outbreaks can spread worldwide via international transportation networks more rapidly than the distribution of vaccines (Ferguson, Fraser, Donnelly, Ghani, & Anderson 2004; Webby & Webster 2003). A key public health question is how to contain infectious diseases at the source or reduce transmission long enough to implement an effective response (Longini Jr et al. 2005).

Network science has been used to study infectious disease transmission and to design effective control (Germann, Kadau, Longini, & Macken 2006; Halloran et al. 2008; Longini Jr et al. 2005). Traditional network models are individual-centric approaches in which nodes

represent individuals and links represent physical contacts (Meyers 2007). Disease simulations based on network models primarily focus on health outcomes (e.g., the number of infected cases) (Mao & Bian 2010; Salathé & Jones 2010), but do not sufficiently consider the role of location in disease transmission and control (Zhong & Bian 2016); for example, what are the critical locations in the transmission? Are location-based control strategies more effective than individual-based control strategies? How do the spatial flows of population mobility among locations influence epidemics and control strategy design? Those are all critical public health questions in terms of spatially informed policies for infectious disease control and prevention (Bian, 2013, Gao & Bian, 2016).

To tackle the above questions, this research treats infectious disease spread as geo-social interaction patterns, in which locations interact with each other via viruses that transmit from one place to another because of human interaction and population movement (Gushulak & MacPherson 2000). Geo-social interaction patterns identified from

human mobility data within cities have the potential to provide valuable insight for designing effective epidemic control measures (Guo 2007). Luo (2016) empirically found that vaccination strategies considering geo-social interaction patterns have a high probability to contain pandemics at the source within a primary school environment. However, human interaction patterns at an urban scale, which are characterized by heterogeneous population interactions and movement in space, have not been used to inform prevention strategies. The spatial heterogeneity of population interaction and movement make the choice of appropriate containment areas for effective vaccination strategies challenging.

This research, therefore, aims to understand the impact of geo-social patterns identified from human movement and interaction data on effective vaccination in the urbanized area of Portland, Oregon. We use a location-based network framework in which nodes represent locations and links represent population flows among locations. The framework explicitly represents spatial dynamics of disease transmission from one location to another as well as design control strategies according to population flows among locations. Based on the location-based network, we simulate influenza transmission dynamics and evaluate the efficacy of different vaccination strategies according to the identified geo-social interaction patterns. Our research questions follow:

(1) Given geo-social interaction patterns, how does population distribution and movement impact the effectiveness of influenza control efforts?
(2) How does the size of the containment area impact the control effectiveness considering human geo-social interaction patterns?

The remainder of this article is organized into the following sections. The second section introduces related work in agent-based epidemic models and prevention strategies. The third section introduces location-based network model and agent-based epidemic simulation model in the urbanized area of Portland. The one section that follow present and discuss the simulation results. The last section concludes the article and discusses its implications.

1.1. Agent-based epidemic models and prevention strategies

Infectious diseases are transmitted from one individual to another through physical contacts. In many agent-based epidemic models, the concepts of networks are used to represent contacts among individuals (Bian 2004; Bian & Liebner 2007). Each individual can have attributes, such as infection status and spatial locations, whereas each link can also have a set of attributes, such as the duration of the contact and the infection rate. Based on human interaction networks, epidemiologists can use agent-based epidemic models to simulate disease transmission and design different control scenarios.

Recently, agent-based models for disease transmission at different spatial scales (e.g., city, nation) have been used to study spatio-temporal patterns of disease spread (Bian et al. 2012; Ferguson et al. 2005; Ferguson et al. 2006; Halloran et al. 2008; Mao & Bain 2011). These models predict that interactions at homes and workplaces could cause local transmission followed by a long distance transmission to an area-wide epidemic (Eubank et al. 2004). Local transmission indicates that early disease transmission is concentrated around the infection sources. Later, population movement and distribution determine the spatio-temporal spread of infection.

Prevention strategies for infectious disease can be classified into three broad categories: antiviral, vaccine, and non-pharmaceutical measures (Ferguson et al. 2006). Antiviral measures require rapid identification of early infections for treatment in order to reduce their infectiousness. Vaccination seeks to protect people from being infected by other infections via physical contacts. Non-pharmaceutical measures (e.g., case isolation, household quarantine, school or workplace closure, restrictions on travel) aim to locally contain the spread of disease

(Ferguson et al. 2005; Ferguson et al. 2006).

Previous research has shown the effectiveness of the above prevention measures in reducing disease transmission with agent-based modeling (Ferguson et al. 2005; Ferguson et al. 2006; Halloran et al. 2008), but those agent-based modelings are individual-centric without capturing explicitly spatial dynamics of the disease transmission for designing control strategy purpose. Spatial locations, one attribute of each agent, are treated as spatial stamps of the transmission process in the individual-centric agent-based modeling (Zhong & Bian 2016). Such representations suffer from a number of drawbacks regarding transmission and control design. First, the importance of locations in understanding disease transmission and control has been missing. For example, antiviral "ring chemoprophylaxis" strategies aim to geographically targeted containment via applying a certain distance threshold (e.g., 5 km) by means of prophylaxis to attempt to prevent spread to unaffected regions (Lee et al. 2010). In fact, infectious disease transmission is the mixing patterns of geographical mobility and social interactions that are much more complicated than a simple distance threshold. Second, the design of control strategies based on individual-based networks (e.g., vaccinate the individual with a large number of contacts) is challenging because it is infeasible to keep track of all social contacts of infections (Cohen, Havlin, & Ben-Avraham 2003; Gómez-Gardenes, Echenique, & Moreno 2006; Holme 2004), but the information to estimate population flows among locations based on intra- and inter-community travelers is widely available in the existing census data and travel survey reports (Mao & Bian 2010). Thus, location-based human interaction network models are required to easily represent explicitly spatial dynamics of the disease transmission and design control strategies to target certain critical locations first rather than prioritizing individuals.

Based on a location-based human interaction network model, Luo (2016) proposed a new framework in terms of an effective disease control strategy that consists of identifying geo-social interaction patterns first, following by designing effective control measures according to those patterns, and ending with control measure evaluation. The work also showed that vaccination strategies considering the geo-social interaction patterns can reduce disease transmission within a primary school environment because they can prevent disease from spreading to other classes.

Though the above research has demonstrated the importance of locations considering human geo-social interaction patterns in designing effective control scenarios, little research has been conducted to evaluate the effectiveness of such approach in a larger scale (e.g., city). Thus, this research develops an integrated two-layer location-based social network approach to investigate the impact of human geo-social interaction patterns on effective vaccination design in an urban environment.

2. Methods

We use a previously built and parameterized network from a synthetic population for a normal day in the city of Portland, Oregon, USA (Eubank et al. 2004) to construct the integrated two-layer location-based social network framework. One layer is a location-based network for the vaccination scenario design and the other layer is an individual-based social network to simulate infectious disease transmission and control to evaluate the effectiveness of those scenarios. Each is detailed next.

2.1. Study area

The synthetic population data includes over 1.5 million individuals and more than 1.5 million locations. Each synthetic individual has their demographic attributes such as income range and age, whereas each synthetic location has UTM coordinates. Based on the individual and location information, the simulated daily activity data also includes the
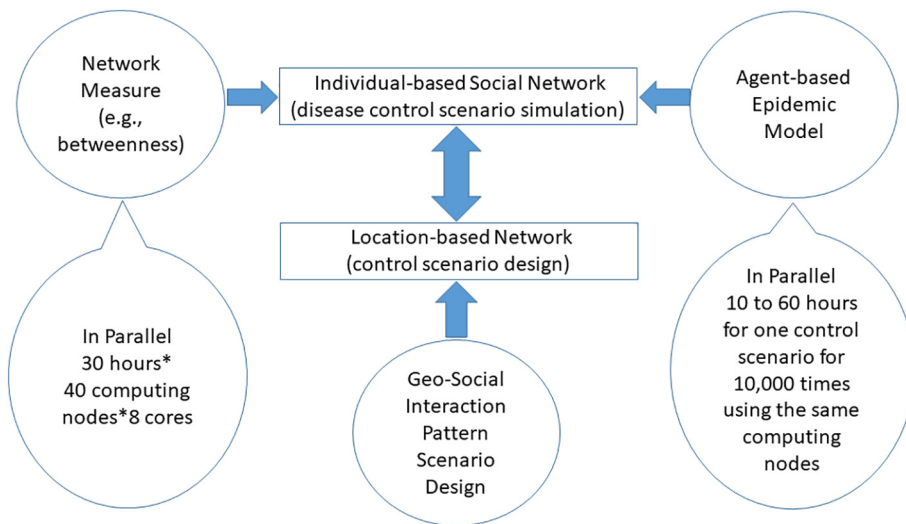
Fig. 1. Methodology framework consists of the integrated two-layer location-based social network: location-based network to design control strategy considering geo-social interaction patterns and individual-based social network to evaluate effectiveness of the above control strategy with the implementation of agent-based epidemic model. The framework is implemented in a high performance computing environment.

purpose of the activities such as work, school, and shopping, in which the starting time and duration of activity are stored in seconds (Fig. 1). The data also include the information representing dynamic person-person contact networks including contact hours, contact type, and location information. The total number of person-person contacts is approximately equal to 20 millions. The simulated data are configured with social surveys, transportation simulations, and census data in the Portland city. The data set is available at http://ndssl.vbi.vt.edu/synthetic-data/.

2.2. Location-based human interaction network

Human interactions and population mobility show spatial and temporal co-occurrence that connect discrete events over the same or different locations to facilitate disease transmission (Zhong & Bian 2016). Spatial and temporal co-occurrence is the basis of the construction of location-based human interaction network, in which nodes represent locations and edges represent population flows between locations. The total number of travel activities by individuals between two locations (population flows) is recorded as the weight of the edge (Fig. 1).The total population flows among locations are more than 6 millions.

2.3. Spatially explicit agent-based epidemic model

2.3.1. Influenza diffusion

A spatially explicit agent-based epidemic model is built on the simulated daily activities of person-person contact network in the city of Portland, Oregon, USA (Eubank et al. 2004) (Fig. 1). The influenza transmission is simulated over the contact network with each individual as one agent changing their infection states over time. Following the natural history of the influenza, each individual can take one of four statuses: susceptible, exposed, infectious, or recovered (SEIR) at a given time (Anderson & May 1992). All individuals in our model are initially considered as susceptible to influenza. After contact with an infection, susceptible individuals have a probability of infection. Those newly infected change their statuses from susceptible to exposed. Exposed individuals may progressively develop their infectious statuses, which further enable them to transmit the influenza viruses to susceptible contacts. Finally, infectious individuals recover from influenza.

First, the exposed and infectious periods are from the established literature (Table 1). Second, the infection probability is calculated based on the R_0 (the basic reproductive number). R_0 is defined as the average number of secondary cases produced by a single infection over the course of its infectious period in a completely susceptible

Table 1
Basic simulation parameters for influenza diffusion.

Parameters	Default values	Data sources
Length of exposed period	2 days	Heymann (2004)
Length of infectious period	4 days	Heymann (2004)
R_0	1.6	Mills et al. (2004) and Ferguson et al. (2005)
Infection probability	0.015	Measured based on R_0

population (Diekmann, Heesterbeek, & Metz 1990). The transmission probability is calculated in order to generate consistent values of R_0 (the basic reproductive number) corresponding to the observed R_0 of pandemic influenza (0.9–2.1) in previous studies (Ferguson et al. 2005; Mills, Robins, & Lipsitch 2004). We use a Monte–Carlo method to randomize the infection probability of individuals to derive R_0. All parameters to describe the natural history of the influenza are in Table 1.

2.3.2. Vaccination strategies

We implement three vaccination strategies: random-based, degree-based, and betweenness-based vaccination strategies over the influenza diffusion in the city of Portland, Oregon, USA. The three strategies are the most representative ones to compare the effectiveness of different vaccination scenarios in the established literature (Mao & Bian 2010; Salathé et al. 2010; Salathé & Jones 2010). The basic idea of vaccination strategies is first to rank the importance of individuals according to their network measures (e.g., degree, betweenness) and then target the individuals from the highest network measures to lowest. The random-based strategy treats all of individuals the same and randomly identifies a certain number of individuals for vaccination. The degree-based strategy prioritizes individuals who have the most contacts within the network for vaccination. The betweenness-based strategy ranks the importance of individuals according to their betweenness centrality that can capture the extent to which a particular node lies on those shortest paths that pass through the nodes (Freeman 1978).

Both degree-based and betweenness-based vaccination strategies have shown much more effective than random-based strategy. For example, betweenness-based vaccination strategy tend to prevent disease outbreak from one community to another via targeting individuals on the shortest paths, so it becomes more effective with stronger community structures (Salathé & Jones 2010). Degree-based vaccination strategy tends to show the most effectiveness with different network structures (Pastor-Satorras & Vespignani 2002; Zanette & Kuperman 2002), especially when community structure is not strong such as a

scale-free network (Albert, Jeong, & Barabási 2000). Though identifying all targeted individuals with high degree or high betweenness may not be practical, information needed to identify these individuals is becoming available (e.g., GPS, social media) (Salathé & Jones 2010).

Using the location-based human interaction network, the vaccination strategy is designed according to population flows among locations with infections and their neighboring locations. Neighboring locations that have direct connections with infected locations are most likely to be infected, followed by locations that have indirect connections with those infected locations. Thus, the vaccination strategy will target those direct neighboring locations first, followed by targeting indirect connected locations. In practice, it is much easier to target a small population size with direct neighboring locations instead of a large population size with indirect connected locations. Thus, this study explores the impact of the size of containment areas on vaccination efficacy in order to provide evidence in public health practice.

The strategy is implemented at the beginning of the influenza epidemic. The simulation will stop when there are no new infections for six consecutive days (length of exposed period + length of infectious period). We randomly picked one infectious individual as the infection source in the study area on the first day of the simulation. Based on the strategy rationale, we considered the locations where the first infection visited as the original case cluster. The original case cluster and all of the rest locations which have the population flows to the original case clusters were used to determine the first level local containment scale. The second level geo-social local containment scale was determined by all of the locations in the first level local containment area and the locations which have the population flows to the locations in the first level local containment area, so were the third, the forth, and the fifth level local geo-social containment scales. Though the location network was dynamic over time in a day, the strategy design was based on the total population flows among different locations within one day for a practice purpose. It is more practical to design vaccination strategy at a daily basis instead of at an hourly basis. In order to understand the impact of the spatially heterogeneous population distribution and mobility on the design of vaccination strategies, we randomly pick the first infection 10 times with varying total number of individuals in the first level local containment area (i.e., 1000, 5000, and 10000). The larger number of individuals within the first level local containment scale is caused by the high local population density which potentially leads to more complicated population interaction and mobility to larger geographical regions.

2.3.3. High computational complexity

Given that the vaccination results are sensitive to the number of available vaccines, we simulate results according to an increasing number of vaccines from 20,000 to 400,000 with an increment of 20,000. There is one network, three different initial infection locations with 10 times' randomness, three strategies (e.g., random-based, degree-based), four different local control regions (from the second to the fifth level local containment area), and twenty different possibilities for the number of available vaccines, yielding 7200 (1*3*10*3*4*20) combinations to simulate (Table 2). The efficacy of vaccination

Table 2
Basic simulation parameters for vaccination strategies.

Parameters	Values
Initial infection number	1
Vaccination strategies	Random-based, degree-based, and betweenness-based
Initial infection locations	1,000, 5,000, and 10,000 individuals in the first level local containment area
Containment areas	Four different geo-social local containment scales
Number of vaccines	20,000 to 400,000 with an increment of 20,000
Simulation runs	10,000

strategies for each combination is estimated for 10,000 simulation runs, resulting in a total of 72,000,000 epidemic simulation runs.

Because of the large number of individuals and locations in the study area and the massive number of simulations involved, the tasks in this research are extremely computation intensive. We tackle this issue by taking the advantage of a high performance platform of CPU clusters. Parallel algorithms are developed to deal with the most computation intensive tasks: 1) the calculation of the betweenness centrality and 2) the simulation of influenza epidemics. For the former, it is widely understood that the definitions of betweenness centrality in very large networks have inherently high computational complexity. The parallel algorithm is adapted from the Brande's algorithm (Brandes 2001), and the actual running time is about 30 hours using 40 computing nodes, each with 8 cores and 16 GB of RAM memory. For the latter, the parallel algorithm is developed following the SEIR rules as discussed earlier, and the running time ranges from 10 to 60 hours for a given first infection and vaccination strategy using the same computing nodes. The parallel design does not only reduce the running time but also improve the scalability of this methodology framework, which are essential for analyzing extremely large network data.

2.4. Simulation results and discussions

We use two metrics to measure the effectiveness of vaccination strategies in the city of Portland: the number of infections and the success rate of local containment. For the first metric, a better vaccination strategy is expected to generate a lower number of infected cases. For the second metric, we measure how well we can contain the outbreak at the source. If the number of infections outside of the containment areas identified by geo-social interaction patterns is zero, we consider the local containment successful. Otherwise we consider the local containment a failure.

2.5. Vaccination effectiveness in terms of infection percentage and local containment success rate

Fig. 2 and Fig. 3 show the containment effectiveness of three typical vaccination strategies in which the first initial infections occurring within the first level local containment scale with approximately 5,000 individuals in the city of Portland according to the two metrics: the number of infections and the success rate of local containment. Fig. 2 illustrates that an increasing number of vaccines can produce a decreasing number of infections. The random-based vaccination strategy generates the largest number of infections, followed by the degree-based and betweenness-based vaccination strategies. The latter two vaccination strategies produce the similar control efficacy in which betweenness-based ones lead to a relatively faster decreasing number of infections. The explanation of such patterns is illustrated in the Fig. 3. Fig. 3 shows that an increasing number of vaccines can produce a decreasing number of local containment failure rates. In other words, betweenness-based vaccination strategies can contain disease outbreak at the local geo-social scales more successfully than the other two vaccination strategies. One reason is that betweenness-based vaccination strategies target high risk individuals who lie on the bridge among different communities, so they can control epidemics more effective through preventing epidemic outbreak from the local geo-social containment areas to outside areas.

From Fig. 2 (a-d), there is an obvious trend that increasing geo-social scales of local containment areas result in an increasing number of infections. Fig. 3 suggests that increasing extents of local geo-social containment areas lead to difficulty in containing outbreaks at the local scale. For example, in order to successfully contain the disease locally larger than 90% with degree-based or betweenness-based approach, it cost 60,000 vaccines in Fig. 3 (a), 180,000 vaccines in Fig. 3 (b), 200,000 vaccines in Fig. 3 (c) and (d). In addition to supporting the importance of reliable detection of early infections to contain outbreaks
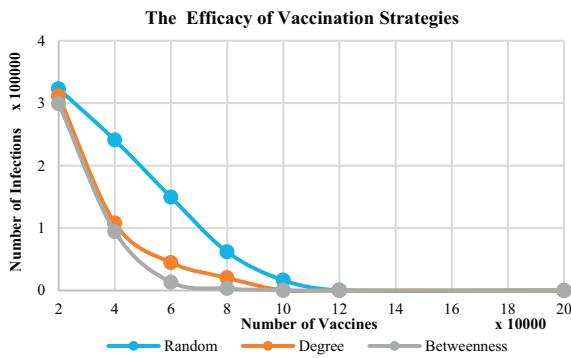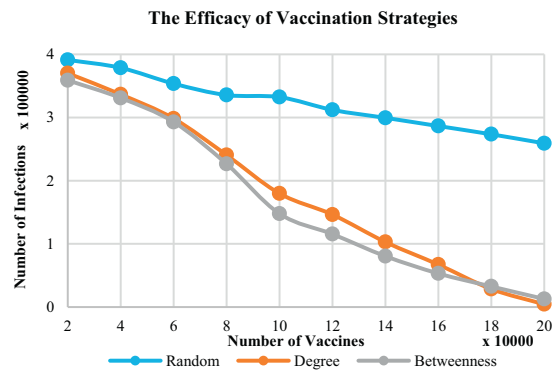
(a) The second level local containment scale
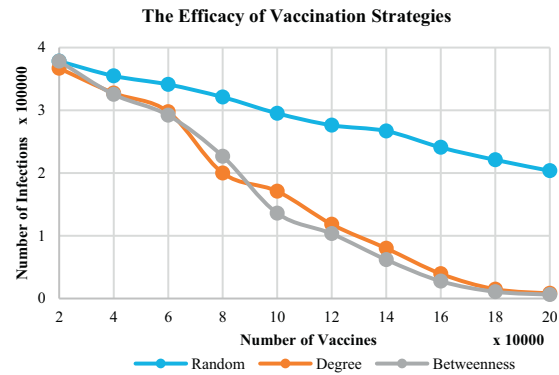
(b) The third level local containment scale
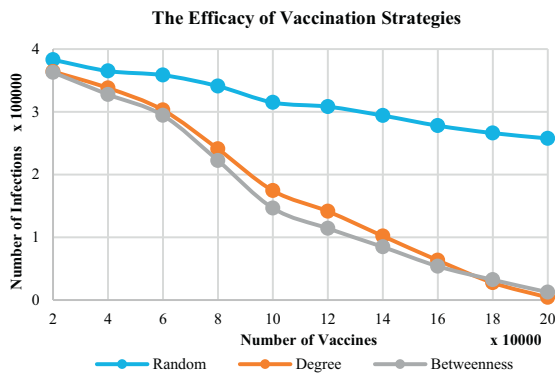
(c) The forth level local containment scale

(d) The fifth level local containment scale

Fig. 2. The efficacy of three vaccination strategies with the first initial infection having the total number of individuals approximately equal to 5000 within the first level local containment scale for 10,000 simulation runs in the city of Portland. The *Y* axis represents the average number of infection for 10,000 simulation runs, and the *X* axis represents the number of vaccines. Three vaccination strategies are represented by three curves in different colors, as the legend shows at the bottom. (a), (b), (c), and (d) represent increasing geo-social scales of local containment areas.

at the source, such patterns also illustrate that identifying an appropriate scale for the local containment is key for designing effective prevention strategies. Fig. 2 (c) and (d) show that when the local containment areas are larger than certain extent, there is no obvious difference in terms of containment efficacy. It is because that the number of available vaccines cannot effectively contain disease outbreak at the local scale in Fig. 3 (c) and (d).

2.6. Vaccination effectiveness given spatially heterogeneous population distribution and movement

This study further explores the impact of spatially heterogeneous population distribution and mobility on the containment efficacy of local geo-social vaccination strategies. This section only displays the simulation results with the random-based vaccination strategies, because the other two strategies display the similar patterns. Fig. 4 and Fig. 5 show the results of the two metrics: the number of infections and the success rate of local containment, respectively. They both display consistent patterns: the initial infections occurring in the higher population density areas can lead to higher number of infections (Fig. 4) because they increase the local containment failure rates with the same number of vaccines (Fig. 5). For example, the disease transmission caused by the first infection from locations with a medium population density can areas be confined locally at a very high percentage (i.e., 80 %) with less than 100,000 infections using 350,000 vaccines (Fig. 4b and 5b), whereas the same amount of vaccines can only successfully contain the disease outbreak from high population density locations less than 20 percent with more than 300,000 infections. Such patterns

imply that we need more vaccine available in order to contain disease outbreak from critical locations with high human interaction density and large numbers of population flows from infected locations to others. Thus, it is essential to place sensors in the hubs of the locations to allow highly efficient outbreak detection.

Fig. 4 and 5 show that increasing scales of local geo-social containment areas lead to decreasing efficacy of the same amount of vaccines. For example, the random-based control scenario with more than 150,000 vaccines can successfully contain the disease outbreak in a high population density area (Fig. 4 a), but it generates less than 600,000 infections (Fig. 4 b) and more than 600,000 infections (Fig. 4 c and d). From the both (c) and (d) of the Fig. 4 and 5, we can tell that there is no obvious difference in terms of containment efficacy when the local containment areas are larger than certain scales in both (c) and (d). One explanation is that a certain number of vaccines have its threshold of susceptible population pool. The number of local control failure increases after the susceptible population pool goes beyond the threshold with the increasing local geo-social containment areas. When the disease outbreak occurs outside of the containment areas in the most simulation runs, the number of total infections finally reach similar amounts. We also notice that it is unlikely to confine the disease outbreaks caused by the first infection from high population density areas, if the local geo-social containment areas are set too large (Fig. 5c and 5d). When the first infection is from low population density areas, the likelihood to contain disease outbreaks within a large local geo-social containment areas is still very high (Fig. 5c and 5d). Thus, we can reach the conclusion that an appropriate geo-social scale for the local containment and the population density where the first infection
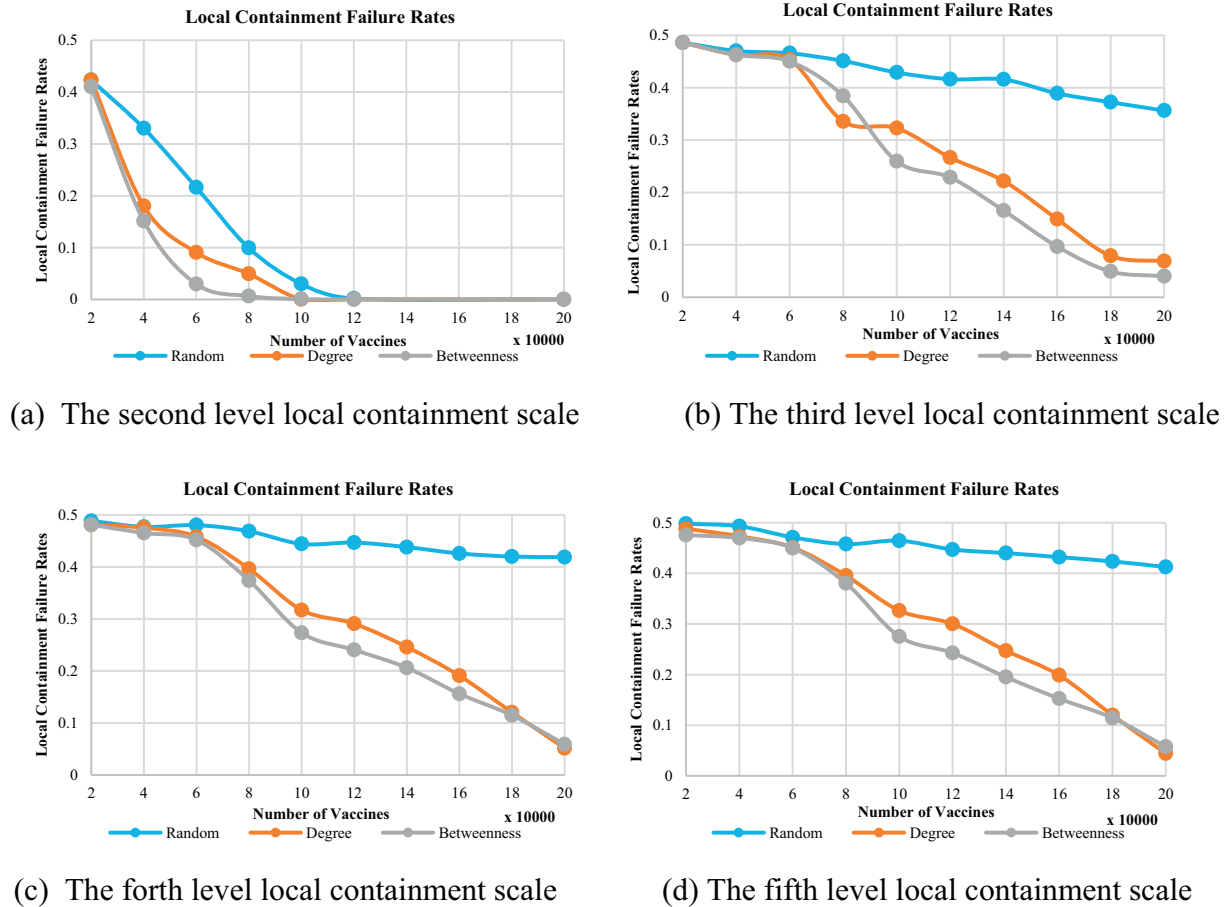
(a) The second level local containment scale

(b) The third level local containment scale

(c) The forth level local containment scale

(d) The fifth level local containment scale

Fig. 3. The local containment failure rates of three vaccination strategies with the first initial infections having the total number of individuals approximately equal to 5000 within the first level local containment scale for 10,000 simulation runs in the city of Portland. The *Y* axis represents the local containment failure rates with 10,000 simulation runs, and the *X* axis represents the number of vaccines. Three vaccination strategies are represented by three curves in different colors, as the legend shows at the bottom. (a), (b), (c), and (d) represent increasing geo-social scales of local containment.

resides are two keys for prevention strategies to achieve their maximum control effectiveness.

### 2.7. Spatial effectiveness of vaccination strategies

Based on the comparison analysis above, broader scales of local geo-social containment areas lead to decreasing efficacy of the same amount of vaccines. We further examine and compare their spatial effectiveness through infection density maps (Fig. 6). We choose degree-based vaccination strategies with the medium number of individuals (approximately equal to 5000) within the first level local containment scale. We pick the number of vaccines equal to 100,000 for the second level local containment scale (Fig. 6a) and the number of vaccines equal to 180,000 for the third, forth, and fifth level local containment scales (Fig. 6 b, c, and d) because they all have the local containment failure rates lower than 10% (Fig. 5).

Fig. 6 (a) induces an extremely low intensity of infections in the whole study area with a relatively higher intensity of infections in the central business district of the study area. It is caused by the densest residential population and business locations in the central business district. Compared to the Fig. 6 (a), Fig. 6 (b) with a larger containment area greatly increases the infection intensities with an additional dense area: the city of Vancouver on the north bank of the Columbia River. Vancouver is the largest suburb of Portland, Oregon with highly concentrated population density. The spatial effectiveness of Fig. 6 (c) and (d) lead to the wide spread of influenza over the study area with three of highest infection density areas. The similar spatial patterns and total

infections in both Fig. 6 (c) and (d) indicate that a certain number of vaccinations have its upper limit of susceptible population pool. When the local containment scale is larger than the upper limit, there is no difference in terms of spatial effectiveness. The spatial effectiveness of Fig. 6 highlights the importance of early detection of infections with an appropriate local containment scale, which is capable of confining the wide spread of influenza over the study area.

2.8. Conclusions and Implications

This research examined the impact of geo-social interaction patterns on effective vaccination strategies in the urbanized area of Portland, Oregon. To achieve this goal, we built a large-scale two-layer location-based social network model including an innovative location-based network to design vaccination strategies and an individual-based spatially explicit disease model to evaluate the efficacy of the above vaccination strategies. We implemented the network model in parallel algorithms to take advantage of a high performance platform of CPU to tackle the challenge of extremely intensive computation with millions of nodes and edges. The simulation results suggested that geo-social interaction patterns can be used to design effective vaccination strategies to contain epidemic outbreaks at the source. Increasing extents of local geo-social containment areas lead to decreasing control effectiveness because it becomes challenging to confine the spread of influenza within local containment areas. Especially when the local containment scale is larger than the upper limit, there is little difference in terms of spatial effectiveness. Vaccination designs should consider
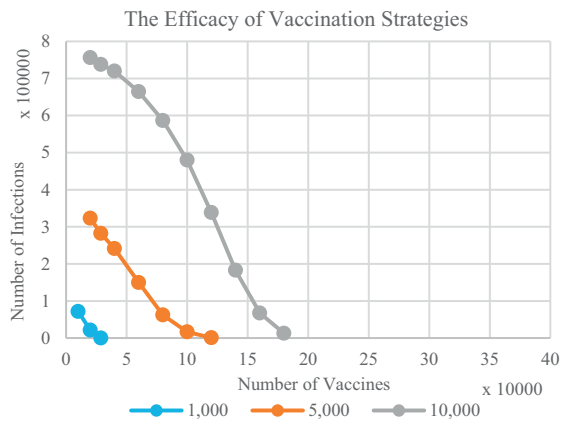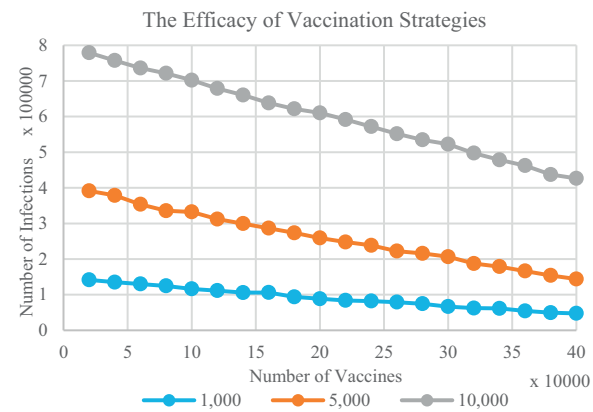
(a) The second level local containment scale
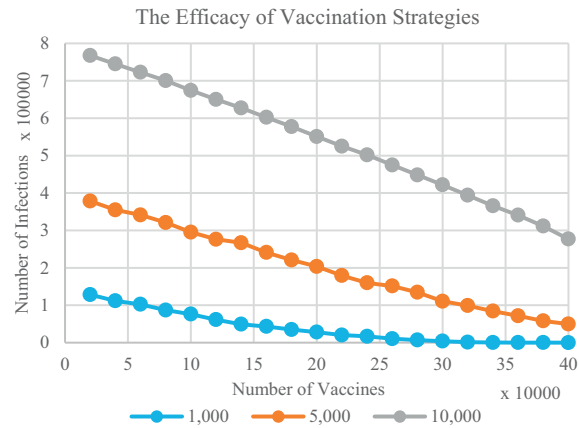
(b) The third level local containment scale
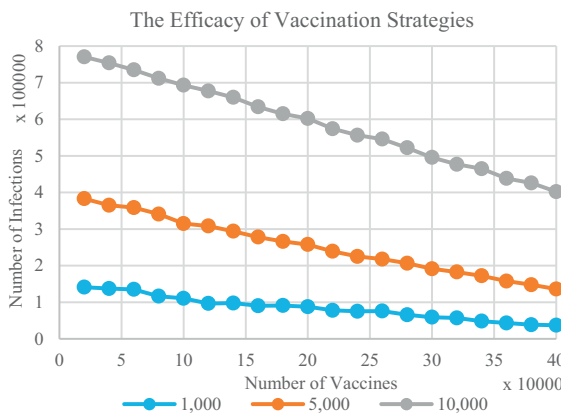
(c) The forth level local containment scale
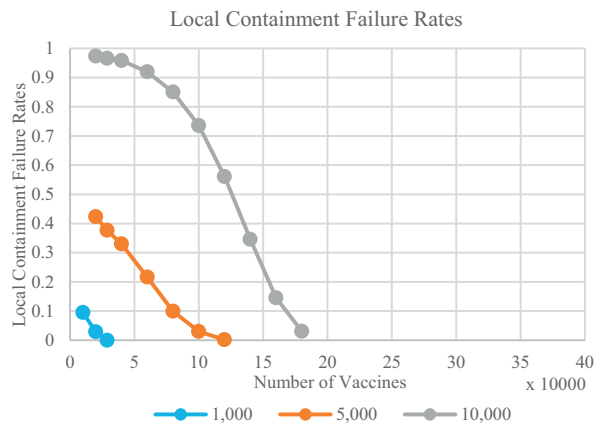

(d) The fifth level local containment scale

Fig. 4. The efficacy of random-based vaccination strategies with the first initial infection having the three different number of individuals (approximately equal to 1000, 5000, 10000) within the first level local containment scale for 10,000 simulation runs in the city of Portland. The *Y* axis represents the average number of infection for 10,000 simulation runs, and the *X* axis represents the number of vaccines. The first initial infections with different number of individuals are represented by three curves in different colors as the legend shows at the bottom. The blue curve indicates a lower number of individuals: 1000. The orange curve indicates a medium number of individuals: 5000. The grey curve indicates a larger number of individuals: 10000. (a) (b), (c), and (d) show increasing geo-social scales of local containment areas.

spatially heterogeneous population distribution and movement in an urban area. It is possible to confine disease outbreaks for early infections in the low population areas, even if local geo-social containment areas have been set too large. It is unlikely to confine disease outbreaks for early infections in the high population areas if local geo-social containment areas are set too large. This study highlighted the importance of identifying the population density for the early infections and an appropriate geo-social scale for local containment in order to achieve maximum control effectiveness with a limited number of vaccines.

A successful vaccine strategy must meet a number of key criteria: (1) early detection of the original case cluster, (2) rapid delivery of treatment to targeted groups, and (3) effective delivery of treatment to high risk individuals (Eubank et al. 2004; Ferguson et al. 2005). The identified geo-social interaction patterns and their effectiveness in designing vaccination strategies provide valuable insights for helping meet the above challenges. They can help identify critical individuals, locations, and clusters of locations for disease control purposes. Traditional prevention strategies, including ring vaccination and contact tracing, can only capture one aspect of infectious disease transmission (i.e., spatial, or social) and both of which have significant limitations. "Contact tracing" strategies require the identification of people who

may have had contact with infectious individuals (Germann et al. 2006), which is difficult to implement in real situations considering the limited resources and time before the disease outbreak. "Ring vaccination" with a simple distance threshold (e.g., 5km) (Lee et al. 2010) is easy to implement, but it may fail to capture the most likely and complicated disease transmission processes. Many research have demonstrated that ring vaccination alone might not lead to containment of disease outbreak with a high R_0 at the source (Kucharski et al. 2016; Wells et al. 2015), but a combined intervention of contact tracing and ring vaccination might contribute increase the ability to contain outbreaks of emerging infectious disease threats (Kretzschmar, Van den Hof, Wallinga, & Van Wijngaarden 2004; Longini Jr et al. 2005; Merler et al. 2016; Diao and Wang, 2014; Diao and Wang, 2016). The proposed geo-social interaction patterns cannot only address the inherent challenges of contact tracing to identify contacts with infections, but also capture the mixed interactions of the social and spatial relationships among individuals that determine infectious disease transmission (Luo & MacEachren 2014).

In addition to taking advantages from both contact tracing and ring vaccination strategies, the proposed location-based network approach provides the potential solution to address the challenge that it is infeasible to identify individuals with high degree or high betweenness
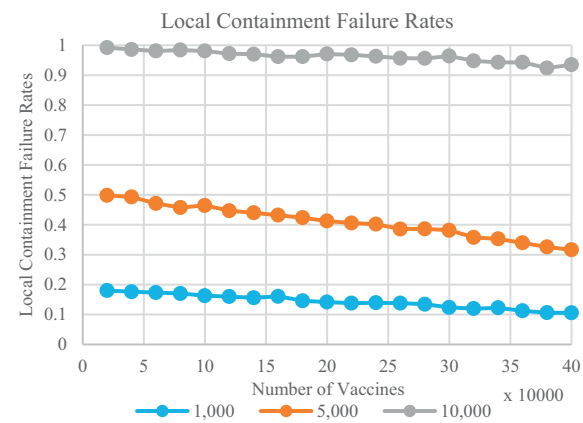
(a) The second level local containment scale
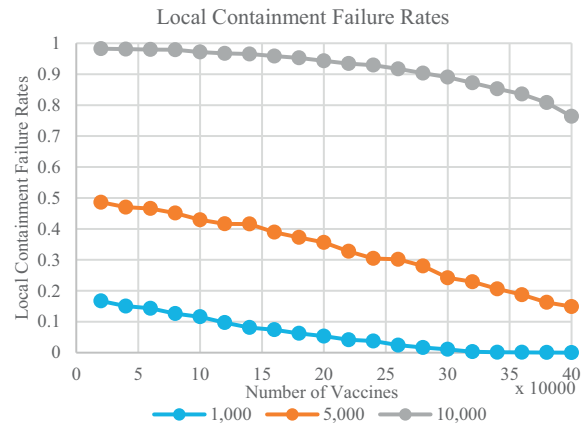
(b) The third level local containment scale
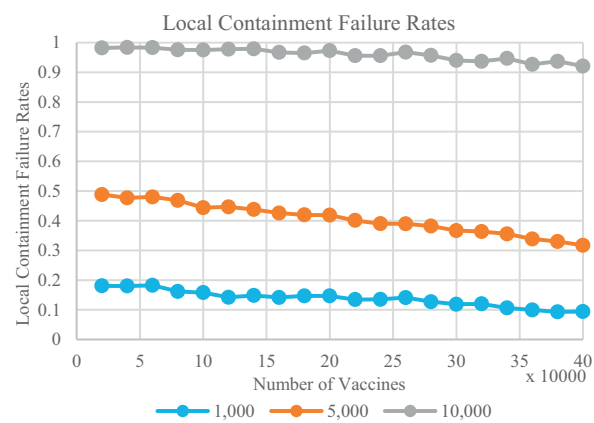
(c) The forth level local containment scale

(d) The fifth level local containment scale

Fig. 5. The local containment failure rates of random-based vaccination strategies with the first initial infection having the three different number of individuals (approximately equal to 1000, 5000, 10000) within the first level local containment scale for 10,000 simulation runs in the city of Portland. The Y axis represents the percentage of the local containment failure with 10,000 simulation runs, and the X axis represents the number of vaccines. The first initial infections with different number of individuals are represented by three curves in different colors as the legend shows at the bottom. The blue curve indicates a lower number of individuals: 1000. The orange curve indicates a medium number of individuals: 5000. The grey curve indicates a larger number of individuals: 10000. (a) (b), (c), and (d) show increasing geo-social scales of local containment areas.

for the disease control purpose in critical and timely situations. The wide available data (e.g., sensor networks, twitter) can provide more reliable estimation on the critical locations than critical individuals with high degree or high betweenness. Thus, public health workers should focus on those critical locations with their close neighbor locations first in a timely manner.

The proposed location-based social network framework also suggests several directions for future research. First, the estimation of population flows among different locations is still based on the existing census data and travel survey reports. Gao et al. (2014) found that large-scale social media data can provide reliable estimates of regional origin-destination trips on weekdays compared with the community survey data in greater Los Angeles area. It shows great promise to build location-based social networks with large-scale social media data to provide real-time estimation and prevention strategy design in a timely manner. Second, the location-based social network framework is built on population interaction and movement patterns in a normal situation, but people may adjust their behaviors accordingly during an epidemic outbreak. Thus, human preventive behaviors caused by inter-personal influence (Mao & Yang 2012) and mass media could be incorporated to refine the framework. Third, this research focuses on the impact of vaccination strategies on geo-social interactions patterns, but has not

evaluated such impact of other prevention strategies such as travel restrictions and case isolation. Forth, this research also has implications for the control of other location-based spreading pheonmena, such as invasive species (Diao & Wang 2016, Diao & Wang 2018), and HIV transmission (Luo et.al 2018).

To summarize, the proposed location-based social network framework can help geographically optimize the design of prevention strategies before epidemic outbreaks, especially for novel viruses with limited resources and time constraints. To understand advantages and disadvantages of the framework in a full picture, it is crucial to apply it for other regions with different population distributions and movements as well as epidemiological settings.

Competing interests

The author declares that he has no competing interests.
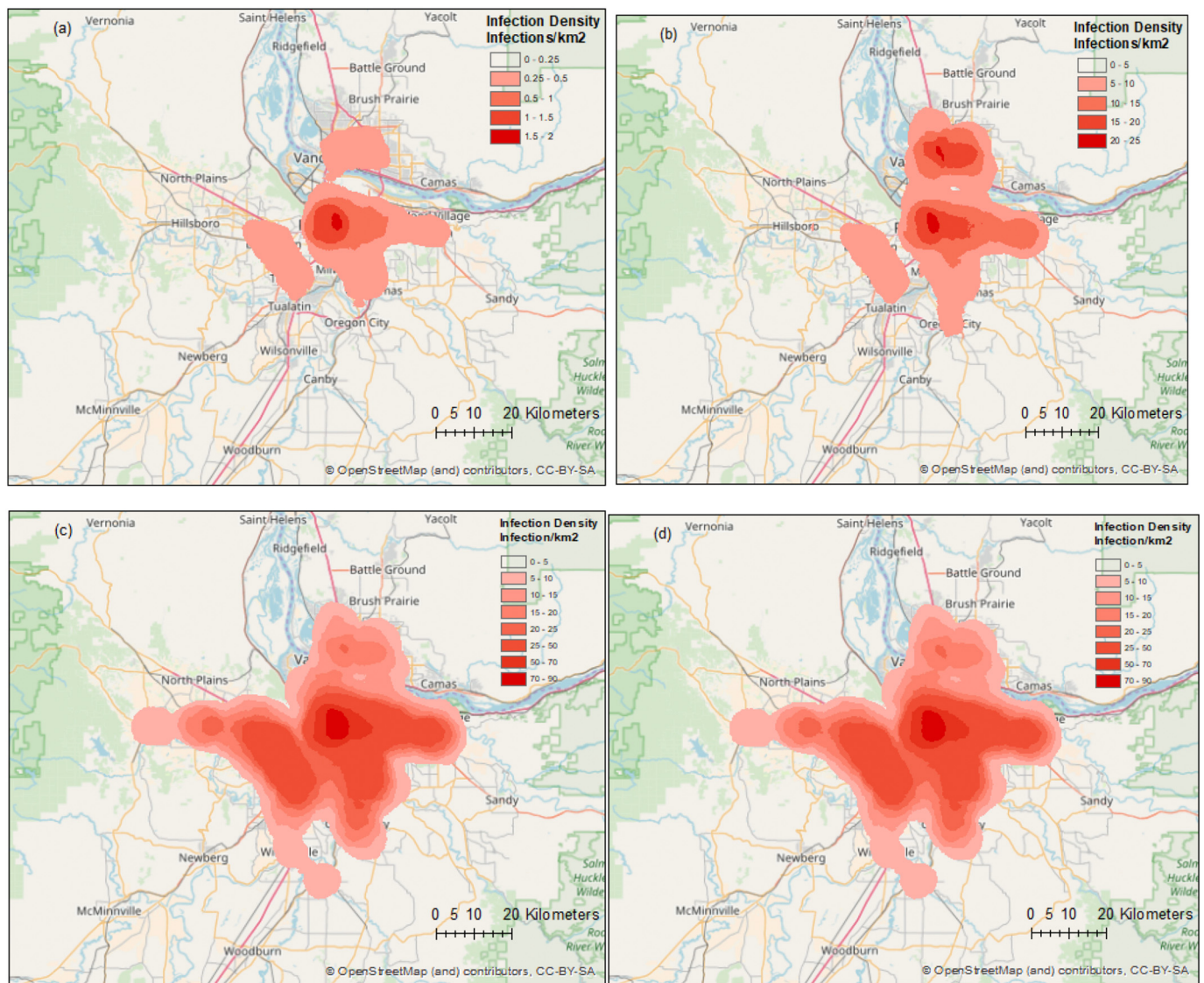
Funding

Fig. 6. Cumulative infections for the entire epidemic with the first initial infection having the medium number of individuals (approximately equal to 5000) within the first level local containment scale for 10,000 simulation runs. (a) Second level local containment scale with the number of vaccines equal to 100,000. (b) Third level local containment scale with the number of vaccines equal to 200,000. (c) Forth level local containment scale with the number of vaccines equal to 200,000. (d) Fifth level local containment scale with the number of vaccines equal to 200,000. The color ramp represents the cumulative number of infections per sq km2. The maps have three different legend scales.

References

Albert, R., Jeong, H., & Barabási, A. L. (2000). Error and attack tolerance of complex networks. *Nature, 406*, 378–382.

Anderson, R., Fraser, C., Ghani, A., Donnelly, C., Riley, S., Ferguson, N., ... Hedley, A. (2004). Epidemiology, transmission dynamics and control of SARS: the 2002-2003 epidemic. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 359*, 1091–1105.

Anderson, R., & May, R. (1992). *Infectious diseases of humans: dynamics and control.* USA: Oxford University Press.

Bian, L. (2013). Spatial approaches to modeling dispersion of communicable diseases–a review. *Transactions in GIS, 17*, 1–17.

Bian, L. (2004). A conceptual framework for an individual-based spatially explicit epidemiological model. *Environment and Planning B, 31*, 381–396.

Bian, L., Huang, Y., Mao, L., Lim, E., Lee, G., Yang, Y., ... Wilson, D. (2012). Modeling individual vulnerability to communicable diseases: A framework and design. *Annals of the Association of American Geographers, 102*, 1016–1025.

Bian, L., & Liebner, D. (2007). A network model for dispersion of communicable diseases. *Transactions in GIS, 11*, 155–173.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology, 25*, 163–177.

Cohen, R., Havlin, S., & Ben-Avraham, D. (2003). Efficient immunization strategies for computer networks and populations. *Physical Review Letters, 91*, 247901.

Diao, C., & Wang, L. (2014). Development of an invasive species distribution model with fine resolution remote sensing. *International Journal of Applied Earth Observation and Geoinformation, 30*, 65–785.

Diao, C., & Wang, L. (2016). Temporal partial unmixing of exotic saltcedar using Landsat time series. *Remote Sensing Letters, 7*, 466–475.

Diao, C., & Wang, L. (2016). Incorporating plant phenological trajectory in exotic salt-cedar detection with monthly time series of Landsat imagery. *Remote Sensing of Environment, 182*, 60–71.

Diao, C., & Wang, L. (2018). Landsat time series-based multiyear spectral angle clustering (MSAC) model to monitor the inter-annual leaf senescence of exotic saltcedar. *Remote Sensing of Environment, 209*, 581–593.

Diekmann, O., Heesterbeek, J., & Metz, J. A. (1990). On the definition and the computation of the basic reproduction ratio R 0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology, 28*, 365–382.

Eubank, S., Guclu, H., Anil Kumar, V., Marathe, M., Srinivasan, A., Toroczkai, Z., & Wang,

N. (2004). Modelling disease outbreaks in realistic urban social networks. *Nature, 429*, 180–184.

Ferguson, N., Cummings, D., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., ... Burke, D. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature, 437*, 209–214.

Ferguson, N., Cummings, D., Fraser, C., Cajka, J., Cooley, P., & Burke, D. (2006). Strategies for mitigating an influenza pandemic. *Nature, 442*, 448–452.

Ferguson, N. M., Fraser, C., Donnelly, C. A., Ghani, A. C., & Anderson, R. M. (2004). Public health risk from the avian H5N1 influenza epidemic. *Science, 304*, 968–969.

Fraser, C., Donnelly, C., Cauchemez, S., Hanage, W., Van Kerkhove, M., Hollingsworth, T., ... Lyons, E. (2009). Pandemic potential of a strain of influenza A (H1N1): early findings. *Science, 324*, 1557–1561.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks, 1*, 215–239.

Gao, S., Yang, J.-A., Yan, B., Hu, Y., Janowicz, K., & McKenzie, G. (2014). Detecting origin-destination mobility flows from geotagged Tweets in greater Los Angeles area. *Eighth international conference on geographic information science (GIScience'14)*.

Gao, P., & Bian, L. (2016). Scale effects on spatially embedded contact networks, Computers. *Environment and Urban Systems, 59*, 142–151.

Germann, T. C., Kadau, K., Longini, I. M., & Macken, C. A. (2006). Mitigation strategies for pandemic influenza in the United States. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 5935–5940.

Gómez-Gardenes, J., Echenique, P., & Moreno, Y. (2006). Immunization of real complex communication networks. *The European Physical Journal B-Condensed Matter and Complex Systems, 49*, 259–264.

Guo, D. (2007). Visual analytics of spatial interaction patterns for pandemic decision support. *International Journal of Geographical Information Science, 21*, 859–877.

Gushulak, B. D., & MacPherson, D. W. (2000). Population mobility and infectious diseases: the diminishing impact of classical infectious diseases and new approaches for the 21st century. *Clinical Infectious Diseases, 31*, 776–780.

Halloran, M. E., Ferguson, N. M., Eubank, S., Longini, I. M., Cummings, D. A. T., Lewis, B., ... Germann, T. C. (2008). Modeling targeted layered containment of an influenza pandemic in the United States. *Proceedings of the National Academy of Sciences, 105*, 4639–4644.

Heymann, D. (2004). *Control of communicable diseases manual* (18th ed.). Washington, DC: American Public Health Associationhttp://dx.doi.org/10.2105/CCDM.2745.

Holme, P. (2004). Efficient local strategies for vaccination and network attack. *EPL (Europhysics Letters), 68*, 908.

Kretzschmar, M., Van den Hof, S., Wallinga, J., & Van Wijngaarden, J. (2004). Ring vaccination and smallpox control. *Emerging Infectious Diseases, 10*, 832.

Kucharski, A. J., Eggo, R. M., Watson, C. H., Camacho, A., Funk, S., & Edmunds, W. J. (2016). Effectiveness of ring vaccination as control strategy for Ebola virus disease. *Emerging Infectious Diseases, 22*, 105.

Lee, V. J., Yap, J., Cook, A. R., Chen, M. I., Tay, J. K., Tan, B. H., ... Lin, R. (2010). Oseltamivir ring prophylaxis for containment of 2009 H1N1 influenza outbreaks. *New England Journal of Medicine, 362*, 2166–2174.

Leung, G., Hedley, A., Ho, L., Chau, P., Wong, I., Thach, T., ... Riley, S. (2004). The epidemiology of severe acute respiratory syndrome in the 2003 Hong Kong epidemic: an analysis of all 1755 patients. *Annals of Internal Medicine, 141*, 662–673.

Longini, I. M., Jr., Nizam, A., Xu, S., Ungchusak, K., Hanshaoworakul, W., Cummings, D. A. T., & Halloran, M. E. (2005). Containing pandemic influenza at the source. *Science, 309*, 1083–1087.

Luo, W. (2016). Visual analytics of geo-social interaction patterns for epidemic control. *International Journal of Health Geographics, 15*, 1–16.

Luo, W., Katz, D. A., Hamilton, D. T., McKenney, J., Jenness, S. M., Goodreau, S. M., Stekler, J. D., Rosenberg, E. S., Sullivan, P. S., & Cassels, S. (2018). Development of an agent-based model to investigate the impact of HIV self-testing programs on men who have sex with men in Atlanta and Seattle. *JMIR Public Health and Surveillance. (forthcoming), 4*(2), e58. http://dx.doi.org/10.2196/publichealth.9357.

Luo, W., & MacEachren, A. M. (2014). Geo-social visual analytics. *Journal of spatial information science*, 27–66.

Mao, L., & Bian, L. (2010). A Dynamic Network with Individual Mobility for Designing Vaccination Strategies. *Transactions in GIS, 14*, 533–545.

Mao, L., & Yang, Y. (2012). Coupling infectious diseases, human preventive behavior, and networks–A conceptual framework for epidemic modeling. *Social Science & Medicine, 74*, 167–175.

Mao, L., & Bain, L. (2011). Agent-based simulation for a dual-diffusion process of influenza and human preventive behavior. *International Journal of Geographical Information Science, 25*, 1371–1388.

Merler, S., Ajelli, M., Fumanelli, L., Parlamento, S., Piontti, A. P. Y., Dean, N. E., ... Halloran, M. E. (2016). Containing Ebola at the source with ring vaccination. *PLoS Neglected Tropical Diseases, 10*, e0005093.

Meyers, L. (2007). Contact network epidemiology: Bond percolation applied to infectious disease prediction and control. *Bulletin of the American Mathematical Society, 44*, 63–86.

Mills, C. E., Robins, J. M., & Lipsitch, M. (2004). Transmissibility of 1918 pandemic influenza. *Nature, 432*, 904–906.

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

Salathé, M., & Jones, J. (2010). Dynamics and control of diseases in networks with community structure. *PLoS Computational Biology, 6*, e1000736.

Salathé, M., Kazandjieva, M., Lee, J. W., Levis, P., Feldman, M. W., & Jones, J. H. (2010). A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences, 107*, 22020–22025.

Webby, R. J., & Webster, R. G. (2003). Are we ready for pandemic influenza? *Science, 302*, 1519–1522.

Wells, C., Yamin, D., Ndeffo-Mbah, M. L., Wenzel, N., Gaffney, S. G., Townsend, J. P., ... Altice, F. L. (2015). Harnessing case isolation and ring vaccination to control Ebola. *PLoS Neglected Tropical Diseases, 9*, e0003794.

Zanette, D. H., & Kuperman, M. (2002). Effects of immunization in small-world epidemics. *Physica A: Statistical Mechanics and its Applications, 309*, 445–452.

Zhong, S., & Bian, L. (2016). A location-centric network approach to analyzing epidemic dynamics. *Annals of the American Association of Geographers, 106*, 480–488.