

城市级新冠肺炎(COVID-19)疫情预测和仿真模型

王金恺¹⁾, 张虎²⁾, 贾鹏²⁾, 权益¹⁾, 陈梁港旭¹⁾, 王长波^{1)*}

¹⁾ (华东师范大学计算机科学与技术学院 上海 200062)

²⁾ (中国航天科工集团第三研究院体系对抗与智能信息系统总体部 北京 100074)
(cbwang@cs.ecnu.edu.cn)

摘要: 传染病的传播分析对于社会公共安全与城市管理有着重要的意义。针对此次新冠肺炎(COVID-19)疫情的传播, 构建了一个城市级结构化疫情预测和仿真模型。该模型基于最新的传染病动力学理论, 综合地理信息网络, 将易感者、潜伏者、感染者和康复者(SEIR)模型与社会网络模型相结合, 对预测区域进行多层次划分; 宏观上使用二分网络模拟公共设施与社区节点的关系; 微观上使用改进的 SEIR 模型模拟节点内部的传染情况; 节点间使用智能体追踪个体的传播过程。采用国家卫生健康委员会公开的 2020 年武汉市与北京市疫情确诊及治愈病例数进行对比实验, 结果表明, 所提模型在疫情仿真上更灵活, 预测结果更准确, 并能直观地体现不同人群的分布情况和流动趋势。

关键词: 2019 新型冠状病毒(COVID-19); 传染病; 易感者、潜伏者、感染者和康复者(SEIR)模型; 复杂网络模型; 智能体

中图分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2022.19146

City-Level Structured Prediction and Simulation Model of COVID-19

Wang Jinkai¹⁾, Zhang Hu²⁾, Jia Peng²⁾, Quan Yi¹⁾, Chen Lianggangxu¹⁾, and Wang Changbo^{1)*}

¹⁾ (School of Computer Science and Technology, East China Normal University, Shanghai 200062)

²⁾ (General Department of System Confrontation and Intelligent Information System, The Third Academic of Aerospace and Technology Group of China, Beijing 100074)

Abstract: It is important for social public security and urban management to explore the spread of infectious diseases. A city-level structured prediction and simulation model for COVID-19 is proposed. This model is consisted of SEIR and social network model on the basis of latest infectious disease dynamics theory and real geographic networks. The prediction region is divided into multiple levels. Specifically, a bipartite network is applied to simulate the relationship between public facilities and community nodes at the macro level, and a modified SEIR is applied to simulate the infection within nodes at the micro level. Besides, intelligent agent is applied to track the individual transmission process. The contrast experimental results based on the confirmed and cured cases of Wuhan and Beijing in 2020 published by National Health Commission, show that the proposed model has better flexibility and higher accuracy, and reflects the distribution and movement of people more directly.

Key words: coronavirus disease 2019 (COVID-19); epidemic; susceptible-exposed-infectious-removed (SEIR) model; complex networks; intelligent agent

收稿日期: 2021-06-09; 修回日期: 2022-04-16. 基金项目: 上海市自然科学基金(19ZR1415800). 王金恺(1991—), 男, 博士研究生, 主要研究方向为轻量化云仿真; 张虎(1986—), 男, 博士, 高级工程师, 主要研究方向为军事智能; 贾鹏(1981—), 女, 博士, 高级工程师, 主要研究方向为机器学习; 权益(1997—), 男, 硕士研究生, 主要研究方向为计算机几何处理; 陈梁港旭(1997—), 男, 博士研究生, 主要研究方向为计算机视觉; 王长波(1976—), 男, 博士, 教授, 博士生导师, CCF 会员, 论文通信作者, 主要研究方向为计算机图形学、数字媒体、虚拟现实。

COVID-19^[1]是一种由严重急性呼吸道综合征冠状病毒 2 型(SARS-CoV-2)所引发的传染病. 由于致病病毒的高传染性和高隐蔽性^[2], 其在短时间内迅速散播至全球并导致大规模的疫情爆发. 为了精确地模拟疫情的传播过程, 需要有针对性地研究现有的传染病模型, 并结合新冠病毒的特性进行改进, 从感染速度、空间范围和传播途径等角度指导对疫情的防控. 目前, 传染病的仿真模型主要以传染病动力学模型与复杂网络模型为主.

常见的传染病动力学模型按照传染病类型分为 SI, SIR, SIRS, SEIR 模型等^[3]. 其中, SEIR 模型, 即易感者(susceptible)、潜伏者(exposed)、感染者(infected)和康复者(recovered)模型, 由 4 种群体组成, 它们的关系由转换率表示. 由于新冠病毒具有 14 天的潜伏期, SEIR 模型不仅考虑了潜伏周期, 还将潜伏者作为模拟人群之一, 满足了新冠病毒的传播特性, 因此大多数关于 COVID-19 的疫情仿真使用 SEIR 模型. 然而, 传统的 SEIR 模型中描述病毒传染过程的内在机理较为简单, 并未引入人员流动、城市间的迁入/迁出等随机因素. Yang 等^[4]在 SEIR 模型的基础上引入输入/输出流, 并基于各省市间的交通数据将原始的传播速度参数分为 2 个流向, 即从易感者至感染者的传播速度和从易感者至潜伏者的传播速度. 改进后的 SEIR(modified SEIR, M-SEIR)模型在推演城市疫情时更符合实际的人口变动情况.

另外, 包括 SEIR 在内的传统传染病动力学模型是基于均质空间的, 其应用场景一般为理想状态的沙盒, 忽略了人口分布和地理位置等空间因素对病毒传播的影响, 无法根据地理信息和社会关系还原当地疫情的传播过程. 因此, 可以在传染病模型的基础上加入智能体^[5], 利用智能体特有的社交属性追踪感染个体的行迹, 以提高建模精度和准确率. 不过, 随着推演区域的扩大和智能体数量的增加, 模型的推演效率也会随之降低, 而智能体的运行也无法脱离真实的地理数据. 因此, 需要按照真实的地理信息构建复杂网络^[6]进一步划分推演区域, 从而明确交互节点(即公共设施与社区节点)间的关系, 并减少智能体的冗余信息和操作.

针对以上问题, 本文提出一种城市级结构化疫情预测和仿真(city-level structured prediction and evaluation, CSPE)模型. 该模型基于真实地理数据, 以社区为基本预测对象, 将研究区域从宏观到微观进行多层次划分(即从城市到社区), 对微观社区进行预测, 融合宏观的信息进行修正, 自底向上汇

总各层级预测结果. 宏观上, 针对整个研究区域, 以社区和公共场所为节点构建二分网络模型, 并以此单独预测社区因社会活动新增的被感染人数, 预测过程还使用智能体近似模拟个体决策行为; 微观上, 利用 M-SEIR 模型独立预测各社区内部的感染情况, 并根据由上述网络模型预测的新增被感染人数修正前者的预测结果.

采用 2020 年北京市和武汉市防疫背景下的真实数据, 在与 M-SEIR 和 GLEaM 模型对比后可知, CSPE 模型能有效地推演城市疫情, 预测结果基本与真实数据拟合, 并能随时更改模型中的防控措施, 以对当前决策进行分析评估.

1 相关工作

1.1 SEIR 传染病动力学模型

传染病动力学模型^[7]是传染病的基本数学模型, 用于研究各种传染病的传播速度、途径、范围和特性等. SEIR 模型^[8]是其中的经典模型之一, 它适用于有潜伏期、治愈后获得终身免疫的疾病. 目前, 大多数对 COVID-19 的研究都是基于该模型的^[4,9-13].

Liu 等^[9]基于 SEIR 模型将潜伏期作为参数建立了微分方程模型. He 等^[10]提出的粒子群算法为针对 COVID-19 的 SEIR 模型参数估计提供了参考. Pandey 等^[11]使用 SEIR 模型和回归模型分析了 COVID-19 在印度的传播范围和速度. Mwalili 等^[12]使用龙格-库塔方法优化了 SEIR 模型的求解过程, 并用矩阵法确定基本再生数 R_0 . 此外, 还有一些深入研究了疫情. 例如, 曹盛力等^[13]建立了可用于评估 COVID-19 疫情的 SEIR 动态模型, 并将其应用于湖北省的疫情推演. Yang 等^[4]提出了改进的 SEIR 传染病动力学模型(M-SEIR), 其基于 2020-01-23 前后的中国迁移数据和最新的流行病学数据, 对 COVID-19 进行了分析和预测, 并基于 2003 年 SARS 冠状病毒的数据使用人工智能证实了 M-SEIR 模型的可行性. 本文的方法论是基于 M-SEIR 模型提出的.

为了分析包括 COVID-19 在内的传染病传播过程和发展规律, 国内有学者用可视化的展示方式对其进行了研究^[14-16].

陈晓慧等^[14]基于地图、时空、时序和关系 4 种分类对传染病传播的可视化方法做了总结. 金思辰等^[15]提出了一种用于分析疾病时空分布的可视化交互系统. 刘建湘等^[16]根据 COVID-19 的多元特性和时空特点, 并综合时间轴交互方式和流

行病数学模型设计了一个疫情可视化模型。

1.2 复杂网络模型与智能体

复杂网络是具有自组织、自相似、吸引子、小世界、无标度中部分或全部性质的网络^[17], 作为一种常见的数据表现形式, 它可用于模拟传染病的传播过程^[18-21]。

Meyers 等^[18]以 SARS 为研究对象, 通过扩展流行病社交网络结合冠状病毒爆发的异质性, 对公共卫生策略做出了评估。范如国等^[19]在复杂网络理论的基础上建立了带有潜伏期的 COVID-19 流行病动力学 SEIR 模型。Liu 等^[20]提出了一种基于社会网络的 SAIR 模型描述 COVID-19 的传播, 并根据武汉市真实疫情数据分析了病毒的传播特性。Balcan 等^[21]提出全球流行病与流动性(global epidemic and mobility, GLEaM)模型, 将人口统计学与流动数据结合在一个空间结构中。本文的网络构建过程参考了 GLEaM 模型。

智能体是具有智能的实体^[22], 因其具备个体的各种属性, 可以对一定数量的群体行为进行精确建模。为了应对 COVID-19, 有不少研究利用了智能体特有的社交属性还原节点间的人员流动和病毒传播^[23-26]。

Perez 等^[23]以荨麻疹为案例研究了智能体模拟传染病传播的方法。刘涛等^[24]建立了基于小世界网络的多智能体传染病时空传播模型。同时, 也有以 COVID-19 为研究对象的智能体模型, 如 Shamil 等^[25]使用了一种集成城市参数和智能体的模型预测新冠病毒在城市居民间的传播; Kerr 等^[26]为精准追踪感染个例, 针对年龄结构和人口信息设计了基于智能体的城市模拟器。

2 本文模型

2.1 概述

为解决城市疫情推演问题, 在细粒度与计算量间取得均衡, 本文提出了一种以社区为基本预测对象并融合宏观信息进行疫情预测和仿真的 CSPE 模型。设研究区域包含 n 个社区, 社区 j 在第 t 天的人口数为 $p_j(t)$, 包含 5 类群体, 即易感者 $p_j^s(t)$ 、潜伏者 $p_j^e(t)$ 、感染者 $p_j^i(t)$ 、康复者 $p_j^r(t)$ 和死亡者 $p_j^d(t)$ 。CSPE 模型可根据输入的第 t 天的各类人口数和预设的模型参数对第 $t+1$ 天的各类人口数进行预测。

与整片研究区域相比, 微观的社区人口规模

较小, CSPE 模型能够独立地使用 M-SEIR 模型精确地计算每个社区在第 $t+1$ 天的各类人口数。由于 M-SEIR 模型必须遵循其迭代公式和预设参数进行计算, 难以宏观地模拟社区间疾病的传播过程。因此, CSPE 模型首先使用 M-SEIR 模型推演各社区因内部活动(如家庭成员间)引发的新增被感染人数, 再根据研究区域的地理信息构建二分网络模拟新增感染个体在社区间由社会活动引起的感染过程, 然后使用感染结果对先前的社区内部感染数据进行修正, 最终输出整片研究区域的各类人口数量。

2.2 基于 M-SEIR 模型的社区内传染预测

由于居住地的人口分布和数量相对稳定, 近似地满足了 SEIR 模型对研究区域人口数所做的恒定假设。CSPE 模型在 M-SEIR 模型的基础上加入死亡率和死亡人口数, 将第 t 天各社区各类人口数作为输入条件, 得出第 $t+1$ 天社区 j 的各类人口数, 其计算公式为

$$\left\{ \begin{aligned} p_j^s(t+1) &= p_j^s(t) + i_j^s - o_j^s - \frac{b_j^1 \cdot r_j \cdot p_j^i(t) \cdot p_j^s(t)}{n_j} - \\ &\quad \frac{b_j^2 \cdot r_j \cdot p_j^e(t) \cdot p_j^s(t)}{n_j} \\ p_j^e(t+1) &= p_j^e(t) + i_j^e - o_j^e + \frac{b_j^1 \cdot r_j \cdot p_j^i(t) \cdot p_j^s(t)}{n_j} + \\ &\quad \frac{b_j^2 \cdot r_j \cdot p_j^e(t) \cdot p_j^s(t)}{n_j} - \sigma_j \cdot p_j^e(t) \\ p_j^i(t+1) &= p_j^i(t) + \sigma_j \cdot p_j^e(t) - c_j \cdot p_j^i(t) - z_j \cdot p_j^i(t) \\ p_j^r(t+1) &= p_j^r(t) + c_j \cdot p_j^i(t) \\ p_j^d(t+1) &= p_j^d(t) + z_j \cdot p_j^i(t) \end{aligned} \right. \quad (1)$$

其中, z_j 为新引入的死亡率, 其余参数定义与 M-SEIR 模型^[4]基本相同, 如表 1 所示。

表 1 CSPE 模型中的参数及其含义

| 参数 | 含义 | 参数 | 含义 |
|------------|--------------|---------|----------------------|
| b_j^1 | 疾病潜伏率 | z_j | 死亡率 |
| b_j^2 | 疾病直接感染率 | i_j^s | 从研究区域外流入社区 j 的未感染者 |
| r_j | 平均接触人数 | o_j^s | 从社区 j 流出研究区域的未感染者 |
| n_j | 社区 j 的总人口数 | i_j^e | 从研究区域外流入社区 j 的潜伏者 |
| σ_j | 潜伏至感染的转化率 | o_j^e | 从社区 j 流出研究区域的潜伏者 |
| c_j | 治愈率 | | |

CSPE 模型通过 M-SEIR 模型计算每天以社区为单位的各类人口数. 但仅靠传染病动力学模型还不足以模拟感染个体在社区间进行的社会活动, 而防控措施的加入也很难与 M-SEIR 模型的经典迭代式结合. 因此, CSPE 模型在 M-SEIR 模型预测结果的基础上, 从宏观层面建立二分网络, 模拟由个体在社区与公共场所往来所产生的感染过程.

2.3 基于二分网络的个体社会活动仿真

人们每日社会活动的典型范式是两点一线, 即白天从住处前往某一公共场所(如单位、学校和娱乐场所), 晚上从该场所返回住处. 模拟研究区域在典型社会活动范式下的感染过程不仅需要计算各公共场所每日新增的被感染者数量, 还需判断其所属社区. 因此, CSPE 模型通过二分网络^[27]构建社区与公共节点的关系, 模拟被感染者在节点间的社会活动.

设研究区域内的所有社区为点集 P , 所有公共场所为点集 Q , 建立二分网络的边集 E 遵循如下规则: (1) P 内的各个节点都与 Q 内的各个节点有边相连; (2) P 与 Q 内的各个节点不相连. 由此可以得到一个二分无向网络, 记为一个二元组 $(P \cup Q, E)$. 个体的社会活动可视为从 P 中某节点到 Q 中某节点的一次往返过程. 基于北京市部分主要节点的二分网络模型如图 1 所示.

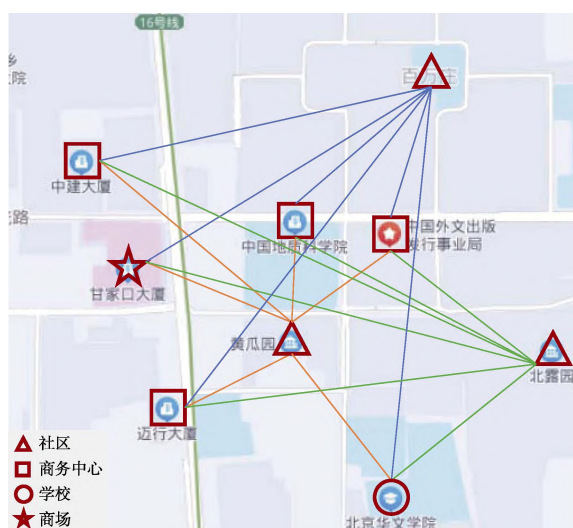


图1 基于北京市部分主要节点的二分网络模型示例

每个公共场所均有来自社区的个体, 这些个体会根据主观因素(如青少年往返于学校)、客观因素(如公共场所与所在社区的距离限制)、社会因素(如当日交通管制情况、公共场所和社区开放情况)以及随机因素(如临时变更出行计划)决定当日前

往的公共场所和返回的社区. 因此, 引入参数建立二分网络可以实现以上因素, 同时提高仿真的准确度. 具体参数定义如表 2 所示.

表2 CSPE 模型中二分网络的参数及其含义

| 参数 | 含义 |
|-------------|-------------|
| γ_j | 社区节点感染者外出率 |
| β_j^k | 节点间畅通率 |
| d_j^k | 节点间归一化距离 |
| t_j^k | 节点间平均交通管制力度 |
| s_j^k | 节点间封锁措施指示参数 |
| α_k | 交互节点疾病感染率 |

已知社区节点 $j \in P$ 在第 t 天有感染者 $p_j^i(t)$ 人, 设该社区感染者外出率为 $\gamma_j \in [0,1]$, 则当日从该社区外出的感染者人数可定义为

$$p_j^o(t) \triangleq \gamma_j \cdot p_j^i(t).$$

设 $p_j^k(t)$ 为 $p_j^o(t)$ 个感染者中计划前往交互节点 $k \in Q$ 的人数, 其值取决于社区内个体的决策, 该过程由一个基于先验信息的智能体选择算法确定, 并满足 $\sum_{k=1}^{|Q|} p_j^k(t) = p_j^o(t)$.

尽管个体能够选择前往的交互节点, 但仍受限于当日的交通条件. 为进一步模拟交通路况对个体社会活动的影响, 网络模型引入了社区节点 j 与交互节点 k 间的畅通率参数 $\beta_j^k \in [0,1]$. 该参数取值与节点间的距离及交通路况相关, 具体定义为

$$\beta_j^k \triangleq (1 - d_j^k \cdot t_j^k) \cdot s_j^k \quad (2)$$

其中, $d_j^k \in [0,1]$ 为归一化后的节点间直线地理距离(根据节点的真实经纬度确定); $t_j^k \in [0,1]$ 为节点间的平均交通管制力度(反映道路管制和公交管制力度); $s_j^k \in \{0,1\}$ 表示节点间是否有封锁措施, 若有则取值为 0, 反之为 1. 由于交通管制力度对畅通率的影响与节点间的距离有关, 距离越近, 管制对其间的通过率影响越小(如交通管制对近距离步行通勤的个体几乎不产生影响), 反之越大. 所以, 式(2)通过 $1 - d_j^k \cdot t_j^k$ 项表示个体在交通管制下实际从社区节点前往交互节点的通过率, 并用 s_j^k 表示是否采取封闭管控措施.

在引入 $p_j^k(t)$ 和 β_j^k 后, 实际从社区节点 j 前往交互节点 k 的感染者人数为 $\hat{p}_j^k(t) \triangleq p_j^k(t) \cdot \beta_j^k$. 那么, 交互节点 k 中具有传染能力的总人数 $\hat{p}^k(t)$ 为从各社区节点实际到达该节点的感染者人数之和, 即 $\hat{p}^k(t) = \sum_{j=1}^{|P|} \hat{p}_j^k(t)$.

在确定具备传染能力的个体数量后, 可得出交互节点的新增被感染人数为 $\tilde{p}^k(t) \triangleq \hat{p}^k(t) \cdot \alpha_k$. 其中, $\alpha_k \in [0, +\infty)$ 为该交互节点的感染率.

由此可得各交互节点的新增被感染人数, 然后通过判断这些感染者返回的社区, 即可推算出各社区因个体社会活动而新增的被感染人数. 设 $\tilde{p}_j^k(t)$ 为 $\tilde{p}^k(t)$ 中计划返回社区节点 j 的人数, 其取值通过智能体算法由个体决策确定, 并满足 $\sum_{j=1}^{|P|} \tilde{p}_j^k(t) = \tilde{p}^k(t)$.

最终, 社区 j 因社会活动而新增的被感染人数 $\tilde{p}_j(t)$ 即为从各交互节点返回该社区的新增被感染者之和, 即 $\tilde{p}_j(t) = \sum_{k=1}^{|Q|} \tilde{p}_j^k(t)$.

通过以上步骤, CSPE 模型便可利用二分网络较为精准地模拟个体在城市中的社会活动过程.

2.4 基于先验概率的智能体选择算法

在 CSPE 模型中, 节点间个体的社会活动由二分网络模拟, 而对个体的决策选择由具备个体属性的智能体决定. 通常, 人们的日常社会活动是相对稳定的, 如青少年偏向去学校学习, 成年人偏向去公司工作, 但个体选择在遵循某种规律的同时也具备一定的随机性. 为模拟个体这种整体稳定偶尔随机的选择过程, CSPE 模型设计了一种基于先验概率的智能体选择算法, 计算每个社区白天前往各交互节点的感染者人数 $p_j^k(t)$ 及每个交互节点返回各社区的新增被感染者人数 $\tilde{p}_j^k(t)$.

由于计算 $p_j^k(t)$ 的过程可视作各社区节点的个体前往各交互节点的分配过程, 同理, 计算 $\tilde{p}_j^k(t)$ 的过程可视作各交互节点的个体返回社区节点的分配过程. 因此, 可将以上计算过程进一步抽象成人员分配问题: 将人员集合 $T = \{t_1, t_2, \dots, t_q\}$ 中的每个

元素 $t_i \in T$ 分配给若干集合 $S = \{S_1, S_2, \dots, S_r\}$ 中的一个集合 $S_j \in S$, 求解每个集合被分配到的人数, 即 $\{|S_1|, |S_2|, \dots, |S_r|\}$.

因此, 智能体选择算法首先需要引入预设参数对个体的先验选择倾向进行建模. 设集合 $W = \{u_1, u_2, \dots, u_r\}$ 为 T 中个体在不受随机扰动的条件下被分配到各集合的人数, 该集合元素之和为 q . 然后利用智能体进行随机采样, 以模拟个体决策的随机过程, 并通过判断采样值和 T 中元素的关系来确定个体被分配到的集合. 该智能体算法的具体步骤如下.

输入. q, W .

输出. $\{|S_1|, |S_2|, \dots, |S_r|\}$.

Step1. 根据 W 建立区间边界向量 $w_{|W|+1}$:

Step1.1. $w_1 = 0$;

Step1.2. $w_{|W|+1} = q$;

Step1.3. 对 $k \in \{2, \dots, |W|\}$, $w_k = w_{k-1} + u_{k-1}$.

Step2. 对 $\{|S_1|, |S_2|, \dots, |S_r|\}$ 中所有元素赋值为 0.

Step3. 进行 q 次智能体随机采样, 记每次的采样值为 v :

Step3.1. 从整数均匀分布采样 $v \sim U_{\text{int}}[0, q]$;

Step3.2. 对于 $k \in \{1, 2, \dots, |W|\}$, 若 $v \in [w_k, w_{k+1})$, 则 $|S_k| += 1$.

Step4. 输出 $\{|S_1|, |S_2|, \dots, |S_r|\}$.

上述步骤中, $\{|S_1|, |S_2|, \dots, |S_r|\}$ 的值恰好是 v 取值落在对应区间的次数. 由大数定理可知, 对于任意 $|S_l| \in \{|S_1|, |S_2|, \dots, |S_r|\}$, 当采样次数趋于正无穷大时, $|S_l| / \sum_{u=1}^r |S_u|$ 将趋于 v 取值落在对应区间的概率, 即 $u_l / \sum_{u=1}^r u_u$. 所以, 当采样次数为 q 时, $|S_l|$ 将趋于 u_l . 通过该算法能够计算各集合被分配到的人员数, 得出 $p_j^k(t)$ 和 $\tilde{p}_j^k(t)$, 既满足了分配的稳定性, 又具备了采样的随机性.

综上所述, CSPE 模型通过二分网络和上述推演过程即可得到社区 j 因个体社会活动而新增的被感染人数 $\tilde{p}_j(t)$, 为得出最后的感染人群数量, 还需将该结果与 M-SEIR 模型预测的数据进行融合. 故式(1)被修正为

$$\left\{ \begin{aligned} p_j^s(t+1) &= p_j^s(t) + i_j^s - o_j^s - \frac{b_j^1 \cdot r_j \cdot p_j^i(t) \cdot p_j^s(t)}{n_j} - \\ &\quad \frac{b_j^2 \cdot r_j \cdot p_j^e(t) \cdot p_j^s(t)}{n_j} - \tilde{p}_j(t) \\ p_j^e(t+1) &= p_j^e(t) + i_j^e - o_j^e + \frac{b_j^1 \cdot r_j \cdot p_j^i(t) \cdot p_j^s(t)}{n_j} + \\ &\quad \frac{b_j^2 \cdot r_j \cdot p_j^e(t) \cdot p_j^s(t)}{n_j} - \sigma_j \cdot p_j^e(t) + \tilde{p}_j(t) \\ p_j^i(t+1) &= p_j^i(t) + \sigma_j \cdot p_j^e(t) - c_j \cdot p_j^i(t) - z_j \cdot p_j^i(t) \\ p_j^r(t+1) &= p_j^r(t) + c_j \cdot p_j^i(t) \\ p_j^d(t+1) &= p_j^d(t) + z_j \cdot p_j^i(t) \end{aligned} \right.$$

由此可见, 在 M-SEIR 模型的基础上, CSPE 模型结合了基于真实地理的宏微观信息, 使用二分网络和智能体对城市疫情进行多层级的细粒度的预测, 不仅能准确地推演疫情发展趋势, 在疫情防控方面也可引入相应措施, 具备一定的可扩展性。

3 实验及分析结果

为验证 CSPE 模型在有防控措施监管下的疫情形势预测准确性, 首先使用 CSPE 模型模拟 2020 年初武汉市的抗疫过程, 再使用 M-SEIR, GLEaM 与 CSPE 模型分别模拟 2020 年北京市的疫情, 并与真实数据进行对比。然后使用 CSPE 模型模拟不同等级防控措施下北京市的疫情发展趋势, 并分析评估疫情下采取不同措施的效果。接着通过热力图表示研究区域不同时间下的感染情况, 进一步证明复杂网络与智能体的加入可以使 CSPE 模型能够直观地展示地域感染状况。最后, 就 CSPE 模型设计了疫情推演可视化系统, 从应用层面上体现其易用性和扩展性。

3.1 2020 年真实抗疫数据对比与分析

本节首先使用 CSPE 模型模拟了武汉市 2019-12-29—2020-03-08 的抗疫过程, 其真实数据来源自文献[28]。CSPE 模型在预测感染情况时所采取的防控措施与实际完全相同, 即 2020 年 01 月 23 日起采取包括交通管制在内的二级响应, 于 2020-02-02—2020-02-16 期间采取一级防控措施, 对应参数设置如表 3 所示。武汉市实际感染人数与 CSPE 模型推演结果如图 2 所示, 其仿真结果与真实数据基本拟合。由此可见, CSPE 模型因其宏微观多层级的仿真及防控措施的加入具有较高的准确度。

表 3 原始参数及 CSPE 模型的防控措施参数

| 参数 | 平均接触人数 | 疾病潜伏率 | 疾病直接感染率 | 潜伏至感染转化率 |
|------|--------|-------|---------|----------|
| 原始 | 7 | 0.109 | 0.126 | 0.148 |
| 一级响应 | 3 | 0.089 | 0.101 | 0.148 |
| 二级响应 | 1 | 0.067 | 0.082 | 0.148 |

| 参数 | 治愈率 | 死亡率 | 社区节点感染者外出率 | 节点间畅通率 |
|------|-------|-------|------------|--------|
| 原始 | 0.163 | 0.015 | 0.20 | 1.0 |
| 一级响应 | 0.325 | 0.011 | 0.10 | 0.5 |
| 二级响应 | 0.462 | 0.005 | 0.05 | 0.2 |

| 参数 | 节点间归一化距离 | 节点间平均交通管制力度 | 节点间封锁措施指示参数 | 交互节点疾病感染率 |
|------|----------|-------------|-------------|-----------|
| 原始 | 0.5 | 0 | 1 | 0.25 |
| 一级响应 | 0.7 | 0.5 | 0 | 0.20 |
| 二级响应 | 0.9 | 0.8 | 0 | 0.10 |

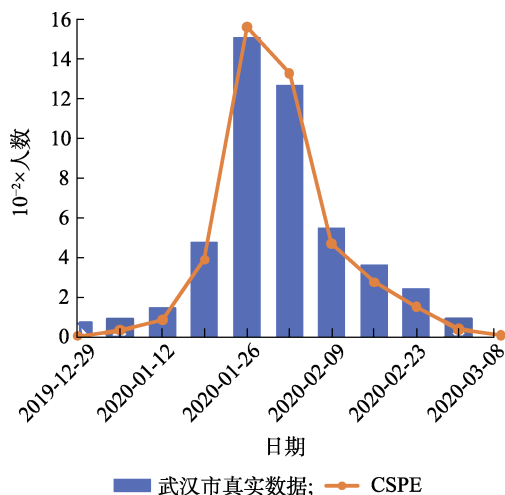


图 2 2020 年初武汉市真实感染人数与 CSPE 模型模拟情况

为进一步验证 CSPE 模型的有效性, 本节继续使用 M-SEIR^[4], GLEaM^[21]与 CSPE 模型分别模拟北京市 2020-06-11—2020-08-06 的疫情发展情况, 真实数据来源为国家卫生健康委员会公开的数据。各模型将 2020 年 6 月 11 日新增的一例本土确诊病例“西城大爷”作为初始输入, 初始治愈人数和死亡人数均设为 0。其中, CSPE 模型在不同时期采取的措施与真实数据保持一致, 如 2020 年 6 月 13 日对海淀区实施二级响应防控措施, 2020 年 6 月 16 日对全市进行社区封闭管理等。对于 M-SEIR 模型, 实验将整个北京市作为研究区域进行模拟, 对于 GLEaM 模型, 实验将引入与 CSPE 模型相同的真实地理数据和交通管控, 其相关参数及 SEIR 的

初始参数如表 3 所示。

M-SEIR, GLEaM 和 CSPE 模型的对比结果与真实数据如表 4 所示. M-SEIR 模型得出的确诊病例呈指数增长, 且无拐点趋向, 直至疫情爆发一个月后才有缓和下降的趋势. 而 M-SEIR 模型中的 4 类人群(即易感者、潜伏者、感染者和康复者)间的转换是由固定参数控制的, 且没有任何措施干预, 所以感染人数和治愈人数远远大于实测结果. GLEaM 模型虽然引用了与 CSPE 模型相同的网络结构, 但对于社区内部只能使用简单的转换率更新各类人群的数量, 且响应措施相关的参数无法全部应用于网络, 因此数据变化幅度较为缓和, 与真实数据仍有差距. CSPE 模型由于加入了措施管控和多层次的地域划分, 更符合实际需求, 因此其确诊、治愈和死亡人数均与实测结果基本吻合. 研究表明, CSPE 模型相比基于传统传染病动力学模型的 M-SEIR 模型和基于复杂网络的 GLEaM 模型, 能从宏微观的角度对城市疫情走势进行更精确的预测, 并能通过调整措施对决策进行分析与评估.

表 4 3 种模型对北京市疫情走势的预测对比

| 模型 | 确诊人数 | | | |
|-----------------------|------------|------------|------------|------------|
| | 2020-06-20 | 2020-06-30 | 2020-07-20 | 2020-08-01 |
| M-SEIR ^[4] | 447 | 54 567 | 24 155 | 5 393 |
| GLEaM ^[21] | 143 | 250 | 93 | 59 |
| CSPE | 198 | 327 | 115 | 4 |
| 真实数据 | 227 | 326 | 104 | 3 |

| 模型 | 治愈人数 | | | |
|-----------------------|------------|------------|------------|------------|
| | 2020-06-20 | 2020-06-30 | 2020-07-20 | 2020-08-01 |
| M-SEIR ^[4] | 249 | 4 964 | 86 300 | 107 929 |
| GLEaM ^[21] | 24 | 47 | 103 | 156 |
| CSPE | 3 | 6 | 246 | 391 |
| 真实数据 | 0 | 2 | 231 | 332 |

| 模型 | 死亡人数 | | | |
|-----------------------|------------|------------|------------|------------|
| | 2020-06-20 | 2020-06-30 | 2020-07-20 | 2020-08-01 |
| M-SEIR ^[4] | 23 | 515 | 11 231 | 11 569 |
| GLEaM ^[21] | 15 | 26 | 53 | 75 |
| CSPE | 0 | 0 | 1 | 4 |
| 真实数据 | 0 | 0 | 0 | 0 |

3.2 不同防控措施下的疫情传播模拟

本节使用 CSPE 模型内置的 4 种响应策略(无措施、一级防控、二级防控和强防控)进行对比测试. 其中, 二级防控所包含的措施有交通检疫、流动人口管理和社区排查; 一级防控增加公共场所封闭; 强防控增加公共交通停运. 而每个措施对

CSPE 模型内的关键参数(如 M-SEIR 模型参数和复杂网络部分的交通管制力度、公共场所防疫管控力度等)都有一定的调整. 测试数据依旧取自 2020-06-11—2020-08-06 的北京市的地理数据, 测试结果为感染者、治愈者和病亡者 3 类人群的数量统计. 具体实验内容如下.

3.2.1 无措施与强防控对疫情传播的影响

为了反映措施对疫情防控的重要性, 实验分别模拟无措施和强防控条件下的北京市疫情传播. 如图 3 所示, 无措施状态下的 3 类人群的变化在高峰期以百万为单位, 且受新冠病毒 14 天潜伏期的影响, 在第 1 例确诊病例出现后的一个月才开始有明显增长, 于 2020 年 7 月 15 日达到巅峰, 此时确认病例为 830 多万例, 直至 8 月初仍有 200 多万的感染者. 而强防控策略在发现初始病例时就采取封城等防控措施, 感染人数于 6 月 28 日达到最多, 仅为 152 人, 新冠病毒的传播仅在 2 周内便得到了遏制, 于 7 月 20 日已基本根除, 所有感染完全获得治疗. 由此可见, 管制措施的介入对于疫情防控作用十分明显.

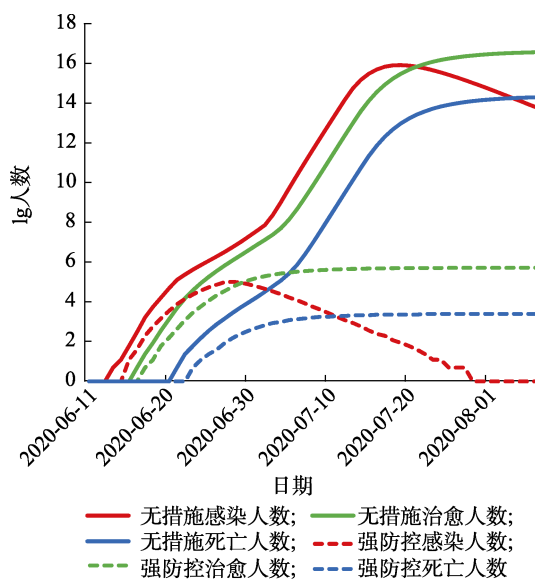


图 3 无措施与强防控对疫情传播影响的对比

3.2.2 有无防控延时效对疫情传播的影响

实际的疫情防控时间是不可控的, 为比较不同时间采取同一防控措施对推演结果的影响, 实验模拟了北京市在第 1 天和第 3 周实施一级防控措施后的疫情发展状况. 对比结果如图 4 所示, 实线表示开始就进入强防控管制, 虚线表示延迟一个月再施加一级防控措施. 直接防控下的感染病例峰值为 360 例, 而有延迟影响防控下的感染病例峰

值高达 7.5 万例。由此可见, 在初始病例刚出现时就采取一级防控, 可以有效地抑制疫情恶化, 公共场所的封闭能在短时间内阻断传染途径, 避免病毒扩散。实验还证明了 CSPE 模型在时间上的可扩展性, 即在任何时间段指定区域添加防控措施, 并能根据对比结果进行多次迭代调整后得出的一套较为完善的防疫体系。

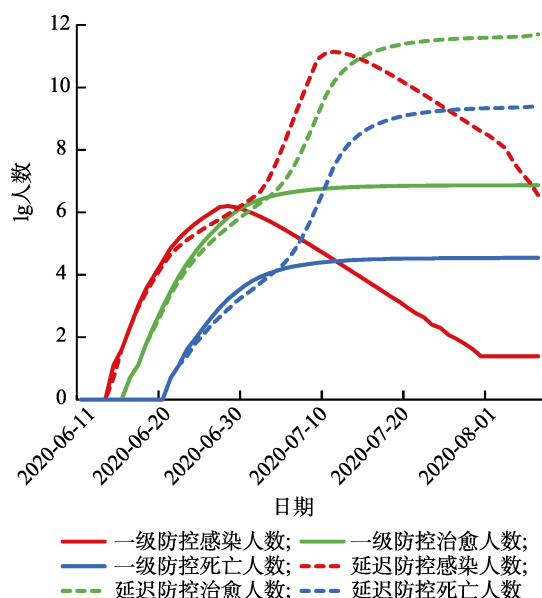


图 4 有无防控延时对疫情传播影响的对比

3.2.3 一级与二级防控对疫情传播的影响

为进一步确认措施选择对疫情防控的影响程度, 实验使用 CSPE 模型分别就一级和二级防控下的北京市的疫情传播过程进行了推演, 通过对比体现有效措施必要性。管控后的模拟结果如图 5 所示, 一级防控下感染曲线的柱点相比二级防控

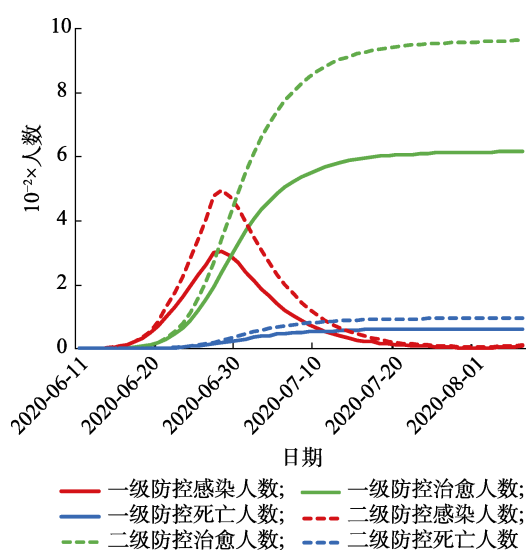


图 5 一级与二级防控对疫情传播影响的对比

虽然仅提前一天, 但峰值却下降了 43%。这是因为一级防控限制了公共场所的人员流动, 在传染途径上减小了外出人群的接触密度, 降低了感染和潜伏人群携带病毒对外传播的可能。因此, 对于人口基数较多的感染城市, 采取多个包括公共交通管制在内的防控措施, 可以避免发生大范围的传染现象。此外, 实验结果也说明了 CSPE 模型在措施管控的可扩展性, 模型可以根据国家颁布的政策对内置措施及其参数作出调整, 以适应城市当前的状况和病毒传播的特性。

总的来说, CSPE 模型相比其他传染病动力学模型, 除了可以暴露更多的复杂网络和智能体相关的参数外, 还提供了多种措施选择, 用户能够根据实际需求进行调整, 经过多方对比后作出决策, 灵活应对疫情。

3.3 基于地理信息系统的疫情传播模拟

本节基于真实地理数据、使用 CSPE 模型以热力图的形式模拟武汉市在施加强防控措施下的病毒传播过程。测试地理位置为武汉市江岸区, 初始感染人数为 100, 仿真天数为 40 天, 并于第 20 天采取强防控措施。

病毒传播过程的预测推演结果如图 6 所示。第 10 天的新冠病毒分散至江岸区各公共和社区节点, 由于病毒的高传染性, 疫情于第 20 天迅速扩散至其他节点。为抑制疫情继续恶化, CSPE 模型对推演区域采用强防控措施。受 14 天潜伏期的影响, 第 30 天已感染区域内的疫情继续恶化, 但得益于强防控的交通管控, 除此之外的地区没有任何新增病例, 直至第 40 天疫情发展有明显好转。由此可见, CSPE 模型能够基于地理信息系统完成对推演区域的疫情模拟, 并能根据智能体的社交属性实现对感染个例的行迹追踪, 从而预测新冠病毒的扩散趋势和影响范围。同时, CSPE 模型还可以实时地干预疫情推演过程并施加措施, 这为用户决策提供了参考。

3.4 基于 CSPE 模型的疫情推演可视化系统

本文基于 CSPE 模型设计了一种用户交互式的可视化系统。该系统能够根据用户输入的信息对某个城市进行疫情仿真, 通过手动添加防控措施对当前疫情形势进行推演, 具备较强的可自定义性和可扩展性。

基于 CSPE 模型的可视化系统主用户界面如图 7 所示。用户通过输入基本疫情初始信息, 即 SEIR 模型 4 类人群数量、推演日期和时长, 预测未来一段时间内某个城市的传染人群分布, 用趋势图、柱

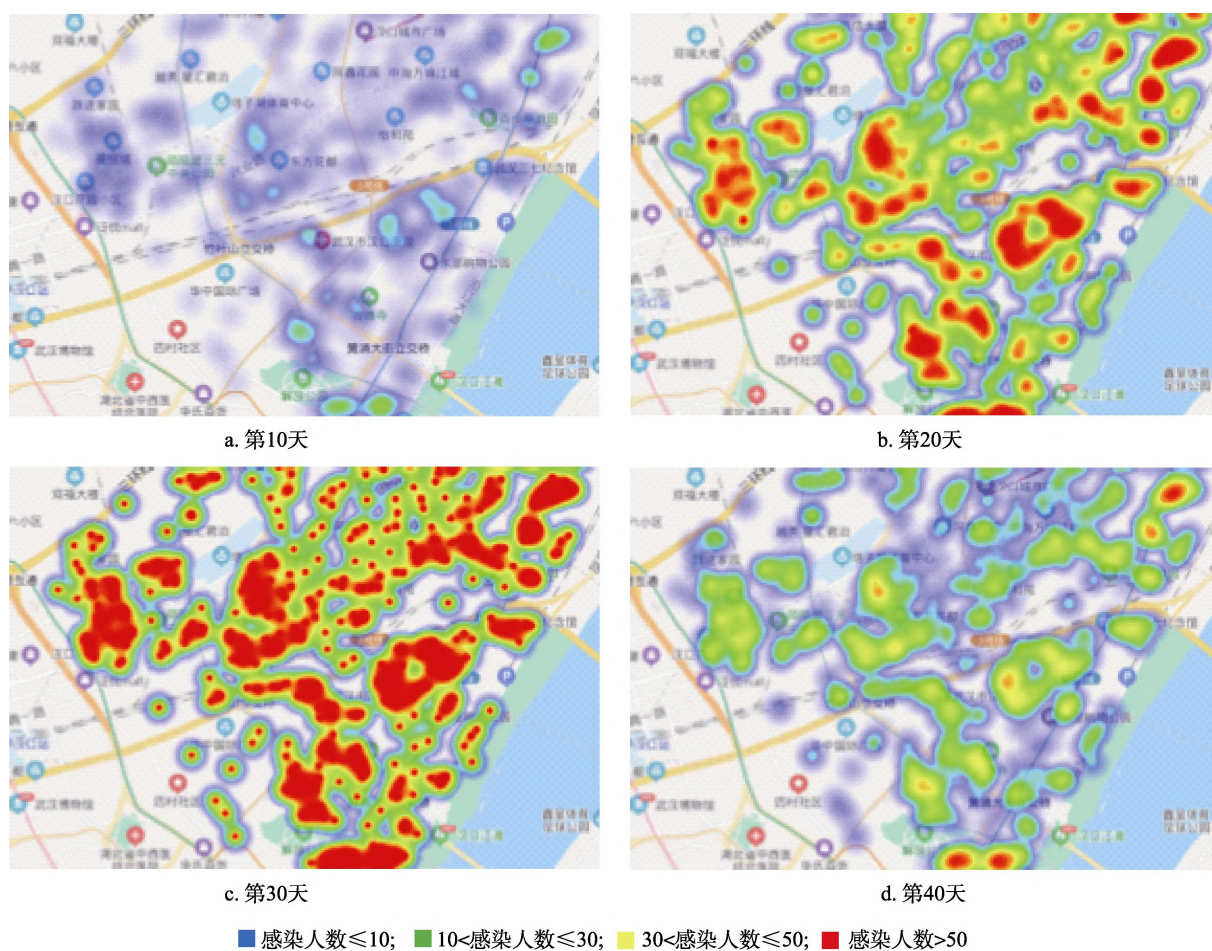


图 6 采取强防控后的 COVID-19 扩散热力图



图 7 基于 CSPE 模型的可视化系统

状图和热力图展示整片区域的发展状况. 用户也可通过自定义预设措施对不同行政区实施防控, 以进行决策从而控制疫情发展.

4 结 语

新冠病毒(COVID-19)的快速蔓延, 对全球公众健康造成了严重威胁. 因此, 本文针对新冠病毒的特性, 在 M-SEIR 模型的基础上引入了复杂网络模型和智能体, 提出了一种城市级结构化新冠疫情预测和仿真模型. 该模型将宏观的网络模型与微观的 SEIR 模型相结合, 并使用智能体模拟个体的社会活动, 能够精确地预测城市在一定时间内的疫情发展状况. 但是, 本文模型在施加措施时依赖参数, 且默认参数只能按现有经验设置, 在没有足够数据的情况下进行调参会缺乏一定的参考性. 此外, 由于模型加入了智能体和复杂网络, 当推演人数或交互节点较多时, 运算量会大幅增加, 导致模型无法在短时间内得出结果, 影响决策效率. 因此, 在后续研究中, 考虑优化智能体的筛选, 尽可能地去除不必要的冗余操作, 或者从分布式计算的角度考虑, 使用多机并行的方式进行推演, 以缩短仿真时间.

参考文献(References):

- [1] Velavan T P, Meyer C G. The COVID-19 epidemic[J]. *Tropical Medicine & International Health: TM & IH*, 2020, 25(3): 278-280
- [2] Wang Linghang. Characteristics and countermeasures of 2019-nCoV infection[J]. *Chinese Journal of Experimental and Clinical Infectious Diseases: Electronic Edition*, 2020, 14(1): 1-5(in Chinese)
(王凌航. 新型冠状病毒感染的特征及应对[J]. *中华实验和临床感染病杂志: 电子版*, 2020, 14(1): 1-5)
- [3] Duan W, Fan Z C, Zhang P, *et al.* Mathematical and computational approaches to epidemic modeling: a comprehensive review[J]. *Frontiers of Computer Science*, 2015, 9(5): 806-826
- [4] Yang Z F, Zeng Z Q, Wang K, *et al.* Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions[J]. *Journal of Thoracic Disease*, 2020, 12(3): 165-174
- [5] Padgham L, Winikoff M. Developing intelligent agent systems: a practical guide[M]. New York: John Wiley & Sons, 2005
- [6] Pastor-Satorras R, Castellano C, van Mieghem P, *et al.* Epidemic processes in complex networks[J]. *Reviews of Modern Physics*, 2015, 87(3): 925-979
- [7] Grassly N C, Fraser C. Mathematical models of infectious disease transmission[J]. *Nature Reviews Microbiology*, 2008, 6(6): 477-487
- [8] Li M Y, Muldowney J S. Global stability for the SEIR model in epidemiology[J]. *Mathematical biosciences*, 1995, 125(2): 155-164
- [9] Liu Z, Magal P, Seydi O, *et al.* A COVID-19 epidemic model with latency period[J]. *Infectious Disease Modelling*, 2020, 5: 323-337
- [10] He S B, Peng Y X, Sun K H. SEIR modeling of the COVID-19 and its dynamics[J]. *Nonlinear Dynamics*, 2020, 101(3): 1667-1680
- [11] Pandey G, Chaudhary P, Gupta R, *et al.* SEIR and regression model based COVID-19 outbreak predictions in India[OL]. [2021-06-09]. <https://arxiv.org/abs/2004.00958>
- [12] Mwalili S, Kimathi M, Ojiambo V, *et al.* SEIR model for COVID-19 dynamics incorporating the environment and social distancing[J]. *BMC Research Notes*, 2020, 13(1): Article No.352
- [13] Cao Shengli, Feng Peihua, Shi Pengpeng. Study on the epidemic development of COVID-19 in Hubei province by a modified SEIR model[J]. *Journal of Zhejiang University: Medical Sciences*, 2020, 49(2): 178-184(in Chinese)
(曹盛力, 冯沛华, 时朋朋. 修正 SEIR 传染病动力学模型应用于湖北省 2019 冠状病毒病(COVID-19)疫情预测和评估[J]. *浙江大学学报: 医学版*, 2020, 49(2): 178-184)
- [14] Chen Xiaohui, Xu Li, Ge Lei, *et al.* A survey of infectious disease transmission data visual analysis[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2020, 32(10): 1581-1593(in Chinese)
(陈晓慧, 徐立, 葛磊, 等. 传染病传播数据可视分析综述[J]. *计算机辅助设计与图形学学报*, 2020, 32(10): 1581-1593)
- [15] Jin Sichen, Tao Yubo, Yan Yuyu, *et al.* Infectious disease patterns analysis based on visualization of multidimensional space-time data[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2019, 31(2): 241-255(in Chinese)
(金思辰, 陶煜波, 严宇宇, 等. 基于多维时空数据可视化的传染病模式分析[J]. *计算机辅助设计与图形学学报*, 2019, 31(2): 241-255)
- [16] Liu Jianxiang, Liu Haiyan, Chen Xiaohui, *et al.* Visual analysis method for COVID-19 epidemic situation[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2020, 32(10): 1617-1627(in Chinese)
(刘建湘, 刘海砚, 陈晓慧, 等. 新冠肺炎疫情数据多维度可视分析方法[J]. *计算机辅助设计与图形学学报*, 2020, 32(10): 1617-1627)
- [17] Amaral L A N, Ottino J M. Complex networks[J]. *The European Physical Journal B-Condensed Matter*, 2004, 38(2): 147-162
- [18] Meyers L A, Pourbohloul B, Newman M E J, *et al.* Network theory and SARS: predicting outbreak diversity[J]. *Journal of Theoretical Biology*, 2005, 232(1): 71-81
- [19] Fan Ruguo, Wang Yibo, Luo Ming, *et al.* SEIR-based COVID-19 transmission model and inflection point prediction analysis[J]. *Journal of University of Electronic Science and Technology of China*, 2020, 49(3): 369-374(in Chinese)
(范如国, 王奕博, 罗明, 等. 基于 SEIR 的新冠肺炎传播模型及拐点预测分析[J]. *电子科技大学学报*, 2020, 49(3): 369-374)

- [20] Liu C Y, Wu X Q, Niu R, *et al.* A new SAIR model on complex networks for analysing the 2019 novel coronavirus (COVID-19) [J]. *Nonlinear Dynamics*, 2020, 101(3): 1777-1787
- [21] Balcan D, Gonçalves B, Hu H, *et al.* Modeling the spatial spread of infectious diseases: the global epidemic and mobility computational model[J]. *Journal of Computational Science*, 2010, 1(3): 132-145
- [22] Abbas H A, Shaheen S I, Amin M H. Organization of multi-agent systems: an overview[OL]. [2021-06-09]. <https://arxiv.org/abs/1506.09032>
- [23] Perez L, Dragicevic S. An agent-based approach for modeling dynamics of contagious disease spread[J]. *International journal of health geographics*, 2009, 8: Article No.50
- [24] Liu Tao, Li Xia, Liu Xiaoping. Integration of small world networks with multi-agent systems for simulating epidemic spatiotemporal transmission[J]. *Chinese Science Bulletin*, 2009, 54(24): 3834-3843(in Chinese)
- (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[J]. *科学通报*, 2009, 54(24): 3834-3843)
- [25] Shamil M S, Farheen F, Ibtehaiz N, *et al.* An agent-based modeling of COVID-19: validation, analysis, and recommendations[J/OL]. *Cognitive Computation*, 2020, 1-17. [2021-06-09]. <https://doi.org/10.1101/2020.07.05.20146977>
- [26] Kerr C C, Stuart R M, Mistry D, *et al.* Covasim: an agent-based model of COVID-19 dynamics and interventions[J]. *PLoS Computational Biology*, 2021, 17(7): e1009149
- [27] Wu Yajing, Zhang Peng, Di Zengru, *et al.* Study on bipartite networks[J]. *Complex Systems and Complexity Science*, 2010, 7(1): 1-12(in Chinese)
- (吴亚晶, 张鹏, 狄增如, 等. 二分网络研究[J]. *复杂系统与复杂性科学*, 2010, 7(1): 1-12)
- [28] Pan A, Liu L, Wang C L, *et al.* Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China[J]. *JAMA*, 2020, 323(19): 1915-1923