科学通报 2009 年 第 54 卷 第 24 期: 3834~3843
论文 www.scichina.com csb.scichina.com
《中国科学》杂志社 SCIENCE IN CHINA PRESS

基于小世界网络的多智能体及传染病时空传播模拟

刘涛, 黎夏*, 刘小平

中山大学地理科学与规划学院, 广州 510275

* 联系人, E-mail: lixia@mail.sysu.edu.cn

2009-04-07 收稿, 2009-08-17 接受

国家自然科学基金重点项目(批准号: 40830532)、国家杰出青年科学基金(批准号: 40525002)、国家高技术研究发展计划(编号: 2006AA12Z206)、资源与环境信息系统国家重点实验室开放课题(批准号: 4106298)资助项目

摘要 提出建立一种新的基于小世界网络的多智能体传染病时空传播模型, 模型中多智能体系统用于反映个体间的时空动态相互作用, 而小世界网络用来表示多智能体之间的社会关系. 模型主要由智能体属性定义、移动规则、邻域、社会关系网络的构建以及智能体状态转换规则几部分所组成. 其中, 智能体社会关系网络的构建以及状态转换规则的确定是模型的核心. 与传统的传染病传播模拟模型相比, 该模型引入了个体对感染"记忆"的时间衰减效应、个体间的空间距离以及社会关系对感染的影响. 利用所提出的模型, 选取广州市为实验区, 在 Swarm 平台上模拟了流感的时空传播过程. 与传统的传染病系统动力学 SEIR 模型以及基于多智能体系统的传染病传播模型相比, 该模型的模拟结果与实际观测数据更为接近. 建立可用于真实地理环境的传染病时空传播模拟模型, 将有助于把握传染病的时空传播规律, 为相关部门控制传染病的蔓延提供科学的决策依据.

关键词
多智能体系统
小世界网络
传染病
时空传播
地理信息系统

2003 年的 SARS 以及 2009 年的甲型 H1N1 流感的全球大范围爆发都给人们的生命财产带来了巨大的危害. 传染病时空传播模拟模型的研究, 将有助于把握传染病的时空传播规律, 为相关部门控制传染病的蔓延提供科学的决策依据.

目前, 传染病传播模型的研究主要包括动力学方程、元胞自动机模型、多智能体系统以及复杂网络等几大类方法. Kermack 等人[1]最早提出了基于微分方程的传染病 SIR 仓室模型, 为传染病传播动力学方程研究奠定了基础. 该模型中, 将人群分为易感者 S(Susceptible) 、感染者 I(Infective) 和移出者 R(Removed)三类. 此后, 许多学者在其基础上不断完善, 陆续提出了 SEIR(考虑疾病的潜伏期)、SIRS(病人治愈后并非永久免疫)等传染病传播动力学方程[2]. 石耀霖[3]基于对每个病人的感染、发病和传染随机过程主要特征的模拟, 构建了 SARS 传播的系统动力学随机模型. 汪海英等人[4]对动力学模型进行改进以评估流行的可能性和控制措施的有效性. 但传染病动力学方程基于均质空间的假设, 大多忽略了人口等因素空间分布的异质性对传染病传播的影响[5].

元胞自动机模型(cellular automata, 简称 CA)首先由 Von Neumann 和 Ulam 于 1948 年提出, 常被用于模拟复杂系统的时空演化过程. Fu[6]进行了两类实验: 第一类实验研究 CA 模型是否能够对传染病传播的宏观行为进行预测, 而第二类实验模拟传染病的空间行为. Mikler 等人[7]针对传统 CA 在模拟较大区域时存在的邻域饱和度问题, 提出了一种全局随机元胞自动机模型 GSCA. 模型中元胞空间的相互作用是全局性的, 即空间任何两个元胞都可能发生作用. Zhang 等人[8]则侧重于利用 CA 模型研究"郊区

引用格式: 刘涛, 黎夏, 刘小平. 基于小世界网络的多智能体及传染病时空传播模拟. 科学通报, 2009, 54: 3834~3843
Liu T, Li X, Liu X P. Integration of small world networks with multi-agent systems for simulating epidemic spatiotemporal transmission. Chinese Sci Bull, doi: 10.1007/s11434-009-0623-3

化”对传染病传播的影响. 但现有的基于 CA 模型的传染病模拟模型忽视了个体间的社会关系在疾病传播过程中的作用. 而相关研究表明社会关系在传染病传播过程中占有着非常重要的角色[9].

多智能体系统(multi-agent system, 简称 MAS)是复杂适应系统、人工生命以及分布式人工智能等多学科的融合. 在多智能体系统中, 通过微观个体的行为以及交互作用, 涌现出全局的非线性特征. 目前该方法已成为进行复杂系统分析及模拟的重要手段, 被广泛用于经济、政治、社会、生态等领域的模拟中[10~13]. 刘颖等人[14]、邓宏钟等人[15]利用多智能体进行计算机模拟的方法, 模拟了 SARS 的爆发流行以及受到控制趋于稳定的过程. 但现有的基于多智能体系统的传染病传播模型较多侧重于研究传染率、治愈率等参数对传染病传播的影响. 事实上, 传染病的传播过程不仅是随时间演变的, 同时也是一种空间过程[16,17]. 部分研究对传染病空间传播过程有所涉及, 譬如, 龚建华等人[18]结合虚拟地理环境与多智能体技术模拟 SARS 在虚拟社区中的传播. 然而, 基于多智能体的传染病模拟模型均未考虑到智能体之间的社会关系网络对疾病传病的影响.

Milgram[19]的“六度分离”实验最早揭示了客观世界中存在着的“小世界”现象. 而小世界网络(small world network, 简称 SWN)模型是 Watts 和 Strogatz 根据人类社会网络提出的网络模型[20]. 小世界网络具有较大的聚类系数和较小的平均路径长度的特点. 小世界网络提出后引起学术界的关注, 在物理、数学、生物等自然科学领域都有较广泛地应用. Moore 等人[9]研究了一维空间中小世界网络中传染病的传播过程. 现有的基于小世界网络的传染病传播模拟模型研究目前仍处于起步阶段, 特别是可用于模拟真实地理环境中传染病的时空传播模型还非常有限.

本文提出建立一种结合多智能体系统、小世界网络模型以及地理信息系统的传染病时空传播模拟新模型. 其中多智能体系统主要用于表现个体的时空动态相互作用过程, 小世界网络被用来描述智能体之间的社会关系, 地理系统系统则负责空间数据的管理以及可视化. 以流感为例, 模拟现实世界中传染病的时空传播过程.

1 模型

1.1 多智能体模型

流感是由病毒所引起的急性呼吸道感染, 主要通过飞沫传播. 流感的传染性强、传播速度快, 具有一定的破坏力, 特别是秋冬季流感高发期, 所引起的并发症和死亡现象非常严重. 绝大部分人容易感染流感病毒, 一旦痊愈后会获得对同型流感病毒的免疫力, 时间约维持 8~12 月左右. 因此在研究时段较短的情况下, 可认为病人痊愈后具有免疫能力, 不会再次被感染[21].

在传染病系统动力学 SEIR 方程中将总人口分为易感类、潜伏类、感染类和移出类 4 类. 其中, $S(t)$, $E(t)$, $I(t)$, $R(t)$分别为 t 时刻易感类、潜伏类、感染类和移出类的数量. 每一时刻 4 类人数之和即为总人数, 可近似认为保持不变. SEIR 传染病传播系统动力学方程可表示为如下微分方程的形式[22]:

$$\begin{aligned}
&\frac{\mathrm{d}S}{\mathrm{d}t} = -\beta\frac{S(t)I(t)}{N},\\
&\frac{\mathrm{d}E}{\mathrm{d}t} = \beta\frac{S(t)I(t)}{N} - \sigma E(t),\\
&\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma E(t) - \eta I(t),\\
&\frac{\mathrm{d}R}{\mathrm{d}t} = \eta I(t),\\
&S(t) + E(t) + I(t) + R(t) = N,
\end{aligned} \tag{1}$$

其中, β 为传染率, 反映了病人的活动能力、环境条件以及病毒的毒力等因素, σ 为由潜伏病例转移为感染病例的比率, $1/\sigma$ 表示流感的潜伏期, η 为感染病例转移为移出病例的比率, $1/\eta$ 表示流感的感染期.

可借鉴 SEIR 模型的人口划分方法将个体分为易感者(S)、潜伏者(E)、感染者(I)、移出者(R)4 类智能体. 可通过智能体的属性定义、移动规则、邻域和状态转换规则几部分构建传染病传播的多智能体模型. 其中, 智能体状态转换规则是基于多智能体系统的传染病时空传播模型的核心.

定义智能体主要包括潜伏天数、感染天数、感染剂量、年龄组和住处等属性. 其中, 潜伏天数表明感染者智能体处于潜伏期第几天, 同理感染天数表明感染者智能体处于感染期的第几天. 传染病传播临界值模型表明, 易感者智能体与感染者智能体接触时, 并非立即被感染, 而是接收到一定量的“毒素”, 即感染剂量, 该值累积超过一定的阈值时个体受到传染[23]. 因此, 可为智能体设定感染剂量属性, 某种程度上, 其数值决定着易感者智能体是否被感染以及何时会被感染. 智能体年龄组属性的取值范围为婴幼儿(0~4 岁)、少年(5~19 岁)、中青年(20~59 岁)

和老年(60 岁以上) 4 个年龄组.

可根据个体的日常行为规律确定智能体的移动规则. 0~8 点大部分智能体都出于休息状态, 可以近似认为传染事件不会发生. 8~16 点智能体外出, 16~24 点智能体回到家中, 这两个时间段智能体间接触的机会较多, 传染的机率也较高. 因此, 每天由 8~16 点和 16~24 点两个时段所组成. 因此, 模拟时每个时间步所代表的实际时间为 8 h. 智能体的移动方式主要为周边移动, 即智能体在居住地附近街区进行工作、生活和上学等日常活动(图 1).

邻域表示智能体的活动所能影响到的范围. 流感的传播规律表明, 当易感者与感染者近距离接触时发生传染的概率较高, 而距离大于一定范围时, 传染事件发生的概率大大降低. 图 2 所示, 阴影部分表示中心位置的感染者智能体的邻域, 邻域内的易感者智能体受中心位置的感染者智能体影响, 且距离越近, 所受影响越大, 邻域外的智能体可认为不受该智能体的影响.

1.2 基于小世界网络的多智能体模型

小世界网络模型反映了个体关系网络的一种特性, 即大部分人的朋友都是和他们住在同一条街道上的邻居或在同一单位上班的同事. 社会网络、互联网和基因网络都呈现小世界网络的特征. 网络的平均路径长度定义为网络中任意两个结点之间距离的平均值. 各结点的聚类系数数值上等于其所有相邻结点之间连边的数目占可能的最大连边数目的比例. 而整个网络的聚类系数等于各结点聚类系数的平均值[24]. 网络是否具有小世界特征主要取决于网络的平均路径距离和聚类系数. 在小世界网络中, 尽管网络的规模很大, 但在网络任意 2 个结点之间的平均路径长度小的惊人. 小世界网络中结点的度定义为与该结点连接的边的数目[24].

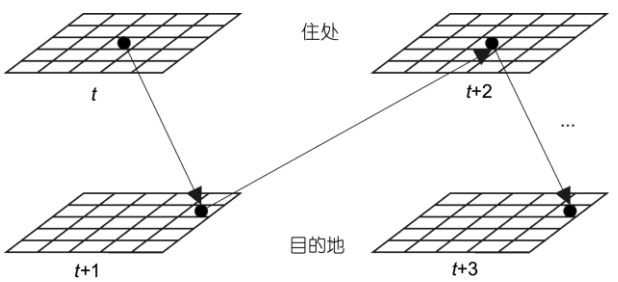


图 1 智能体移动规则

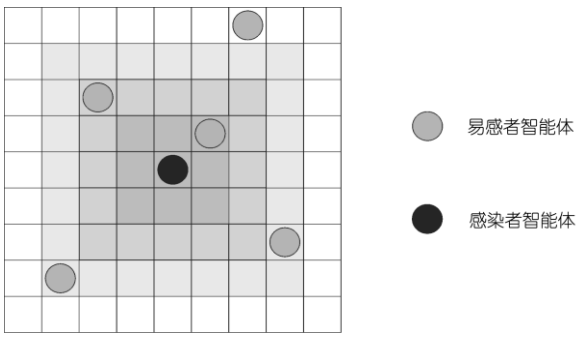


图 2 智能体邻域规则

$$U_v = \sum_{l \in E} \delta_l^v , \tag{2}$$

其中, U_v 表示结点 v 的度, E 为所有边的集合, 当边 l 包含结点 v 时, δ_l^v 取值为 1, 否则为 0. 某个结点的度越大就意味着这个结点在某种意义上越“重要”. 网络中所有结点度的平均值称为网络的结点平均度.

由于多智能体系统可用于描述小世界网络的结构, 而小世界网络中的结点可看作是多智能体系统中的智能体, 结点之间的连接可用于表示智能体之间的社会联系. 因此, 可建立基于小世界网络的多智能体模拟模型, 明确考虑智能体之间的社会关系在模拟过程中的作用. Newman 等人[25]提出的 NW 小世界模型通过“随机化加边”的方法构造小世界网络, 是目前使用最为广泛的小世界网络构造方法. 本研究中将智能体之间的社会关系抽象为家庭、同学、同事、邻居和朋友等几种类型. 通过以下几个步骤构建智能体之间的小世界网络:

(1) 按照就近原则建立智能体之间的家庭关系. (2) 根据学校的空间分布情况, 将少年年龄组智能体按照就近原则分配到附近的学校上. 在同一学校上学的智能体之间随机加边, 生成智能体的同学关系. (3) 为同一位置附近工作的中青年年龄组智能体随机加边, 生成智能体的同事关系. (4) 在 10×10 网格窗口内的智能体之间随机加边, 生成智能体的邻居关系. (5) 在全部智能体之间随机加边, 生成智能体的朋友关系. 取上述图的并集, 即得到智能体的社会关系网络.

若智能体对应的社会关系网络中有边相连接, 则两个智能体有社会关系, 否则定义没有社会关系.

$$w = \delta_l^{ij} , \tag{3}$$

其中, i, j 分别对应网络的结点 i 和结点 j, 若结点 i, j 之间有边 l 通过, 则 w 值为 1, 否则取值为 0.

1.3 智能体状态转换规则的确定

智能体之间的状态可能发生改变. 易感者智能体被感染者智能体传染后成为潜伏者智能体; 潜伏期内潜伏者智能体没有传染性, 一旦潜伏期结束, 潜伏者智能体转换为感染者智能体; 而感染者智能体具有传染性, 感染期内每天与其有亲密接触关系的人都有一定的概率被传染; 感染期结束后, 感染者智能体一旦痊愈后就获得了对该传染病的免疫能力, 成为移出者智能体, 不会再次被传染[26].

从易感者到感染者的状态转换, 许多研究都基于感染概率加以确定, 并假定该参数在模拟过程中始终保持不变[14,15]. 事实上, 在传染病传播过程中, 存在着"时空"两个方面的影响. 一方面, 如果易感者前一两天曾与感染者接触, 则该时刻易感者受感染的概率也就越高. 另一方面, 易感者与感染者的空间距离越近, 感染的概率也就越高, 反之距离越远, 感染的概率就越低. 因此, 本模型利用"时空"两个方面来确定易感者是否受到感染以及何时受到感染.

(1) 计算各易感者智能体所接收到感染者的感染剂量总和. 假设 t'时刻易感者智能体 i 与感染者智能体接触而收到感染剂量 $d_{ik}(t')$, 其计算公式如下:

$$d_{ik}(t')=\begin{cases} d & w=1 \\ d/R_s & w=0 \end{cases} \quad \alpha \leqslant p, \tag{4}$$

其中, d 为单位感染剂量值, R_s 为易感者智能体与感染者智能体之间的距离, α 为随机变量, p 为易感者智能体与感染者智能体接触时接收到感染剂量的概率.

而 t'时刻易感者智能体所受到的总的感染剂量 $D_i(t')$在数值上等于与其有接触的所有智能体对其施加感染剂量的和[23].

$$D_i(t')=\sum_{k=1}^{K} d_{ik}(t'), \tag{5}$$

式中 K 为 t'时刻与该易感者智能体接触的感染者智能体个数.

(2) 计算各易感者智能体 T时间内的累积感染剂量. Dodds 等人[23]提出的推广模型中将易感者过去一段时间内所接收到的感染剂量作为个体对感染的"记忆"引入到模型中. 本研究中则进一步考虑个体对感染"记忆"的时间衰减效应. 可利用下式计算易感者智能体 t 时刻的累积感染剂量:

$$D_i(t)=\sum_{t'=t-T+1}^{t} \gamma^{t-t'} D_i(t'), \tag{6}$$

式中 γ 为引入的感染剂量的衰减率, $\gamma\in[0,1]$, 当 $\gamma=1$ 时为不考虑感染的衰减效应, T 为易感者对感染"记忆"的时间长度.

(3) 比较易感者的累积感染剂量与感染临界值. 根据传染病传播的临界值模型, 若易感者智能体在 T 时间内的累积感染剂量大于临界值, 则该易感者智能体受到感染, 否则不受感染:

$$D_i(t)\begin{cases} >=d^*, \text{感染} \\ <d^*, \text{不感染} \end{cases}, \tag{7}$$

式中 d^*为判断是否感染的临界值.

基于小世界网络的多智能体传染病时空传播模型结构如图 3 所示.

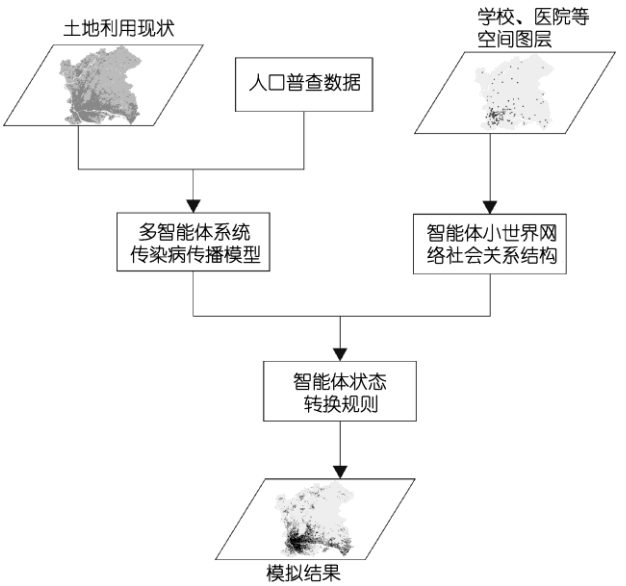


图 3 基于小世界网络的多智能体传染病时空传播模型

2 应用

2.1 研究区及数据处理

选择广州市区作为研究区, 主要包括广州市 2005 年行政区划调整前的老八区即越秀区、东山区、荔湾区、天河区、芳村区、海珠区、黄埔区和白云区. 其中, 越秀区、东山区和荔湾区为老城区, 人口稠密. 白云区位于研究区的北半部, 面积较大, 大部分为林地、耕地和园地所覆盖, 人口密度相对较小.

模型中所使用的数据主要分为人口数据、GIS 空间数据以及流感病例数据. 人口数据主要包括以街道办(镇)为单位的广州市第 5 次人口普查数据, 内容主要是个街道办(镇)的人口以及各年龄段的人数.

GIS 数据主要是广州市区 2006 年土地利用现状图, 中小学、住宅用地空间分布矢量图层. 流感病例数据主要为 2006 年 1~3 月逐日报告的流感病例数据, 前两个月的流感观测数据用于参数估计, 3 月份的流感观测数据则用于模型精度检验. 所有空间数据都转成 Grid 栅格数据格式, 并通过空间配准后叠加在一起, Grid 数据的分辨率为 100 m.

2.2 模拟平台

Swarm 是由 Santa Fe 研究所于 1994 年为多智能体系统建模而设计的软件平台. Swarm 的标准类库里面包括许多可重用的类以支持计算机模拟实验的进程控制、参数调整、数据分析以及图形显示. 用户可以更专注于具体研究领域的问题和算法, 利用 Swarm 提供的类库构建多智能体模拟系统, 系统中的各个智能体通过离散事件相互作用, 相应的调度机制保证了模拟过程中各个时间步骤依次得到执行, 并且支持按照随机的序列方式调用各种多智能体类的多个实例. 本研究中利用 java 语言在 Swarm 标准类库的基础上开发了模拟所需要的 GIS 数据读入、动态显示等功能.

2.3 模型简化

为了便于实现, 有必要作如下假设, 对模型进行简化: (1) 潜伏者智能体在潜伏期结束后全部转变为感染者智能体. (2) 所有智能体分布在网格空间中, 并且每个网格空间单元中最多只容纳一个智能体. (3) 由于研究区的人口较多, 如果将每个个体都表示为智能体, 程序所要占用的内存空间将大大超出计算机的承受能力, 最终无法运行. 而相关研究表明, 在传染病的实际传播过程中, 感染者的密切接触者在整个研究区人口中所占比例是非常小的[27]. 因此我们可以进行这样的近似处理, 生成一定数目的智能体, 用这些智能体代表模型运行过程中所有感染者智能体所可能接触到的易感者智能体.

2.4 模型实现

(ⅰ) 智能体属性初始化

由于全国性的人口普查一般每 10 年进行一次, 当需要使用两次普查之间年份的人口年龄结构数据时, 只能进行人口预测. 本研究主要采用生命表法进行人口年龄结构的预测, 该方法是人口统计分析与应用的重要内容和分析工具, 是人口研究过程中最常用、最基本的方法之一[28]. 利用中国人口预测软件计算得到 2006 年研究区年龄结构人口数据[29].

通过人口普查并进行逐级汇总所得到的人口数据是典型的统计型属性数据, 通过预测所得到的 2006 年人口数据也同样具备这一特点. 用这种数据进行人口、资源环境与社会经济发展的宏观分析是比较方便的, 进行微观尺度的研究时这种数据就不能反映人口空间分布的真实状况. Tian 等人[30]研究了土地利用数据与 DEM、坡度、温度等影响人口分布影响因素的相关性, 其结果表明土地利用类型与这些因素有着高度的相关性. 因此可以把土地利用类型作为控制变量, 将人口按行政单元统计的结果, 按照土地是否可居住的原则进行分解, 使得人口分布更能接近于实际.

(1) 计算每个网格的人口密度. Clark[31]于 1951 年对美国 20 多个城市的人口密度空间分布进行研究发现, 随着从城市中心向外围距离的增加, 人口密度与距离之间呈负指数关系. 谢守红等人[32]的验证结果表明广州市的人口分布同样呈现典型的 Clark 分布. 可用 Clark 人口密度模型计算研究区人口密度[31]:

$$F_i = F_0 e^{-\lambda r_i}, \tag{8}$$

其中, r_i为网格 i到城市中心的距离, 利用该网格的坐标与城市中心的坐标计算可得, F_i为该网格的人口数, F_0 为比例系数, 它数值上等于城市中心处的人口数, λ 为距离衰减效应的速率. F_0 可由下式计算:

$$F_0 = P_0 / A_0, \tag{9}$$

其中, P_0 为城市中心城区某街道办的总人口数, A_0 为该街道办居住用地的网格数. 广州市越秀区属于广州市老城区, 人口非常稠密, 而解放南街道大致位于越秀区的中心位置, 人口密度最大, 可将其作为研究区中心. 计算得到 F_0 取值约为 2221 人/网格.

对某一街道办(镇)内所有网格的人口数进行累加求和, 其人口可用如下公式表示:

$$\sum_{i=1}^{m} F_0 e^{-br_i} = P, \tag{10}$$

其中, P 为街道办(镇)的人口, m 为该区域“可居住”网格的总数. 对于同一街道(镇)内的各网格可近似认为 λ取值相同. 利用“二分法”不断缩小待求解的范围, 直到该范围足够小为止, 即可计算出距离衰减速率 λ 的可接受的近似解. 将该值代入(8)式即可得到该街道办(镇)中每个“可居住”网格的人口数 F_i(图 4).

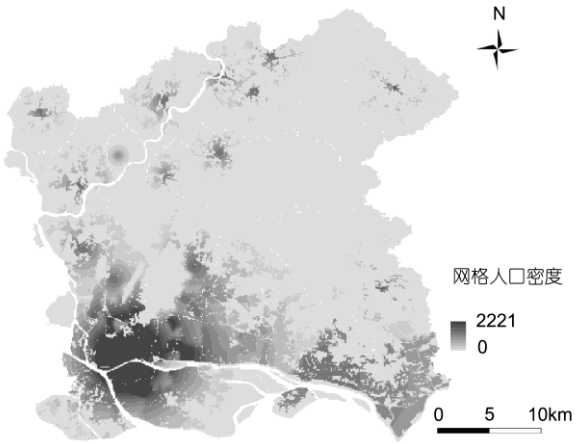


图 4 研究区网格人口密度图(单位: 人/网格)

(2) 计算每个网格上的相对人口密度 P_r. 通过步骤(1)所得到的结果是每个网格的人口密度, 而事实上模拟中由于实现的需要假设每个网格上最多只有 1 个智能体. 因此需要把网格人口密度进行标准化为[0,1]之间的数值, 得到网格的相对人口密度 P_r, 近而将该相对人口密度看作是智能体出现的概率.

(3) 确定某个网格是否会出现智能体. 在[0,1]区间内产生一个均匀分布的随机数 τ, 若 $\tau \leqslant P_r$, 则该网格分布有智能体, 否则智能体在该网格中不出现. 模拟中所生成的智能体总数为 2 万个. 利用该方法可以近似模拟人口的空间分布, 这样每个智能体的住处属性就得到了初始化.

此外, 还需要对智能体的年龄组属性进行初始化. 在人口预测所得到的人口年龄结构数据中每 5 岁为一个年龄段, 需要根据模型的要求对其进行合并, 分为 0~4 岁, 5~19 岁, 20~59 岁, 60 岁以上 4 个年龄组, 分别对应婴幼儿、少年、中青年和老年, 进而可得到每个年龄组所占的比例. 具体对某一个智能体的年龄组进行初始化时, 可利用赌轮选择机制确定. 主要包括以下几个步骤:

(1) 先按照每种年龄组所占比例创建赌轮. 第 k 种年龄组被选择的概率 P_k 可根据如下公式计算.

$$P_k = \frac{A_k}{\sum_{k=1}^{4} A_k} (1 \leqslant k \leqslant 4), \tag{11}$$

其中 A_k 为第 k 种年龄组在整个年龄组中所占比例;

(2) 在[0,1]区间内产生一个均匀分布的随机数 θ;

(3) 若 $\theta \leqslant P_1$,则智能体的年龄组属性取值为 1, 即为婴幼儿;

(4) 若 $Q_{k-1} < \theta \leqslant Q_k (2 \leqslant k \leqslant 4)$, 则智能体的年龄组属性取值为 k, 其中, Q_k 被称为年龄组 k 的累积概率, 计算公式为:

$$Q_k = \sum_{j=1}^{4} P_j . \tag{12}$$

(ⅱ) 智能体小世界网络关系的构建

按照 1.2 中的方法得到下图所示的小世界网络关系图(图 5). 利用公式(2)计算得到小世界网络的平均度值为 13.01, 即平均每个智能体的社会关系数约为 13 个. 计算得到网络平均路径长度为 8.62, 即任意两个智能体之间平均通过不到 9 个智能体就能够发生联系, 网络的聚类系数为 0.23.

2.5 结果及分析

流感潜伏期取值为 3 d, 传染期取值为 5 d[7]. 根据参数的实际意义, 对参数的范围加以界定 $p \in [0, 0.5]$, $\gamma \in [0.5, 1]$, $T \in [1, 5]$, $d^* \in [d, 4d]$. 考虑到婴幼儿和老人对流感的抵抗能力较弱, 因此其感染阈值应小于少年和中青年的感染阈值, 故婴幼儿和老人的感染阈值 $d^* \in [d, 3d]$. 对易感者与感染者接触受到感染剂量的概率 p, 距离衰减系数 γ, 感染的 “记忆”长度 T, 感染阈值 d^*4 个参数进行估计, 定义 4 个参数的变化步长分别为 0.05, 0.02, 1, d. 当 p, γ, T, d^* 取值分别为 0.20, 0.78, 3, 3d(婴幼儿和老人为 2d)时, 模拟结果与流感实际观测数据的拟合效果较好. SEIR 模型的传染率 β 取值范围为[0.24, 0.26][21], 本研究中取 β=0.25.

分别利用基于小世界网络的多智能体模型(SWN-MAS 模型)、传统的 SEIR 模型以及多智能体系统(MAS 模型)模拟研究区中的流感传染过程, 并将其与 2006 年 3 月份流感实际观测数据进行比较.

图 6 所示, 在多智能体系统中加入小世界网络社会关系时, 计算得到模拟结果的精度为 78.2%, SEIR 模型的模拟精度为 67.8%, 而多智能体模型的模拟结果的精度为 69.3%. 可见, 智能体之间的小世界网络关系在传染病传播过程中是不可忽视的因素, 将其引入到多智能体模型中将使模拟结果的精度得到提高, 其结果也更接近于实际观测数据.

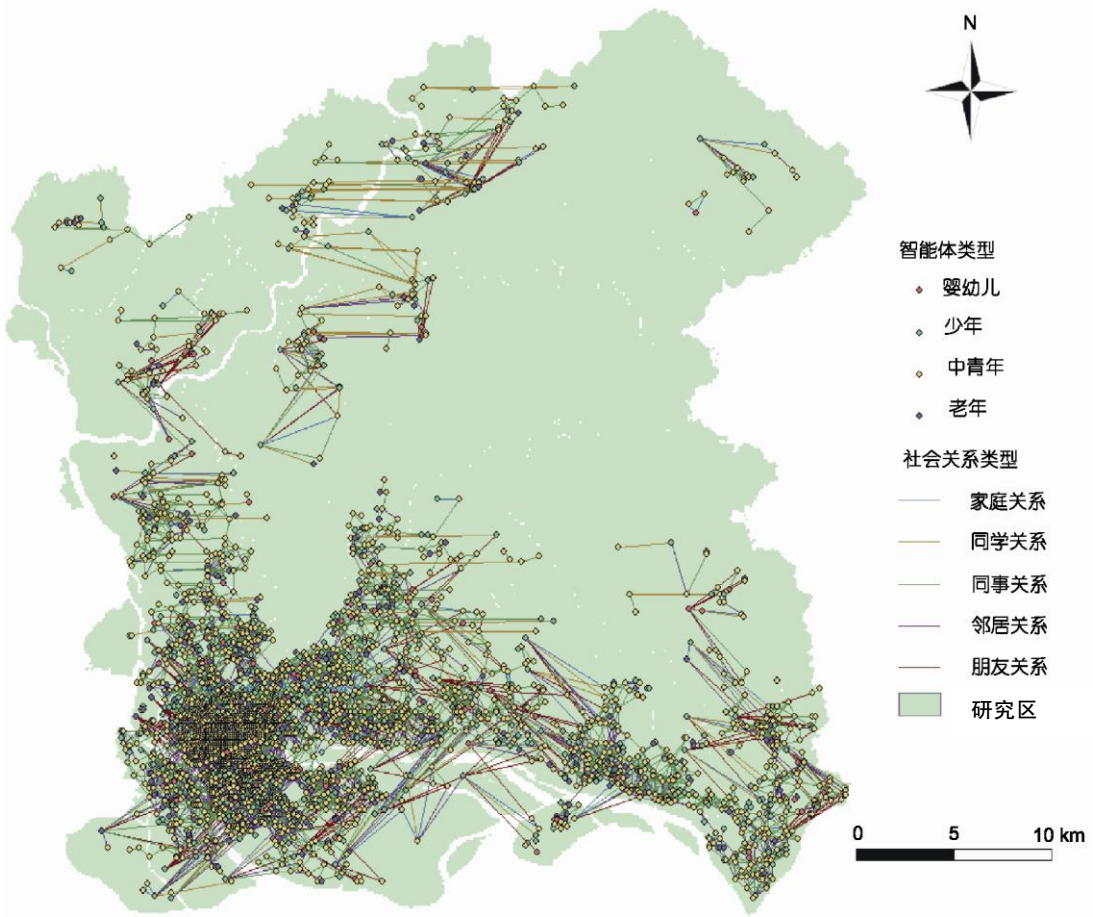


图 5 智能体间的社会关系图

在全部智能体中随机抽取 20%显示

在基于小世界网络的多智能体模型的模拟过程中，分别记录下 T=0, T=0~20, T=0~40, T=0~60 四个时段每个网格感染事件发生的次数.

图 7 所示 T=0 时为模拟的初始时刻，此时只有易感者、潜伏者和感染者三类智能体，其数目可根据实际观测数据估算得到. 此时感染事件发生次数可认为是感染者智能体出现的位置. T=20 时流感扩散，

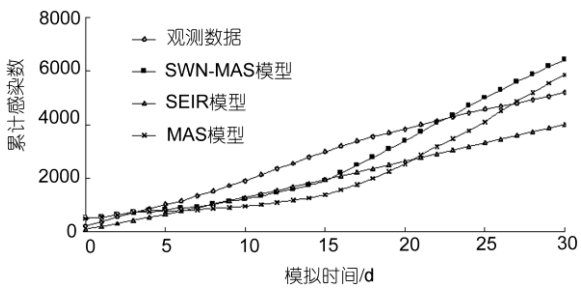


图 6 不同模型的模拟结果与实际观测数据比较

但此时已被感染的智能体数量还很少，大部分位于 T=0 时刻已被感染智能体的周边. T=40 时刻被感染的智能体数量显著增多，并且被感染的智能体大多分布于研究区中人口密度较大的越秀区以及海珠区的西北部人口稠密的区域. T=60 时流感扩散规模较大，在 T=40 的基础上荔湾区、东山区被感染的数目也大为增加，整体上呈空间聚集的格局，越往中心感染事件发生次数越大.

传染病的发病率是用于衡量疾病的扩散程度的定量指标. 其计算方法为一段时间段内某区域疾病发生的次数在该区域整个人口中所占的比例，再乘以 10 万. 由于本研究的时段为 2006 年 3 月，因此计算结果为该月的流感发病率. 研究区各街道办(镇)的流感发病率如图 8 所示.

由图 8 看出，城市中心流感发病率较高，而白云区、海珠区的东南部流感发病率较低. 由此不难

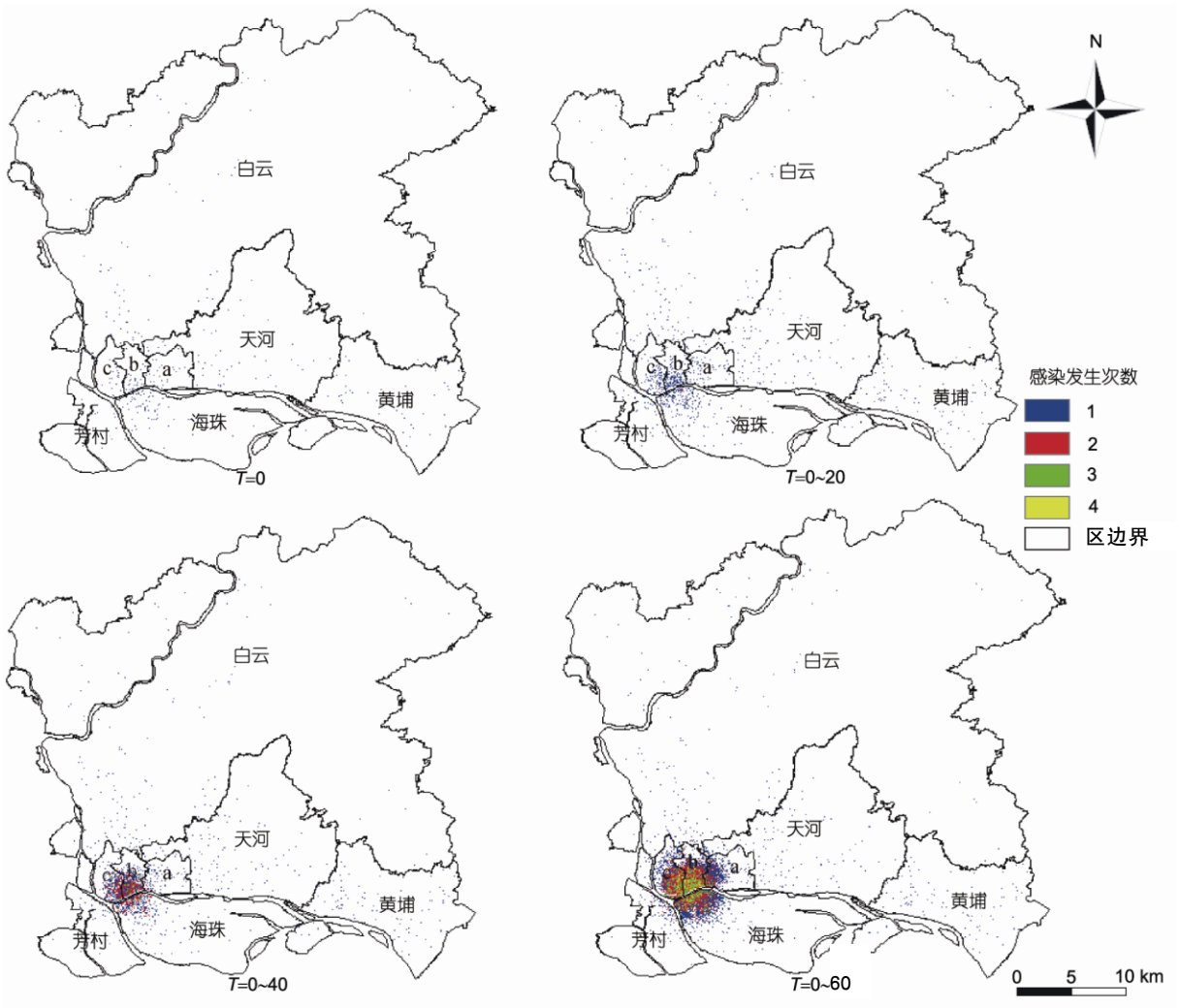


图 7　感染发生次数空间分布图

a 为东山区, b 为越秀区, c 为荔湾区

看出流感发病率与人口密度有着较为密切的关系. 城市中心由于交通的拥挤、经济活动的频繁, 使得人

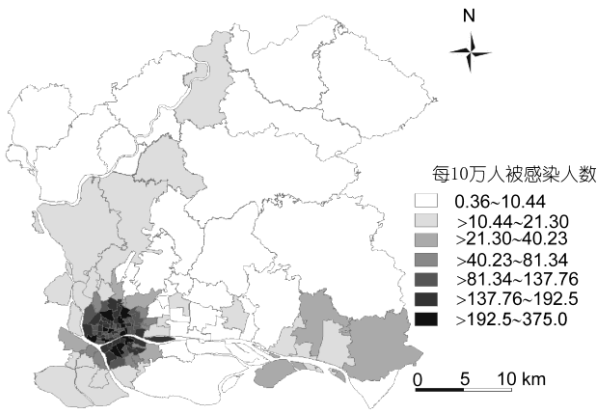


图 8　流感发病率图

与人之间有着较高的接触机会, 因此流感的发病率也就较高.

分别统计出易感者智能体以及目前为止包括潜伏者、感染者以及移出者在内的已感染者智能体(累计结果, 如 T=20 时的已感者智能体数目为 T=0 到 T=20 时刻所有的潜伏者、感染者以及移出者的平均度)平均度变化曲线, 得到如图 9 所示的曲线.

如图 9 所示, T=0~30 时易感者智能体的平均度有所降低, 但幅度不大, 主要是由于此前被感染的智能体数目较少, 在整个易感者智能体中所占比例还相对较小. 随着传染病的迅速传播, 易感智能体中度值较大的智能体陆续被感染, 易感者智能体的平均度曲线不断下降. 已被感染的智能体平均度曲线呈现先升高后降低的趋势, 在 T=10 到 T=30 间,

平均度值较大. 这说明在模拟过程中, 那些度值较高, 也就是社会关系较多的智能体被感染的概率较高. 随着模拟的进行, 度值较高的智能体大部分均已被感染, 已被感染的智能体的平均度曲线达到最大值. 此后, 传染由这些度值较高的智能体向与其联系的度值较低的智能体扩散, 因此已被感染的智能体的平均度逐渐下降.

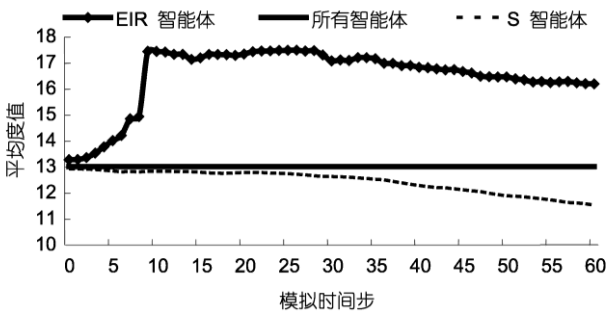


图 9 智能体平均度变化曲线

3 结论

在传染病的传播机理基础上, 提出一种基于小世界网络的多智能体传染病传播模型. 模型中明确考虑了个体对感染"记忆"的时间衰减效应、个体间的距离以及社会关系对感染的影响. 选择广州市区为研究区在 Swarm 平台上模拟了流感的时空传播过程. 该模型的模拟精度达到 78.2%, 高于传统的系统动力学 SEIR 模型以及多智能体模型的模拟精度, 模拟结果更为接近于实际观测病例数据.

通过流感时空传播的模拟我们不难发现, 流感感染事件发生的次数与人口有着较为密切的关系, 人口越集中的区域, 感染事件发生的频率也越高. 此外, 疾病开始扩散时社会关系较多的智能体被感染的概率要远远高于社会关系较少的智能体, 随着传染病的蔓延, 那些社会联系较少的智能体被感染的概率逐渐增高.

致谢 审稿专家提出宝贵意见和建议, 本文利用了广州市疾病预防控制中心提供的流感数据, 在此一并致谢.

参考文献

1 Kermack W O, McKendrick A G. A contribution to the mathematical theory of epidemics. In: Proceedings of the Royal Society of London(Series A). London: The Royal Society, 1927. 700—721
2 马知恩, 周义仓, 王稳地, 等. 传染病动力学的数学建模与研究. 北京: 科学出版社, 2004. 3—9
3 石耀霖. SARS 传染扩散的动力学随机模型. 科学通报, 2003, 48: 1373—1377
4 汪海英, 荣峰, 柯孚久, 等. 严重急性呼吸综合征(SARS)的传播与控制动力学. 科学通报, 2003, 48: 1933—1935
5 Stefano B D, Fukś H, Lawniczak A. Object-oriented implementation of CA/LGCA modelling applied the spread of epidemics. In: Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering. Halifax: IEEE, 2000. 26—31
6 Fu S C. Modelling epidemic spread using cellular automata. Master Dissertation. Perth: University of Western Australia, 2002. 27—49
7 Mikler A R, Venkatachalam S, Abbas K. Modeling infectious diseases using global stochastic cellular automata. J Biol Syst, 2005, 13: 421—439[doi]
8 Zhang P, Atkinson P M. Modelling the effect of urbanization on the transmission of an infectious disease. Math Biosci, 2008, 211: 166—185[doi]
9 Moore C, Newman M E J. Epidemics and percolation in small-world networks. Phys Rev E, 2000, 61: 5678—5682[doi]
10 Chebeane H, Echalier F. Towards the use of a multi-agents event based design to improve reactivity of production systems. Comput Ind Eng, 1999, 37: 9—13[doi]
11 黎夏, 叶嘉安, 刘小平, 等. 地理模拟系统: 元胞自动机与多智能体. 北京: 科学出版社, 2006. 250—258
12 刘小平, 黎夏, 叶嘉安. 基于多智能体系统的空间决策行为及土地利用格局演变的模拟. 中国科学 D 辑: 地球科学, 2006, 36: 1027—1036
13 张鸿辉, 曾永年, 金晓斌, 等. 多智能体城市土地扩张模型及其应用. 地理学报, 2008, 63: 869—881
14 刘颖, 陈禹. 复杂适应系统理论对控制 SARS 疫情的模拟分析. 复杂系统与复杂性科学, 2004, 1: 74—79
15 邓宏钟, 迟妍, 谭跃进. 基于多智能体的疾病传染仿真研究. 计算机仿真, 2004, 21: 167—170
16 Morrill R L, Angulo J J. Spatial aspects of a smallpox epidemic in a small brazilian city. Geogr Rev, 1979, 69: 319—330[doi]
17 Bian L. A conceptual framework for an individual-based spatially explicit epidemiological model. Environ Plann Plann Des, 2004, 31:

381—395
18 龚建华, 周洁萍, 徐珊, 等. SARS 传播动力学模型及其多智能体模拟研究. 遥感学报, 2006, 10: 829—835
19 Milgram S. The small world problem. Psych Today, 1967, 1: 60—67
20 Watts D J, Strogatz S H. Collective dynamics of 'small-world' networks. Nature, 1998, 393: 440—442[doi]
21 王劲峰. 空间分析. 北京: 科学出版社, 2006. 222—225
22 Anderson R, May R. Infectious Diseases of Humans: Dynamics and Control. Oxford: Oxford University Press, 1991. 467—630
23 Dodds P S, Watts D J. Universal behavior in a generalized model of contagion. Phys Rev Lett, 2004, 92: 218701[doi]
24 汪小帆, 李翔, 陈关荣. 复杂网络理论及其应用. 北京: 清华大学出版社, 2006. 9—23
25 Newman M E J, Watts D J. Renormalization group analysis of the small-world network model. Phys Lett A, 1999, 263: 341—346[doi]
26 Bailey T C, Norman T J. The Mathematical Theory of Infectious Diseases and Its Applications. New York: Oxford University Press, 1975. 6—42
27 Wang J F, McMichael A J, Meng B, et al. Spatial dynamics of an epidemic of severe acute respiratory syndrome in an urban area. Bull World Health Organ, 2006, 84: 965—968[doi]
28 曾毅. 人口分析方法与应用. 北京: 北京大学出版社, 1993. 47—124
29 王广州. 中国人口预测软件培训手册, 2002
30 Tian Y, Yue T, Zhu L, et al. Modeling population density using land cover data. Ecol Model, 2005, 189: 72—88[doi]
31 Clark C. Urban population densities. J R Stat Soc A Stat, 1951, 114: 490—496
32 谢守红, 宁越敏. 广州市人口密度分布及演化模型研究. 数理统计与管理, 2006, 25: 518—522