Computers & Industrial Engineering 129 (2019) 563–577

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Modeling learning and forgetting processes with the corresponding impacts on human behaviors in infectious disease epidemics

Kaiming Bi, Yuyang Chen, Songnian Zhao, David Ben-Arieh, Chih-Hang (John) Wu*

Department of Industrial & Manufacturing Systems Engineering, Kansas State University, Manhattan, KS 66502, USA

ARTICLE INFO

Keywords:
Forgetting and learning
Epidemics
Behavior changes
Agent-based simulation

ABSTRACT

This article presents two new mathematical models, an information forgetting curve (IFC) model and a memory reception fading and cumulating (MRFC) model, to examine forgetting and learning behaviors of individuals during an infectious disease epidemic. Both models consider how epidemic prevalence and community behavior-change information may affect agent emotions and subsequently influence an individual's behavior changes during an epidemic. The IFC model utilizes a forgetting curve to process epidemic information, and the MRFC model formulates disease information variations using the Itô diffusion process. Sensitivity analysis and simulation comparisons showed that the MRFC model more accurately describes the epidemic with high lethal rate gets high attention. The author also demonstrated that MRFC model has higher sensitivity parameters and is more flexible on wide ranges of infection rates than the IFC model. However, the IFC model is a better suited for widespread, low-risk mortality epidemics, such as seasonal influenza, the infection information and protective behavior have close relationships among the susceptible population. An agent-based simulation model also developed to mimic the epidemic prevalence of the 2009 Chicago H1N1 using public available historical data sets by IFC model.

1. Introduction

Human disease awareness and related behavior changes during disease epidemics have recently attracted considerable research attention (Polgar, 1962). In order to be more accurately predict a disease epidemic and estimate its potential impacts, however, a comprehensive understanding of information dissemination within human contact networks and the effects of this information on human emotions, awareness, and behavior must increase. Extensive literature and studies have investigated how information affects human behaviors, but minimal research has focused on human memory and forgetting/learning processes related to disease information, or the process in which information may be forgotten and relearned during an epidemic episode. This paper proposes two new mathematical models to investigate the effects of information in disease transmission, including the forgetting and learning phenomenon.

The human brain cannot store an infinite amount of retrieved information. In 1968, Shiffrin and Atkinson (1969) first classified memory as long term and short term. Engle, Tuholski, Laughlin, and Conway (1999) defined short-term memory as retainable for a short period of time (usually from 6 to 600 s) but unable to be manipulated; however, they did not detail how long-term or short-term memory relates to people forgetting information. Ebbinghaus (1913) experimentally investigated how the process of forgetting proceeds with influences of time or daily events, hypothesizing that, although a memory series is gradually forgotten, memories that have been learned twice fade more slowly compared to memories that have been learned once. Wingfield and Byrnes (2013) proposed a “forgetting curve” to show the process of memory loss over a period from 20 min to 31 days. A recent paper has indicated that the people learn language also following the forgetting curve (Weltens & Cohen, 1989), and experiments have been conducted to increase understanding of the learning and forgetting phenomenon (Badiru, 1992; Bailey, 1989). In 1985, Brainerd, Kingma, and Howe (1985) concluded that forgetting is governed by various laws and therefore requires unique theoretical assumptions.

Notable learning and forgetting mathematical models have been proposed. In 1976, Carlson and Rowe (1976) introduced the variable regression variable forgetting (VRVF) model, and, in 1990, Elm'Aghraby (1990) proposed the variable regression invariant forgetting (VRIF) model, which corrected errors in previous forgetting models and accommodated a finite horizon. In 1996, Jaber and Bonney (1996) proposed the learn-forget curve model (LFCM), which showed that forgetting is dependent on some factors such as the

* Corresponding author at: Kansas State University, 1701B Platt St, Manhattan, KS 66502, USA.
E-mail addresses: bikaiming@ksu.edu (K. Bi), cyy@ksu.edu (Y. Chen), songnian@ksu.edu (S. Zhao), davidbe@ksu.edu (D. Ben-Arieh), chw@ksu.edu (C.-H.J. Wu).

https://doi.org/10.1016/j.cie.2018.04.035

Available online 20 April 2018
0360-8352/ Published by Elsevier Ltd.

learning slope, the quantity produced and the minimum production breaks. In 1997, Jaber and Bonney (1997) compared these three models, and in recent years, Jaber, Kher, and Davis (2003) reviewed factors that influence forgetting and incorporated the job similarity factor into the LFCM. In 2002, Sikström and Jaber (2002) provided an elaborate review of forgetting curves in psychology and industrial engineering literature. Because the previous papers did not consider the forgetting phenomenon in disease, however, this paper proposes an information forgetting curve model (IFC) to describe how disease information fades over time during an epidemic, following a forgetting curve through time and thereby influencing final disease memory. In 2012, Sikström & Jaber updated their research in the modeling of learning and forgetting area (Sikström & Jaber, 2012). They proposed a Depletion-Power-Integration-Latency (DPIL) Model. This model considered the depletion of the encoding resource as forgetting and learning behavior when the system repetitively performs a task. More than fitting the historical dataset and calculate the settings of optimal performance, this model discussed how learning can interact with the forgetting by modeling the repetitively encoding can increase the memory strength.

Stochastic factors can influence the information perception process and new information can diversely affect human memory when agents receive new information and forget previous information. Researchers (Slovic, 2016) have shown that biased media coverage, misleading personal experiences, and anxieties can cause people to process information with unwarranted confidence or uncertain judgment. Zhao, Wu, Kuang, Bi, and Ben-Arieh (2018) also discussed the stochastic change rate of perception to infectious disease risk. This paper proposes a memory reception fading and cumulating (MRFC) model to describe the stochastic phenomenon of human memory as it pertains to disease information based on learning and forgetting.

Agents have unique understandings based on identical information, and they react with distinct switch behaviors. In 2009, Chen (2009) concluded that agents learn prevalence through the spread of information and can adjust human behavior during a disease epidemic. Funk, Gilad, Watkins, and Jansen (2009) found that the spread of disease awareness significantly decreases the infection rate, and Kiss, Cassell, Recker, and Simon (2010) proved that the diffusion of disease information increases risk awareness and causes the host population to take infection prevention measures. In 2015, Zhao, Wu, Kuang, and Ben-Arieh (2015) proposed a disease model using a spatial evolutionary game to illustrate the impact of information dissemination on human behavior in an epidemic, proving that how an agent feels depends on information content and context (Nahl & Bilal, 2007). In other words, agents demonstrate unique perspectives for the same disease information, resulting in diverse emotional responses. When disease information is positive, agents may have minimal concerns about the disease; negative information, however, may increase agent's awareness. Hence, certain types of information could alter agents' moods or emotions (Pessoa, 2008). Chen, Bi, Zhao, Ben-Arieh, and Wu (2017) also modeled how disease information, such as the number of infected individuals and the number of susceptible individuals who choose the switching behaviors, impact agents' fears about the epidemic. This paper applies an agent-based model to determine how disease information can cause diverse human behaviors, as well as use of the one-factor-at-a-time method (OFAT) to conduct sensitivity analysis for various parameter settings in the IFC and MRFC models, highlighting the influence of parameter settings for the model and comparing agents' epidemic behaviors using the IFC, MRFC, and no-memory models.

This paper also discusses the 2009 H1N1 influenza epidemic. By combining historical infection data in affected cities with corresponding population characteristics, the authors restored the 2009 H1N1 prevalence in Chicago, a typical H1N1-affected city. This paper also investigates how the phenomenon of memory fading and behavior-switched protection influence epidemic spreading.

2. Information forgetting curve model

2.1. Contact network and disease information

Contact networks have been widely applied to many implementations of disease transmission (Altizer et al., 2003; Bansal, Grenfell, & Meyers, 2007; Lloyd-Smith, Schreiber, Kopp, & Getz, 2005). Zhao et al. (2015) introduced the concept of disease information dissemination and its effect on epidemic disease transmission by proposing that agents gain disease information from two layers: local and global contact networks. Local information is gained from neighboring agents, while global information is acquired from all locations. The researchers in Zhao et al. (2015) used a spatial evolutionary game to figure out if agents switch behavior based on payoff. In general, local contact networks contain many social cliques and more readily transmit pathogens, while global contact networks are usually dominated by non-face-to-face contacts and random long-distance connections (Klovdahl, 1985).

Rapidly increasing advanced technology and popularity of the internet, social media, and new-media broadcasting channels have revolutionized the ways about information transmission and human communication. Sahneh and Scoglio (2013) changed traditional global and local contact network divisions to disease transmission and information transmission divisions. A disease transmission contact network (DTCN) is comprised of daily face-to-face contacts, such as family members, neighbors, and colleagues, while an information transmission contact network (ITCN) includes all contacts in an agent's social media. In general, contacts in a DTCN are a subset of contacts in an ITCN (Fig. 2.1).

Recent advancements in the information technology industry and the Internet of Things are spurring a rapid transition into an information era. Subsequently, disease information and its dissemination via modern information systems, such as social media, virtual communities, alternative media, and traditional media broadcasting, have begun to significantly influence disease transmission (Fishbein, Middlestadt, & Hitchcock, 1994). However, current research on the effects of disease information are primarily limited to the infected information (i.e., disease prevalence) (Grassly & Fraser, 2008; Hadeler & Castillo-Chávez, 1995; Van den Driessche & Watmough, 2002). In 2009, Chen (2009) found that self-protection qualities in a disease can positively influence disease transmission. Agents frequently choose to use protective measures in an epidemic to reduce infection risks; these behaviors are known as switching behaviors. When agents choose not to take any protective measures in an epidemic, their behavior is referred to as normal behavior. In order to most accurately describe an individual's perception of information related to an ongoing epidemic, this research refers to a measure as perceived disease information (PDI), as defined in Eq. (2.1), which can be divided into infected information (denoted as $I_i(t)$) and switch behavior information (denoted as $sw_i(t)$):

$$PDI_i(t) = \alpha I_i(t) + (1-\alpha) sw_i(t), \quad \text{where } 0 \leqslant \alpha \leqslant 1 \tag{2.1}$$

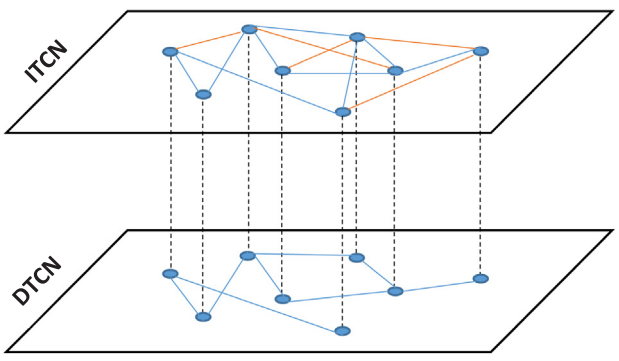

Fig. 2.1. Schematics of a two contact networks: ITCN and DTCN.

In Eq. (2.1), $PDI_i(t)$ represents new disease information of agent i at time t; $I_i(t)$ denotes the estimation of infected population percentages of agent $i's$ contact network at time t, which represents the infected information; $sw_i(t)$ is the estimation of switch behavior population percentages of agent $i's$ contact network at time t, which represents behavior-switched information; weighted parameter α denotes how an individual agent weights the proportion of infected information.

Since ITCN and DTCN can be utilized to collect information, $I_i(t)$ and $sw_i(t)$ can be divided based on the information source. $I_i^{ITCN}(t)$ and $I_i^{DTCN}(t)$ represent infected information collected by ITCN and DTCN, $sw_i^{ITCN}(t)$ and represent behavior-switch information collected by ITCN and DTCN, and β is the weight parameter to determine the proportion of information sources. Thus, $I_i(t)$ and $sw_i(t)$ can be described as

$$I_i(t) = \beta I_i^{ITCN}(t) + (1-\beta) I_i^{DTCN}(t) \tag{2.2}$$

$$sw_i(t) = \beta sw_i^{ITCN}(t) + (1-\beta) sw_i^{DTCN}(t) \tag{2.3}$$

Then,

$$PDI_i(t) = \alpha\beta I_i^{ITCN}(t) + \alpha(1-\beta) I_i^{DTCN}(t) + (1-\alpha)\beta sw_i^{ITCN}(t) + (1-\alpha)(1-\beta) sw_i^{DTCN}(t) \tag{2.4}$$

Since I_i^{ITCN}, $I_i^{DTCN}(t)$, $sw_i^{ITCN}(t)$, and are four percentages; also we have $\alpha\beta + \alpha(1-\beta) + (1-\alpha)\beta + (1-\alpha)(1-\beta) = 1$, the range of perceived disease information $PDI_i(t)$ is [0, 1]. The high $PDI_i(t)$ indicates increased seriousness of the disease.

2.2. Information forgetting in a disease epidemic

Human memory is a system that can store and retrieve information (Baddeley, 1997). In 1913, Ebbinghaus (1913) introduced the forgetting curve to describe the information process of forgetting over time. In 1991, Wixted and Ebbesen (1991) presented mathematical functions to represent the process of forgetting, including experiments to demonstrate and analyze the forgetting curve based on those forgetting functions. The authors considered recall, recognition, and saving as measures of memory. The article also studied several materials to be remembered, including present words, faces, nonsense syllables, and graphic images. In addition, serval different subjects were used in the experiments, and the experiments were carried out at various time intervals. The study found that the process of forgetting can be represented mathematically using the following simple power function of time:

$$y = at^{-b} \tag{2.5}$$

In Eq. (2.5), y is a memory performance measure for the strength of the memory trace, or the proportion recalled by memory, and t represents time (one day as a unit period) (Anderson & Tweney, 1997). Parameter a is the degree of learning, which represents the estimated level of performance after one unit of time, and parameter b is the rate of forgetting, where a and b range from 0 to 1 (Wixted & Ebbesen, 1997). Therefore, this power function, y, ranges from 0 to infinity; that is, as t tends to zero, y tends to infinity, and as t tends to infinity, y tends to zero. Fig. 2.2(a) and (b) show the trend of function y when the degree of learning ($0 \le a \le 1$) and rate of forgetting ($0 \le b \le 1$) differ.

Disease parameters such as infectious periods, numbers of infections, and disease susceptibility are considered disease knowledge or information (Becker, 1989) and can be remembered or recalled by an individual agent. This paper assumes that, similar to normal daily information, disease information acquired via disease parameters incorporates the learning and forgetting phenomenon. When agents receive similar information that they received before, corresponding memory is reinforced, causing slow memory fading. However, when signal of prevalence aggravation has not stimulated human memory, memory about the information follows the forgetting curve.

$$w_t = \frac{\int_t^{t+1} at^{-b}dt}{\int_0^{t_f} at^{-b}dt} \tag{2.6}$$

$$w_0 = \frac{\int_0^1 at^{-b}dt}{\int_0^{t_f} at^{-b}dt} = \frac{1}{t_f^{b+1}-1} \tag{2.7}$$

Eq. (2.6) defines memory residual weight of disease information. In the equation, w_t represents the memory performance proportion for integration of y_t in one day at total memory performance on t days before current day, where the integration of at^{-b} from 0 to t_f represents the sum of memory portions; integration of at^{-b} from t to $t+1$ represents the memory performance proportion in the t−th day. Eq. (2.7) shows formulation of the special case (performance proportion of memory in the day before current day) w_0, meaning feasible longest memory epoch t_f and forgetting rate b. Therefore, the mathematical equation of final disease information (FDI) in the IFC model can be described as

$$FDI_i(t) = w_0 PDI_i(t) + w_1 PDI_i(t-1) + \cdots + w_t PDI_i \tag{2.8}$$

where $FDI_i(t)$ represents FDI of agent i at time t. Since $I(t)$ and $sw(t)$ are percentages from 0 to 1, $PDI_i(t)$ has the range [0, 1]. In addition, $\sum_{j=0}^{t} w_j = 1$, meaning the range of $FDI_i(t)$ is also [0, 1]. For agent i, $FDI_i(t)$ reflects disease severity cognition at time t:

$$FDI_i(t) = w_0 PDI_i(t) + (1-w_0) FDI_i(t-1) \tag{2.9}$$

FDI also could be exponential smoothing, as shown in Eq. (2.9). Specifically, $w_0 PDI_i(t)$ could be new information learning, and $(1-w_0)FDI_i(t-1)$ could be past information forgetting. w_0 is a crucial adjective parameter in the IFC model. In the case of $w_0 = 1$, the agent is assumed to be memoryless, so FDI is equal to PDI and the agent chooses behavior based only on current information.

2.3. Fear factor and human behavior in disease

Epidemic information is often disseminated in correlation with the spread of disease. In addition to acquiring disease information through social networks, newspaper, or TV news, agents also gain disease information via communication with colleagues and families in their contact networks. This method of attaining disease-related information often includes an emotional response that influences agents' immediate behavior changes. For example, if agents know many individuals have become infected or expired at the outbreak of an infectious disease, they may fear the disease and take protective measures, such as decreasing travel, wearing masks, or becoming vaccinated, to prevent infection. The process of transforming information to emotion and then to action is illustrated in Fig. 2.3.

Chen et al. (2017) proposed a model to describe how disease information can affect individuals' emotions by introducing an individual fear factor ($IFF_i(t)$). Where t represents the current time, i represents the sequence number of the agent. Their model showed that individuals demonstrate diverse disease perceptions when they know infected individuals or switch individuals, thereby altering their emotions. For example, an increasing number of infected individuals in a neighborhood and increasingly negative media about the disease enhance, an individual's concern about the disease, consequently increasing the individual's fear factor. If fewer individuals become infected in neighboring areas, an individual tends to feel safer rather than fearful, resulting in a minimal individual fear factor. This paper assumes that disease information and an individual's emotions have significant correlation, as shown in Eq. (2.10). When the number of final disease information $FDI_i(t)$ is large (more close to 1), an agent has strong fear emotion about the disease; when the number of $FDI_i(t)$ is small (more close to 0), an agent has weak fear emotion about the disease.

$$FDI_i(t) \propto IFF_i \tag{2.10}$$

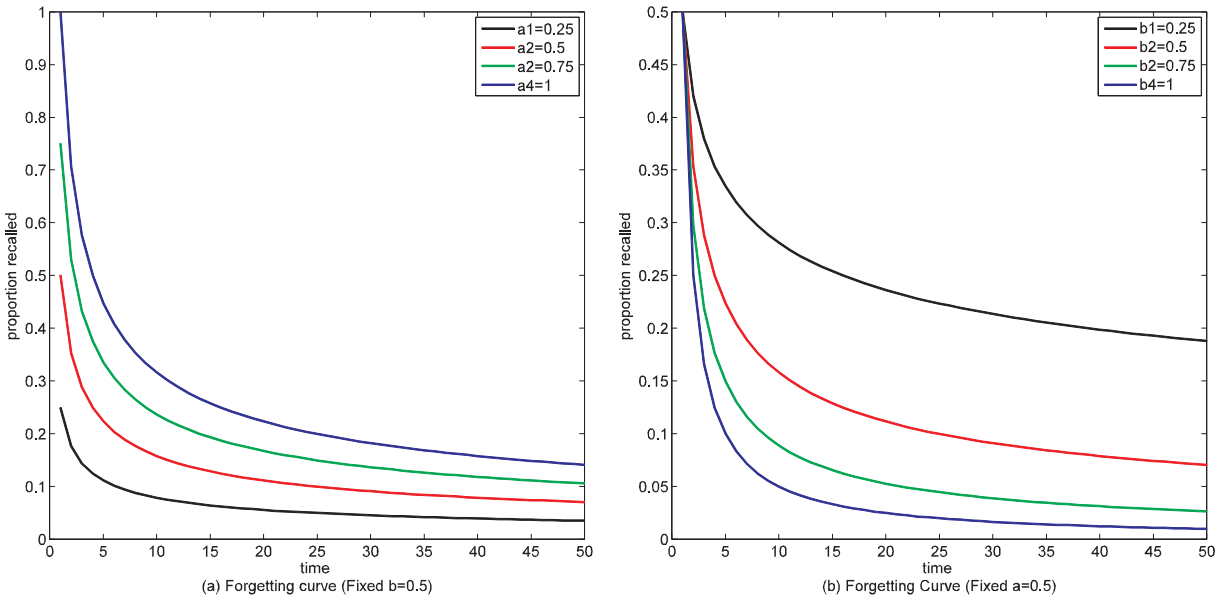

Fig. 2.2. Forgetting curves.

Steimer (2002) stated that emotions such as fear can result in defensive behaviors. That is, when agents feel fear, they tend to demonstrate self-protective behavior. Chen et al. (2017) described the relationship between emotion and human behavior using a logistic function. An individual with a large fear factor tends to choose switch behavior with large probability, and an individual with a small fear factor tends to choose switch behavior with a small probability. Because emotion and switch behavior have a positive correlation, this paper assumes that when $FDI_i(t)$ of agent i at time t is large (closer to 1 than 0), agent i possesses a strong fear emotion and he/she will likely choose switch behavior. However, when $FDI_i(t)$ of agent i at time t is low (closer to 0 than 1), agent i has a weak fear emotion, so agent i will likely choose to do nothing (i.e., normal behavior).

3. Memory reception, fading, and cumulating model

3.1. Memory reception and fading

Based on the study by Shiffrin and Atkinson (1969), the memory fading process occurs at the same time as the information reception process. The main objective of this section is to establish the memory fading model for received disease information. Similar to Eq. (2.4), gained information $\gamma_i(t)$ of agent i over time period t can be divided into infected information and surrounding switching behavior information collected on the ITCN and DTCN, $\gamma_i(t)$ represents all information transmitted from the information sources to an agent. However, the collected information varies with time, so $\gamma_i(t)$ is defined as changes in disease information of agent i from time $t-1$ to t. For example, $\Delta I_i^{ITCN}(t)$ is equal to $I_i^{ITCN}(t)-I_i^{ITCN}(t-1)$, which represents the change of infected information in ITCN from time $t-1$ to time t.

$$\gamma_i(t) = \alpha\beta\Delta I_i^{ITCN}(t) + \alpha(1-\beta)\Delta I_i^{DTCN}(t) + (1-\alpha)\beta\Delta sw_i^{ITCN}(t) + (1-\alpha)(1-\beta)\Delta sw_i^{DTCN}(t) \tag{3.1}$$

$$\gamma_i'(t) = (\gamma_i(t) + 1)/2 \tag{3.2}$$

In real-world scenarios, individuals constantly perceive new information, and they tend to respond to and receive new cognition that occur infrequently (Reeves & Nass, 1996). During a disease epidemic, agents tend to receive strong stimulations of fresh information about the disease, thereby creating distinct contrasts in their memories. For example, when an agent receives new that breast cancer can be contagious (even it is not true), the perception regarding this new information is strong, so the agent will diligently consider and remember this information. In addition, models presented in this paper incorporate a discount factor between received disease information $\gamma_i(t)$ and an individual's perception of disease information, defined as $PDI_i(t)$, since people typically overlook minor trifles and overemphasize crucial issues in their minds. Therefore, this research used the Hill equation as the discount between $\gamma_i(t)$ and $PDI_i(t)$. Since free ligand concentration must be positive, a transformation was applied in Eq. (3.2) to maintain positive received disease information $\gamma_i(t)$.

The Hill equation (as shown in Fig. 3.1), widely used in biochemistry and pharmacology, describes the fraction variation of a macromolecule in the molecular binding process (Coval, 1970). Zhao et al. (2018) first introduced the Hill equation to describe memory perception rate in the disease transmission process. In their assumption, the combination of new disease information and memory of past information is similar to the macromolecule binding process. The reasonable boundary of the Hill equation is between 0 and 1, this number from the hill equation represents the rate of information learning. The memory fading process is shown in Eq. (3.4), where ε is the forgetting rate, meaning that an individual could lose his/her memory of the disease over time. Disease information processing of agent i $\mu_i(t)$ represents adjusted disease information after memory fading. c is an adjustive constant ($c = 0.5$) with a range of $-0.5 \leqslant H(\gamma_i(t),n) \leqslant 0.5$ since $\mu_i(t)$ is the parameter to represent new information about an epidemic and can be positive or negative, indicating whether the disease prevalence could be aggravated or mitigated, respectively.

$$H(\gamma_i(t),n) = \frac{(\gamma_i(t))^n}{(K)^n + (\gamma_i(t))^n} - c \tag{3.3}$$

$$\mu_i(t) = H(\gamma_i(t),n) - \varepsilon \tag{3.4}$$

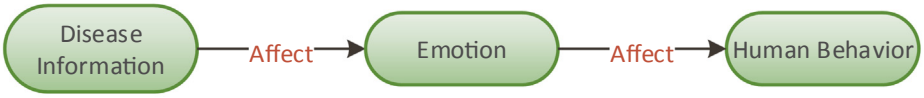

Fig. 2.3. Process of disease information affects human behavior.

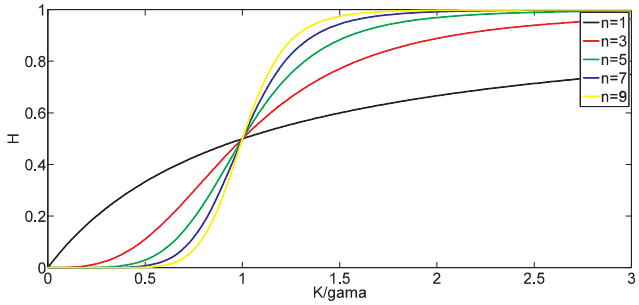

Fig. 3.1. Hill equation.

$H(\gamma_i(t),n)$ in Eqs. (3.3) and (3.4) represents newly gained disease information via a revised Hill equation; $\mu_i(t)$ is the disease information processing of agent i, which is a process variable between newly gained information $\gamma_i(t)$ and perceived disease information $PDI_i(t)$. K is the equilibrium constant and $n > 1$ is the Hill coefficient assumed to have a positive cooperative influence (binding) property in this application.

3.2. Information cumulation

Because disease-related information may vary throughout an epidemic, an individual's perceived disease information, $PDI_i(t)$ also changes during the epidemic. As mentioned, an individual's perception of disease information contains both learning and forgetting processes; that is, the perception process is affected by an agent's previous memory and newly acquired disease information. Wakefield, Loken, and Hornik (2010) found that information reporting by mass media leads to the behavior change because media can disseminate information to many agents, subsequently influencing social networks and agents' decisions (Alvarez-Zuzek, La Rocca, Iglesias, & Braunstein, 2017; Pires & Crokidakis, 2017). In addition, agents typically pay minimal attention to switch-protective behaviors in an epidemic when people have access to media coverage of the disease (Bagnoli, Lio, & Sguanci, 2007). Zhao et al. (2018) proposed that memory accumulation and fading for disease information can be found using a stochastic differential equation of Itô drift-diffusion process, including a drifting factor and random walk, to predict when an individual switches to a protective behavior. This research identifies perceived disease information $PDI_i(t)$ as disease prevalence information remembered by agent i at time t. As described in Section 3.1, the processing disease information $\mu_i(t)$ represents the difference between gained information and perceived disease information. The larger the difference of the current disease information, the larger the value of the processing disease information $\mu_i(t)$, meaning that $\mu_i(t)$ is the change magnitude of disease information used to represent a drifting factor in the Itô drift-diffusion process. An uncertain factor, also present during an epidemic due to population diversity and uncertainties, can be modeled as a random walk in the Itô drift-diffusion process. Similarly, in this paper, we assumed that the stochastic process of perceived disease information $PDI_i(t)$ in the MRFC model is an Itô drift-diffusion process as

$$dPDI_i(t) = \mu_i(t)dt + \sigma dZ_t \tag{3.5}$$

where $Z = \{Z_t: t \in [0,\infty)\}$ is standard Brownian motion with a mean of 0 and standard deviation of 1. $\mu_i(t)$ is a drafting factor that represents the processing disease information, and $\sigma(t)$ represents the variance of randomness in the population. The initial value $PDI_i(0)$ is calculated based on the definition in Eq. (2.4); the definition of $PDI_i(t)$ is shown in Eq. (3.6).

$$PDI_i(t) = PDI_i(t-1) + \int_{t-1}^{t} (\mu_i(t)dt + \sigma dZ_t) \tag{3.6}$$

Because an individual's memory and disease information vary over time, PDI can change due to memory fading of prior information and continuous new information updates. Therefore, final disease information $FDI_i(t)$ is not only the sum of all new information acquired by an individual i at time t, but it is also affected by faded memory of prior information. Considering the exponential smoothing memory fading method in Eq. (2.9), final disease information $FDI_i(t)$ in the MRFC model can be defined as

$$FDI_i(t) = w_0 PDI_i(t) + (1-w_0)FDI_i(t-1) \tag{3.7}$$

The entire process of disease information is shown in Fig. 3.2.

4. Agent-based modeling

4.1. Agent-based modeling

This section discusses agent-based modeling simulation of epidemic transmission to determine if IFC and MRFC models are effective in the real world. Because individuals have unique memories, moods, and behaviors, the unit of simulation must be individual, hence the use of the agent-based model. Agents are typically categorized into four types: switch susceptible, normal susceptible, infected, and recovery. An infected agent can contaminate a nearby switch/normal susceptible agent with a varying infection rate based on the switching behavior of susceptible agents. Switch susceptible agents have lower infection rates than normal susceptible agents because they choose behaviors to protect themselves from disease. An infected agent has a probability of recovering after reaching the recovery period, and an infected agent becomes a recovery agent after completing the recovery process (as shown in Fig. 4.1).

The agent-based epidemic model in this research references the basic susceptible, Infectious and Recovered (SIR) model framework built by Kermack and McKendrick (1927). Kermack-McKendrick theory assumes no births, deaths, or travel into or out of the population and that every agent has an equal chance of contacting with any other agent. In terms of an epidemic virus, the model assumes no dormant and latent periods in the disease. Viral mutation is not considered.

The agent-based simulation model primarily defines information reception, memory fading, and behavior switching through the IFC and MRFC models. Upon model initialization all agents are randomly arranged spatially in two dimensional (2-D) simulation space. When the simulation begins agents randomly move to nearby areas or remain at their current location. Disease information is calculated based on the

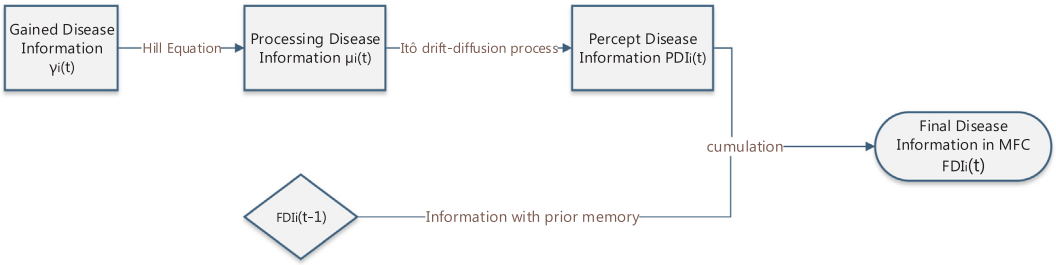

Fig. 3.2. Flowchart of disease information in the MRFC model.

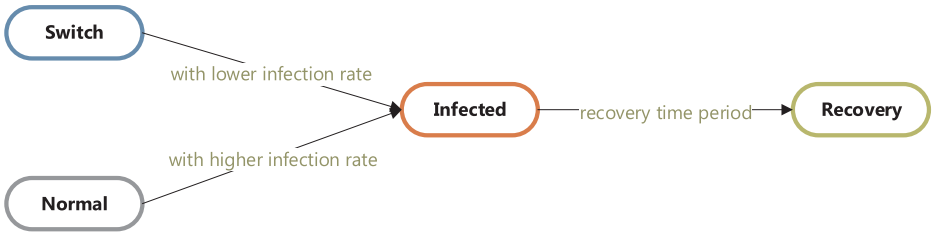

Fig. 4.1. Flowchart of disease transmission.

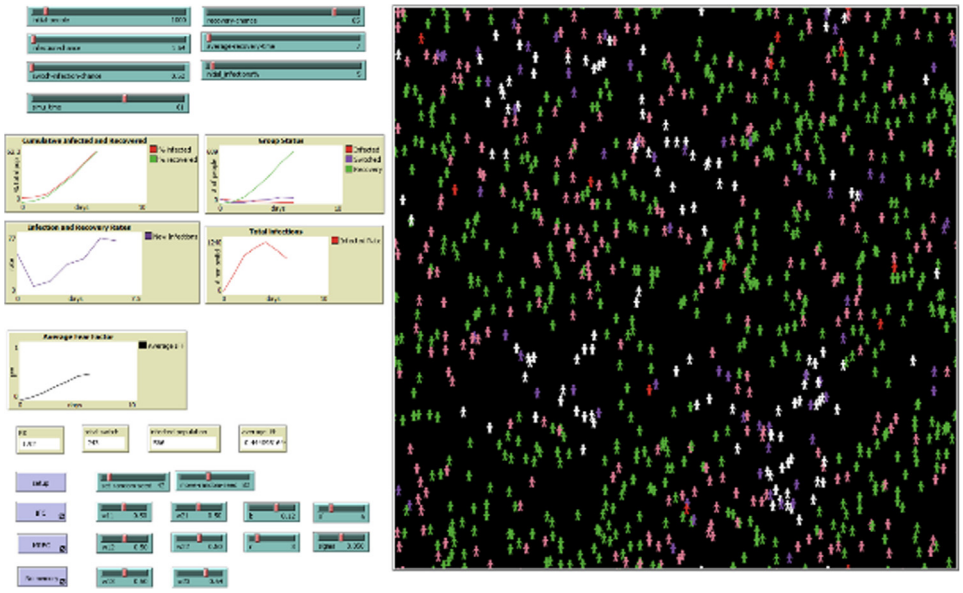

Fig. 4.2. Netlogo model and simulation GUI.

prevalence and switch-behavior information at each time epoch following agent movement, and then susceptible agents reconsider whether or not to choose switching behavior based on the updated $PDI(t)$ calculated by the IFC or MRFC models. Infected agents can infect nearby susceptible individuals at a certain infection rate. The simulation terminates when all infected agents reach recovery, signaling the end of the epidemic. Otherwise simulation continues until the given maximum simulation time.

Agent-based models have gradually become mainstream due to rapid advancements in hardware and software computing power. Several agent-based simulation and modeling environments, such as Swarm, Mason, and NetLogo, were developed to help researchers study detail behaviors of their models. Netlogo is the most researcher-recommended software because of the friendly programming interface and ease to code (Wilensky, 1999). Netlogo software also includes a library with a large amount of example models. Based on these advantages, this research utilized Netlogo as the simulation platform.

The simulation framework in this paper is based on the epiDEM framework (Yang & Wilensky, 2011), an existing example model in the Netlogo library. The simulation is innovative because it defines switch-susceptible agents, which were not considered in the original Netlogo library Moreover, the IFC and MRFC models are embedded into the model to determine if susceptible agents will switch to a protective behavior. This model also highlights susceptible population average fear factors that can be used to describe psychological emotion variations during epidemics.

Simulation results reported total switch populations and accumulated infections results. The initial testing simulation model was set in a 51 × 51 2-D grid. Each agent could randomly move to the grid nearby or stay at the current position, and each susceptible agent could be infected if and only if exist infections existed in his/her surrounding grids, referred to as the DTCN. In addition, the agents were assumed to receive disease information (infection rate and switch rate) from the entire simulation map, referred to as the ITCN. The recovery time for each agent followed a normal distribution (assume $\mu = 30$, $\sigma = 7.5$) (Spencer & Jones, 2003), and following the recovery time, each infected individual became a recovered individual and was not infected nor reinfected. This simulation also assumed that the infected rates for normal and switched-susceptible agents are 15% and 5% to clearly embody effects of the protective measures.

The simulation contained an initial population of 1000 agents randomly placed on the simulation map. Among those agents, 5% was randomly selected to be infected individuals based on the binomial distribution and 10% was randomly selected to be the switched population based on binomial distribution. No overlapping occurred between switched and infected agents. The white agent in the graphical user interface (GUI) represented non-switched susceptible agents, purple agents represented switched susceptible agents, red agents (also orange and pink agents in simulation of Section 5) represented infections, and green agents represented recovered agents.

Simulations were run on a workstation equipped with an Intel-based central processing unit i7-6700 K and 32 GB memory (RAM). System parameters such as infection rates or specific model parameters could be set manually prior to the simulation runs. The primary objective of the simulation was to monitor the trends or changes of crucial system variables (e.g., switched population, infected population, average fear factor) (see Fig. 4.2.).

4.2. Sensitivity analysis

Although Sections 2 and 3 of this paper detail IFC and MRFC models and Section 4 presents Netlogo models and simulation setups to study model behaviors, the significant influences of parameters such as infected information weight α, were not yet discussed. This section

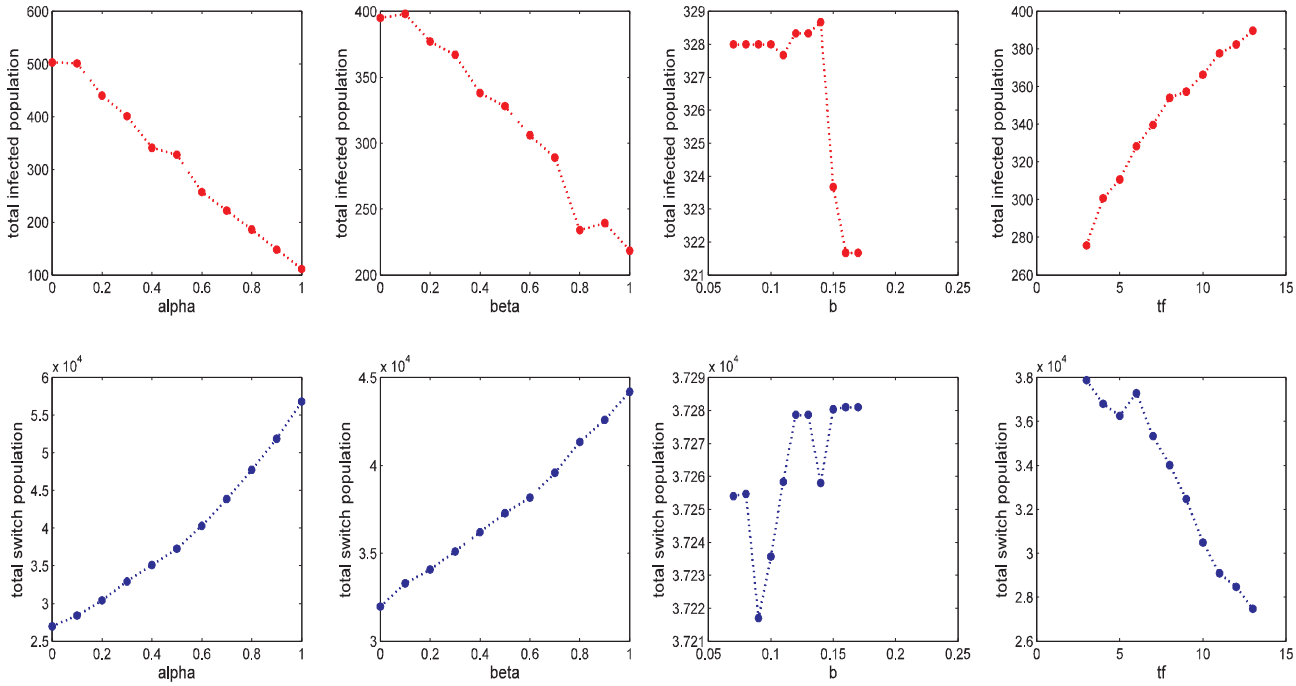

Fig. 4.3. IFC model sensitivity analysis.

utilizes the OFAT method to analyze the sensitivity by changing the setting of each parameter in the IFC and MRFC models. The main goal is to analysis model behaviors caused by altered parameter settings. In addition, parameter ranges also reflect the limit application range of crucial variables. The application range can be used to determine model flexibility.

Three fixed random seed values were used in each simulation to eliminate effects from the random process and ensure the presence of only one variable in each sensitive analysis simulation. Figs. 4.3 and 4.4 show the average number of three example simulation results (with common random seeds). The common random seeds setting ensures the analyzed parameter is the only variable.

The IFC model contains three crucial parameters: infected information weight α, behavior switch information weight β (Eq. (2.4)), and memory performance proportion w_t (Eq. (2.5)), which is determined by the rate of forgetting b and longest memory epoch t_f (Eq. (2.10)). Therefore, this section considers parameters α, β, b, and t_f (set as variables of information reception and forgetting process in the Netlogo simulation environment) by changing the value of parameters in their corresponding reasonable range and then analyzing the varying tendencies of infected rate and a total population of switch.

Sensitivity analysis results of the IFC model are shown in Fig. 4.3.

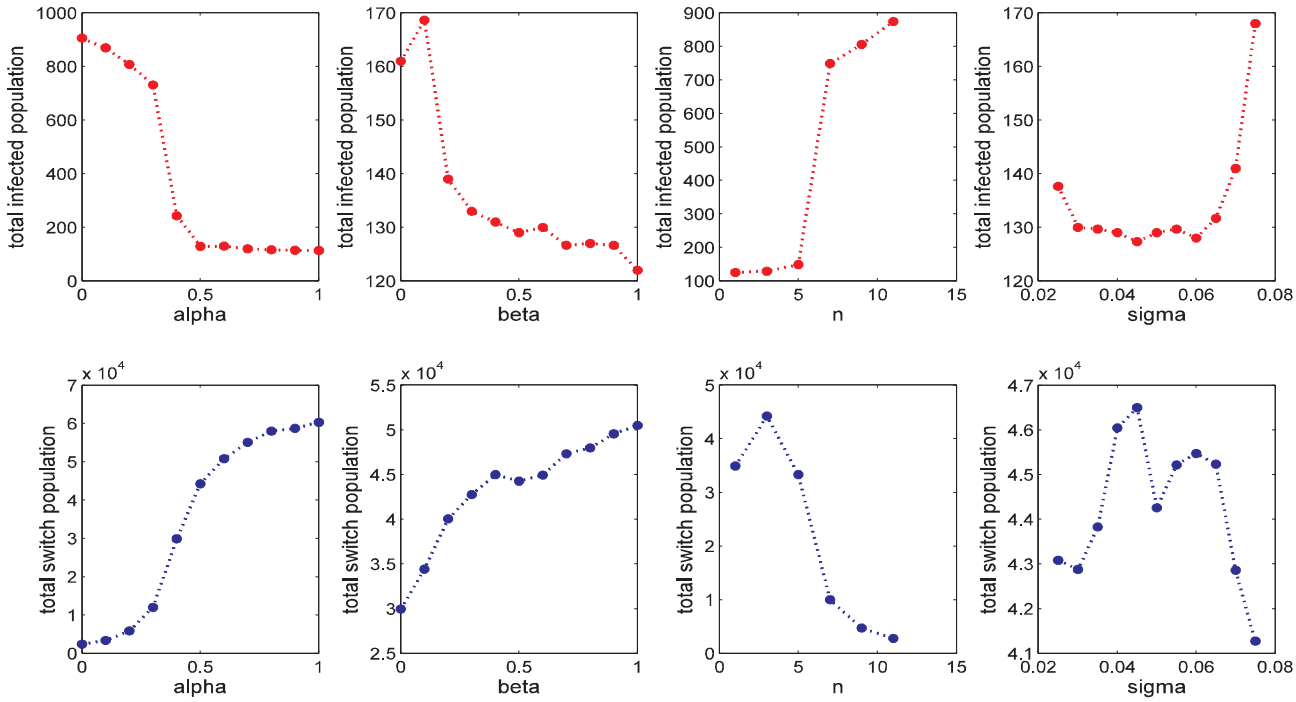

Fig. 4.4. MRFC model sensitivity analysis.

The red and blue lines illustrate variations of cumulated infections and total switch population, respectively, corresponding to parameter setting variations. As shown, cumulated infections decreased and switch population increased with the growth of α,β, and b. In addition, the variation range resulting from changing β was more pronounced than the variation ranged when α was changed, proving that β is more sensitive to total switch population than α. For the longest memory epoch t_f, however, the t_f cause the lower total switched population and higher infections. In general, α,β,andt_f are sensitive, but b is not sensitive.

The MRFC model contains six parameters: infected information weight α, behavior switch information weight β (Eq. (2.4)), the power of Hill equation n, equilibrium constant K (Eq. (3.2)), forgetting process constant ε, and maximum random range σ (Eq. (3.4)). In addition to α and β, n and σ are also advantageous for sensitivity analysis because n determines the information reception level and σ determines the randomness level. However, K and ε identify the learning and forgetting process, which means K and ε should be fixed at the simulation initiation.

MRFC model sensitivity analysis results for the total infected population and the total switched population are shown in Fig. 4.4. Similar to the IFC model, cumulated infections tended to decrease and switch population increased with the growth of α and β. However, the MRFC model was more sensitive to α than β, contrary to results from the IFC model. For the parameter n, cumulated infections and switch population were very sensitive when $n \in [3,7]$. For the variation parameter (σ), the cumulated infections and switch populations became more random as σ increased, resulting in a population diversity parameter σ below 0.6, which was not a significant parameter for system sensitivity.

In order to determine whether the IFC or MRFC model is more sensitive to parameter changes, this study compared sensitivity analysis results of the models shown in Figs. 4.3 and 4.4. Comparison showed that increasing α and β resulted in a decrease of total infections and total switching populations in both models. In addition, parameter α demonstrated greater sensitivity than β in the models. Table 4.1 presents the covering range of total infected and total switched populations when common parameters α and β were changed in both models. Sensitivity levels of the MRFC model were generally higher than the IFC model, the only exception is the total switch population changing by β. The infected population in the MRFC model varied from 11.3% to 90.6% (95% confidence interval [6.5%, 93.9%]), and the switched population ranged from 2.4 to 60.3 thousands (95% confidence interval [1.2, 79.7] thousands). Infected population variation was much smaller for the IFC model, ranging between 11.1% and 52.2% (95% confidence interval [8.2%, 67.7%]). Similarly, the switched population varied only between 26.9 and 56.8 thousands (95% confidence interval [16.2, 68.3] thousands). The MRFC model demonstrated less robust to the parameter changes than the IFC model when modeling general epidemics because the model is more sensitive to parameter settings. Therefore, the MRFC model could be used to model epidemics with high variations of infected and switched population ranges.

The second part of sensitivity analysis focuses on infection track variations when crucial parameters are changed. Using OFAT sensitivity analysis, the number of infections was inversely correlated to the number of switch populations, eliminating the need to track percentages of both the infected population and the switched population. Therefore, this portion of analysis focused on the percentage of infection population, which more accurately reflects the severity level of epidemics. Simulation comparison is shown in Fig. 4.5. For the IFC model, all shapes of tracks with different parameter settings were similar. The total infected population throughout the epidemic decreased when α and β increased, thereby corresponding to OFAT sensitivity analysis results. In comparison, the tracks of infection population percentage with various parameter settings demonstrated multiple shapes in the MRFC model. Parameter α was the key to controlling the midterm (approximately Day 25) epidemic performance of the model. Therefore, if agents focus more on infection information rather than switch behavior information (with higher α), total infections will decrease. However, parameter β significantly determines epidemic performance of late periods (approximately Days 25–50), meaning that increased attention to local disease information rather than global information will decrease total infections in the second half of the simulation. The second part of the sensitivity analysis proved that the MRFC model has greater flexibility than the IFC model in the range of infections and switched populations and in the shape variations in infection population percentage tracks.

4.3. Simulation comparisons

This section describes simulation runs to compare IFC and MRFC model performances, including use of a no memory model as the baseline in each comparison. Populations of each type of agent and percentages of switched population were considered for each simulation run. Common random seed and other parameters ($\alpha = \beta = 0.5$) were used to ensure that variations occurred due to different modeling methods only. Simulation run times were set to 70 days, although in most cases in our simulations, the epidemic ended prior to that limit. All populations, percentages of infected/switched populations, and average fear factors were compared, as illustrated in Fig. 4.6. Populations contain the number of infected, susceptible, switched, and recovery agents, and infected/switched rates are defined as the infections/switch-susceptible populations divided by the total non-cured population. The average fear factor represents the average number of fear factors for the entire susceptible population.

As shown in Fig. 4.6, infected individuals in the IFC model totaled 914, 847 in the MRFC model, and 902 in the no-memory model. The total switched population in the IFC model was 3299, 3978 for the MRFC model, and 4642 for the no-memory model. In general, the IFC and MRFC models demonstrated similar performances, although the MRFC model showed relatively more switched-susceptible agents in the mid stage of the epidemic (10–20 days). Because in the MRFC model, the process to gain information can immediately detect epidemic variation information. Therefore, the MRFC model can be applied to highly lethal or infectious epidemics, such as the SARS epidemic in 2003. Agents typically pay close attention to new epidemic developments, thereby causing rapid, elevated fear emotions in the population and motivating a majority of individuals to take protective measures even if only a minimal number of infections has been reported.

Both the IFC model and the no-memory model showed relatively lower switched behavior than the MRFC model, making them better-suited to model epidemics with higher numbers of infectious and

Table 4.1
Comparison of model flexibilities based on sensitivity analysis data.

	Range of infections by changing α	Range of infections by changing β	Total switch population range by changing α	Total switch population range by changing β
IFC model	11.1–52.2%	21.8–41.4%	26.9–56.8 thousands	31.9–44.2 thousands
95% confidence interval	8.2–67.7%	4.2–63.4%	16.2–68.3 thousands	20.2–53.4 thousands
MRFC model	11.3–90.6%	12.2–16.8%	2.4–60.3 thousands	29.9–50.5 thousands
95% confidence interval	6.5–93.9%	9.4–20.8%	1.2–79.7 thousands	17.1–61.6 thousands

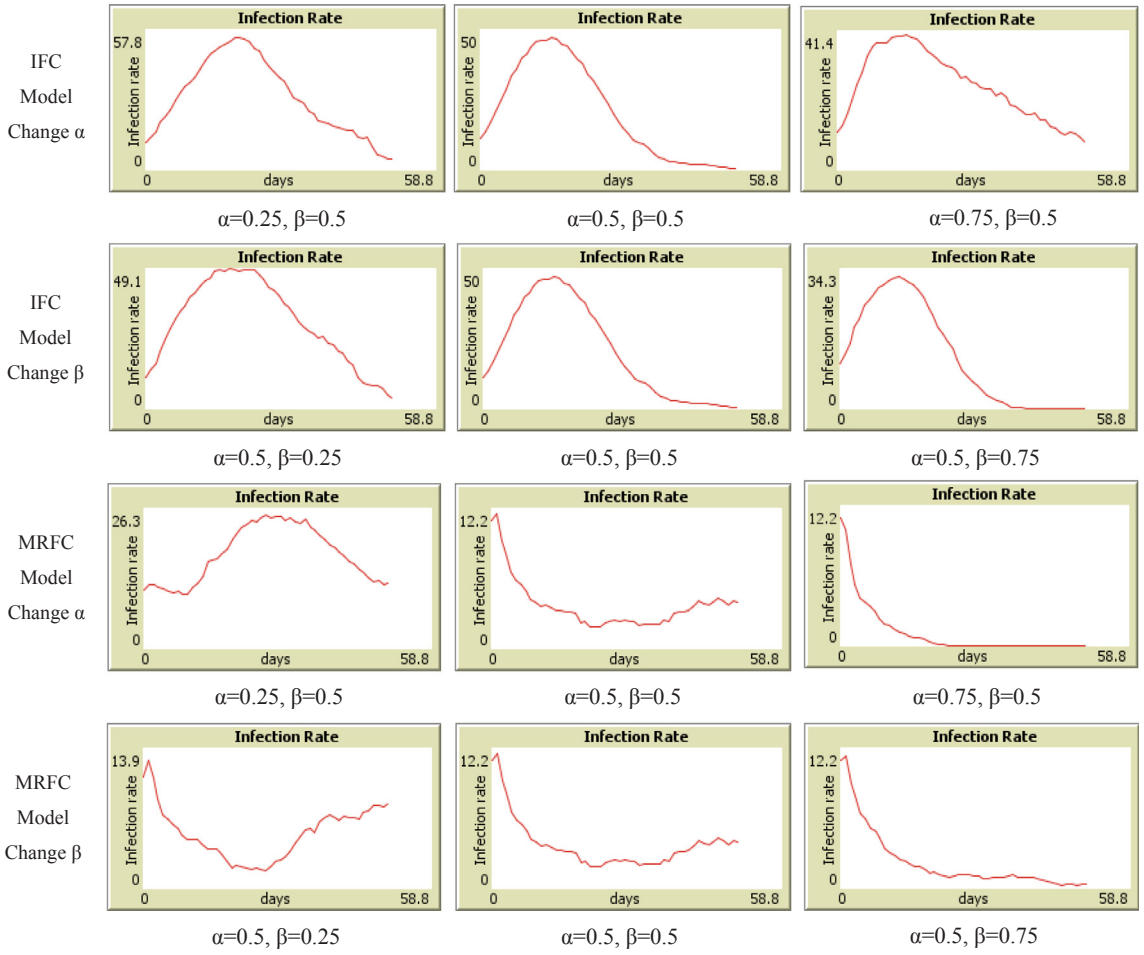

Fig. 4.5. Sensitivity analysis of infection population percentage tracks for IFC and MRFC models.

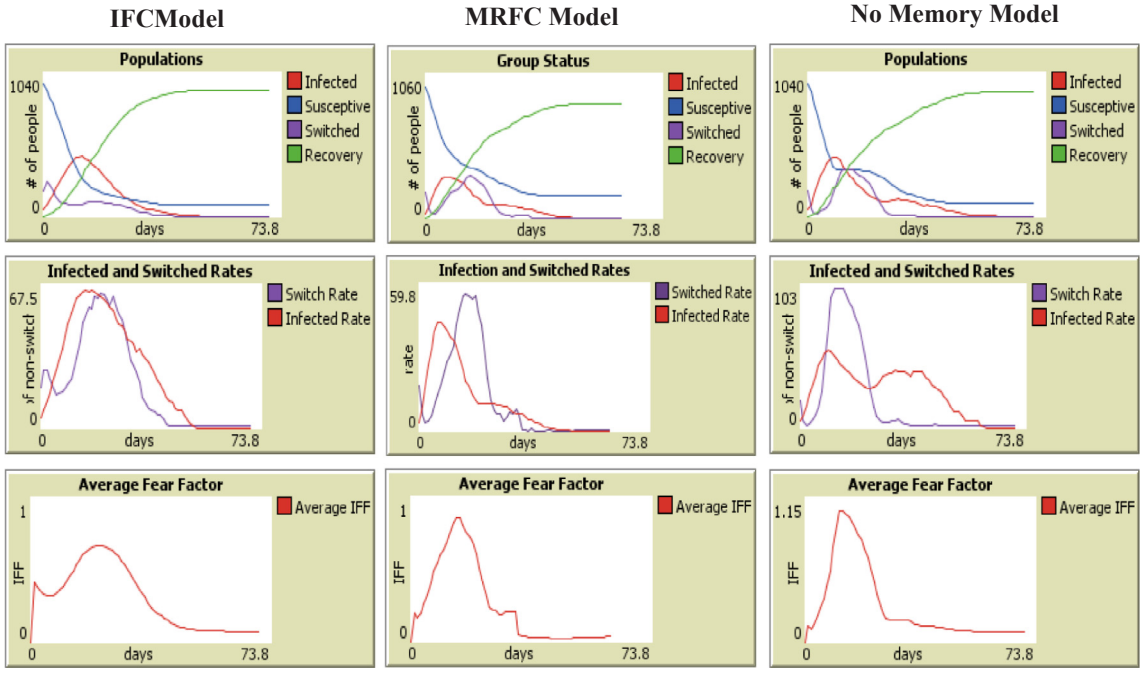

Fig. 4.6. Comparison of IFC, MRFC, and no-memory models.

minimal lethal situations, such as an influenza epidemic, in which people frequently pay less attention to new disease developments because the risk of death or other adverse consequences are relatively low. Although overall performances of the IFC model and no-memory model were similar, the tendencies of infected population percentages differed significantly. The switched population in the IFC model followed the decline of the infected population approximately 30 days after the onset of the epidemic. However, in the no-memory model, the infected population demonstrated a second peak around 40 days due to the quick decline of the switched population. Because agents in the IFC model retain memory of past epidemic information, they maintained their switching behavior for an extended period even as the epidemic neared completion. Nevertheless, the susceptible population in the no-memory model does not retain past disease information, so a majority of individuals switched behavior back to normal immediately, allowing the epidemic to spread.

5. H1N1 case study

5.1. H1N1 pandemic in 2009

H1N1, also known as swine flu (Peiris, Poon, & Guan, 2009), is a form of influenza in pigs that can be transmitted to people by exposure to infected droplets. H1N1 is an orthomyxovirus, a subtype of influenza A that is the most common cause of seasonal human flu. H1N1 viruses attack the human immune system, attaching and replicating within infected cells. A person infected by the H1N1 virus will develop a progressive lower respiratory tract disease that could result in respiratory failure (Rello et al., 2009). The first human case of H1N1 virus was reported in Mexico in 2009, quickly spreading to the United States and the world and resulting in a pandemic outbreak (Archer et al., 2009). Determination of H1N1 infection is difficult based on symptoms because influenza symptoms are nonspecific, typically lasting four to six days with an effective infection period continuing for approximately seven days (Centers for Disease Control, 2009).

Prior to the outbreak of the H1N1 virus pandemic in 2009, minimal information was available about the disease and people had limited or no disease awareness or past memory. Immediately following the initial disease outbreak, people began to understand and focus on the new strain of influenza A. Therefore, in order to more accurately describe how people learned about and forgot H1N1 virus pandemic information, this research utilized data from the pandemic from several days (5 days) after the outbreak since people began to have memories about the disease after that time.

Many reports documented the H1N1 virus pandemic in 2009 (Plennevaux, Sheldon, Blatter, Reeves-Hoché, & Denis, 2010). The first graph in Fig. 5.1 illustrates pandemic H1N1 virus infection among New York City residents hospitalized from May 29 to July 1, 2009. At the preliminary stage of the pandemic, people had minimal memories of the disease and paid limited attention to it, resulting in an increase in infection cases. After several days of infection, people developed increased awareness and memory about the disease and began to take preventive measures; at the end of June 2009 the number of infected cases gradually declined. Pandemic prevalence showed a similar trend in Chicago from June 7 to July 10, 2009, and in Shanghai, China, from June 29 to July 29, 2009, the number of infection cases initially increased and then decreased during July. Mexico City showed a similar trend from April 17 to May 17, 2009. This research used the graphs in Fig. 5.1 to determine whether the learning and forgetting phenomenon affected the pandemic trend in the four cities.

Table 5.1 shows the proportional distribution of infection cases based on age group and total population in four cities. From May 1 to June 30, 2009, the highest infection rate occurred in students 5–19 years old and adults 20–64 years old. The infection rate for students was approximately three times higher than the infection rate for infants under 4 years old, with the exception of Mexico City. The numbers of infected cases in urban areas were significantly higher than non-urban areas.

In order to determine if receiving and forgetting information during an epidemic can influence the spread of disease, this paper assumes that people can both learn new information and forget prior information over time. The IFC and MRFC models were applied to simulate an outbreak of the H1N1 virus pandemic in 2009 for the Chicago using report data to determine if the IFC model or MRFC model can better reproduce infection population trends throughout the pandemic.

5.2. Infection rate calculation with historical epidemic data

Infection rate is the probability of a susceptible agent to be infected at time t. It can be defined mathematically as the new infected population divided by the susceptible population (Utah Department of Health, 2017):

$$\text{Infection rate} = \frac{new\ infections\ (t)}{susceptible\ agents\ in\ risky\ (t)} * K \tag{5.1}$$

In Eq. (5.1), *new infections*(t) represents infections caused by the original infection after last time epoch, and K is a constant used to adjust the infection rate within different time periods. For example, if the infection rate is calculated based on a one-day period, the infection rate in 10 days can be determined by assigning a value of 10 to K. *Susceptible agents in risky*(t) have face-to-face contact with the infected population in their DTCN. Even if the data of the original infected population are known, however, accurate recording of the number of potential contacts for each infected agent is nearly impossible.

This section uses three methods to estimate the average number of contacts for an infected individual. The first method, the average-contact-based method, utilizes the estimated average number of contacts from the report. Mossong and et al. (2008) recorded physical contact behavior for 7290 participants from eight countries with various age ranges. The average number of contacts for these participants was 13.4 per day (standard deviation was 10.6).

The second method, the age-ranges-based method, is based on the average number of contacts from various age ranges. Valle, Sara, et al. (2007) utilized U.S. Census Bureau data from 2000 to analyze contacts per person. They found that the adult group (between 20 and 60 years old) had the highest number of contacts (approximately 20) per day. Children and elderly groups had the lowest number of contacts (approximately 10). Therefore, the susceptible population was defined as $\sum_i I_i(t)c_i$, where i is the corresponding age group I, $I_i(t)$ represents how many infected individuals are present in age group i in time t, and c_i represents the average number of contacts in age group i.

The third method, the urban-population-percentage-based method, considers the average number of contacts in areas with various population densities. Read and et al. (2014) researched the relationship between daily contacts and population densities. They found that urban citizens have a higher number of contacts than rural citizens. This method most accurately reflects the epidemic performance difference between rural and urban populations.

Infection rates were calculated at each time epoch in each city using these three methods. Table 5.2 presents the average numbers with 95% confidence intervals of the infection rates. Among the four cities, Shanghai had the highest infection rate and Chicago had the lowest infection rate, a difference that could be attributed to the areas' different population densities and epidemic outbreak locations. However, the infection rate of H1N1 in 2009 remained approximately 1–2%. This study used the average-contact-based method to calculate the infection rate in the case simulation.

This research also investigated whether behavior switching reduces the chances for infection. Unfortunately, no data reported how many agents protected themselves from H1N1 infection using switch behavior such as protective masks, vaccinations, or staying at home. To most accurately determine the H1N1 infection rate for behavior switched

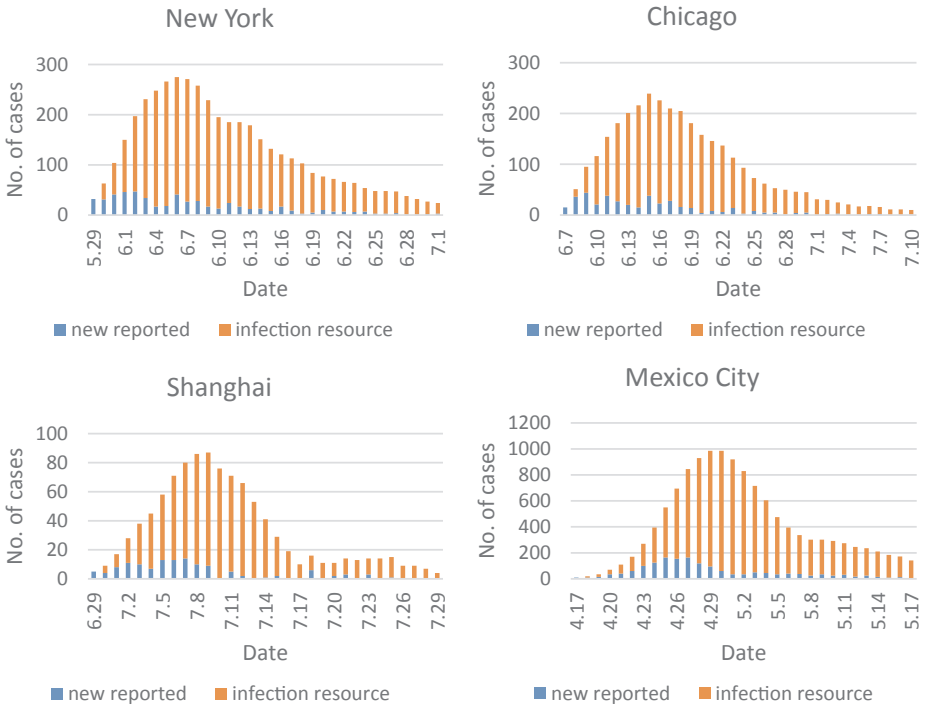

Fig. 5.1. H1N1 infection data in 2009.

population, therefore, this research used existing reference data showing switch behavior effectiveness to be 50–86% (Bridges, 2000; Hurwitz, 2000; Osterholm, 2012). This case study used the median of the range as the effectiveness of the switched behavior, meaning the infection rate of behavior-switched agents was approximately 33% of the normal infection rate.

5.3. Simulation of 2009 Chicago H1N1 case

As mentioned, Chicago was one of the cities which severely afflicted by H1N1 in the 2009 virus pandemic. This mega city has an approximate population of 2.705 million people (third largest in the United States) and a population density of 4582.3 people per km^2 (fourth larges in the United States) (Gibson, 1998). Chicago has 588.26 km^2 of the land area, which has a strip-shaped, but the northeast portion of Chicago has the highest population density. The fast-paced urban lifestyle and highly concentrated population in this area contributed to the extensive H1N1 outbreak. This section uses Chicago as a special H1N1 case to conduct epidemic simulation using historical data and analyzing the fear emotion with learning/forgetting behavior.

The Chicago H1N1 virus epidemic simulation was set in a $60 * 120$ grid 2-D space with each grid representing $0.2858 * 0.2858 = 0.0817$ km^2. A total of 2705 agents were located in the simulation space, and each agent represented a group of 1000 individuals. Since June 7, 2009, there were 25 H1N1 reported infections, or the initial infected agents (red) were 1% of the total agent groups; each infected group had one infected person. Initial switched-behavior agents were assumed to comprise approximately 3% of the total population. All infected, switched, and susceptible agents were randomly placed according to population densities in Chicago, as shown in Fig. 5.2.

Table 5.2
Calculated infected rates using three methods in four cities.

Regions	Average-contact-based method	Age-ranges-based method	Urban-population percentage-based method
New York	1.30% [1.05–1.54%]	1.08% [0.87–1.28%]	1.22% [0.99–1.45%]
Shanghai	2.39% [1.71–3.05%]	1.88% [1.36–2.41%]	2.22% [1.60–2.85%]
Chicago	1.08% [0.91–1.25%]	1.01% [0.85–1.16%]	1.16% [0.98–1.34%]
Mexico City	1.90% [1.12–2.69%]	1.59% [0.93–2.24%]	1.86% [1.09–2.62%]

Table 5.1
Infection characteristic percentages by cities.

Regions	Age group of infection percentage				Urban/non-urban percentage	
	Infant aged 0–4	Student aged 5–19	Adult aged 20–64	Elder aged > 64	Urban	Rural
New York	6.7% (Lee & Wong, 2010)	20% (Lee & Wong, 2010)	58.8% (Lee & Wong, 2010)	14.5% (Lee & Wong, 2010)	87.9 (Iowa State University, 2010)	12.1 (Iowa State University, 2010)
Shanghai	16% (Shen & Lu, 2010)	51% (Shen & Lu, 2010)	30% (Shen & Lu, 2010)	3% (Shen & Lu, 2010)	59 (Chan, 2007)	41 (Chan, 2007)
Chicago	3.8% (Centers for Disease Control, 2009)	34.2% (Centers for Disease Control, 2009)	60% (Centers for Disease Control, 2009)	2% (Centers for Disease Control, 2009)	89.3 (Newgeography, 2011)	10.7 (Newgeography, 2011)
Mexico City	16% (Perez-Padilla & et al., 2009)	20% (Perez-Padilla, 2009)	56% (Perez-Padilla, 2009)	8% (Perez-Padilla, 2009)	78.8 (Ruiz-Rivera, Suárez, & Delgado-Campos, 2016)	21.2 (Ruiz-Rivera et al., 2016)

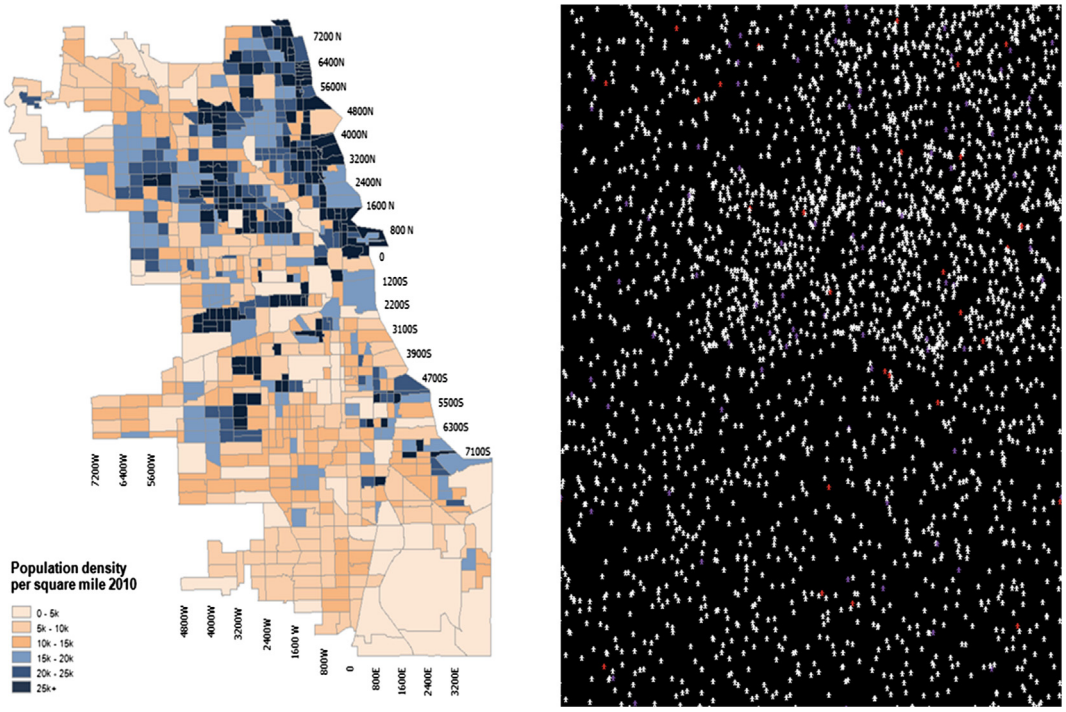

Fig. 5.2. Comparison of population density map with simulation initial setting in Chicago.

Simulations runs were set up with identical initial settings and random seeds, but with two models (the IFC and MRFC models). Simulation results related to infection source, average fear factor, and switch groups in the two models are shown in Fig. 5.3. The simulated time period was June 7–10, 2009 (33 days). In the IFC model simulation, although the total number of infections (492) was slightly higher than the real situation (410), the tendency of the epidemic was very similar to the historical data. In addition, the average fear factor (not shown in Fig. 5.3 due to space limitations) in the susceptible population and the number of individuals in the switched-behavior group (Fig. 5.4) were similar to the prevalence data reported in 2009. However, simulation results of the MRFC model did not correlate well with the historical data. The peak of the H1N1 outbreak occurred around June 25 in the MRFC simulation, when the real-world epidemic was nearly finished, and the switched-group number was nearly half of the total groups (Fig. 5.4) at the beginning of the prevalence, which is not realistic. In conclusion, results of the IFC model simulation runs fits better to the 2009 Chicago H1N1 historical data.

Fig. 5.4 uses GUI to show epidemic tendency in the IFC and MRFC models. In the interface, red, orange, and pink agents represent the infected group with 1, 2, and 3 infected people, respectively, white agents represent normal-behavior susceptible agents, violet agents are switch-behavior susceptible agents, and green agent groups contain recovered people.

On Day 6 of IFC model simulation runs only a few agent groups in the north part of the city were infected, and almost no agents switched their behavior to prevent infection. On Day 14 the epidemic area expanded into the entire downtown and northern areas of the city. With the increasing number of newly infected individuals, more and more agents gained awareness of the H1N1 virus epidemic from the information learning process. The number of switched-agent groups also grew significantly due to increasing fear of the H1N1 virus. On Day 22 the infected population was primarily located at the edge of highly populated areas, and the number of infected individuals began to decrease from the overall peak. Since memory fading requires several days (usually around 5 days in our simulation), the number of switched groups did not show significant decline. The prevalence nearly ended on Day 30 when the number of infected agents approached zero. Very

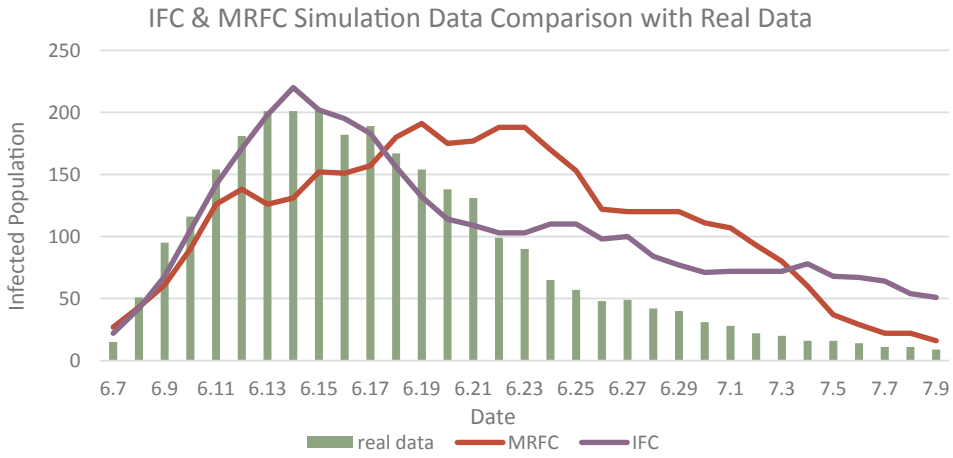

Fig. 5.3. Simulation results comparison of the IFC and MRFC models.

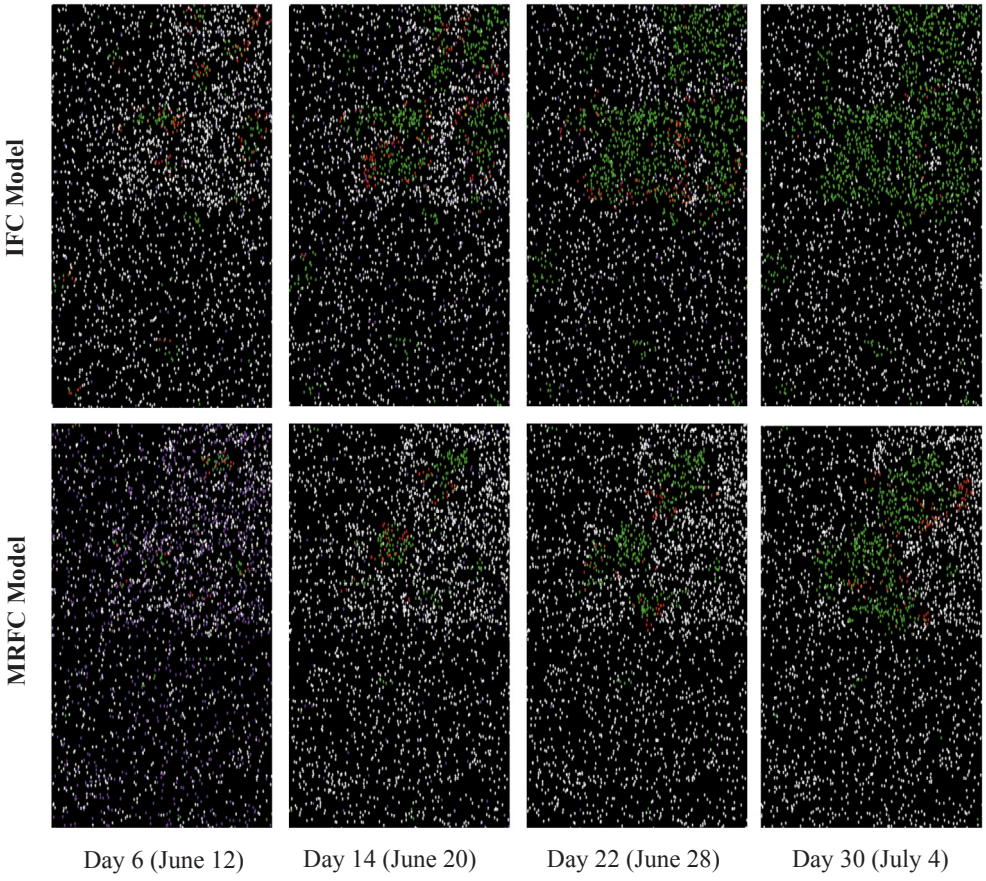

Fig. 5.4. IFC and MRFC model topographic chart for 2009 H1N1 in Chicago.

Table 6.1
Comparison of model characteristics.

	IFC Model	MRFC Model
Information source	ITCN (local) and DTCN (global)	ITCN (local) and DTCN (global)
Information types	Infection and behavior	Infection and behavior
Information cumulative process	Learning and forgetting	Learning and forgetting
Randomness process	No randomness	Randomness in Itô stochastic process
Information completeness	Based on complete disease information for each agent	Only based on Δ information at each time epoch
Information transformation	No information transformation	Information transformation by hill equation
Sensitivity and model flexibility	Relative normal sensitivity and flexibility	Relative higher sensitivity and flexibility
Relationship between infections and switched behavior	Relative higher relationship	Relative normal relationship
Real data simulation performance	Fit the tracks of epidemics	Obviously different from historical data
Potential scope of application	Highly infectious and low lethal epidemics	Highly lethal epidemics

few agent groups (around 10 percent) decided to switch their behavior, meaning that most agents had already forgotten that the H1N1 virus is a highly pathogenic influenza.

Results from the MRFC model simulation showed that most of the susceptible population decided to switch their behavior at the beginning of the epidemic (around Day 6) due to fears caused by sudden disease prevalence, thereby preventing the early epidemic diffusion. In the metaphase of the epidemic fewer people protected themselves by switching their behavior even though there was no strong signal of prevalence deterioration. The result of returning to normal behavior caused the epidemic outbreak to become more serious near the end of the simulation. However, simulation results from the MRFC model may not actually reflect characteristics of 2009s H1N1 virus scenario in Chicago in regards to epidemic infection tracks and behavior patterns of susceptible agents.

In general, the IFC model accurately recreated the complete process of Chicago's H1N1 epidemic in 2009. The time zone and infection tendency from this simulation resembled real historical data. Prevalence began in the northeast portion of the city and spread through downtown. Suburb areas were not influenced significantly from the H1N1 virus. The average fear factor and number of switched agents also followed the tendency of epidemics with a short delay (2–3 days), potentially indicating the time needed to process new disease information and allow for memory fading.

6. Summary

This paper investigated the assumption that disease information can influence an individual's fear emotion and that agents' emotions potentially affect behavior during an epidemic. This study used two mathematical models (IFC and MRFC models) to discuss disease information fading and learning processes. Both models synthesized disease information on local and global levels with infection information and switched-behavior information, thereby providing comprehensive

disease information to agents. However, the performances of the two models are different in modeling techniques, sensitivity levels and simulation performances (show in Table 6.1).

Modeling technique is the primary difference between the IFC and MRFC models. The IFC model assumes agents can obtain a complete picture of the epidemic via information from local daily contacts or global news coverage. This prevalence information affects the protective behavior of agents by changing their fear emotion level. However, the MRFC model assumes agents can detect discrepancies in disease information. The Hill equation transformation showed that agents usually ignore minor discrepancies and pay attention to major inconsistencies. Cumulative information transformation to knowledge was modeled mathematically by an Itô stochastic diffusion process.

The differences of modeling methods reduced the significant imparity in simulation results. The IFC model more accurately describes epidemics with high infectious ability and low lethality. Although the IFC model has less sensitivity and flexibility than the MRFC model, it more precisely restored the tracks of 2009s H1N1 virus epidemic in Chicago. Therefore, the MRFC model should be applied for highly lethal epidemics. Moreover, the simulation results show a weak relationship between infections and switched behavior in the MRFC model, which results that did not correlate to current data sets.

In conclusion, a dynamic agent-based model can be used to mimic real-world epidemic situations and explain disease transmission, behavior changes, and distribution of prevalence panic. Moreover, agent-based simulation with real data restored the historical H1N1 virus influenza data from Chicago in 2009. Therefore, the future work of this research should explore more in the areas of the potential applications of both agent-based epidemic model and forgetting and learning model in following areas: (1) Health Organizations and Disease Control Centers can utilize the model presented in this research to evaluate and experiment the possible impacts and influences using various control strategies. For example, the agent-based epidemic model can be used to forecast the effects of information dissemination through media and broadcast. The model presented in this research can also reflect the effectiveness of public health education regarding the underlining disease in the researched area. (2) Our model can also be applied to examine the self-protection ability of the general public during a spontaneous or unannounced epidemic. By simulation experiments aimed to the specific area, the disease outbreaks in a localized area can be detected and subsequently mitigated by the public health agencies. (3) The forgetting and learning model presented in this research can be applied to other areas of application to reflect the diversified opinions or spontaneous behaviors within a heterogeneous population. For example, *the* investor's diversified perspectives and investing patterns on the stock market or the popularity or the electability of a public figure in the commercial or political campaigns can be modeled using the forgetting and learning curve presented in this research."

References

Altizer, Sonia, et al. (2003). Social organization and parasite risk in mammals: Integrating theory and empirical studies. *Annual Review of Ecology, Evolution, and Systematics, 34*(1), 517–547.

Alvarez-Zuzek, L. G., La Rocca, C. E., Iglesias, J. R., & Braunstein, L. A. (2017). Epidemic spreading in multiplex networks influenced by opinion exchanges on vaccination. *PLoS ONE, 12*(11), e0186492.

Anderson, R. B., & Tweney, R. D. (1997). Artifactual power curves in forgetting. *Memory & Cognition, 25*(5), 724–730.

Archer, B. N., Cohen, C., Naidoo, D., Thomas, J., Makunga, C., Blumberg, L. H., … Schoub, B. D. (2009). Interim report on pandemic H1N1 influenza virus infections in South Africa, April to October 2009: epidemiology and factors associated with fatal cases.

Baddeley, Alan D. (1997). *Human memory: Theory and practice.* Psychology Press.

Badiru, A. B. (1992). Computational survey of univariate and multivariate learning curve models. *IEEE Transactions on Engineering Management, 39*(2), 176–188.

Bagnoli, F., Lio, P., & Sguanci, L. (2007). Risk perception in epidemic modeling. *Physical Review E, 76*(6), 061904.

Bailey, C. D. (1989). Forgetting and the learning curve: A laboratory study. *Management Science, 35*(3), 340–352.

Bansal, S., Grenfell, B. T., & Meyers, L. A. (2007). When individual behaviour matters: Homogeneous and network models in epidemiology. *Journal of the Royal Society Interface, 4*, 879–891.

Becker, N. G. (1989). *Analysis of infectious disease data, Vol. 33.* CRC Press.

Brainerd, C. J., Kingma, J., & Howe, M. L. (1985). On the development of forgetting. *Child Development,* 1103–1119.

Bridges, Carolyn Buxton, et al. (2000). Effectiveness and cost-benefit of influenza vaccination of healthy working adults: A randomized controlled trial. *JAMA, 284*(13), 1655–1663.

Carlson, J. G., & Rowe, A. J. (1976). How much does forgetting cost. *Industrial Engineering, 8*(9), 40–47.

Centers for Disease Control and Prevention, 2009. Interim guidance on infection control measures for 2009 H1N1 influenza in healthcare settings, including protection of healthcare personnel. Avaiable from: < https://www.cdc.gov/h1n1flu/guidelines_infection_control.htm > (last access: August 17th 2017).

Centers for Disease Control and Prevention, 2009. Pandemic Influenza A (H1N1) virus infections – Chicago, Illinois. < https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5833a1.htm > (last access: August 21st 2017).

Chan, K. W. (2007). Misconceptions and complexities in the study of China's cities: Definitions, statistics, and implications. *Eurasian Geography and Economics, 48*(4), 383–412.

Chen, F. H. (2009). Modeling the effect of information quality on risk behavior change and the transmission of infectious diseases. *Mathematical Biosciences, 217*(2), 125–133.

Chen, Y., Bi, K., Zhao, S., Ben-Arieh, D., & Wu, C. (2017). Modeling individual fear factor with optimal control in a disease-dynamic system. *Chaos, Solitons & Fractals, 104*(November), 531–545.

Coval, Myer L. (1970). Analysis of Hill interaction coefficients and the invalidity of the Kwon and Brown equation. *Journal of Biological Chemistry, 245*(23), 6335–6336.

Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology. Vol. 3.* University Microfilms.

Elm'Aghraby, S. E. (1990). Economic manufacturing quantities under conditions of learning and forgetting (EMQ/LaF). *Production Planning & Control, 1*(4), 196–208.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309.

Fishbein, Martin., Middlestadt, Susan E., & Hitchcock, Penelope J. (1994). *Using information to change sexually transmitted disease-related behaviors.* Springe, US: Preventing Aids61–78.

Funk, S., Gilad, E., Watkins, C., & Jansen, V. A. (2009). The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences, 106*(16), 6872–6877.

Gibson, C. (1998). *Population of the 100 largest cities and other urban places in the United States: 1790–1990.* Washington, DC: US Bureau of the Census.

Grassly, Nicholas C., & Fraser, Christophe. (2008). Mathematical models of infectious disease transmission. *Nature Reviews Microbiology, 6*(6), 477.

Hadeler, Karl P., & Castillo-Chávez, Carlos. (1995). A core group model for disease transmission. *Mathematical Biosciences, 128*(1), 41–55.

Hurwitz, Eugene S., et al. (2000). Effectiveness of influenza vaccination of day care children in reducing influenza-related morbidity among household contacts. *JAMA, 284*(13), 1677–1682.

Iowa State University, 2010. Urban percentage of the population for states, historical. < http://www.icip.iastate.edu/tables/population/urban-pct-states > (last access: August 21st 2017).

Jaber, M. Y., & Bonney, M. (1996). Production breaks and the learning curve: The forgetting phenomenon. *Applied Mathematical Modelling, 20*(2), 162–169.

Jaber, M. Y., & Bonney, M. (1997). A comparative study of learning curves with forgetting. *Applied Mathematical Modelling, 21*(8), 523–531.

Jaber, M. Y., Kher, H. V., & Davis, D. J. (2003). Countering forgetting through training and deployment. *International Journal of Production Economics, 85*(1), 33–46.

Kermack, William O., & McKendrick, Anderson G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 115*(772).

Kiss, I. Z., Cassell, J., Recker, M., & Simon, P. L. (2010). The impact of information transmission on epidemic outbreaks. *Mathematical Biosciences, 225*(1), 1–10.

Klovdahl, A. S. (1985). Social networks and the spread of infectious diseases: The AIDS example. *Social Science and Medicine, 21*, 1203–1216.

Lee, S. S., & Wong, N. S. (2010). Reconstruction of epidemic curves for pandemic influenza A (H1N1) 2009 at city and sub-city levels. *Virology Journal, 7*(1), 321.

Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E., & Getz, W. M. (2005). Superspreading and the effect of individual variation on disease emergence. *Nature, 438*, 355–359.

Mossong, Joël, et al. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine, 5*(3), e74.

Nahl, D., & Bilal, D. (2007). *Information and emotion: The emergent affective paradigm in information behavior research and theory.* Information Today Inc.

Newgeography, 2011. The evolving urban form: CHICAGO. < http://www.newgeography.com/content/002346-the-evolving-urban-form-chicago > (last access: August 21st 2017).

Osterholm, Michael T., et al. (2012). Efficacy and effectiveness of influenza vaccines: A systematic review and meta-analysis. *The Lancet Infectious Diseases, 12*(1), 36–44.

Peiris, J. M., Poon, L. L., & Guan, Y. (2009). Emergence of a novel swine-origin influenza A virus (S-OIV) H1N1 virus in humans. *Journal of Clinical Virology, 45*(3), 169–173.

Perez-Padilla, Rogelio, et al. (2009). Pneumonia and respiratory failure from swine-origin influenza A (H1N1) in Mexico. *New England Journal of Medicine, 361*(7), 680–689.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience, 9*(2), 148–158.

Pires, M. A., & Crokidakis, N. (2017). Dynamics of epidemic spreading with vaccination: Impact of social pressure and engagement. *Physica A: Statistical Mechanics and its Applications, 467*, 167–179.
Plennevaux, E., Sheldon, E., Blatter, M., Reeves-Hoché, M. K., & Denis, M. (2010). Immune response after a single vaccination against 2009 influenza A H1N1 in USA: A preliminary report of two randomised controlled phase 2 trials. *The Lancet, 375*(9708), 41–48.
Polgar, S. (1962). Health and human behavior: Areas of interest common to the social and medical sciences. *Current Anthropology, 3*(2), 159–205.
Read, Jonathan M., et al. (2014). Social mixing patterns in rural and urban areas of southern China. *Proceedings of the Royal Society of London B: Biological Sciences, 281*(1785), 20140268.
Reeves, B., & Nass, C. (1996). *How people treat computers, television, and new media like real people and places.* CSLI Publications and Cambridge.
Rello, J., Rodríguez, A., Ibañez, P., Socias, L., Cebrian, J., Marques, A., et al. (2009). Intensive care adult patients with severe respiratory failure caused by Influenza A (H1N1) v in Spain. *Critical Care, 13*(5), R148.
Ruiz-Rivera, N., Suárez, M., & Delgado-Campos, J. (2016). Urban segregation and local retail environments. Evidence from Mexico City. *Habitat International, 54*, 58–64.
Sahneh, F. D., & Scoglio, C. (2013). May the best meme win!: new exploration of competitive epidemic spreading over arbitrary multi-layer networks. arXiv preprint arXiv:1308.4880.
Shen, Y., & Lu, H. (2010). Pandemic (H1N1) 2009, Shanghai, China. *Emerging Infectious Diseases, 16*(6), 1011.
Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review, 76*(2), 179.
Sikström, S., & Jaber, M. Y. (2002). The power integration diffusion model for production breaks. *Journal of Experimental Psychology: Applied, 8*(2), 118.
Sikström, S., & Jaber, M. Y. (2012). The Depletion–Power–Integration–Latency (DPIL) model of spaced and massed repetition. *Computers & Industrial Engineering, 63*(1), 323–337.
Slovic, P. (2016). *The perception of risk.* Routledge.
Spencer, Sally, & Jones, Paul W. (2003). Time course of recovery of health status following an infective exacerbation of chronic bronchitis. *Thorax, 58*(7), 589–593.
Steimer, T. (2002). The biology of fear-and anxiety-related behaviors. *Dialogues in Clinical Neuroscience, 4*, 231–250.
Utah Department of Health, 2017. *Calculation of infection rates (PDF)*. Utah Department of Health (retrieved 2017-01-09).
Valle, Del., Sara, Y., et al. (2007). Mixing patterns between age groups in social networks. *Social Networks, 29*(4), 539–554.
Van den Driessche, Pauline, & Watmough, James (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences, 180*(1), 29–48.
Wakefield, M. A., Loken, B., & Hornik, R. C. (2010). Use of mass media campaigns to change health behaviour. *The Lancet, 376*(9748), 1261–1271.
Weltens, B., & Cohen, A. D. (1989). Language attrition research: An introduction. *Studies in Second Language Acquisition, 11*(02), 127–133.
Wilensky, U. (1999). NetLogo. < http://ccl.northwestern.edu/netlogo/ > (last access: May 25th 2017). Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.
Wingfield, A., & Byrnes, D. L. (2013). *The psychology of human memory*. Academic Press.
Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science, 2*(6), 409–415.
Wixted, J. T., & Ebbesen, E. B. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory & Cognition, 25*(5), 731–739.
Yang, C., & Wilensky, U. (2011). NetLogo epiDEM basic model. http://ccl.northwestern.edu/netlogo/models/epiDEMBasic (last access: May 25th 2017). Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.
Zhao, S., Wu, C., Kuang, Y., & Ben-Arieh, D. (2015). Information dissemination and human behaviors in epidemics. *Proceedings of the IIE annual conference* (pp. 1907). Institute of Industrial and Systems Engineers (IISE).
Zhao, S., Kuang, Y., Wu, C. H., Bi, K., & Ben-Arieh, D. (2018). Risk perception and human behaviors in epidemics. *IISE Transactions on Healthcare Systems Engineering,* 1–34 (submitted for publication).