Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Modeling triple-diffusions of infectious diseases, information, and preventive behaviors through a metropolitan social network—An agent-based simulation

Liang Mao*

Department of Geography, University of Florida, P.O. Box 117315, 3142 Turlington Hall, Gainesville, FL 32611, USA

Keywords:
Infectious diseases
Preventive behavior
Information diffusion
Social network
Agent-based simulation
GIS

A B S T R A C T

A typical epidemic often involves the transmission of a disease, the flow of information regarding the disease, and the spread of human preventive behaviors against the disease. These three processes diffuse simultaneously through human social networks, and interact with one another, forming negative and positive feedback loops in the complex human-disease systems. Few studies, however, have been devoted to coupling all the three diffusions together and representing their interactions. To fill the knowledge gap, this article proposes a spatially explicit agent-based model to simulate a triple-diffusion process in a metropolitan area of 1 million people. The individual-based approach, network model, behavioral theories, and stochastic processes are used to formulate the three diffusions and integrate them together. Compared to the observed facts, the model results reasonably replicate the trends of influenza spread and information propagation. The model thus could be a valid and effective tool to evaluate information/behavior-based intervention strategies. Besides its implications to the public health, the research findings also contribute to network modeling, systems science, and medical geography.

© 2014 Elsevier Ltd. All rights reserved.

Introduction

Recent outbreaks of infectious diseases, such as the H1N1 flu, bird flu, and severe acute respiratory syndrome (SARS), have brought images of empty streets and people wearing face masks to television screens and web pages, as fear of unknown diseases swept around the globe (Funk, Salathé, & Jansen, 2010). These images depict three basic components of epidemics, namely infectious diseases, information about diseases, and human preventive behavior against diseases. From a perspective of diffusion theory, each of the three components can be viewed as a spreading process throughout a population. The disease could be transmitted through person-to-person contact, the information is circulated by communication channels, and the preventive behavior can spread via the ‘social contagion’ process, such as the observational learning. The interactions among these three diffusion processes shape the scale and dynamics of epidemics (Funk & Jansen, 2013; Lau et al., 2005; Mao & Yang, 2011).

Mathematical and computational models have been extensively used by health policy makers to predict and control disease epidemics. A majority of existing models have been focused on the diffusion of diseases alone, assuming a ‘passive’ population that would not respond to diseases (Bian et al., 2012; Eubank et al., 2004; Longini, Halloran, Nizam, & Yang, 2004). This is rarely the case because it is natural for people to protect themselves when realizing disease risks (Eames, Tilston, Brooks-Pollock, & Edmunds, 2012; Ferguson, 2007). To improve, there has been much recent interest in modeling two diffusion processes in an epidemic, either a behavior-disease diffusion (House, 2011; Mao & Bian, 2011; Vardavas, Breban, & Blower, 2007), or an information (awareness)-disease diffusion (Funk, Gilad, Watkins, & Jansen, 2009; Kiss, Cassell, Recker, & Simon, 2010). These ‘dual-diffusion’ models have made a remarkable progress toward the reality, but none of them consider all the three diffusion processes together. The third diffusion process has often been neglected or simplified.

In the current literature, few modeling efforts have been devoted to explicitly representing all the three components, their spreading processes, and interactions. The lack of such models prevents researchers from unveiling a full picture of an epidemic, and inevitably introduces biases into the deep understanding on
human-disease systems. For epidemiologists, it is of difficulty to explore how one diffusion process influences the other two, and what key factors govern the three diffusion processes. Without a complete model, health policy makers would not be able to systematically evaluate social-network interventions for disease control, such as mass-media campaigns and behavior promotion strategies. As in the age of information, the fusion of disease-behavior-information in epidemic modeling becomes a pressing task in public health.

To fill the knowledge deficit, this research proposes a conceptual framework to integrate the three diffusion processes, and develops a triple-diffusion model in a realistic urban area. Following sections discuss the conceptualization, formulation, and implementation of the model, and evaluate the simulation results.

Methods and materials

Conceptual framework

The proposed model conceptualizes a typical epidemic as one network structure, three paralleling diffusion processes, and three external factors (Fig. 1). First, the contacts among individuals form a network structure as a basis for diffusion and interaction. Second, infectious diseases are transmitted through direct contacts among individuals (the middle layer). Disease control strategies, such as vaccination program, case treatment and isolation, pose external effects on the disease diffusion. Third, the diffusion of diseases prompts the “word-of-mouth” discussion among individuals, which disseminates the information concerning diseases and prevention (the upper layer). The outbreak of diseases may also stimulate various mass media, such as TV, newspapers, and radio, to propagate relevant information, thus accelerating the diffusion of information. Fourth, people being informed start to consider and make a decision toward the adoption of preventive behaviors. The adoptive behavior of individuals also influences their network neighbors to adopt, widely known as the “social contagion” effects (the lower layer). The diffusion of preventive behaviors, in turn, limits the dispersion of diseases and speeds the diffusion of information. Behavioral interventions, as an external factor, can be implemented by health agencies to promote preventive behaviors, such as educational, incentive and role-model strategies. During an epidemic, these three diffusion processes interact with one another and form negative/positive feedbacks loops in the human-disease system, shown as arrows between layers in Fig. 1. Manipulated by the three external factors, these three diffusion processes, hereinafter named as the triple-diffusion process, determine the spatial and temporal dynamics of an epidemic.

Model formulation

Social network

The conceptual model is formulated by an agent-based approach, which has gained its momentum in epidemic modeling during the last decade (Burke et al., 2006; Huang, Sun, Hsieh, & Lin, 2004). Different from classic population-based models, each individual in a population is a basic modeling unit, associated with a number of attributes and events that change the attributes. To represent the contact network, individuals are modeled as nodes and are linked to one another through their daily contacts (as network ties). The individualized contacts are assumed to take place during three time periods in a day at four types of locations (Mao & Bian, 2010), namely the daytime at workplaces, the nighttime at homes, and the pastime at service places or neighbor households (Fig. 2). Individuals travel between the three time periods and the four types of locations to carry out their daily activities, thus having contact with different groups of individuals and exposing themselves to disease infection. These contacts link all individuals into a population-wide network.

Two types of individual contacts are modeled in terms of the contact duration and closeness. One type is the close contacts

![Fig. 1. Conceptual framework of the triple-diffusion model as a representation to an epidemic. Three layers illustrate three diffusion processes upon a contact network, where nodes represent individuals and links represent their relationships. Colors of nodes indicate the states of individuals in each diffusion process. Arrows between layers show the interactions between the three diffusions. Red arrow shows a negative effect, while the green arrow denotes a positive effect. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image-url)
(solid-line ovals in Fig. 2) that happen at homes (with family members), workplaces (with co-workers), and neighbor households (with friends). This type of contacts last for sufficient time to enable disease transmission. The other type refers to the occasional contacts (dash-line ovals in Fig. 2) that only happen at service places (with clerks and other consumers). In this case, an individual encounters only a limited number of individuals for a short time period, and thus the contact is less effective for infection.

**Diffusion of infectious diseases**

The diffusion of infectious diseases is formulated following the concept of classic susceptible-infectious-recovered (SIR) model (Anderson & May, 1992). Each individual possesses a serial of infection states and events as shown in Fig. 3 (the red dash-line box, in the web version). The progress of an infectious disease starts with a “susceptible” individual, who may receive infectious agents if having contact with an infectious neighbor in the network. The receipt event triggers a “latent” state, during which the disease agents develop internally in the body and are not emitted. The end of latent period initiates an “infectious” state, in which this individual is able to infect other susceptible neighbors and sustain the cascade of infection in the network. During the infectious period, individual may manifest disease symptoms (“Symptomatic”) or not (“Asymptomatic”). For either state, this individual remains infectious but would be unaware of the infection if asymptomatic. After the infectious period, this individual gets recovered and is assumed to be immune to infection during this epidemic.

Two disease events connect the disease diffusion with other diffusions. First, the event of symptom manifestation will motivate individuals to discuss disease information, and prompt their social contacts to adopt preventive behavior by posing infection risks. The second event is the receipt of disease agents, which is affected by the diffusion of preventive behavior. Specifically, the adoption of preventive behavior reduces the probability of disease transmission $P$, as specified in Equation (1):

$$P = E_{contact} \times I_{age} \times (1 - E_{prevention})$$

$E_{contact}$, $I_{age}$, and $E_{prevention}$ are three model parameters varying in [0, 1]. $E_{contact}$ indicates the effectiveness of a contact to transmit diseases, dependent on the physical closeness of the contact. Its value can be calibrated based on the observed characteristics of a disease, such as the basic reproductive number $R_0$. $I_{age}$ is an age-specific infection rate, specifying the likelihood of receiving disease agents by age group, such as children, adults and seniors. $E_{prevention}$ indicates the efficacy of a preventive behavior in reducing infection.
The parameterization is discussed later in the model implementation (Simulating the diffusion of disease section and Simulating the diffusion of preventive behavior section) when a specific disease and a specific preventive behavior are selected.

**Diffusion of information**

Regarding to the diffusion of information (blue dash-line box in Fig. 3, in the web version), individuals are initially “unaware” of the disease, but can be “informed” through two channels: the word-of-mouth discussion and the mass media. The former circulates the information locally through the contact network, while the latter disseminates the information globally in the population, both modeled as probabilistic events. First, an informed or symptomatic individual will discuss the disease with each network neighbor at a rate $\gamma_{\text{discussion}}$, as formulated in Equation (2):

$$\gamma_{\text{discussion}}(t) = \frac{\gamma_{\text{discussion}}(0)}{t - t_0}$$

where $t$ is the current time step and $t_0$ is the starting time of being informed or manifesting symptoms. The discussion rate $\gamma_{\text{discussion}}$ decays nonlinearly as time proceeds, i.e., an individual is more likely to talk about the disease within a few days after being informed or feeling sick.

Turning to the mass media, the probability of an individual being informed $\gamma_{\text{mass}}$ is formulated as a function of total symptomatic case number $N_i(t)$ at time step $t$ (Equation (3)):

$$\gamma_{\text{mass}}(t) = \frac{1 - e^{-N_i(t)/\beta}}{1 + e^{-N_i(t)/\beta}}, \quad 0 \leq N_i(t) \leq 0, \quad \beta > 0$$

The more individuals get sick during the time $t$, the higher the intensity of mass-media propagation, and thus the greater chance for an individual being informed. The constant $\beta$ is a scaling parameter that controls the intensity of mass-media propagation, and a small $\beta$ results in a large $\gamma_{\text{mass}}$. A mass-media campaign then can be modeled by varying the $\beta$, the timing of campaign (when to start), and the frequency of campaign (time intervals between two broadcastings). Once informed by either the discussion or the mass media, individuals will become decision makers toward the adoption of preventive behavior. In such a manner, the diffusion of information is coupled with the diffusion of disease and that of preventive behavior.

**Diffusion of preventive behavior**

Individuals being informed start to evaluate and make a decision toward the adoption of preventive behavior (Green dash-line box in Fig. 3, in the web version). The decision depends on individuals’ own characteristics and inter-personal influence from their social networks. This research uses a threshold behavioral model (Granovetter & Soong, 1983) to formulate the decision process. Specifically, each individual has two adoption states, and the change of state is calculated based on Equation (4):

$$\text{Adoption State}_i = \begin{cases} \text{Adopted,} & \text{if } \alpha_i(t) \geq T_{pi} \text{ or } \mu_i(t) \geq T_{ri} \\ \text{Not Adopted,} & \text{Otherwise} \end{cases}$$

where

$$\alpha_i(t) = \frac{\text{Number of adopters in } i\text{'s social contacts at } t}{\text{Total number of } i\text{'s social contacts}}$$

$$\mu_i(t) = \frac{\text{Number of infections in } i\text{'s social contacts at } t}{\text{Total number of } i\text{'s social contacts}}$$

For a given time step $t$, individual $i$ will evaluate the proportion of adopters in $i$’s personal network, as the peer pressure of adoption $\alpha_i(t)$. Once the peer pressure reaches a threshold $T_{pi}$ (called the threshold of adoption pressure), an individual will decide to adopt. Meanwhile, individual $i$ also evaluates the proportion of symptomatic individuals in the personal network, as the perceivable risks of infection $\mu_i(t)$. If the perceived risk exceeds another threshold $T_{ri}$ (termed the threshold of infection risk), an individual will also adopt. The individualized thresholds ($T_{pi}$ and $T_{ri}$) reflect personal characteristics of individuals, while the events of evaluation represent the inter-personal influence between individuals. In such a way, the diffusion of disease elevates the perceivable risks of individuals, and stimulates them to adopt preventive behavior. In turn, the adoption of preventive behavior impedes the diffusion of disease, forming a negative feedback loop in the human-disease system.

**Model implementation**

**Simulating individuals and their contact network**

The proposed triple-diffusion model is implemented in the Greater Buffalo Metropolitan Area, NY, USA, with a population of 985,001 (according to Census 2000). Each individual is programmed into a software agent with attributes and events (Table 1). Besides a unique identifier, each individual has 6 groups of attributes, including the network, demographic, spatio-temporal, infection, adoption, and information attributes. The events change the values of corresponding attributes.

| Attributes, values, and events of an individual as a software agent (revised based on Mao & Bian, 2011). |
|---|
| **Attributes** | **Descriptions** | **Values** |
| ID | Personal Identification | Integer [1, $N$] |
| Contacts | A list of IDs of contacts | Integer [1, $N$] |
| Effectiveness of contacts | Close contact or occasional contact | Real number [0, 1] |
| Demographics | Age group, gender, occupation, etc. | Integer |
| Current time | Daytime, pastime, or nighttime | Integer |
| Current location | Home, workplace, service place, or friend household | $(x, y)$ coordinates |
| Infection state | Susceptible, latent, infectious, recovered | Integer |
| Infection time | Day of receiving infection | Integer |
| Threshold of infection risk | The risk of infection | 0–100% |
| Threshold of adoption pressure | Necessary for adoption | 0–100% |
| Adoption state | Adopted or not adopted | 1 or 0 |
| Adoption time | Day of adoption | Integer |
| Information state | Informed or not informed | 1 or 0 |
| Informed time | Day of being informed | Integer |

| Events | **Descriptions** |
|---|
| Movement | Change $(x, y)$ coordinates over time |
| Having contact | Trace individuals who have contact with an infectious individual |
| Having discussion | Change the information state of the contacts |
| Being infected | Change infection state to latent |
| Evaluation of infection risk | Calculate the proportion of influenza cases in the contacts |
| Evaluation of social standard | Calculate the proportion of adopters in the contacts |
| Adoption of preventive behavior | Change adoption state |
| Reduction of transmission rate | Adjust the transmission probability of influenza after adoption |
The social network is realized by a previously developed algorithm that assigns values to the demographic and network attributes of individuals (Mao & Bian, 2010). The value assignment involves a large amount of geo-referenced data, including census data, business location data, land parcel data, transportation network, and results of a household travel survey. Statistical distributions derived from these datasets, such as distributions of family size, workplace size, and household daily trips, are used to ensure the validity of value assignments. To differentiate the weekdays and weekends, individuals are not assigned to work (or schools) at weekends except those who work in service-oriented businesses (such as restaurants and grocery stores). Those who do not work during the weekends would have increased trips to service-oriented businesses. The completion of assignments forms three linked populations, including a nighttime population at homes, a daytime population at workplaces, and a pastime population at service places or neighbor households. The three populations represent the same set of individuals, but at different locations and time periods of a day. Individuals have contact with a number of other individuals at a same time period and same location, forming a spatio-temporally varying network. The simulated network has an average of 16.9 daily contacts per person, consistent with the observed number (16.8) from empirical studies (Beutels, Shkedy, Aerts, & Van Damme, 2006; Edmunds, 1997; Fu, 2005).

**Simulating the diffusion of disease**

The seasonal influenza is selected as an example due to its natural history has been well understood. A number of influenza parameters are either adopted or calibrated from existing literature as shown in Table 2. The product of $I_{age}$ and $E_{contact}$ determines the transmission probability through one contact (Equation (1)), which is used to simulate individuals’ transition from susceptible to latent state as a stochastic branching process. The latent, incubation, and infectious periods control the sequential transitions from latent to infectious, symptomatic, and recovered states.

**Simulating the diffusion of information**

Two groups of parameters are set to simulate the word-of-mouth discussion and the mass-media effects, respectively. The parameter values are calibrated from the model evaluation later but are reported here. For the word-of-mouth discussion, the initial discussion rate $\gamma_{discussion}(0)$ in Equation (2) is set to 0.001, and then the $\gamma_{discussion}(t)$ is updated as the time goes by. For the mass media, the scaling parameter $\beta$ in Equation (3) is specified as 5000, based on which the probability of being informed by mass media $\gamma_{mass}$ can be computed at every time step. The mass-media campaign is assumed to be triggered when the total symptomatic individuals exceed 1% of the total population, and the frequency of broadcasting follows a weekly basis. With these two probabilities, a Monte-Carlo simulation is used to determine whether an unaware individual will be informed or not at each time step.

**Simulating the diffusion of preventive behavior**

The use of flu prophylaxis (e.g., Oseltamivir) is taken as a typical example of preventive behavior, because its efficacy is more conclusive than other preventive behaviors, such as hand washing and facemask wearing. Three parameters are specified to simulate the behavioral diffusion and couple it with the diffusion of influenza. First, the model assumes that symptomatic individuals has a 75% likelihood to adopt flu prophylaxis to mitigate the symptoms and reduce their infectivity (Mills, Levine, & Goel, 1998). Second, the preventive efficacy of flu prophylaxis ($E_{prevent}$ in Equation (1)) is set to 70% and 40% for susceptible and infectious individuals, respectively, indicating that their likelihood of being infected or infecting others can be reduced by such amount (Hayden, 2001; Longini et al., 2004). Third, the two adoptive thresholds of individuals $T_{pi}$ and $T_{ri}$ (in Equation (4)) are generated from their statistical distributions using a Monte-Carlo method. These statistical distributions were derived based on a health behavioral survey, whose details are provided in the Supplementary Document. Each individual is assigned to random numbers from those statistical distributions as their adoptive thresholds. For each time step, the model computes the peer pressure and perceived risk for every informed individual, and updates his/her adoption state using the threshold model.

**Model initialization and evaluation**

The triple-diffusion model is simulated over 150 days, covering a general flu season (from December to May). At the beginning of simulation, all individuals are assigned susceptible and unaware states. To consider a background immunity before the epidemic, the model randomly selects 62.7% of seniors, 15.6% of adults, and 17.9% of children according to the national immunization coverage (Euler et al., 2005; Molinari et al., 2007), and directly moves them to the adopted and recovered states. All unselected individuals are set as non-adopters. To initialize the disease diffusion, five infectious individuals are randomly introduced into the study area at the first day. The simulation takes a tri-daily time step and runs the three diffusion processes concurrently in each time step. To stabilize the final outcomes, the model has been implemented by 50 realizations. In each realization, the background immunity, the first five infectious individuals, their contacts, and the infection, awareness, and adoption of these contacts are randomized. The final outcomes are three diffusion curves, namely the epidemic curve (the weekly number of new cases), the adoption curve (the weekly number of new adopters), and the awareness curve (the weekly number of newly informed), all averaged from 50 model realizations.

Two independent data sources are used to evaluate the model results and calibrate model parameters. One is the weekly reports of laboratory confirmed specimens in the 2004–2005 in Buffalo, NY, issued by the New York State Department of Health (NYSDOH, 2005). The simulated epidemic curve is compared to the weekly reported data to show the validity of modeling disease diffusion. The other data source is the weekly statistics from Google Flu Trends of the study area (Google, 2011), which summarizes the number of online flu-related inquiries as a tool for monitoring influenza outbreaks (Ginsberg et al., 2008). Relevant to this research, the weekly flu trend data could be a reliable representation to the real diffusion of influenza-related information (Fenichel, 2011).

---

**Table 2**

| Parameters                  | Values                  | Literature                     |
|-----------------------------|-------------------------|--------------------------------|
| Basic reproduction number $R_0$ | 1.2–1.3                | Mills, Robins, & Lipstitch, 2004 |
| Infection rates $I_{age}$    | Children: 0.1 Youth and adults: 0.08 Seniors: 0.09 | Calibrated based on $R_0$ |
| Effectiveness of contacts $E_{contact}$ | $E_{close,contact}$: 1 $E_{casual,contact}$: 0.2 | Calibrated based on $R_0$ |
| Likelihood of symptom manifestation | 0.5                    | Ferguson et al., 2005 |
| Latent period               | 2 days                  | Heymann, 2004                  |
| Incubation period           | 3 days                  | Heymann, 2004; Halloran et al., 2008 |
| Infectious period           | Children: 7 days Youth and adults: 4 days Seniors: 4 days | Heymann, 2004; CDC, 2008 |
Kuminoff, & Chowell, 2013), and is compared to the simulated awareness curve from the model.

Results and discussion

Model result evaluation

Fig. 4 displays the simulated weekly number of newly infected individuals, compared to the actual number of weekly lab-confirmed cases during the 2004–2005 influenza season. The shape and peak time of the predicted curve correspond well with those of the reported epidemic, although the magnitude of simulated cases is much larger than the reported data. The first possible reason is that many sick people may choose self-care instead of seeing a doctor, and thus cannot be reported. Second, for those who seek healthcare, only a small portion of their specimens were submitted for laboratory testing. Therefore, the number of influenza cases is often highly under-reported, and a complete data is rather difficult to collect. The laboratory data, so far, is the best available touchstone for model validation. In this sense, the model performs well in predicting the trend, and at least allows the estimate of a worse case result.

Fig. 5 compares the simulated weekly number of newly informed people to the excessive weekly Google Flu search statistics that indicate the amount of online searching behavior relevant to influenza epidemic during the 2004–2005 flu season. The excessive weekly search statistic (the left Y axis) is the difference between the observed statistics and its long-term average (1296), which removes searches of influenza in a normal day but not caused by the epidemic. The temporal course of the information diffusion is well predicted. Since the two measurements are not in a same unit (primary and secondary axis in Fig. 5), their magnitudes are not comparable but are highly correlated.

To my knowledge, both comparisons show a level of consistency that has rarely been achieved by other epidemic models (Ferguson et al., 2006; Funk et al., 2009; Kiss et al., 2010; Vardavas et al., 2007). A majority of previous models cannot validate themselves by observed facts, particularly for the information diffusion. The triple-diffusion model, thus, could provide a reliable foundation to devise much-needed control and intervention strategies for infectious diseases, such as behavior promotion strategies and mass-media campaigns.

Fig. 6 shows how the diffusion of influenza motivates people to adopt preventive behavior. The two diffusion processes take a similar shape, but there is a time lag about 1 week between their peak times. As the number of influenza cases rises, individuals perceive increasing risks, which motivate them to adopt preventive behavior. Their adaptive behavior further influences surrounding individuals to adopt. The time lag between the epidemic and adoption curves is possibly the time individuals need to be informed and take preventive actions. Fig. 6 also suggests that monitoring the real-time flu prophylaxis sales could detect the epidemic peak about 1 week ahead of the traditional disease surveillance networks, such as the CDC sentinel network, which take up to 2 weeks to collect, process, and report disease cases registered at health centers.

The diffusions of influenza and its related information also take a similar bell-shape, but the information peaked approximately one week earlier than the disease (Fig. 7). A possible reason is the wide coverage of modern mass media over the population, enabling a faster spread of information than the disease. At the beginning of the epidemic, only a few influenza cases occurred and a vast majority of people were unaware of the disease. As the diffusion of influenza took off (January 2nd–9th), individuals started to notice the disease problem and discuss with each other. When the epidemic became more sensational and drawn attention from the mass media (January 10th–16th), the awareness curve climbed steeply and reached the peak in four weeks. Fig. 7 implies that overseeing the diffusion of disease information, such as the Google flu trend, could warn the public 1–2 weeks earlier before the epidemic peak actually occurs.

In addition to forecasting the timeline of the three diffusion processes, this spatially explicit model allows the prediction of geographic distributions over time. Fig. 8 maps the spatial distributions of simulated infection (square), adoption (circle), and awareness (triangle) are mapped. For the purpose of clarity, only the downtown area is presented. The spatial distributions on Days 50, 75, and 100 are displayed in order to present different stages of the epidemic. On Day 50 the epidemic is on the rise, Day 75 is around the peak time, and on Day 100, the epidemic is in decline.

Still in its infancy, the triple-diffusion model has several limitations in its design and implementation. First, the contact network
Fig. 5. A comparison of the simulated weekly number of newly informed people (as an average of 50 realizations) to the weekly excessive Google Flu Trend statistics during the 2004–2005 influenza season.

Fig. 6. Temporal progress of influenza diffusion vs. the behavioral diffusion. The Y axis on the left shows the number of weekly new cases, while the Y axis on the right shows the number of weekly new adopters.

Fig. 7. Temporal progress of influenza diffusion vs. the information diffusion. The Y axis on the left shows the weekly number of new infected cases, while the Y axis on the right shows the weekly number of newly informed individuals.
used in the model could be further refined into three different but partially overlapping networks, namely an infection network, information network, and influential network, each channels a diffusion process. Admittedly, the model would be more realistic, but building these networks requires extensive social survey to collect relevant data, which is often costly and time consuming. Recent work on extracting social networks from Facebook and Twitter may be a promising method to address this issue (Lewis, Kaufman, Gonzalez, Wimmer, & Christakis, 2008). Second, the effects of mass media is formulated as a simple formula, but could be more complicated if further considering various types of media and their corresponding coverage. Probably, a sophisticated function could depict the human communications better, such as an exponential decay or a power-law decay function. Life style data of individuals may be helpful to identify their preferred media and delineate the media coverage. Third, the discussion rate is assumed homogenous over the study area. This rate may vary between age groups, occupations, and personalities. A health behavior survey may need to estimate discussion rates for different groups of individuals. Fourth, there would be a certain amount of uncertainty if the model was used to predict the future influenza outbreaks. If new census data and travel survey data were filled in, this model could reasonably predict the timeline and scale of a future outbreak of seasonal influenza. However, for a pandemic influenza, such as the new H1N1, most of the model parameters should be adjusted to account for the highly infectious virus, the faster circulation of information, and possibly distinct response of individuals toward preventive behaviors. All these limitations warrant a future study. After all, the goal of modeling is not to predict what exactly happen during an epidemic, but rather to observe how the epidemic may proceed and encourage appropriate questions. In this sense, the model results provide valuable knowledge regarding city-wide epidemics.

Fig. 8. Simulated distribution of infection (red square, left), adoption (green circle, middle), and awareness (blue triangle, right) on Days 50, 75, and 100, respectively. On day 50 the epidemic is on the rise and the peak is reached around day 75, and on day 100, the epidemic is in decline. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Conclusions

This article presents an original triple-diffusion model for epidemiology, and discusses its conceptual framework and design. The conceptual framework integrates three interactive processes: the diffusion of influenza, the diffusion of information, and that of preventive behavior, upon a human social network. The agent-based approach, network model, theories from epidemiology, information and behavioral sciences are used to formulate the conceptual framework into a working model. For illustration purposes, the model is implemented in an urbanized area with a large population. Compared to the reported data, the proposed model reasonably replicates the observed trends of influenza infection and online query frequency. The model, thus, could be a valid and effective tool for exploring various control policies.

There are two key contributions of this research to the literature body of network diffusion theory, public health, and agent-based modeling. First, the proposed triple-diffusion framework is a significant advancement to previous disease-only models and dual-diffusion models, such as Bian and Mao’s work. The fusion of disease, information, and human behavior allows for a more comprehensive 3D cubic view of the human-disease system, which can be studied from a 3D perspective. The increase of one dimension exposes much more details of an epidemic and thus enables a deeper understanding of this complex system, for example, the interactive mechanisms among the three diffusion processes. The proposed modeling framework can flexibly accommodate the mobile phone tracking data and the latest census data to improve the accuracy of modeled daytime and nighttime populations. The online social networking data (from Facebook and Twitter) can also be included to modify the way of communications between individuals, and as well as the personal influence between them.

Second, this model can be further developed into a virtual platform for health decision makers to test disease control policies in many other metropolitan areas. Particularly, since the model explicitly represents the diffusion of information and human preventive behavior, it permits a systematic evaluation of disease control policies that have not been well studied before, such as the mass-media campaigns and behavioral incentive strategies. The evaluation results will enrich the family of disease control policies, and help the public health overcome the socio-economic challenges posed by potential influenza outbreaks.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.apgeog.2014.02.005.

References

Anderson, R. M., & May, R. M. (1992). Infectious diseases of humans: Dynamics and control. Oxford University Press.
Beutels, P., Shkedy, Z., Aerts, M., & Van Damme, P. (2006). Social mixing patterns for transmission models of close contact infections: exploring self-evaluation and diary-based data collection through a web-based interface. Epidemiology and Infection, 134(6), 1158–1166.
Bian, L., Huang, Y., Mao, L., Lim, E., Lee, G., Yang, Y., et al. (2012). Modeling individual vulnerability to communicable diseases: a framework and design. Annals of the Association of American Geographers, 102(5), 1016–1025.
Burke, D. S., Epstein, J. M., Cummings, D. A. T., Parker, J. I., Cline, K. C., Singa, R. M., et al. (2006). Individual-based computational modeling of smallpox epidemic control strategies. Academic Emergency Medicine, 13(11), 1142–1149.
CDC. (2008). Key facts about seasonal influenza (flu). Retrieved January 2008, from http://www.cdc.gov/flu/about/disease/index.htm.
Eames, K. T. D., Tilston, N. L., Brooks-Pollock, E., & Edmunds, W. J. (2012). Measured dynamic social contact patterns explain the spread of H1N1v influenza. PLoS Computational Biology, 8(3), e1002425.