



Article

Matrix-Based Formulation of Heterogeneous Individual-Based Models of Infectious Diseases: Using SARS Epidemic as a Case Study

Wei Duan

College of Systems Engineering, National University of Defense Technology, Changsha 410073, China; duanwei@nudt.edu.cn

Abstract: Heterogeneities of individual attributes and behaviors play an important role in the complex process of epidemic spreading. Compared to differential equation-based system dynamical models of infectious disease transmission, individual-based epidemic models exhibit the advantage of providing a more detailed description of realities to capture heterogeneities across a population. However, the higher granularity and resolution of individual-based epidemic models comes with the cost of increased computational complexities, which result in difficulty in formulating individual-based epidemic models with mathematics. Furthermore, it requires great effort to understand and reproduce existing individual-based epidemic models presented by previous researchers. We proposed a mathematical formulation of heterogeneous individual-based epidemic models using matrices. Matrices and vectors were applied to represent individual attributes and behaviors. We derived analytical results from the matrix-based formulations of individual epidemic models, and then designed algorithms to force the computation of matrix-based individual epidemic models. Finally, we used a SARS epidemic control as a case study to verify the matrix-based formulation of heterogeneous individual-based epidemic models.

Keywords: public health; epidemic modeling; agent-based models; heterogeneity; algorithms



Citation: Duan, W. Matrix-based Formulation of Heterogeneous Individual-Based Models of Infectious Diseases: Using SARS Epidemic as a Case Study. *Int. J. Environ. Res. Public Health* **2021**, *18*, 5716. <https://doi.org/10.3390/ijerph18115716>

Academic Editor: Paul B. Tchounwou

Received: 8 February 2021

Accepted: 12 April 2021

Published: 26 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Heterogeneity among the units composing a system is a very generic feature, which means that one never finds two units behaving in exactly the same way in most real systems [1]. In human societies, individual variability widely exists in social settings, cultural backgrounds, ages, genders, and occupations. Human beings have diverse features, physical characteristics, activities, and contact patterns. In epidemic outbreaks, individuals exhibit various health states, infectiousness, epidemic progresses, and behavior response. Despite these general facts, it is usually assumed that a population is homogeneous so as to build analytically tractable epidemic models, such as differential equation-based system dynamical models [2,3].

However, recent studies have shown that individual heterogeneity plays an important role in the epidemic spreading process [4–6]. The impact that individual heterogeneity has on epidemic diffusion has increasingly received more attention. Some researchers have concentrated on the effect of heterogeneous populations on epidemic dynamics [6–8]. They incorporated the heterogeneities of host populations as state variables, including age, stage, size [8], infection threshold [7], and resistant hosts [6]. Some other researchers have focused on epidemic models with spatial heterogeneities [9–11]. They utilized a patch structure where an epidemic system was divided into a number of patches to represent different spatial characteristics. Many researchers also investigated the impact of individuals' heterogeneous behaviors and activities on epidemic diffusion [12–15]. Yang et al. [12] studied the impact of heterogeneous human activities on epidemic spread and found that the heterogeneity of activities affects spreading velocity remarkably. Merler et al. also studied

how human motilities and population heterogeneity affect the course of an epidemic [13]. They found that the cumulative attack rate is positively correlated with average household size and the fraction of students in a population, and this was negatively correlated with the fraction of inactive population. Shang addressed the impact of awareness on epidemic outbreaks by proposing a mean-filed approach accommodating heterogeneous transmission rates [14]. He found that both contact awareness and local awareness can raise the epidemic threshold, while global awareness cannot. Cui et al. applied an edge-based compartmental model to study epidemic spreading dynamics with heterogeneous contacts [15]. In their models, contacts were grouped into two classes: strong contacts and normal ones. Additionally, the super spreading events that emerged in the severe acute respiratory syndrome (SARS) epidemic outbreaks of 2003 highlighted the necessity of a better understanding of social heterogeneities and individual-based models [16]. When studying super spreading events of the SARS epidemic using differential equation models, Mkhathshwa et al. [17] divided infectious individuals into two subgroups: super spreaders and regular spreaders. They then assigned a higher infection rate to super spreaders. Duan et al. employed heterogeneous and stochastic agent-based models to analyze the characteristics of infectious diseases' super spreaders [18]. Researchers also paid attention to epidemic dynamics in heterogeneous networks with different connectivities [19–21], and went forward and investigated epidemic spread in weighted networks representing the heterogeneity of individual interaction strengths [22–25].

Differential equation models and individual-based models are the two most widely used approaches of representing epidemic diffusion [26]. The former makes some reasonable assumptions and simplifications so as to represent epidemic spread at a macro level [2]. A population is assumed to be homogeneous, well-mixed, and aggregated into several compartments according to people's health states. The transitions of population between compartments are described by using differential equations with variables, such as the infection rate, onset rate of symptoms, and recovery rate. These assumptions and simplifications endow differential equation models with the advantage of performing theoretical analysis of macroscopic regularities of epidemic spread, such as the epidemic threshold and the final epidemic size. However, the advantage also comes with the limitation in representing epidemic diffusion in detail [27]. Firstly, the assumption of homogeneous and well-mixed population results in difficulties in representing the variants of individual microscopic attributes and behaviors. Secondly, a small set of variables parameterized with average quantities and mean values are inadequate to capture a variety of factors associated with the epidemic spreading process.

The individual-based model is a promising modeling paradigm to represent individual heterogeneities [28,29]. This approach can provide a more detailed depiction of realities to capture heterogeneity across individuals and incorporate the stochastic nature of infectious disease transmission. However, the higher granularity and resolution provided by individual-based models comes with the cost of increased computational complexities, growing computational power, and the development of intelligent computational algorithms [30]. The complexity of heterogeneous individual-based models leads to difficulty in mathematical formulation. Sometimes, heterogeneous individual-based epidemic models do not provide an insight into technical details, but only some qualitative representations. Consequently, we need to make great efforts to understand and reproduce existing individual-based epidemic models.

Our purpose is to realize the mathematical formulation of heterogeneous individual-based epidemic models, as well as make individual-based epidemic models much easier to be understood and reproduced. We used matrices and vectors to represent individuals, activity locations, social organizations, and relationship networks. Matrix-based representation of individual-based epidemic models reserves the capability of describing the heterogeneity of individuals and epidemic spreading process [8]. Furthermore, all of factors or elements related to epidemic diffusion are described as specific variables in matrices or vectors. According to the specific variables, it is feasible to build a mathe-

mathematical formulation of individual-based epidemic models and algorithms. In addition, we can derive analytical results from the matrix-based representation of individual-based epidemic models. We applied the SARS epidemic as a case study to verify the matrix-based formulation of heterogeneous individual-based epidemic models.

2. Heterogeneous Individual-Based Models of SARS Epidemic

We here presented an improved agent-based model of the SARS epidemic based on our previous works, found in reference [18], which were used as a case study of matrix-based formulations of heterogeneous individual-based epidemic models. The main improvement of the model here is to increase the presentation of individuals' daily commute activities among spatial locations into our previous works so as to study the heterogeneous spatial transmission pattern of the SARS epidemic. Consequently, we designed individuals' daily activities at spatial locations, and introduced a bipartite network to describe individuals' commute behaviors. We assumed a closed human society to study the spread and control of the SARS epidemic. We applied agent-based models to represent human individuals, spatial activity locations, social organizations in the closed society, and their behaviors. In addition, we employed complex networks to describe the multi-relationships among these social entities. We designed individuals' daily activities, such as sleeping, working, shopping, and recreation, which take place at different spatial activity locations. Agents commute among spatial locations to execute daily activities and contact each other. The contacts between susceptible individuals and infectious individuals cause the transmission of the SARS pathogen. Social organizations take some non-pharmaceutical interventions to mitigate the transmission of SARS epidemic, such as contact tracing, spatial activity location closure, individual quarantine, which may change the states of individuals and spatial activity locations. We then computed the discrete events and updated the states of social entities in time sequence to simulate the spread and control of the SARS epidemic.

2.1. Individuals' Daily Commute Patterns

We represented individuals' daily activities as mobility events, and randomly scheduled these discrete events in queuing models. We assumed individuals executed k^m mobility activity events in each day.

We used a bipartite network with time-varying weights to model individuals' daily commute patterns. There are two types of vertices in the bipartite network, including individual vertices and spatial location vertices. Network edges only exist between individual vertices and spatial location vertices. Edge weights denote the propensity that individuals visit spatial locations.

The bipartite network was designed in term of individuals' living regularities. On account of the power-law characteristics of human dynamics [31], the edge weights of bipartite network were generated by a truncated power-law random variable described as

$$w(r) = \left(w_{\min}^{-\rho+1} + \left(w_{\max}^{-\rho+1} - w_{\min}^{-\rho+1} \right) r \right)^{\frac{1}{-\rho+1}}, \quad (1)$$

where $w(r) \in [w_{\min}, w_{\max}]$ is a truncated power-law random variable. $r \in [0, 1]$ is a uniform random number. ρ is the exponent of the power-law distribution. The probability that an individual moves to a spatial location in a commute activity is defined as

$$p_{ij}(t) = \frac{w_{ij}(t)}{\sum_{n_i(t)} w_{ik}(t)}, \quad (2)$$

where $p_{ij}(t)$ is the probability that individual i move to spatial activity location j . $w_{ij}(t)$ is the edge weight between individual i and spatial location j at time t . $n_i(t)$ is the number of spatial locations individual i could visit at time t . It means spatial locations closed in the SARS epidemic outbreak is not allowed to be visited by individual i at time t .

2.2. Individuals' Structured Contact Patterns

We also represented individuals' contact activities as contact events, and randomly scheduled contact events in queuing models. According to survey results of human daily contacts [32], we used a normal random variable to represent the number of individuals' daily contacts, which is described as

$$k^c(r_1, r_2) = \mu + \sigma(-2 \ln r_1)^{\frac{1}{2}} \cos(2\pi r_2), \quad (3)$$

where k^c is the random variable of individuals' daily contact number. r_1 and r_2 are both uniform random numbers in the range $[0, 1]$. In addition, considering the statistical features of human behaviors, we designed the time duration of each contact activity as a power-law waiting time distribution described as

$$d(r) = \left(d_{\min}^{-\rho+1} + \left(d_{\max}^{-\rho+1} - d_{\min}^{-\rho+1} \right) r \right)^{-\frac{1}{\rho+1}}, \quad (4)$$

where $d(r)$ is a random variable denoting the time duration of individuals' contacts. $r \in [0, 1]$ is a uniform random number. ρ is the exponent of the power-law distribution.

We built a contact network in term of individuals' social relationships to represent individual contact patterns. The edge weights of contact network were generated by the formulation described as

$$w_{ij} = w_0(n_i n_j)^\theta, \quad (5)$$

where w_{ij} is the edge weight between individual i and individual j . w_0 is the proportional parameter. n_i and n_j are the node degrees of individual i and individual j , respectively. Once individual i initiates a contact activity, it selects a neighbor in its personal contact network as a contact object according to the following rule

$$q_{ij}(t) = \frac{w_{ij}(t)}{\sum_{n_i(t)} w_{ik}(t)}, \quad (6)$$

where $q_{ij}(t)$ is the probability that individual i selects individual j as a contact object. $w_{ij}(t)$ here is the edge weight between individual i and individual j in the contact network. $n_i(t)$ here is the number of neighbors in personal contact networks of individual i , which could be contacted by individual i . It means individuals who are quarantined or are not at the same spatial activity location cannot be contacted by individual i .

2.3. Infection Probability of SARS

SARS pathogens are transmitted from infectious individuals to susceptible individuals in their face-to-face contacts. Infectious individuals shed SARS pathogens through aerosols and droplets to susceptible individuals. Once susceptible individuals inhale a sufficient dose of SARS pathogens, they are infected by the SARS pathogen. Consequently, the infection probability of the SARS epidemic in a single contact is related to contact intensity and duration, infectivity of infectious individuals, and immunity of susceptible individuals. Here, we only consider contact duration and infectivity. We assumed the infection probability of SARS epidemic in a single contact is

$$Pro = Dur \times In, \quad (7)$$

where Dur is the time duration of a single contact between a susceptible individual and an infectious individual. In is the infectivity of the infectious individual. The infectivity is related to pathogen load and shedding rates of infectious individuals, and evolves along

with the infectious period. According to consensus documents [33], we used a triangle distribution to model the infectivity (*Inf*) of infectious individuals, which is described as

$$f^{Inf}(t) = \begin{cases} \frac{2(t-t_0)}{(t^*-t_0)(t_1-t_0)}, t_0 \leq t \leq t^* \\ \frac{2(t_1-t)}{(t_1-t^*)(t_1-t_0)}, t^* < t \leq t_1 \end{cases} \quad (8)$$

where t_0 and t_1 are the time points of beginning and ending infectious period. t^* is the time point when infectious individuals have the max infectivity, which is set as $t^* = t_0 + 10$ day. According to Reference [18], we fitted the infection probability (*Inf-Pro*) of SARS in case infectious individuals have the max infectivity as

$$f^{Inf-Pro}(d) = (1 - e^{-0.019d}), t_0 \leq t \leq t_1, \quad (9)$$

where d is the time duration of a contact between an infectious individual and a susceptible individual. We then designed infection probability by the proportion of infectivity with the max infectivity as

$$f^{Inf-Pro}(t, d) = \begin{cases} (1 - e^{-0.019d}) \frac{t-t_0}{t^*-t_0}, t_0 \leq t \leq t^* \\ (1 - e^{-0.019d}) \frac{t_1-t}{t_1-t^*}, t^* < t \leq t_1 \end{cases} \quad (10)$$

where t is the time of infectious period.

2.4. Heterogeneous Epidemic Progress of SARS

We assumed each infected individual went through four health state period, including susceptible period, latent period, infectious period, and recovered period. A susceptible individual is infected by the SARS epidemic in a single contact with an infectious individual. Then, the health state of the susceptible individual become latent. Several days later, the individual goes over the latent period, and become in infectious state. The infectious individual exhibits the capability of transmitting SARS pathogens to other susceptible individuals. On account of the onset of symptoms, infectious individuals will be admitted as patient cases of the SARS epidemic in hospitals. Admitted cases receives treatment in hospitals, and are quarantined until they recover from infectious state. Recovered individuals could be discharged from hospitals as they will not be infected by SARS epidemic again.

The duration of health state periods, admission time, and discharge time are fitted as Gamma distributions, of which the means (standard deviation) were 6.37 (16.69), 23.5 (62.1), 4.85 (12.19), and 23.1 (62.1) days, respectively [34]. To build the heterogeneous presentation of SARS progress, we applied Gamma random variables to scale the SARS epidemic progress.

3. Matrix-Based Formulation of the Models

To make heterogeneous individual-based models of SARS epidemic above much easier to be understood and reproduced, we applied matrices and vectors to describe objects and entities in the closed society. More technical details related to models and algorithms are represented in the matrix-based formulation. We firstly defined the models' framework as a group of sets formulated as $\{M(t), C(t), S(t), A(a_k)\}$. $M(t)$ is the set of matrices representing entities and their attributes, such as individuals, spatial locations, health states. $C(t)$ is the set of constraint conditions for entity matrices set $M(t)$. $S(t)$ is the set of analytical results derived from $M(t)$. $A(a_k)$ is the set of algorithms advancing the computation of heterogeneous individual-based epidemic models formulated by matrices.

3.1. Matrix-Based Formulation of Entities

3.1.1. Formulation of Individuals and Their Attributes

We assumed a closed population consisted of N^{Ind} ($N^{Ind} \in \mathbb{Z}^+$) individuals (Ind) in epidemic system. Individuals were represented as

$$M^{Ind} = [a_1^{Ind} \quad a_2^{Ind} \quad \dots \quad a_n^{Ind}]^T, \tag{11}$$

where M^{Ind} is individual matrix with n rows and 1 column. a_i^{Ind} is individual i in the population. The constraint condition of individual matrix M^{Ind} is defined as $C^{Ind} = \{ n = N^{Ind} \}$, which means the row of individual matrix M^{Ind} equals to the number of individuals in the population. Similarly, we used matrices to formulate individuals' attributes, such as health states, epidemic progresses, infectiousness, and immunity.

We formulated individuals' health states (HS) as

$$M^{HS}(t) = \begin{bmatrix} a_{11(x)}^{HS}(t) & \dots & a_{1h(x)}^{HS}(t) \\ \vdots & \ddots & \vdots \\ a_{n1(x)}^{HS}(t) & \dots & a_{nh(x)}^{HS}(t) \end{bmatrix}, \tag{12}$$

where $M^{HS}(t)$ is health state matrix with n rows and h columns. h is the number of individual health states in epidemic progresses. $a_{ij(x)}^{HS}(t) = 1$ means individual a_i^{Ind} is in health state $j(x)$ at time t . Susceptible state (S), latent state (L), infectious state (I), and recovered state (R) are marked as $x \in \{S, L, I, R\}$, respectively. The constraint conditions $C^{HS}(t)$ of health state matrix $M^{HS}(t)$ include $n = N^{Ind}$, $a_{ij(x)}^{HS}(t) \in \{0, 1\}$ denoting the label of health state j of individual a_i^{Ind} at time t , $\sum_j a_{ij(x)}^{HS}(t) = 1$ meaning individual a_i^{Ind} have to be in only one health state at time t , and $\sum_i \sum_j a_{ij(x)}^{HS}(t) = N^{Ind}$ denoting all individuals are in one of health states at time t .

Epidemic progresses describe the transitions of individuals among different health periods. Individuals go through different periods of health states, which occupy different time durations, such as latent period, infectious period, and recovered period. We described epidemic progresses (EP) based on heterogeneous time scales as

$$M^{EP} = \begin{bmatrix} a_{11(x)}^{EP} & \dots & a_{1m(x)}^{EP} \\ \vdots & \ddots & \vdots \\ a_{n1(x)}^{EP} & \dots & a_{nm(x)}^{EP} \end{bmatrix}, \tag{13}$$

where M^{EP} is epidemic progress matrix with n rows and m columns. $a_{ij(x)}^{EP}$ is the time duration of individual a_i^{Ind} in the period of health state $j(x)$. We described infection time (IT), latent period (LP), infectious period (IP), admission time (AT), and discharge time (DT) of individual a_i^{Ind} as $a_{i1(IT)}^{EP}$, $a_{i2(LP)}^{EP}$, $a_{i3(IP)}^{EP}$, $a_{i4(AT)}^{EP}$, and $a_{i5(DT)}^{EP}$, respectively. The constraint conditions C^{EP} of epidemic progress matrix M^{EP} include $n = N^{Ind}$, $a_{ij(x)}^{EP} = f^{EP}(x, t)$ meaning the time duration of health state period comes from a certain function, which is a stochastic Gamma distribution or a constant value, and $a_{i1(S)}^{HS} = 1 \Rightarrow a_{ij(S)}^{EP} = 0$ meaning in case individual a_i^{Ind} is not infected, all time durations of health state period are 0.

We applied a vector to represent the infectivity (Inf) of infectious individuals as

$$M^{Inf}(t) = [a_1^{Inf}(t) \quad a_2^{Inf}(t) \quad \dots \quad a_n^{Inf}(t)]^T, \tag{14}$$

where $M^{Inf}(t)$ is infectivity matrix with n rows and 1 column. $a_i^{Inf}(t)$ is the infectivity of individual a_i^{Ind} at time t . The constraint conditions $C^{Inf}(t)$ of infectivity matrix $M^{Inf}(t)$

include $n = N^{Ind}$, $a_i^{Inf}(t) = f^{Inf}(t)$ meaning the infectivity of infectious individuals is a function defined in Equation (8), and $a_{i3(I)}^{HS} = 0 \Rightarrow a_i^{Inf}(t) = 0$ denoting individuals have infectivity only in case they are in infectious state.

Recovered individuals exhibit diverse abilities to resist pathogen. We formulated the immunity (*Imm*) of individuals as

$$M^{Imm}(t) = [a_1^{Imm}(t) \quad a_2^{Imm}(t) \quad \dots \quad a_n^{Imm}(t)]^T, \tag{15}$$

where $M^{Imm}(t)$ is immunity matrix with n rows and 1 column. $a_i^{Imm}(t)$ is the immunity of individual a_i^{Ind} at time t . The constraint conditions $C^{Imm}(t)$ of immunity matrix $M^{Imm}(t)$ include $n = N^{Ind}$, $a_i^{Imm}(t) = f^{Imm}(t)$ meaning the immunity of individuals is a certain function, and $a_{i4(R)}^{HS} = 0 \Rightarrow a_i^{Imm}(t) = 0$ denoting individuals exhibit immunity in case they are in recovered state.

Once individuals are admitted as patients in hospitals, they are quarantined and prohibited to move outside and contact with others. We formulated individuals' quarantine (*Qua*) state as

$$M^{Qua}(t) = [a_1^{Qua}(t) \quad a_2^{Qua}(t) \quad \dots \quad a_n^{Qua}(t)]^T, \tag{16}$$

where $M^{Qua}(t)$ is quarantine matrix with n rows. $a_i^{Qua}(t)$ is label denoting whether individual a_i^{Ind} is quarantined. The constraint conditions $C^{Qua}(t)$ of quarantine matrix include $n = N^{Ind}$ and $a_i^{Qua}(t) \in \{0, 1\}$.

3.1.2. Formulation of Spatial Locations and Their Attributes

Human mobility patterns depend on the transmission routes of infectious diseases. We assumed that a closed environment consists of many spatial locations. The number of spatial locations (*SL*) was assumed as $N^{SL} \in Z^+$. The spatial location matrix is formulated as

$$M^{SL} = [a_1^{SL} \quad a_2^{SL} \quad \dots \quad a_l^{SL}]^T, \tag{17}$$

where a_i^{SL} is spatial location i in the scene. The constraint condition C^{SL} of spatial location matrix M^{SL} is $l = N^{SL}$, which means the number of rows of spatial location matrix equals to the number of spatial locations in the closed environment.

During epidemic outbreaks, non-pharmaceutical interventions may be employed to close some spatial locations, such as campus, restaurants, and cinemas. We formulated the closure (*Clo*) state of spatial locations as

$$M^{Clo}(t) = [a_1^{Clo}(t) \quad a_2^{Clo}(t) \quad \dots \quad a_l^{Clo}(t)]^T, \tag{18}$$

where $M^{Clo}(t)$ is closure state matrix with l rows. The constraint conditions $C^{Clo}(t)$ of closure state matrix $M^{Clo}(t)$ include $l = N^{SL}$ and $a_i^{Clo}(t) \in \{0, 1\}$.

3.1.3. Formulation of Social Organizations and Their Attributes

We assumed there were $N^{Org} \in Z^+$ social organizations (*Org*) in the closed environment, and formulated social organizations as

$$M^{Org} = [a_1^{Org} \quad a_2^{Org} \quad \dots \quad a_g^{Org}]^T, \tag{19}$$

where M^{Org} is social organization matrix with g rows. a_i^{Org} is social organization i . The constraint condition C^{Org} of social organization matrix is $g = N^{Org}$, which means the rows of social organization matrix equals to the number of social organizations.

Considering social organizations own spatial locations, such as office buildings, we formulated the ownerships between social organizations and spatial locations (*Org-SL*) as

$$M^{Org-SL} = \begin{bmatrix} a_{11}^{Org-SL} & \cdots & a_{1l}^{Org-SL} \\ \vdots & \ddots & \vdots \\ a_{g1}^{Org-SL} & \cdots & a_{gl}^{Org-SL} \end{bmatrix}, \quad (20)$$

where M^{Org-SL} is called as organization-location relationship matrix with g rows and l columns. a_{ij}^{Org-SL} is the label denoting whether social organization a_i^{Org} is sited in spatial location a_j^{SL} . The constraint conditions C^{Org-SL} of organization-location relationship matrix including $g = N^{Org}$, $l = N^{SL}$, and $a_{ij}^{Org-SL} \in \{0, 1\}$. Furthermore, we also formulated the relationship between social organizations and individuals (*Org-Ind*) as

$$M^{Org-Ind} = \begin{bmatrix} a_{11}^{Org-Ind} & \cdots & a_{1g}^{Org-Ind} \\ \vdots & \ddots & \vdots \\ a_{n1}^{Org-Ind} & \cdots & a_{ng}^{Org-Ind} \end{bmatrix}, \quad (21)$$

where $M^{Org-Ind}$ is organization-individual relationship matrix with n rows and g columns. $a_{ij}^{Org-Ind}$ is the label denoting whether individual a_i^{Ind} belongs to social organization a_j^{Org} . The constraint conditions $C^{Org-Ind}$ of organization-individual relationship matrix include $n = N^{Ind}$, $g = N^{Org}$, and $a_{ij}^{Org-Ind} \in \{0, 1\}$.

3.1.4. Formulation of Weighted Bipartite Networks

We applied a weighted bipartite network to represent individuals' commute patterns among spatial locations. There are two types of vertices in the bipartite network, which denote individuals and spatial locations, respectively. Network edges only exist between individuals and spatial locations. Edge weights represent the propensity that individuals visit spatial locations. We formulated the bipartite network (*BN*) as

$$M^{BN} = \begin{bmatrix} a_{11}^{BN} & \cdots & a_{1l}^{BN} \\ \vdots & \ddots & \vdots \\ a_{n1}^{BN} & \cdots & a_{nl}^{BN} \end{bmatrix}, \quad (22)$$

where M^{BN} is bipartite network matrix with n rows and l columns. a_{ij}^{BN} is the label denoting whether individual a_i^{Ind} is linked to spatial location a_j^{SL} . The constraint conditions C^{BN} of bipartite network matrix include $n = N^{Ind}$, $l = N^{SL}$, and $a_{ij}^{BN} \in \{0, 1\}$. The edge weights of bipartite network (*BNW*) are formulated as

$$M^{BNW}(t) = \begin{bmatrix} a_{11}^{BNW}(t) & \cdots & a_{1l}^{BNW}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}^{BNW}(t) & \cdots & a_{nl}^{BNW}(t) \end{bmatrix}, \quad (23)$$

where $M^{BNW}(t)$ is the edge weights matrix of bipartite network with n rows and l columns. $a_{ij}^{BNW}(t)$ is the edge weight between individual a_i^{Ind} and spatial location a_j^{SL} at time t . The constraint conditions $C^{BNW}(t)$ of edge weights matrix of bipartite network include $n = N^{Ind}$, $l = N^{SL}$, $a_{ij}^{BNW}(t) = f^{BNW}(t, i, j) \Leftarrow \omega(r)$ meaning the edge weights of bipartite network are generated by a function defined in Equation (1), and $a_{ij}^{BN} = 0 \Rightarrow a_{ij}^{BNW}(t) = 0$ meaning the edge weight is 0 in case individual a_i^{Ind} is not connected to spatial location a_j^{SL} .

Individuals' current spatial positions (CSL) at time t are recorded as

$$M^{CSL}(t) = \begin{bmatrix} a_{11}^{CSL}(t) & \cdots & a_{1l}^{CSL}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}^{CSL}(t) & \cdots & a_{nl}^{CSL}(t) \end{bmatrix}, \tag{24}$$

where $M^{CSL}(t)$ is individual current spatial location matrix with n rows and m columns. $a_{ij}^{CSL}(t)$ is the label denoting whether individual a_i^{Ind} is at spatial location a_j^{SL} at time t . The constraint conditions $C^{CSL}(t)$ of individual current spatial location matrix include $n = N^{Ind}, l = N^{SL}, a_{ij}^{CSL}(t) \in \{0, 1\}, \sum_j a_{ij}^{CSL}(t) = 1$ meaning individual a_i^{Ind} have to stay at only one spatial location, and $\sum_i \sum_j a_{ij}^{CSL}(t) = N^{Ind}$ meaning all individuals stay at the spatial locations.

3.1.5. Formulation of Weighted Contact Networks

We represented individuals' heterogeneous contact patterns by using a weighted contact network, where nodes stood for individuals, edges represented individuals' relationships, and edge weights described individuals' interaction strength. We formulated the contact network (CN) as

$$M^{CN} = \begin{bmatrix} a_{11}^{CN} & \cdots & a_{1n}^{CN} \\ \vdots & \ddots & \vdots \\ a_{n1}^{CN} & \cdots & a_{nn}^{CN} \end{bmatrix}, \tag{25}$$

where M^{CN} is contact network matrix with n rows and n columns. a_{ij}^{CN} is the label denoting whether individual a_i^{Ind} is connected to individual a_j^{Ind} . The constraint conditions C^{CN} of contact network matrix include $n = N^{Ind}, a_{ii}^{CN} = 0$, and $a_{ij}^{CN} \in \{0, 1\}$. Edge weights of the contact network (CNW) were formulated as

$$M^{CNW}(t) = \begin{bmatrix} a_{11}^{CNW}(t) & \cdots & a_{1n}^{CNW}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}^{CNW}(t) & \cdots & a_{nn}^{CNW}(t) \end{bmatrix}, \tag{26}$$

where $M^{CNW}(t)$ is the edge weight matrix of contact network with n rows and n columns. $a_{ij}^{CNW}(t)$ is the edge weight between individual a_i^{Ind} and individual a_j^{Ind} at time t . The constraint conditions C^{CNW} of edge weight matrix of contact network include $n = N^{Ind}, a_{ij}^{CNW}(t) = f^{CNW}(t, i, j) \Leftarrow \omega_{ij}$ meaning the edge weights of contact network is generated by a certain function defined in Equation (5), and $a_{ij}^{CN} = 0 \Rightarrow a_{ij}^{CNW}(t) = 0$ meaning the edge weight between two individuals is 0 in case they are not connected with each other. We marked who were contacting with whom as

$$M^{CO}(t) = \begin{bmatrix} a_{11}^{CO}(t) & \cdots & a_{1n}^{CO}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}^{CO}(t) & \cdots & a_{nn}^{CO}(t) \end{bmatrix}, \tag{27}$$

where $M^{CO}(t)$ is contact object (CO) matrix. $a_{ij}^{CO}(t)$ is the label denoting whether individual a_i^{Ind} is contacting with individual a_j^{Ind} at time t . The constraint conditions C^{CO} of contact object matrix include $n = N^{Ind}, a_{ij}^{CO}(t) \in \{0, 1\}$, and $a_{ii}^{CO}(t) = 0$ meaning individuals cannot contact with themselves.

3.1.6. Formulation of Transmission Networks

We recorded the transmission networks of epidemics to analyze transmission routes of epidemics and derive the reproduce number of infected cases. The epidemic transmission network (TN) is represented as

$$M^{TN} = \begin{bmatrix} a_{11}^{TN} & \cdots & a_{1n}^{TN} \\ \vdots & \ddots & \vdots \\ a_{n1}^{TN} & \cdots & a_{nn}^{TN} \end{bmatrix}, \tag{28}$$

where a_{ij}^{TN} is the label denoting whether individual a_i^{Ind} is infected by epidemic from individual a_j^{Ind} . The constraint conditions C^{TN} of transmission network matrix include $n = N^{Ind}$, $a_{ij}^{TN} \in \{0, 1\}$, and $a_{ii}^{TN} = 0$ meaning individuals cannot transmit epidemics to themselves.

3.1.7. Formulation of Individuals' Mobility Activities

We represented the heterogeneous mobility patterns of individuals, and scheduled individuals' mobility activities as discrete events in queue models. Individual mobility event (ME) matrix is formulated as

$$M^{ME} = \begin{bmatrix} a_{11}^{ME}(t_1, l_1) & \cdots & a_{1k^m}^{ME}(t_{k^m}, l_{k^m}) \\ \vdots & \ddots & \vdots \\ a_{n1}^{ME}(t_1, l_1) & \cdots & a_{nk^m}^{ME}(t_{k^m}, l_{k^m}) \end{bmatrix}, \tag{29}$$

where M^{ME} is mobility event matrix with n rows and k^m columns. $a_{ij}^{ME}(t_j, l_j)$ is the mobility event j of individual a_i^{Ind} , which denotes individual a_i^{Ind} will move to spatial location $l_j \in M^{SL}$ at time t_j . The constraint conditions C^{ME} of mobility event matrix include $i < j \Rightarrow t_i < t_j$ meaning mobility events are scheduled in time sequence, $t_j = f^{METP}(t)$ meaning the time point when mobility event (METP) $a_{ij}^{ME}(t_j, l_j)$ happens is generated by a certain function, $l_j = f^{MESL}(t) \Leftarrow p_{ij}(t)$ meaning which spatial location individual a_i^{Ind} moves (MESL) to at time t_j is generated by a certain function defined in Equation (2).

3.1.8. Formulation of Individuals' Contact Activities

We also represented individual heterogeneous contact patterns and scheduled contact activities as discrete events in queue models. Individual contact activity matrix is formulated as

$$M^{CE} = \begin{bmatrix} a_{11}^{CE}(t_1, b_1, d_1) & \cdots & a_{1k}^{CE}(t_k, b_k, d_k) \\ \vdots & \ddots & \vdots \\ a_{n1}^{CE}(t_1, b_1, d_1) & \cdots & a_{nk}^{CE}(t_k, b_k, d_k) \end{bmatrix}, \tag{30}$$

where M^{CE} is contact event (CE) matrix with n rows and k columns. $a_{ij}^{CE}(t_j, b_j, d_j)$ is the contact event j of individual a_i^{Ind} , which denotes individual a_i^{Ind} initiates a contact event with individual a_j^{Ind} at time t_j . Furthermore, the contact activity sustains a duration time d_j . The constraint conditions C^{CE} of contact event matrix include $i < j \Rightarrow t_i < t_j$ meaning contact events are scheduled in time sequence, $t_j = f^{CETP}(t)$ meaning the time point (CETP) when contact events happen is generated by a certain function, $b_j = f^{CEO}(t) \Leftarrow q_{ij}(t)$ denoting individuals select contact objects (CEO) according to a function defined in Equation (6), and $d_j = f^{CED}(t) \Leftarrow d(r)$ meaning the duration time of contact activities (CED) is generated by a certain function defined in Equation (4).

3.2. Statistical Properties of Epidemic Diffusion

To conduct theoretical analysis of epidemic spreading regularities, we derived analytical results from matrix-based formulation of entities above. The numbers of susceptible, latent, infectious, and recovered individuals at time t were formulated as

$$\begin{cases} s(t) = \sum_k a_{k1(S)}^{HS}(t) \\ l(t) = \sum_k a_{k2(L)}^{HS}(t) \\ i(t) = \sum_k a_{k3(I)}^{HS}(t) \\ r(t) = \sum_k a_{k4(R)}^{HS}(t) \end{cases}, \tag{31}$$

The number of new infections (NI) at time t is derived as

$$n^{NI}(t) = \sum_k a_{k4(R)}^{HS}(t) - \sum_k a_{k4(R)}^{HS}(t-1) + \sum_k a_{k3(I)}^{HS}(t) - \sum_k a_{k3(I)}^{HS}(t-1). \tag{32}$$

The accumulated number of infected individuals (AI) is derived as

$$n^{AI}(t) = N^{Ind} - \sum_k a_{k1(S)}^{HS}(t), \tag{33}$$

where $n^{AI}(t)$ is the accumulated number of infections at time t . Then, we can get final epidemic size as $n^{AI}(t_{end})$.

To analyze the spatial patterns of epidemic diffusion, we derived the distribution of infectious individuals at spatial locations (SL-I) as

$$p_k^{SL-I}(t) = \frac{\sum_j (a_{jk}^{CSL}(t) a_{j3(I)}^{HS}(t))}{\sum_k a_{k3(I)}^{HS}(t)}, \tag{34}$$

where $p_k^{SL-I}(t)$ is the proportion of infectious individuals at spatial location a_k^{SL} at time t . $a_{jk}^{CSL}(t) a_{j3(I)}^{HS}(t)$ is the label denoting whether individual a_j^{Ind} is in infectious state and stays at location a_k^{SL} at time t . $\sum_k a_{k3(I)}^{HS}(t)$ is the total number of infectious individuals at time t . Furthermore, we derived the density of infectious individuals at spatial locations as

$$d_k^{SL-I}(t) = \frac{\sum_j (a_{jk}^{CSL}(t) a_{j3(I)}^{HS}(t))}{\sum_j a_{jk}^{CSL}(t)}, \tag{35}$$

where $d_k^{SL-I}(t)$ is the density of infectious individuals at spatial location a_k^{SL} . $\sum_j a_{jk}^{CSL}(t)$ is the number of individuals who are at spatial location a_k^{SL} at time t . Similarly, we got the distribution of infectious individuals at social organizations (Org-I) as

$$p_k^{Org-I}(t) = \frac{\sum_j (a_{jk}^{Org-Ind} a_{j3(I)}^{HS}(t))}{\sum_k a_{k3(I)}^{HS}(t)}, \tag{36}$$

where $p_k^{Org-I}(t)$ is the proportion of infectious individuals at social organization a_k^{Org} at time t . $a_{jk}^{Org-Ind} a_{j3(I)}^{HS}(t)$ is the label denoting whether individual a_j^{Ind} is in infectious state

and belongs to social organization a_k^{Org} . The density of infectious individuals at social locations is formulated as

$$d_k^{Org-I}(t) = \frac{\sum_j (a_{jk}^{Org-Ind} a_{j3(I)}^{HS}(t))}{\sum_j a_{jk}^{Org-Ind}}, \tag{37}$$

where $d_k^{Org-I}(t)$ is the density of infectious individuals at social locations a_k^{Org} . $\sum_j a_{jk}^{Org-Ind}$ is the number of individuals who belong to social organization a_k^{Org} . We derived the reproduce number of epidemics according to transmission network matrix as

$$\begin{cases} R = \frac{\sum_i \sum_j a_{ij}^{TN}}{\sum_i n_i^{TN}}, \sum_i n_i^{tn} \neq 0 \\ n_i^{TN} = 1, \sum_j a_{ij}^{tn} \geq 1 \\ n_i^{TN} = 0, \sum_j a_{ij}^{tn} = 0 \end{cases}, \tag{38}$$

where R is the reproduce number of epidemics. $\sum_i n_i^{TN}$ is the number of infected individuals who produce more than one of the second-generation cases.

3.3. Algorithms of Computation

We designed algorithms to force the computation of heterogeneous individual-based epidemic models formulated by matrices and vectors. The computational framework of our models is illustrated in Figure 1.

There are four computational processes invoked in time ticks, which contain eight major modules composed of algorithms in the framework. These modules are described as follows.

1. Computing engine: an engine that forces the computation of models. It advances the computation of models in time order. In each time tick, the engine invokes the computing of mobility activity progress, contact activity progress, and epidemic progress. Furthermore, it realizes time management service that calculates date list. Besides, it invokes events schedules to generate and update mobility events and contact events at the beginning of each day.
2. Mobility events schedule: a container that generates, updates, stores mobility events. At the beginning of each day, it generates new mobility events for each individual to update event schedule.
3. Contact events schedule: a container that generates, updates, store contact events. At the beginning of each day, it generates new contact events for each individual to update event schedule.
4. Mobility activity progress: a progress that executes the computation of a mobility activity including choosing expected location, calculating individual's position, etc. A weighted spatial network is applied to build the model of individual mobility.
5. Contact activity progress: a progress that executes the computation of a contact activity including choosing expected contact object, calculating contact duration, calculating infectious probability, etc. A weighted contact network is applied to build the model of individual contact.
6. Epidemic progress: a progress that represents the evolution of health states of infected individuals. It fulfills the computation of individual transitions among different health states by using the infected time and the duration of each health state.

7. A weighted temporal-spatial network: a weighted network that combined a spatial network with a contact network. The weighted temporal-spatial network is applied to build the model of individual mobility and contact behaviors.
8. Random number generator: a generator that produces random numbers subject to various distributions. These random numbers are utilized to represent the stochastic nature of epidemic spreading and individual behaviors (e.g., infection probability, mobility time, contact time, contact duration, the durations of health states, etc.).

These modules implement different services in order to execute the computation of epidemic models. It is essential to design algorithms to realize the capabilities of providing these services in these modules. We summarize four major algorithms represented as follow.

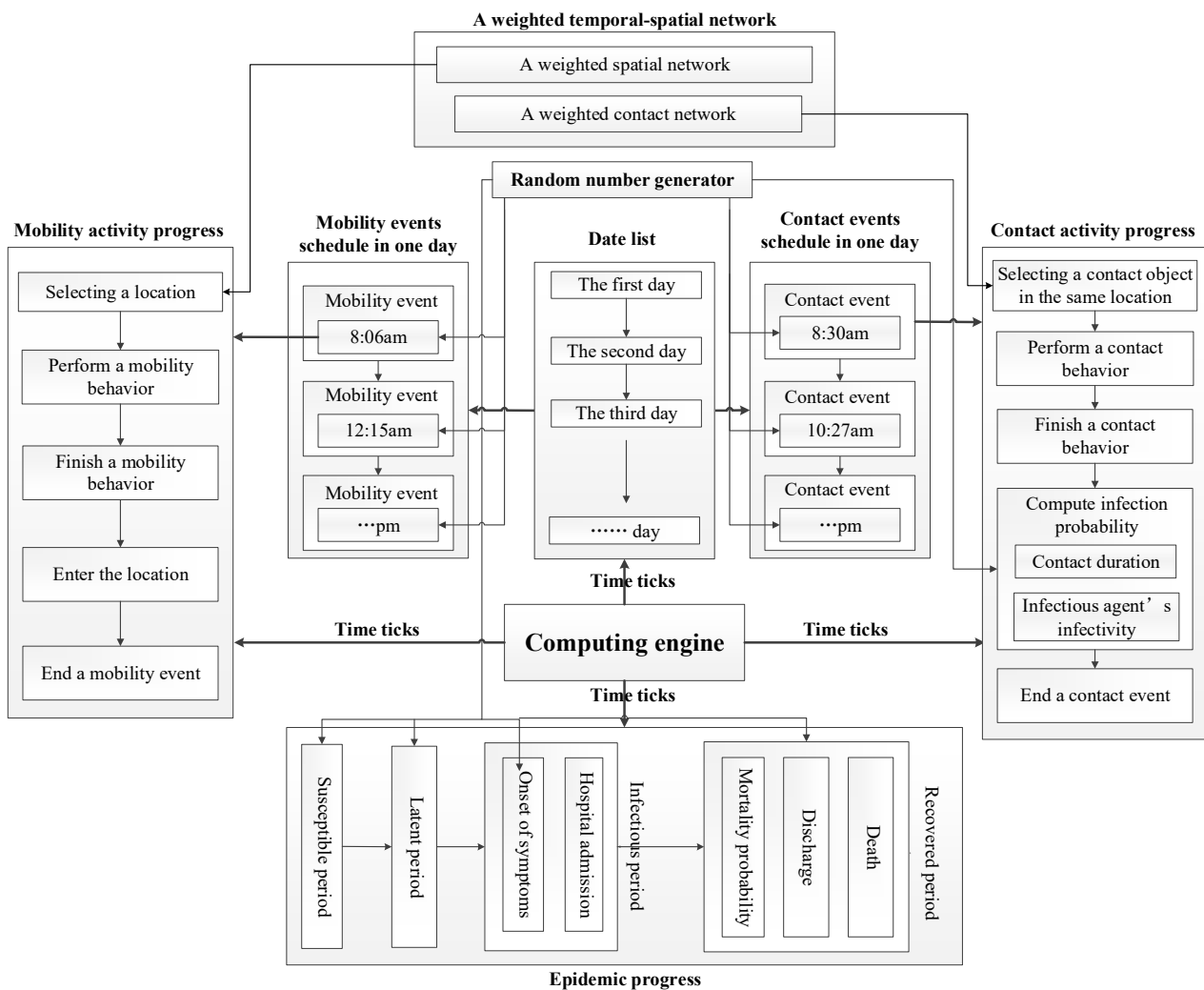


Figure 1. The computational framework of algorithms.

3.3.1. Algorithm of Computing Engine

Computing engine implements three jobs described in Figure 2. Computing engine firstly initializes models and parameters described as matrices and constraints. Then, computing engine enters the main loop of calculations and advances the computation of models in time ticks. In each time tick, computing engine invokes epidemic progress, mobility activity progress, and contact activity progress. Meanwhile, computing engine record the state of epidemic diffusion and update mobility and contact event schedules. Computing engine finally outputs analytical results and terminates the computation of models.

```

1  begin
2  for each matrix  $M^i(t)$  and its' constraints  $C^i(t)$  do
3       $M^i(t_0) \rightarrow M^i(t), C^i(t_0) \rightarrow C^i(t)$ ; // initialize matrices and constraints
4  end for
5   $t_0 \rightarrow t$ ; // initialize simulation time
6  while  $t < t_{end}$  do // judge whether simulation time reaches to the end
7       $t = t + 1$ ; // advance simulation time
8      if a new day begins then
9          statisticInDay( $t$ ); // record statistical data by this function
10         update  $M^{ME}$ ; // reschedule individuals' daily mobility events
11         update  $M^{CE}$ ; // reschedule individuals' daily contact events
12     end if
13     enter epidemic progress; // compute epidemic progress by another algorithm
14     enter mobility activity progress; // compute mobility activities by another algorithm
15     enter contact activity progress; // compute contact activities by another algorithm
16 end while
17 save results; // save experimental results in a file
18 end

```

Figure 2. Algorithm of computing engine.

3.3.2. Algorithm of Mobility Activity Progress

We represent mobility activity progress in Figure 3, which contains three jobs: querying a mobility activity in mobility event schedule; selecting an activity location according to weighted spatial mobility network; and updating the current location matrix.

```

1  begin
2  for each  $a_i^{ind}$  in matrix  $M^{ind}$  do
3      if  $a_i^{Qu} = 0$  then // judge whether individuals are quarantined
4          for each  $a_{ij}^{ME}(t_j, l_j)$  in matrix  $M^{ME}$  do
5              if  $t = t_j$  then // judge whether mobility events happen
6                   $f^{ME^{SL}}(t) \Rightarrow l_j = a_j^{SL}$ ; // select a spatial location by the function
7                  for each  $a_{ij}^{CSL}(t)$  in matrix  $M^{CSL}(t)$  do
8                      if  $j = k$  then
9                           $a_{ij}^{CSL}(t) = 1$ ; // update individuals' current spatial loca-
10                         tions
11                      else
12                           $a_{ij}^{CSL}(t) = 0$ ; // update individuals' current spatial loca-
13                         tions
14                      end if
15                  end for
16              end if
17          end for
18      end if
19  end for
20  end

```

Figure 3. Algorithm of computing mobility activities.

The function $f^{MESL}(t)$ in Figure 3 decides which spatial location is selected as individuals' destination in their daily mobility activities according to the probability formulated as

$$p_{ij}(t) = \frac{(1 - a_j^{Clo}(t)) a_{ij}^{BN} a_{ij}^{BNW}(t)}{\sum_k [(1 - a_k^{Clo}(t)) a_{ik}^{BN} a_{ik}^{BNW}(t)]}, \tag{39}$$

where $p_{ij}(t)$ is the probability that individual a_i^{Ind} selects spatial location a_j^{SL} as his destination in a daily mobility activity. The probability $p_{ij}(t)$ is defined by using the proportion of the edge weights of weighted bipartite network. $1 - a_j^{Clo}(t)$ denotes spatial location a_j^{SL} is not closed in epidemic outbreaks. This equation is a matrix-based formulation of Equation (2).

3.3.3. Algorithm of Contact Activity Progress

We described contact activity process in Figure 4, which contains four jobs: querying a contact activity in contact event schedule; selecting a contact object according to weighted contact network; updating current contact object matrix; and computing the infection at the end of contact activities.

The function $f^{CEO}(t)$ in Figure 4 decides who are going to contact with whom in individuals' contact activities according to the probability described as

$$q_{ij}(t) = \frac{(1 - a_j^{Qua}(t)) \left(\sum_m [a_{im}^{CSL}(t) a_{jm}^{CSL}(t)] \right) a_{ij}^{CN} a_{ij}^{CNW}(t)}{\sum_k \left[(1 - a_k^{Qua}(t)) \left(\sum_m [a_{im}^{CSL}(t) a_{km}^{CSL}(t)] \right) a_{ik}^{CN} a_{ik}^{CNW}(t) \right]}, \tag{40}$$

where $q_{ij}(t)$ is the probability that individual a_i^{Ind} selects individual a_j^{Ind} as a contact object when individual a_i^{Ind} initiates a contact event. The probability $q_{ij}(t)$ is defined by using the proportion of edge weights of weighted contact network. $1 - a_j^{Qua}(t)$ denotes individual a_j^{Ind} is not quarantined at time t . $\sum_m [a_{im}^{CSL}(t) a_{jm}^{CSL}(t)]$ ensures individual a_i^{Ind} and individual a_j^{Ind} are simultaneously at the same spatial location at time t . This equation is a matrix-based formulation of Equation (6).

The function $f^{Inf-Pro}(a_j^{Inf}(t), a_i^{Imm}(t), d_j)$ decides whether susceptible individual a_i^{Ind} is infected by epidemics from infectious individual a_j^{Ind} in a contact activity for a time duration d_j . The infection function is formulated as

$$f^{inf-pro}(t, d_j) = \begin{cases} \left(1 - e^{-0.019d_j} \right)^{\frac{t-t_0}{t^*-t_0}}, & t_0 \leq t \leq t^* \\ \left(1 - e^{-0.019d_j} \right)^{\frac{t_1-t}{t_1-t^*}}, & t^* < t \leq t_1 \end{cases}, \tag{41}$$

where $t_0 = a_{j1(IT)}^{EP} + a_{j2(LP)}^{EP}$ is the time point when the infectious period of individual a_j^{Ind} begins. $t_1 = a_{j1(IT)}^{EP} + a_{j2(LP)}^{EP} + a_{j3(IP)}^{EP}$ is the time point when the infectious period of individual a_j^{Ind} ends. $t^* = t_0 + 10$ (days) is the time point when the infectivity of infections individual a_j^{Ind} reaches the maximum. This equation is a matrix-based formulation of Equation (10).

```

1  begin
2  for each  $a_i^{hid}$  in matrix  $M^{hid}$  do
3      if  $a_i^{Qis} = 0$  then // judge whether individuals are quarantined
4          for each  $a_{ij}^{CE}(t, b_j, d_j)$  in matrix  $M^{CE}$  do
5              if  $t = t_j$  then // judge whether contact events happen
6                   $f^{CED}(t) \Rightarrow b_j = a_j^{hid}$ ; // select a contact object by the function
7                   $f^{CED}(t) \Rightarrow d_j = t_d$ ; // generate the time duration of a contact
8                   $a_{ij}^{CO}(t) = 1$ ; // record contact object
9                   $a_{ji}^{CO}(t) = 1$ ;
10                 Else if  $t = t_j + d_j$  then
11                      $a_{ij}^{CO}(t) = 0$ ; // record contact object when contact activity ends
12                      $a_{ji}^{CO}(t) = 0$ ;
13                 if  $a_{i|l(s)}^{HS} = 1$  and  $a_{j|k(t)}^{HS} = 1$  then // judge whether infection happens
14                      $f^{inf-Pre}(a_{ij}^{CO}(t), a_{ji}^{CO}(t), d_j) \Rightarrow a_{i|l(s)}^{HS}(t) = 1$  or  $a_{j|k(t)}^{HS}(t) = 1$ ;
15                      $a_{ij}^{TN} = 1$ ; // record transmission network
16                 update  $a_{\text{set}(t)}^{EP}$ ; // update epidemic progress parameters
17
18
19
20
21
22

```

Figure 4. Algorithm of computing contact activities.

3.3.4. Algorithm of Epidemic Progress

We represented epidemic progress in Figure 5, which contains three jobs: checking the health state of individuals according to health state matrix $M^{HS}(t)$; judging whether the current health state of individuals is over according to epidemic progress matrix M^{EP} ; and then updating health state matrix.

```

1  begin
2  for each  $a_i^{hid}$  in matrix  $M^{hid}$  do
3      if  $a_{i|2(L)}^{HS}(t) = 1$  then // judge whether individuals are in latent state
4          if  $t = a_{i|1(TR)}^{EP} + a_{i|2(LP)}^{EP}$  then // judge whether latent period ends
5               $a_{i|2(L)}^{HS}(t) = 0$ ;
6               $a_{i|3(I)}^{HS}(t) = 1$ ; // become infectious state
7          else if  $a_{i|3(I)}^{HS}(t) = 1$  then
8              if  $t = a_{i|1(TR)}^{EP} + a_{i|2(LP)}^{EP} + a_{i|3(IP)}^{EP}$  then // judge whether infectious period ends
9                   $a_{i|3(I)}^{HS}(t) = 0$ ;
10                  $a_{i|4(R)}^{HS}(t) = 1$  // become recovered state
11             end if
12         end if
13         if  $t = a_{i|1(TR)}^{EP} + a_{i|2(LP)}^{EP} + a_{i|4(AT)}^{EP}$  then // judge whether infected cases are admitted
14              $a_i^{Qis} = 1$ ;
15         end if
16         if  $t = a_{i|1(TR)}^{EP} + a_{i|2(LP)}^{EP} + a_{i|4(AT)}^{EP} + a_{i|5(DT)}^{EP}$  then // judge whether infected cases are dis-
17             charged from hospitals
18              $a_i^{Qis} = 0$ ;
19         end if
20     end if
21 end for
22 end

```

Figure 5. Algorithm of computing epidemic progress.

4. Simulating the Spread and Control of SARS Epidemic

To test and verify matrix-based formulation of heterogeneous individual-based transmission models of the SARS epidemic, we used the matrix-based models and algorithms to simulate the spread and control of the SARS epidemic.

4.1. Experiment Design

We assumed a closed society containing $N^{Ind} = 1000$ individuals and $N^{SL} = 250$ spatial activity locations. We used the method of randomly adding edges to build a bipartite network and a contact network. The edge weights of bipartite network were generated by a truncated power-law random variable $w(r) \in [1, 1000]$. The exponent ρ of power-law distribution was set as 1.5 in Equation (1) according to reference [31]. The edge weights of contact network were generated according to Equation (5) whose proportional parameter was set as 0.1, and scaling parameter was set as 1.

We assumed each individual executed five mobility activities in a day. Individuals' daily contact number approximates to a normal distribution with mean 8.6 (standard deviation 4.25) [34]. The time duration of contact activities was represented as a truncated power-law random variable in the range [1, 120] minute with an exponent 1.5 [31].

We studied the effect of non-pharmaceutical interventions on the SARS epidemic diffusion, including contact tracing and workplace closure. In case susceptible individuals ever contacted with infectious individuals were found out, they were quarantined. Furthermore, once quarantined individuals had the onset of symptoms, they were admitted to be infected by the SARS epidemic. The admission time of traced infectious individuals was changed to be a Gamma distribution with a mean (standard deviation) of 1 (0.25) day. In case workplaces were closed, individuals stayed at home or had activities in other spatial locations. The changes were realized by revising bipartite network matrix, contact network matrix.

4.2. Experimental Results

We randomly chosen an individual as the infection source of the SARS epidemic, and executed the epidemic control policies of contact tracing and workplace closure at various time in simulation experiments. We conducted a set of 100 experiments for each group of control policies, and illustrated experimental results in Figure 6.

We observed the evolution of the density of infectious individual over time in Figure 6a. The slower we executed a workplace closing policy, the higher peak the curves got. Moreover, the curves quickly drop down after the peak, and may have serious oscillations in the second half of time. In Figure 6b, we found that with the increase of time when we began to use workplace closing policies, the data points denoting the final epidemic size in experiments scattered at the positive direction of Y axis. The final epidemic size has a higher mean value. In Figure 6c,d, we found that the densities of home places and workplaces approximated to power distributions over the number of infected individuals and the range of infected individuals, respectively. Furthermore, in case we adopted a workplace closing policy earlier, the epidemic invaded a fewer families and workplaces.

We found the correlation effect caused by contact activity heterogeneity and activity location heterogeneity in Figure 6. Individuals' contact events are randomly scheduled in a single day according to a uniform distribution. Moreover, individuals conduct five mobility events in a single day and have activities at different locations in the same moment. When executing a contact event, an individual is only able to contact a neighbor that is also in his current activity location. So, the heterogeneities in the beginning time of contact events, the time duration of contacts, the contact probabilities between neighbors, the beginning time of mobility events, and the probabilities of activity locations bring up temporal-spatial constraints for individuals' contact behaviors. At the same time, the contact patterns of individuals lead to the temporal-spatial characteristics of epidemic diffusion. Due to workplace closure interventions, the epidemic diffusion across different workplaces is mitigated. It is shown in Figure 6d that workplace closure interventions

stop most workplaces from epidemics. However, individuals are likely to stay at home when they cannot go to workplaces. The increase of contact behaviors taking place at home enhances epidemic spread among families. This explains why the tails of curves raise up in Figure 6c.

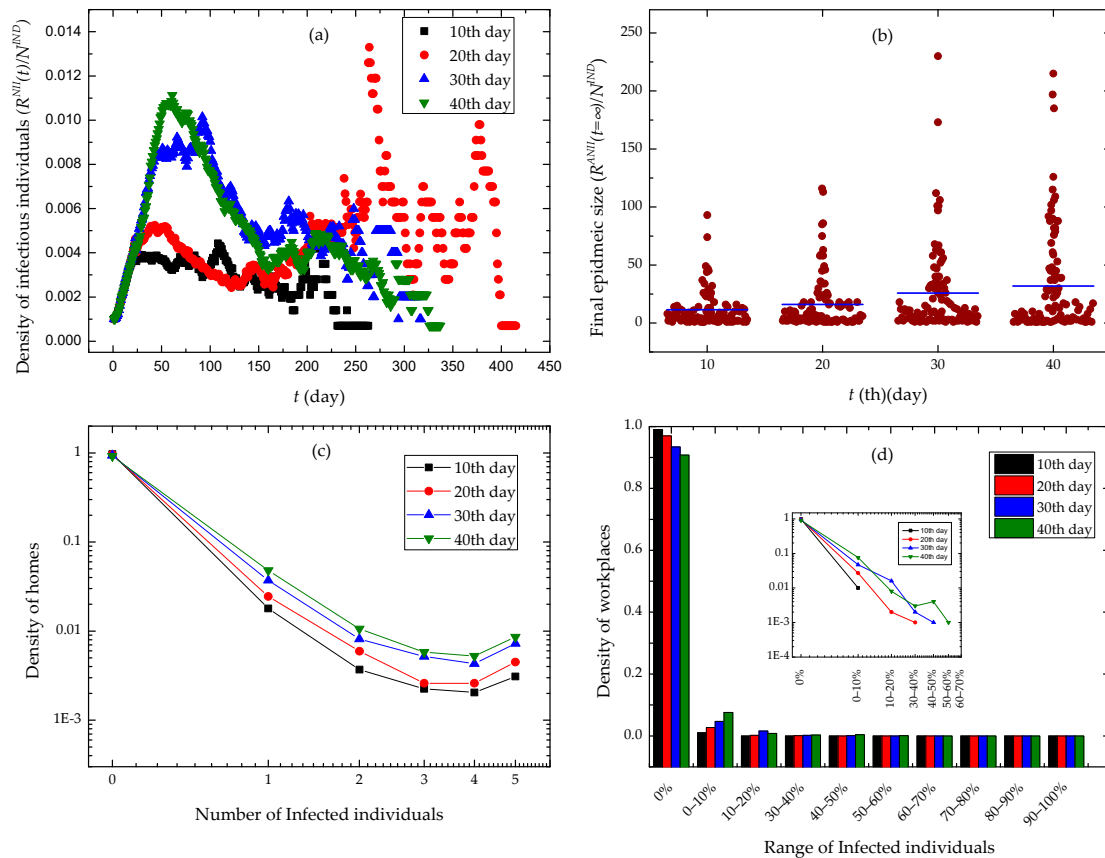


Figure 6. Experimental results. The different days in the legends of subfigures denote the days until workplace closure and contact tracing were implemented. (a) The time evolution of the density of infectious individuals. (b) Final epidemic sizes with different days until contact tracing and workplace closing policies were implemented. (c) The distribution of the density of homes via the number of infected individuals. (d) The distribution of the density of workplaces via the range of infected individuals.

5. Conclusions

Individual heterogeneity is a key factor bridging the link between micro behaviors and macro phenomena. As a bottom-up method, individual-based epidemic models exhibit the capability of representing the heterogeneities in epidemic spreading process. However, the complexity of individual-based models results in the limitation of mathematical formulation.

In this paper, we used matrices to formulate heterogeneous individual-based epidemic models. We adopted matrices and vectors to represent entities and their attributes in epidemic systems, such as individuals, individuals' health states, the infectivity of infectious individuals, spatial locations, and bipartite networks. Then, we derived analytical results of epidemic diffusion from the matrix-based formulation of entities, and designed algorithms to force the computation of heterogeneous individual-based epidemic models. Finally, we applied the SARS epidemic as a case study to test and verify the heterogeneous individual-based epidemic models in matrix-based formulation. In this case study, we investigated the effect of non-pharmaceutical interventions on the SARS epidemic control, including contact tracing and workplace closure. Experimental results indicate that in case of executing workplace closing interventions early, we can obtain a better effect on mitigating SARS epidemic diffusion. Moreover, we also analyzed the correlation effect

caused by contact activity heterogeneity and mobility location heterogeneity, as well as the temporal-spatial patterns of epidemic spread.

Heterogeneous individual-based epidemic models formulated by matrices and vectors are scalable and flexible. We could add more matrices into the models to describe more entities in epidemic systems. Furthermore, we could improve the algorithms or design new algorithms to expand the capability of the models. More analytical results could be formulated from matrix-based representations of entities and their attributes. In addition, heterogeneous individual-based epidemic models formulated by matrices and vectors not only can capture individual heterogeneity, but also have the ability of performing theoretical analysis of epidemic diffusion.

Funding: This work was supported by the National Natural Science Foundation of China under Grant. 72071207, 61673388, 61503402.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The author declares that there are no conflicts of interest.

References

- Kim, M.; Paini, D. Modeling stochastic process in disease spread across a heterogeneous social system. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 401–406. [[CrossRef](#)]
- Chen, T.-M.; Rui, J.; Wang, Q.-P.; Zhao, Z.-Y.; Cui, J.-A.; Yin, L. A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. *Infect. Dis. Poverty* **2020**, *9*, 24. [[CrossRef](#)]
- Hao, X.; Cheng, S.; Wu, D.; Wu, T.; Lin, X.; Wang, C. Reconstruction of the full transmission dynamics of COVID-19 in Wuhan. *Nature* **2020**, *584*, 420–424. [[CrossRef](#)] [[PubMed](#)]
- Woolhouse, M.E.J.; Dye, C.; Etard, J.-F.; Smith, T.; Charlwood, J.D.; Garnett, G.P.; Hagan, P.; Hii, J.L.; Ndhlovu, P.D.; Quinnell, R.J.; et al. Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 338–342. [[CrossRef](#)] [[PubMed](#)]
- Peng, H.; Peng, W.; Zhao, D.; Wang, W. Impact of heterogeneity of adoption thresholds on behavior spreading in complex networks. *Appl. Math. Comput.* **2020**, *386*, 125504. [[CrossRef](#)]
- Chabas, H.; Lion, S.; Nicot, A.; Meaden, S.; van Houte, S.; Moineau, S.; Wahl, L.M.; Westra, E.R.; Gandon, S. Evolutionary emergence of infectious diseases in heterogeneous host populations. *PLoS Biol.* **2018**, *16*, e2006738. [[CrossRef](#)]
- Zhu, S.-S.; Zhu, X.-Z.; Wang, J.-Q.; Zhang, Z.-P.; Wang, W. Social contagions on multiplex networks with heterogeneous population. *Physica A* **2019**, *516*, 105–113. [[CrossRef](#)]
- Caswell, H.; Vindenes, Y. Demographic variance in heterogeneous populations: Matrix models and sensitivity analysis. *OIKOS* **2018**, *127*, 648–663. [[CrossRef](#)]
- Tong, Y.; Lei, C. An SIS epidemic reaction-diffusion model with spontaneous infection in a spatially heterogeneous environment. *Nonlinear Anal. Real World Appl.* **2018**, *41*, 443–460. [[CrossRef](#)]
- Wang, X.; Zhao, X.-Q.; Wang, J. A cholera epidemic model in a spatiotemporally heterogeneous environment. *J. Math. Anal. Appl.* **2018**, *468*, 893–912. [[CrossRef](#)]
- Kuniya, T.; Inaba, H.; Yang, J. Global behavior of SIS epidemic models with age structure and spatial heterogeneity. *Jpn. J. Ind. Appl. Math.* **2018**, *35*, 669–706. [[CrossRef](#)]
- Yang, Z.; Cui, A.X.; Zhou, T. Impact of heterogeneous human activities on epidemic spreading. *Physica A* **2011**, *390*, 4543–4548. [[CrossRef](#)]
- Merler, S.; Ajelli, M. The role of population heterogeneity and human mobility in the spread pandemic influenza. *Proc. R. Soc. B* **2009**, *277*, 557–567. [[CrossRef](#)] [[PubMed](#)]
- Shang, Y. Modeling epidemic spread with awareness and heterogeneous transmission rates in networks. *J. Biol. Phys.* **2013**, *39*, 489–500. [[CrossRef](#)] [[PubMed](#)]
- Cui, P.-B.; Gao, L.; Wang, W. Epidemic spreading dynamics with drug resistance and heterogeneous contacts. *J. Theor. Biol.* **2018**, *441*, 19–27. [[CrossRef](#)]
- Lloyd-Smith, J.O.; Schreiber, S.J.; Kopp, P.E.; Getz, W.M. Super spreading and the effect of individual variation on disease emergency. *Nature* **2005**, *438*, 355–359. [[CrossRef](#)] [[PubMed](#)]
- Mkhatshwa, T.; Mummert, A. Modeling super-spreading events for infectious disease: Case study SARS. *IAENG Int. J. Appl. Math.* **2011**, *41*, 82.

18. Duan, W.; Qiu, X.; Cao, Z.; Zheng, X.; Cui, K. Heterogeneous and stochastic agent-based models for analyzing infectious diseases' super spreaders. *IEEE Intell. Syst.* **2013**, *28*, 18–25. [[CrossRef](#)]
19. Chen, S.; Small, M.; Tao, Y.; Fu, X. Transmission dynamics of an SIS model with age structure on heterogeneous networks. *Bull. Math. Biol.* **2018**, *80*, 2049–2087. [[CrossRef](#)]
20. Chen, S.; Wang, K.; Sun, M.; Fu, X. Spread of competing viruses on heterogeneous networks. *Philos. Trans. R. Soc. A* **2017**, *375*. [[CrossRef](#)]
21. Lee, M.J.; Lee, D.-S. Understanding the temporal pattern of spreading in heterogeneous networks: Theory of the mean infection time. *Phys. Rev. E* **2019**, *99*, 032309. [[CrossRef](#)]
22. Tolić, D.; Kleineberg, K.-K.; Antulov-Fantulin, N. Simulating SIR processes on networks using weighted shortest paths. *Sci. Rep.* **2018**, *8*, 6562. [[CrossRef](#)]
23. Unicomb, S.; Iñiguez, G.; Karsai, M. Threshold driven contagion on weighted networks. *Sci. Rep.* **2018**, *8*, 3094. [[CrossRef](#)]
24. Shao, F.; Chang, Z. Suppress traffic-driven epidemic spreading in weighted network. *Clust. Comput.* **2019**, *22*, 14201–14206. [[CrossRef](#)]
25. Shen, D.; Cao, S. An efficient immunization strategy based on transmission limit in weighted complex networks. *Chaos Solitons Fractals* **2018**, *114*, 1–7. [[CrossRef](#)]
26. Rahmandad, H.; Sterman, J. Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Manag. Sci.* **2008**, *54*, 998–1014. [[CrossRef](#)]
27. Perkins, T.A.; Scott, T.W.; Menach, A.L.; Smith, D.L. Heterogeneity, mixing, and the spatial scales of mosquito-borne pathogen transmission. *PLoS Comput. Biol.* **2013**, *9*, e1003327. [[CrossRef](#)] [[PubMed](#)]
28. Waldrop, M.M. Free agents: Monumentally complex models are gaming out disaster scenarios with millions of simulated people. *Science* **2018**, *360*, 144–148. [[CrossRef](#)]
29. Duan, W.; Cao, Z.; Wang, Y.; Zhu, B.; Zeng, D.; Wang, F.Y.; Qiu, X.; Song, H. An ACP approach to public health emergency management: Using a campus outbreak of H1N1 influenza as a case study. *IEEE Trans. Syst. Man Cybern. Syst.* **2013**, *43*, 1028–1041. [[CrossRef](#)]
30. Duan, W.; Fan, Z.; Zhang, P.; Guo, G.; Qiu, X. Mathematical and computational approaches to epidemic modeling: A comprehensive review. *Front. Comput. Sci.* **2015**, *9*, 806–826. [[CrossRef](#)] [[PubMed](#)]
31. Vázquez, A.; Oliveira, G.; Dezső, Z.; Goh, K.I.; Kondor, I.; Barabási, A.L. Modeling bursts and heavy tails in human dynamics. *Phys. Rev. E* **2006**, *73*, 036127. [[CrossRef](#)] [[PubMed](#)]
32. Edmunds, W.J.; O'Callaghan, C.J.; Nokes, D.J. Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. *Proc. R. Soc. Biol. Sci.* **1997**, *264*, 949–957. [[CrossRef](#)] [[PubMed](#)]
33. World Health Organization. Consensus Document on the Epidemiology of Severe Acute Respiratory Syndrome (SARS). 2003. Available online: www.who.int/csr/sars/en/WHOconsensus.pdf (accessed on 10 January 2021).
34. Donnelly, C.A.; Ghani, A.C.; Leung, G.M.; Hedley, A.J.; Fraser, C.; Riley, S.; Abu-Raddad, L.J.; Ho, L.M.; Thach, T.Q.; Chau, P.; et al. Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. *Lancet* **2003**, *361*, 1761–1766. [[CrossRef](#)]