

SHANGHAIUNIVERSITY

毕业论文（设计）

**UNDERGRADUATE THESIS (PROJECT)**

**题目:地点影响力下的个体空间移动与传播行为研究**

|  |  |
| --- | --- |
| **学院** | **中欧工程技术学院** |
| **专业** | **信息工程** |
| **学号** | **18124703** |
| **学生姓名** | **夏艺颖** |
| **指导教师** | **王冰** |
| **起讫日期** | **2022.02.21– 2022.06.03** |

目录

[摘要 III](#_Toc10091)

[ABSTRACT IV](#_Toc9734)

[第1章 绪论 1](#_Toc12899)

[§ 1.1 背景及意义 1](#_Toc10449)

[§ 1.2 国内外研究现状与发展趋势 2](#_Toc25569)

[§ 1.2.1 流行病传播模型的研究现状 2](#_Toc17620)

[§ 1.2.2 强化学习的研究现状 3](#_Toc25601)

[§ 1.2.3 发展趋势 3](#_Toc5913)

[§ 1.3 研究内容及目标 4](#_Toc2022)

[§ 1.3.1 研究内容 4](#_Toc9071)

[§ 1.3.2 研究目标 4](#_Toc8821)

[§ 1.4 本文组织结构 4](#_Toc7733)

[第2章 地点影响力下的个体空间移动的实现 6](#_Toc14194)

[§ 2.1 地点影响力的定义 6](#_Toc8564)

[§ 2.2 二维网络的构建 6](#_Toc28851)

[§ 2.3 基于强化学习的个体空间移动的实现 7](#_Toc4052)

[§ 2.3.1强化学习的基本概念 7](#_Toc11716)

[§ 2.3.2 强化学习的分类 8](#_Toc15408)

[§ 2.3.3 Q-learning算法介绍 9](#_Toc2466)

[§ 2.3.4 结合Q-learning算法的地点影响力下的个体空间移动模型 11](#_Toc25304)

[§ 2.4 本章小结 13](#_Toc20359)

[第3章 基于个体移动的集合种群传播模型的建立 14](#_Toc10680)

[§ 3.1 流行病传播模型介绍 14](#_Toc24280)

[§ 3.1.1 SIR传播模型介绍 14](#_Toc18186)

[§ 3.1.2 集合种群传播模型介绍 15](#_Toc16305)

[§ 3.2 结合个体空间移动模型的集合种群模型的建立 17](#_Toc10806)

[§ 3.3 本章小结 19](#_Toc29861)

[第4章 仿真实验结果与分析 20](#_Toc8460)

[§ 4.1 实验一：不同增长系数下的流行病传播情况。 20](#_Toc15276)

[§ 4.1.1 不同感染率增长系数对流行病传播的影响 20](#_Toc898)

[§ 4.1.2 不同恢复率增长系数对流行病传播的影响 22](#_Toc16212)

[§ 4.2 实验二：特定增长系数下的流行病传播情况 24](#_Toc656)

[§ 4.2.1 实验具体设置 24](#_Toc692)

[§ 4.2.2 实验结果及分析 26](#_Toc23282)

[§ 4.3 本章小结 28](#_Toc26029)

[第5章 仿真实验的可视化 29](#_Toc29384)

[§ 5.1 个体空间移动轨迹的可视化 29](#_Toc19584)

[§ 5.1.1 实现方法 29](#_Toc24946)

[§ 5.1.2 可视化展示 30](#_Toc3263)

[§ 5.2 流行病传播过程可视化 31](#_Toc6941)

[§ 5.2.1 实现方法 31](#_Toc4900)

[§ 5.2.2 可视化展示 31](#_Toc1495)

[§ 5.3 本章小结 33](#_Toc4316)

[第6章 总结与展望 34](#_Toc11300)

[§ 6.1 本文总结 34](#_Toc13366)

[§ 6.1.2 本文主要工作 34](#_Toc8560)

[§ 6.1.3 本文主要创新点 34](#_Toc20452)

[§ 6.2 展望 34](#_Toc3790)

[致谢 36](#_Toc11567)

[参考文献 37](#_Toc21214)

[附录：部分源程序清单 38](#_Toc26823)

地点影响力下的个体空间移动

与传播行为研究

**摘要**

流行病作为一类感染性极强的传染病，千百年来一直威胁着人类的生活与健康，为有效预测和控制流行病传播，建立数学模型是一个很好的途径；在现实生活中，人类的移动行为常常受到地点影响力的影响，而城市规模和医疗资源是决定地点影响力大小的主要因素。本文为研究地点影响力下的个体空间移动与流行病传播的关系，利用Q-learning算法，使个体进行强化学习下的移动，并结合个体空间移动，建立基于SIR模型的集合种群模型。通过对感染率增长系数、恢复率增长系数、城市规模和医疗资源数量的研究，本文发现：感染率增长系数的增长不利于流行病的控制，而恢复率增长系数的增长有助于控制流行病传播。另外，城市规模的均匀分布以及医疗资源的分配能够有效控制感染规模，其中医疗资源的不均匀分配比均匀分配更能减少感染规模。此外，本文运用可视化方法，直观地展现了个体的移动轨迹和传播过程。

关键词：SIR模型，集合种群模型，个体空间移动，地点影响力，Q-learning

Research on individual mobility and disease spreading based on location influence

**ABSTRACT**

Epidemic disease is a highly contagious disease, which has been threatening human health for thousands of years and has brought great challenges to human society, economy and medical treatment. In order to effectively predict and control the spread of epidemic disease, human beings began to study the mathematical model of epidemic spread from the 20th century. At present, there have been many mathematical models such as warehouse model, complex network model and Compound population network model.

In real life, human mobility is often influenced by the Place influence. A place's medical resources, economic activity, education level, employment and wages are all factors that determine the Place influence. If the Place influence of a city is higher than that of other cities, individuals will tend to move to this city. In this study, city size and medical resources are selected as two factors to determine the Place influence.

In this paper, Q-learning algorithm is used to enable individuals to carry out reinforcement learning and move in space under the influence of place. The relationship between individual movement and disease spreading is studied by using Metapopulation model based on SIR model.

By studying the increasing coefficient of infection rate, the increasing coefficient of recovery rate, the city size and the number of medical resources, we found that the increasing coefficient of infection rate is not conducive to the control of epidemic, while the increasing coefficient of recovery rate is helpful to the control of epidemic spread. In addition, the uniform distribution of city size and the distribution of medical resources can effectively control the scale of infection, and the uneven distribution of medical resources can reduce the scale of infection more than the uniform distribution.In addition, this paper uses visualization method to visually show individual movement track and epidemic transmission process.

**Keywords:** SIR model，Individual mobile，Place influence，Q-learning

# 绪论

本章主要描述了流行病传播的背景、意义，分析了流行病传染模型和强化学习的国内外研究现状，进而提出了本文所要研究的内容及目标。

## **1.1 背景及意义**

流行病是一种可以在短时间内广泛传播、并且使大量人口感染的传染病。从古至今，流行病的问题一直困扰着人类，几乎每个时代、每个国家都不能幸免于流行病的侵扰。

公元二世纪到四世纪，瘟疫在罗马帝国横行，对帝国的农业、军事、人口带来了巨大的冲击，使一个繁盛的帝国走向分崩离析[1]。1347年，黑死病在欧洲爆发，短短六年时间就剥夺了欧洲三分之一人口的生命，也在无形之中改变了世界文明的发展走向[2]。公元十六世纪到十八世纪，天花肆虐，被称为“人类史上最大的屠杀”，全球有超过10%的人口死于天花的魔爪[3]。自1817年以来，霍乱在全球范围内爆发过不下10次，直至今日，海地等地区依然遭受着霍乱的侵蚀，却束手无策[4]。1918年，始于西班牙的流感通过回国的士兵传播开来，造成五亿人感染、五千万人死亡，感染人数占了全球人口的20%[5]。随后SARS病毒、甲型H1N1流感病毒、MERS病毒和Zika病毒也相继在全球范围内两百多个国家爆发，造成了大量死亡。直至今日，新冠肺炎病毒依旧在全球肆虐，累积确诊病例超过五亿人口，六百多万人死于新冠，还有许多人受到后遗症的困扰。每一次流行病的爆发都是对人类生活、经济、医疗、健康的重大打击。

早在1760年就已经有人开始关注流行病的传播特征，试图通过了解流行病的特点来找到干预流行病传播的方法。但是直到20世纪人类才正式开始系统地研究流行病传播模型。著名的“仓室模型”就是流行病传染领域中最早使用的数学模型，此后不断地在仓室模型的基础上出现新的模型。随着流行病传播模型的研究，人类对流行病有了一定的了解，加之公共卫生基础设施的发展、医学的进步，许多诸如天花、麻疹、黑死病等疾病已经逐渐退出了流行病的舞台。然而人类在进步，病毒也在不断地进化，传统的流行病传播模型已经不能用来模拟所有的流行病传播，人类依旧需要研究更加多元、更加完善、更符合实际的数学模型，来模拟流行病的传播，有针对性地遏制流行病的肆虐。

在现实生活中，人类的空间移动行为受到地点影响力的影响，地点的医疗资源、经济活跃度、教育水平等等因素都是地点影响力的一个组成部分。如果一个地点的影响力比其他地点的大，那么人们就会倾向于向地点影响力大的地点移动，这种现象对流行病的传播有很大的影响。

本课题借助计算机建立数学模型，模拟地点影响力下的个体空间移动与流行病的传播行为。将基于地点影响力的个体移动与流行病的传播相结合，是一个新的研究方向。与个体的随机移动相比，研究强化学习下的个体空间移动与流行病传播行为的影响，更符合人类移动规律，能有效地帮助人们观察并了解流行病传播过程，进而对研究流行病传播机理、选择干预措施和预测传染趋势有很大帮助。

1.2 国内外研究现状与发展趋势

* 1.2.1 流行病传播模型的研究现状

自人类文明存在以来，流行病就一直伴随着人类、影响着人类文明发展进程。为战胜流行病，人类发展了许多研究流行病的方法：观察法、实验法和数理法[6]。其中，数理法就是运用真实数据建立数学模型，并借助计算机进行模拟仿真，以此来研究流行病的特性。

最早的流行病传播模型是仓室模型，该模型是由W. Kermack和McKendrick在研究黑死病时提出的[7]。仓室模型是单一群体模型，是最经典的模型，该模型的建立基于一个均匀分布的种群，且认为种群内的个体无差异，即每个个体的行为相同、接触概率相等。仓室模型主要有SI（易感-感染）模型、SIS（易感-感染-易感）模型和SIR（易感-感染-免疫）模型，后又出现了SEIR（易感-暴露-感染-免疫）模型、SEIRS（易感-暴露-感染-恢复-易感）模型等更复杂的模型。作为经典的流行病传播模型，仓室模型的应用极其广泛，后来的很多模型，都是在仓室模型的基础上进行改进的。

然而内部均匀分布、个体无差异的仓室模型忽略了个体的异质性，与现实世界有很大的差异，于是后来出现了微观个体模型[9]。微观个体模型利用复杂网络，将单个个体看成一个结点，个体与个体间的社会关系看成一条边，侧重研究个体的异质性（即每个个体有不同的行为，并不是完全无差异的）对流行病传播的影响。

上述仓室模型和复杂网络下的微观个体模型都是静态的模型，完全忽略了个体的动态移动。在很多情况下，依旧不能反映真实的流行病传播情况。于是近几年又出现了动力学模型[9]和集合种群网络模型[8]。动力学模型是由Gonzalezd基于牛顿动力学提出的，该模型使个体一直做永无止境的无规则随机移动，侧重研究个体随机移动下的流行病传播情况。而由Colizza提出的集合种群网络模型，顾名思义，就是由多个子种群组成一个网络的模型。通常模型中的子种群间是异质的，而子种群内部是均匀混合的。集合种群网络模型研究的是个体在不同种群之间的移动对流行病传播过程的影响[9]。

流行病传播模型不仅可以模拟流行病的传播，还能模拟信息的传递与网络病毒的扩散[16]。

* 1.2.2 强化学习的研究现状

强化学习是机器学习的一个分支，与它并列的还有监督学习、非监督学习和进化学习。强化学习，亦称“评价学习”，是一种融合了监督学习与非监督学习的机器学习。

最早出现的强化学习方法，是Bellman在1856年提出的动态规划方法。该方法是将原有的复杂问题分解成相互联系又相对简单的小问题，用倒推的方法从后到前求解问题，即后面一个问题的最优解就是前面一个问题的最优解。动态规划方法将每个小问题的效益结合起来求出最优解，大大减少了计算量、提高了效率，但是这种方法只有在模型已知的前提下才能使用。1977年Werbos在Bellman动态规划法的基础上，提出了自适应动态规划法[11]。该方法利用函数近似结构成功地解决了维数的问题。

动态规划方法是基于已知模型的算法，然而通常情况下模型是未知的。1946年由三位美国科学家提出的蒙特卡罗法，就是基于未知模型的强化学习算法。蒙特卡罗法将问题与概率统计结合起来获得近似解，但是这种算法往往需要走完一个完整序列才能更新数据。1988年Sutton提出了一种无模型的时间差分算法。该算法将动态规划方法与蒙特卡罗方法结合起来，利用后继函数的近似解估计当前函数。这种算法是单步更新的，与蒙特卡罗法相比速度更快，更灵活。

在时间差分算法中，最为常见的是Q-learning算法和Saras算法。前者是在1992年由Watkins提出的离线策略算法，而后者是在1996年由Bersekas提出的在线策略算法。两种算法在选择行为时都遵循一定的策略，但在更新行为值函数时，Saras遵循的是与选择动作时一样的策略，而Q-learning算法则直接采取了值最大的行为。因此在计算周期上，Q-learning能更快地找到最优解，但也因此需要承担更大的学习失败的风险。

而后，Silver在2014年提出了确定性策略梯度算法，成功地将强化学习的算法划分为基于值函数和基于策略探索两种类型。随后一年，就有人提出了利用深度学习中的卷积神经网络代替Q函数的Deep-Q-Network算法[12]，轰动了整个强化学习领域。

强化学习作为热门领域，其应用非常广泛，在道路规划、无人机、工业自动化、金融贸易、语言处理、新闻推荐、游戏等各种领域都能发光发热[20]。

* 1.2.3 发展趋势

流行病传染模型发展至今，逐渐从静态的模型转变为动态的模型；从单一群体模型，转变为集合种群网络模型。在目前的动态模型中，个体大多是随机移动的，或者是按照一定的移动策略移动的，鲜有研究者将强化学习与个体移动结合起来，研究流行病的传播。未来流行病模型的发展趋向于研究个体与种群的异质与交互。除了与时下热门的强化学习相结合，研究个体在集合种群间的移动对流行病传播的影响外，还将趋向于研究多种流行病在动态网络中的传播。

1.3 研究内容及目标

本课题运用强化学习算法，使个体在地点影响力的作用下进行种群间的移动，并结合个体移动建立流行病传播模型。在不同城市规模和医疗资源的分布下，研究地点影响力下的个体空间移动对流行病传播的影响。

* 1.3.1 研究内容

研究地点影响力下的个体空间移动与流行病的传播，主要完成以下内容：

1. 地点影响力的定义与数学化描述；
2. 二维网络的构建与强化学习下个体空间移动的实现；
3. 结合个体移动的流行病传播模型的实现；
4. 不同地点影响力分布下的实验与数据分析；
5. 可视化实现。

* 1.3.2 研究目标

本文通过强化学习的算法，实现地点影响力作用下的个体空间移动，并建立移动种群模型。运用建立好的模型，研究不同地点影响力分布时，个体空间移动对流行病传播的影响。针对本课题的研究内容，制定了以下几项目标：

1. 完成整体模型框架的建立；
2. 实现个体移动的可视化和传播过程的可视化；
3. 完成两个实验，并且对实验数据进行分析：
   1. 研究感染率增长系数和恢复率增长系数对流行病传播的影响；
   2. 设计四种地点影响力分布情况，研究个体移动对流行病传播的影响。

通过完成上述目标，理解地点影响力下的个体空间移动对流行病传播过程的影响，从而为深入了解流行病特点、研究干预措施和预测传播过程提供有力支持。

1.4 本文组织结构

整篇论文分为六章。

第一章介绍课题的研究背影、研究意义，分析流行病模型和强化学习的研究现状及发展趋势，并提出课题的研究内容以及研究目标。

第二章首先对地点影响力进行定义，描述二维网络的构建方法，介绍强化学习的基本概念、分类和强化学习中Q-learning算法的原理，并且介绍地点影响力下个体空间移动的具体实现方法。

第三章介绍SIR传播模型和集合种群模型的基本概念，并且实现结合了第二章中个体空间移动的集合种群模型。

第四章提出了两个实验，对实验的内容进行介绍，并且对结果进行分析。

第五章展示个体空间移动和传播过程两个方面的可视化方法和结果。

第六章对整篇论文进行总结，归纳课题的主要内容与创新点，并提出需要进一步研究的问题。

# 地点影响力下的个体空间移动的实现

本章具体描述了地点影响力下的个体空间移动的实现方法：介绍了地点影响力的定义，描述了二维网络的构建方法，介绍了强化学习的基础知识与分类，并且详细介绍了强化学习中的Q-learning算法，结合地点影响力实现强化学习下的个体空间移动。

## **2.1 地点影响力的定义**

影响力是一种能够以被人接受的方式改变他人想法和行为的能力。因此，地点影响力就是一个地点对个体的吸引力。一个地点影响力越高，它对个体的吸引力就越强，越能使个体改变行为向这个地点移动。

在现实生活中，每个城市的地点影响力受各种因素的影响而不同，比如：城市的医疗资源、经济活跃度、人口密度、就业提供水平、薪资收入、气候环境、教育水平、生活节奏和人际关系等[13][14]。本课题以流行病的传播为背景，选择了两个与地点影响力有关的主要因素：城市规模大小和医疗资源数量。这两个因素与地点影响力的关系满足公式（2-1）：



（2-1）

其中，CI代表城市的地点影响力大小，CS代表城市的规模，MR代表城市拥有医疗资源的数量，CS和MR都属于自然数集。k1是常数，其含义是当医疗资源总数不变时，城市规模每扩大1，地点影响力就增加k1，此时地点影响力与城市规模成正比。k2也是常数，其含义是当城市规模大小不变时，医疗资源每增加1，地点影响力就增加k2，此时城市的地点影响力与医疗资源数量成正比。在本课题中k1=k2=2。

## **2.2 二维网络的构建**

本课题建立了一个大小为L\*L的二维网络，并且假设每个城市的大小与形状都相同且紧密相连。具体模型如图2-1所示。

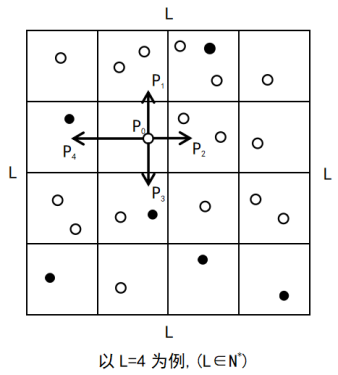


图2-1 二维网络模型（以L=4为例）

图2-1中每一个网格代表了一个城市，在二维网络上一共有L\*L个城市。每个城市都有一个城市编号i、感染率βi、恢复率μi和地点影响力CIi。图中黑色或白色的圆点代表一个个体。在流行病开始传播之前，将N个个体按照初始人口密度的大小随机分配到L\*L个城市中，城市i的初始人口密度 ρi 满足公式（2-2），其中CS是城市规模，Nsum是二维网络总人数，CS属于自然数集，Nsum属于正整数集。每个个体都有它所在城市的编号i和个体当前的感染状态s。在流行病传播之初，即传播时刻t=0时，假设有n个个体已经被感染（黑色的点表示已经感染的个体），则有Nsum-n个个体未被感染（白色的点表示未感染的个体）。在本课题中n=5。



（2-2）

由于这是一个规则的二维网络，城市的大小形状相同，且城市与城市之间唯一的连接就是四条边，因此假设在二维网络中，个体只能在城市之间进行五种简单的移动，分别为向上移动、向下移动、向左移动、向右移动和待在原来的城市不向其他城市移动。个体每一次移动的方向都由概率Paction\_i（0≤Paction\_i≤1）决定（2.3节详细说明），且每次最多移动一个城市。为了移动方便，有时城市i的编号会与城市的坐标（Xi,Yi）进行转换，它们的关系满足方程组（2-3），其中i是城市编号，L是二维网络边长。

Xi = i // L

（2-3）

Yi = i % L

## **2.3 基于强化学习的个体空间移动的实现**

### **2.3.1强化学习的基本概念**

强化学习从属于机器学习，是一个集数学、工程、心理学、计算科学、神经科学、经济学等众多学科为一体的大分支。实际上，强化学习源于模仿动物学习新技能的过程[15]，是计算机自动地通过不断地尝试，根据环境反馈，强化好的行为，排除错误的行为，最终获得最优策略的算法。

强化学习通常包含七个基本要素：智能体agent、环境environment、策略policy、状态s、行为a、奖励R和状态转移概率P。其中，智能体是学习者和行为的决策者，只能通过选择行为去执行来改变环境状态，通常智能体都是通过传感器感知环境并且执行行为的；环境，是指除了智能体之外，一切能与智能体交互的事物；策略，是从状态到行为的映射，是智能体选择行为的依据。策略分为确定和随机两种，确定策略就是状态与行为一一对应的策略，而随机策略的状态对应的行为不唯一，通常用概率描述，即某一状态下选择这一动作行为执行的概率；状态，就是智能体当前所处的状态，是智能体从环境中获取的信息；行为，是智能体改变环境和状态的媒介；奖励，是环境对于行为结果的反馈，智能体通过奖励来判断行为的好坏，以此更新策略；状态转移概率，是智能体采取行为后进入下一个状态的概率。

强化学习的基本实现过程如图2-2所示：智能体根据现有策略选择一个行为a，通过执行行为a使环境进行一定改变，环境在接受行为a之后通过状态转移概率P改变状态，并把及时奖励R和状态s反馈给智能体，智能体接受环境信号之后，更新状态，并且按照累积奖励最大化的原则更新策略。智能体和环境不断交互学习，当算法收敛时，学习结束，此时的策略就是智能体达到目的状态的最优策略。

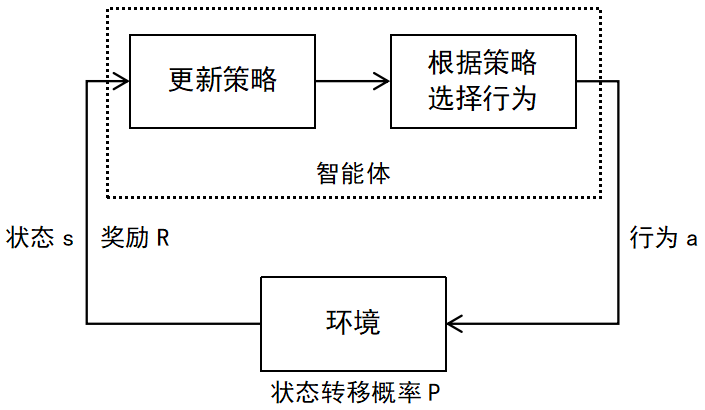


图2-2 强化学习基本原理示意图

* 2.3.2 强化学习的分类

强化学习（RL）算法的分类有很多种，根据环境模型是否已知，分为基于模型（model-based）的算法和不基于模型的算法（model-free）；根据直接输出结果是行为还是价值，分为基于策略（policy-based）的算法和基于价值（value-based）的算法；根据更新方法，分为回合更新算法（Monte-Carloupdate，MC）和单步更新算法（Temporal-Difference update，TD）；根据行为策略和决策策略是否一致，分为在线学习（On-Policy）算法和离线学习（Off-Policy）算法。下面详细介绍每种类型的算法特征[21]。

①基于模型的算法：该算法的环境对于智能体来说是已知的，智能体选择行为实施时，事先通过已知模型预判到所有可能会发生的情况，然后从中选取最好的一种。这种基于模型的强化学习在模拟环境中是完美的，然而现实环境和模型环境必定会存在差异，在真实环境中使用基于模拟环境学习到的策略，其效果往往不能达到预期。AlphaZero是常见的基于模型的算法。

②不基于模型的算法：对于智能体来说，该模型的环境是未知的。智能体直接在环境中探索，通过环境的反馈，更新状态、策略和行为。与model-based RL相比，该模型更容易实现和调整，不会出现模拟环境与真实环境误差的问题，因此被广泛使用。常用的Q-learning算法、Sarsa算法[10]、Policy Gradients策略梯度算法都是不基于模型的算法。

③基于策略的算法：该算法通过对环境的学习，根据策略直接输出要采取的行为。通常所有行为都有一个概率，智能体的决策是根据行为发生概率的大小来选择概率最大对应的行为，因此该算法可以连续选择多个行为。策略梯度就是典型的基于策略的算法。

④基于价值的算法：该算法通过已知的直接奖励与对环境的学习，不断更新每个状态下所有可能行为的价值回报，智能体的决策是通过选取价值回报最高的行为来实现的。由于没每采取一个行为就要更新价值回报，因此该算法无法连续选择多个行为。Q-learning、Sarsa和Deep-Q-Network就是典型的基于价值的算法。

⑤回合更新算法：该算法的策略更新是基于回合的，即从开始到结束，智能体需要走完一个回合才能对策略进行总结和更新。这种算法的效率不高，典型的代表是蒙特卡洛算法和基础的Policy Gradients。

⑥单步更新算法：与回合更新算法不同，智能体每采取一次行为，单步更新算法就更新一次策略，因此这种算法的效率高。目前大部分算法都是单步更新算法，如Q-learning和Sarsa等。

⑦在线学习算法：该算法必须由智能体通过与环境的交互，来采取行动、更新策略。典型的在线学习算法是Sarsa。

⑧离线学习算法：该算法除了通过智能体与环境的交互来选取行为、更新策略外，还可以直接使用他人的经验或者历史经验进行学习。典型的离线学习算法有Q-learning和Deep-Q-Network。

本课题为研究地点影响力下的个体空间移动，需要假设不同的地点影响力分布情况，环境对于智能体来说是未知的，因此本课题应该选用不基于模型的算法。另外，个体在二维网络中并不是完全按照最佳路线移动的，也会随机移动探索环境，因此课题需要选用离线学习的算法。综上，本课题选用Q-learing算法作为强化学习的算法。

### **2.3.3 Q-learning算法介绍**

在强化学习中，Q-learning算法是一种基于价值函数和时间差分（TD）控制的算法，它将学习到的行为价值函数Q直接用来近似最优行为价值函数，而与当前遵循的策略无关[12]。

1. learning的行为价值函数Q是记录在一张行为-价值表Q-table中的，计算机通过智能体与环境的不断交互更新行为-价值表，正向的奖励能增大行为-价值表中的Q值，负向奖励会减小行为-价值表中的Q值。Q-learning算法通过选择指定状态下行为-价值表中所有可能会执行的行为对应的Q值中的最大值，来筛选出最优行为。行为-价值表结构如表2-1所示。第一行代表所有可能的行为a，第一列代表所有可能的状态s。每一行数据记录的都是在同一状态下采取对应行为能获得的Q值。

表2-1 行为-价值表Q-table结构

|  |
| --- |
| 状态\行为 a1 a2 … an |
| s1 Q(s1,a1)Q(s1,a2)Q(s1,…)Q(s1,an)  s2 Q(s2,a1)Q(s2,a2)Q(s2,…)Q(s2,an)  s3 Q(s3,a1)Q(s3,a2)Q(s3,…)Q(s3,an)  …Q(…,a1)Q(…,a2)Q(…,…)Q(…,an)  sm Q(sm,a1)Q(sm,a2)Q(sm,…)Q(sm,an) |

Q-learning算法的核心公式如公式（2-4）所示，其作用是更新行为-价值表中的Q值：当前状态Q值的更新，源于下一个状态中所有可能行为对应的最大Q值。式中 α (0 ≤ α ≤ 1) 是学习率，γ (0 ≤ γ ≤ 1) 是衰减因子。学习率 α 的大小决定学习到的新Q值能替换旧Q值的比例：当α = 0时，智能体什么都没有学习到；当α = 1时，智能体将学习到的新Q值全部替换原有的Q值。衰减因子 γ 的大小决定智能体的“眼界”，即未来行为的奖励对当前行为的影响：当γ = 0时，智能体只注重眼前行为的奖励，而不考虑未来行为的奖励；当γ = 1时，智能体将未来行为与当前行为看得一样重要。Rt是智能体在状态st时获得的直接奖励。



（2-4）

Q-learning算法的具体实现步骤如表2-2所示。学习率 α 、衰减因子 γ 和奖励矩阵R是已知的，在强化学习开始之前，初始化行为-价值表，使得表中的Q值全为零。之后进入循环开始强化学习：每一次循环都给定一个初始状态s和目标状态s0。在整个强化学习过程中，初始状态和目标状态始终都是不变的。在循环中，智能体每一次学习，都在当前状态所有可能发生的行为中选择一个行为a，这个行为的选择，通常由贪婪度 λ 决定，即有 λ 的概率选择最大Q值对应的行为，有(1-λ)的概率随机选择其他行为。通过执行行为a，智能体可以获得下一个状态s’，运用Q-learning的核心公式计算出新的Q值并且更新，同时更新状态，然后进行下一个循环。通过智能体与环境的交互，不断更新Q值直到策略收敛，此时的策略，就是智能体学习到的最佳策略。

表2-2 Q-learning算法步骤

|  |
| --- |
| **算法：Q-learning算法** |
| Step1 给定参数：学习率α、衰减因子γ和奖励矩阵R  Step2 初始化行为-价值Q表，使表中Q值全为0  Step3 For each episode:  3.1 给定一个初始状态s  3.2 若未达到目标状态s0，则执行以下几步:  （1）在当前状态s的所有可能行为中选取一个行为 a（贪婪度λ）  （2）利用选定的行为 a ，得到下一个状态s’  （3）按照公式：Q’(s,a)=R(s,a)+γ∙max{Q(s’,a’)} ,求出Q’(s,a)  （4）更新Q表中的Q值：Q(s,a)=Q(s,a)+α[Q’(s,a)−Q(s,a)]  （5）令 s = s’ ̃ |

### **2.3.4 结合Q-learning算法的地点影响力下的个体空间移动模型**

传统的Q-learning算法在计算最优行为时，初始状态和终止目标状态始终都是不变的，通过多次循环更新行为-价值表中的Q值，能得到从起点到终点的最优路径。然而在本课题中，每一个城市都有一个地点影响力，每个城市的地点影响力都不同。个体因为受到城市地点影响力的影响而进行空间移动，也就意味着每一个城市都有可能是个体移动的目标。因此本课题在对Q-learning算法上有了一些改进：

**①初始状态的选择：**在本课题中，强化学习的状态指的是个体所在城市的坐标。在二维网络中，Q-learning的初始状态s不再是固定不变的城市坐标，而是随机选择的，即城市i被选为起始状态的概率满足公式（2-5），其中L是二维网络的大小，L\*L表示网络中的城市个数。



（2-5）

**②目标状态的选择：**在本课题中，每一次循环的目标状态s0也不再是一个不变的城市坐标，而是根据每个城市的地点影响力大小有概率地选择的。城市i被选为目标终点的概率满足公式（2-6），其中CI代表每个城市的地点影响力大小。



（2-6）

**③奖励矩阵的确定：**传统的Q-learning算法只有在到达目标状态时才会反馈一个正奖励，在遇到陷阱或者是障碍时反馈一个负奖励。然而本课题没有固定的目标状态，且每个城市都对个体有地点影响力的作用，因此智能体每执行一次行为都能获得一个奖励。本课题假设智能体到达一个城市i的奖励与该城市的地点影响力成正比，即Ri=CIi，(0≤i≤L\*L)，其中i是城市编号，R是到达城市的直接奖励值，L是二维网络的大小，L\*L表示网络中的城市个数。

通过改进后的Q-learning算法，可以得到一个行为-价值表Q-table。Q-table的结构如图2-3所示，该图是某地点影响力分布下的行为-价值表。最左边两列，分别是个体所在城市的横坐标和纵坐标，最上面一行是个体能采取的所有行为。

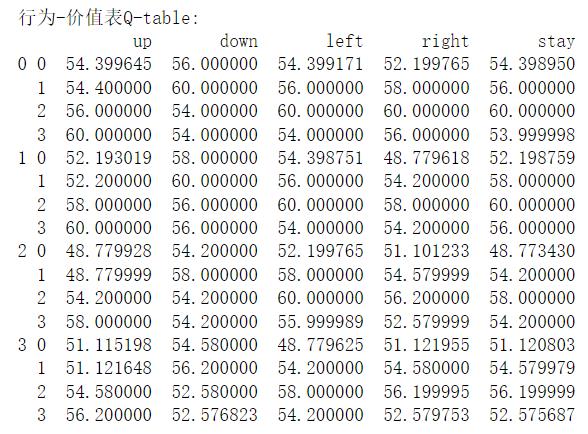


图2-3 行为-价值表Q-table（以L=4为例）

本课题使用Q-learning算法进行强化学习的目的，不是为了计算个体向地点影响力大的城市移动的最优路径，而是为了利用Q-learning算法得到的行为-价值表，计算个体处于不同坐标状态时，采取上、下、左、右和不动这五种行为的概率。在每个状态下，个体按照概概率的大小选择行为，根据选择的行为更新状态，达到在地点影响力下进行空间移动的目的。具体实现流程如图2-4所示。

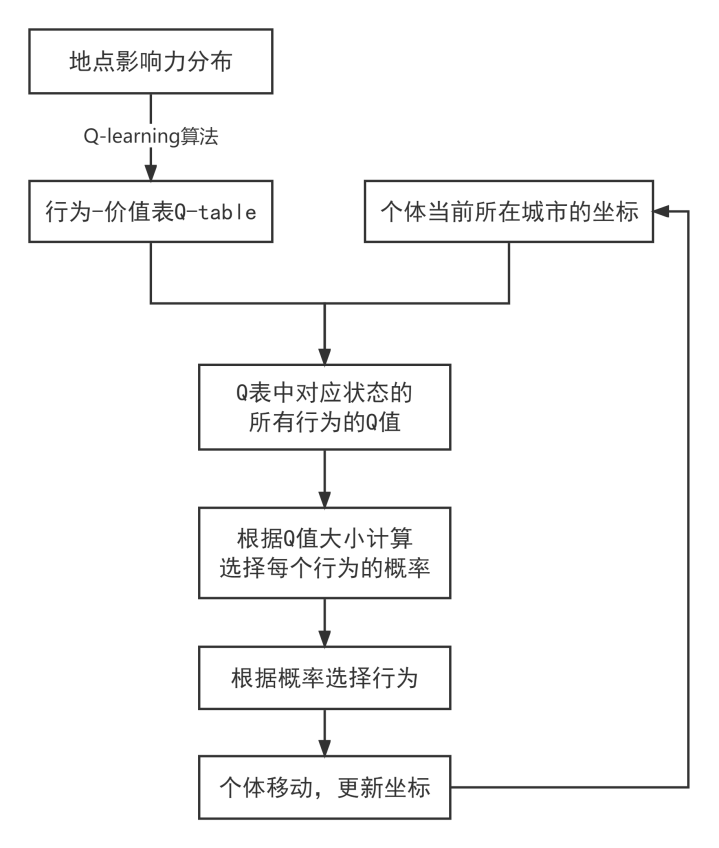


图2-4 结合Q-learning算法的个体空间移动流程图

根据二维网络中城市的地点影响力分布情况，用Q-learning算法进行强化学习，得出一张记录着每一个坐标状态下所有可能动作的Q值的行为-价值表，表中的Q值越大，表明个体越倾向于选择这个Q值对应的行为。当个体要选择行为时，在行为-价值表中寻找个体当前位置对应的状态，并将该状态下所有可能的行为和行为对应的Q值提取出来。运用公式（2-7）计算出每个动作被选择的概率，根据概率选择行为，更新坐标，个体的一次移动就完成了。



（2-7）

2.4 本章小结

与地点影响力有关的因素有很多，本章选取了城市规模大小和医疗资源数量两个因素，并且用数学语言描述了这两个因素与城市地点影响力大小的关系。通过构建二维网络、改进强化学习中的Q-learning算法，使个体在城市地点影响力的作用下进行空间移动，建立移动模型为之后与流行病传播模型相结合做好准备。

# 基于个体移动的集合种群传播模型的建立

本章简单介绍流行病传播模型的分类，详细介绍其中两种与课题密切相关的模型，并将第2章中地点影响力下的个体空间移动与基于SIR模型的集合种群模型相结合，实现课题所需的流行病传播模型。

3.1 流行病传播模型介绍

流行病肆虐以来，人类已经研究了许多数学传播模型。从传播类型来看，常见的流行病传播模型可以分为基于SI、基于SIS、基于SIR、基于SIRS和基于SEIR的模型。SI、SIS、SIR、SIRS和SEIR这五种模型是最基本的流行病传播模型，无论是后来的复杂网络模型，还是集合种群网络模型，所有的模型都是在这五种模型的基础上进行改进的；从传播机理来看。流行病可以分为常微分模型、偏微分模型和网络动力学模型。典型的仓室模型和集合种群模型就是常微分模型，一些常微分方程经过改进后拥有了多变量就成了偏微分模型，而复杂网络动力学模型、无标度网络动力学模型和集合种群网络动力学模型等都是网络动力学模型。

接下来将详细介绍与本课题有关的两种流行病传播模型。

* 3.1.1 SIR传播模型介绍

SIR传播模型最早是由W. Kermack和McKendrick[7]在1927年建立的。该模型是仓室模型的一种，也是经典的动力学传播模型，被人们广泛应用于研究流行病的传播、信息的传播和计算机病毒的传播[16]。SIR模型将人群分为三类（即三个仓室）：

**①易感者（Susceptible）：**用符号S表示，指未感染流行病但是有被感染风险的人。t时刻易感者的总人数用符号S（t）表示。

**②感染者（Infective）：**用符号I表示，指已经感染流行病且有能力使其他人感染的人。t时刻感染者的总人数用符号I（t）表示。

**③免疫者（Recovered）：**亦称“移出者”或“恢复者”。用符号R表示，指已经从感染态恢复健康的人。t时刻免疫者的总人数用符号R（t）表示。通常情况下免疫者对感染过的流行病完全免疫，即R态免疫者不会再成为S态易感者和I态感染者。

假设t时刻种群中有总人数N(t)，则模型满足式子S( t ) + I( t ) + R( t ) = N( t )。SIR模型是均匀混合的单一群体模型，有以下三个特点：

**①群体中总人数不变：**即SIR传播模型忽视了人口的出生、流动和死亡（包括自然死亡、意外死亡和流行病导致的死亡）。流行病是一种拥有很强的传播能力、能在短时间内迅速爆发，并使大量人口感染的疾病。在一定时间范围内，流行病带来的变化，远比自然的人口变化显著得多，因此SIR传播模型将后者产生的变化忽略不计，认为总人数始终是一个不变的常数K，即满足式子（3-1）。

*S( t ) + I( t ) + R( t ) ≡ K* （3-1）

**②一旦成为感染者就会开始传播流行病：**在SIR模型中，感染者I一旦与易感者S接触，就能以一定传染力将流行病传染给易感者。在单位时间内被传染的人数满足式子β S(t)I(t)，其中 β 是流行病的感染率，即在 t 时刻内，一个感染者I将流行病传染给易感者S的人数与当前时刻的易感者S人数成正比。

**③一旦成为感染者就会立即开始痊愈：**在SIR模型中易感者S一旦被感染，就能以一定概率恢复。在单位时间内从感染者I变为免疫者R的人数满足式子 μI(t)，其中 μ 是恢复率，即 t 时刻内，从感染者I变为免疫者R的人数与当前时刻的感染者I成正比。

SIR传播模型三个仓室之间的关系可以用图3-1描述。

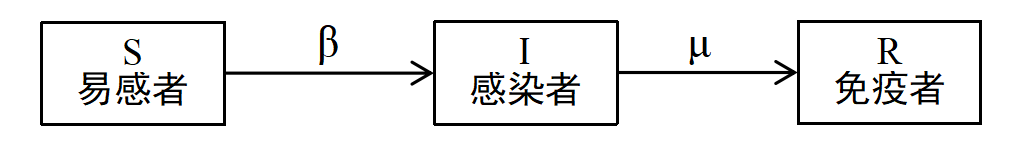


图3-1 SIR传播模型框图

SIR模型是常微分模型，每个仓室的人口变化都满足微分方程组（3-2），其中S是易感者人数 ，I是感染者人数，R是免疫者人数，β为感染率，表示易感者人群与感染者人群接触被感染的概率，μ是恢复率，为平均感染时间的倒数。上述就是SIR模型，可以通过模型推算出对应时间的感染情况。



（3-2）

* 3.1.2 集合种群传播模型介绍

集合种群（metapopulation），亦称“异质种群”或“多斑块”，是1969年由美国生态学家Levins Richard首先提出的[17]。他在研究害虫繁殖时，将集合种群定义为“由常常局部灭绝，却又重新再生的种群集合而成的种群”[18]。随着集合种群理论的发展，出现了许多集合种群数学模型，例如 Levins 模型、斑块模型、元胞自动机模型、联合映射网格模型、概率转移模型和关联函数模型等[19]。

在流行病传播中，集合种群是指一个由多个子种群组合而成的目标种群。集合种群中的每个子种群都是一个传统意义上的种群，可以将其看成一个节点。每个子种群节点包含一定数量的个体，这些个体拥有一些相同的属性，通常也被分为多个类型，例如：S态易感者、I态感染者和R态免疫者。子种群的内部和仓室模型一样，是均匀混合分布的，但是子种群之间既可以是异质的，也可以是同质的。节点与节点之间的连边，可以看成子种群之间连通的桥梁，个体通过这些连边在子种群与子种群之间移动。与单一种群不同，集合种群既要考虑每个子种群内部的流行病传播情况，还要考虑由个体在子种群间移动造成的流行病的扩散。集合种群的具体结构如图3-2所示。

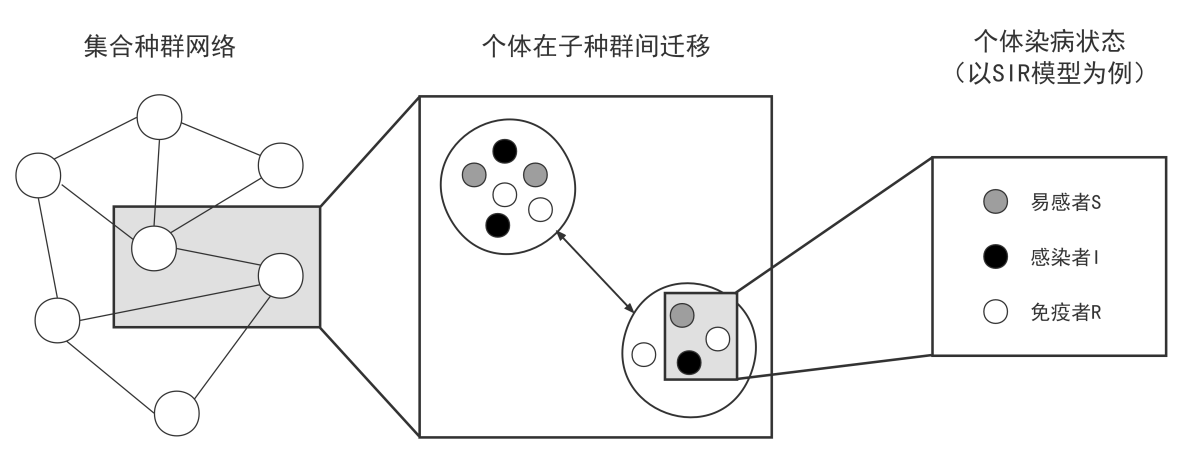


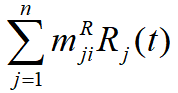
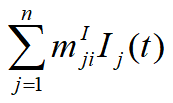
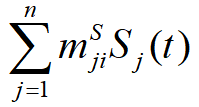
图3-2 集合种群的结构（子种群以SIR模型为例）

集合种群的实现方法主要有两种，一种是交叉耦合集合种群，另一种是移动集合种群[8]。交叉耦合集合种群的特点是不直接显示个体在子种群之间的移动，而是通过一个连接矩阵来具体记录哪个子种群中的个体被哪个子种群中的个体传染的情况。连接矩阵是一个二维数组，通过连接矩阵，子种群之间的连接强度可以清晰地被表示出来。当集合种群中一共有k个子种群时，连接矩阵的大小为k\*k，wij表示在连接矩阵中第i个子种群与第j个子种群之间的连接强度大小（i，j≤k且i，j∈{1,2,3,…}）。单位时间内子种群i和子种群j中的易感者被传染为感染者的概率就是连接强度，wij越大，被传染的概率越大，连接矩阵中的wij就是传统意义上的感染率。通常连接矩阵是对称的，因为两个子网络之间的个体移动是双向的，即集合种群网络是无向图。以SIR模型为基本模型，交叉耦合集合种群的微分方程组如方程组（3-3）所示。其中Si(t)、Ii(t)、Ri(t)和Ni(t)分别表示t时刻编号为i的子种群包含的易感者人数、感染者人数、免疫者人数和总人数，且Ni(t) = Si(t) + Ii(t) + Ri(t)，μ是恢复率。



（3-3）

对于移动集合种群网络，个体在子种群之间的移动直接由移动矩阵显示。移动矩阵记录的是个体从一个子种群移动到另一个子种群的迁移率。流行病的传播仅在子种群内部进行，当个体移动后，子种群内部除了原有的个体外还有新来的个体。与连接矩阵类似，移动矩阵也是一个二维数组，当集合种群中一共有k个子种群时，移动矩阵的大小为k\*k，mij表示在移动矩阵中第i个子种群移动到第j个子种群的迁移率大小（i，j≤k且i，j∈{1,2,3,…}），通常移动矩阵不是对称的，因为在现实生活中个体总是倾向于从一个城市移动到另一个更好的城市。以SIR模型为基本模型，移动集合种群的微分方程组如方程组（3-4）所示。其中Si(t)、Ii(t)、Ri(t)和Ni(t)表示的含义与交叉耦合集合种群方程组（3-3）中的相同，μ表示流行病的恢复率，所有子种群的恢复率相同，βi表示编号为i的子种群的感染率大小。另外，mijT表示单位时间内编号为i的子种群中感染状态为T的个体移动到编号为j的子种群去的迁移率（T∈{S，I，R}）。因此 表示t时刻所有移动到i子种群中去的S态个体的总数， 表示t时刻所有移动到i子种群中的I态个体总数， 表示t时刻所有移动到i子种群的R态个体总数。当子种群数量足够大时，集合种群模型就演变为集合种群网络模型，本文不过多论述。[8]





（3-4）

3.2 结合个体空间移动模型的集合种群模型的建立

本课题结合地点影响力，重点研究个体移动行为对流行病传播的作用，因此对于基础的传播模型，选用的是最基本、最常用的SIR模型。由于在二维网络中存在多个种群，且个体在种群中的移动已经通过第2章中的个体空间移动模型已知，因此选择建立移动集合种群模型。在基于SIR模型的移动集合种群模型的基础上，本课题进行了一些改进：

**①子种群的感染率和恢复率：**本课题建立了一个边长为L的二维网络，网络中存在L\*L个城市，因此整个二维网络是一个集合种群，网络中的每个城市是一个子种群，流行病在每个子种群内部传播，通过个体在子种群间的移动扩散。子种群与子种群之间是异质的，因为每个城市的地点影响力不同。不同的城市规模大小CS和医疗资源数量MR，使得每个城市i的感染率 βi和恢复率 μi 不同。

城市i的感染率满足公式（3-5）。其中 β0是一个常数，表示流行病的基础感染率；CSi是城市i的规模，CSi是自然数；γβ是感染率增长系数。城市的感染率增长量与城市的规模成正比，城市规模每增加1，感染了率 βi 就增加一个感染率增长系数γβ。在整个集合种群中，感染率增长系数是一个定值。



（3-5）

城市i的恢复率满足公式（3-6）。其中μ0是一个常数，表示流行病的基础恢复率；MRi是城市i拥有的医疗资源的数量，MRi是自然数；γμ 是恢复率增长系数。城市的恢复率增长量与城市的医疗资源成正比，医疗资源每增加1，恢复率μi 就增加一个恢复率增长系数γμ。与γβ相同，在整个集合种群中，恢复率增长系数是一个定值。



（3-6）

**②移动集合种群的移动矩阵。**移动矩阵是用来记录一个子种群中所有个体从原先的子种群移动到另一个新的子种群的迁移率。对于传统的移动集合群体，移动矩阵中所有的迁移率大小都是确定已知的。然而在本课题中，个体的空间移动是基于地点影响力的作用的，当二维网络中城市的地点影响力分布情况改变时，个体的移动策略也会相应改变，即迁移率改变，因此使用移动矩阵来描述地点影响力下的个体空间移动会比较麻烦。为了解决这个问题，本课题用上一章实现的个体空间移动模型替代移动矩阵，使个体在子种群间移动，在子种群内感染。

基于个体空间移动模型的集合种群传播模型框架图如图3-3所示。t时刻，二维网络中的所有个体在地点影响力的作用下，通过个体空间移动模型进行一次移动，更新坐标之后，网络中的所有子种群（城市）内部进行流行病的传播：S态易感者按照所在子种群的感染率 β 被感染为I态感染者，感染者又按照所在子种群的恢复率 μ 痊愈成为R态免疫者。当所有子种群内部的个体状态更新之后，t时刻流行病的传播完成，进入下一个时刻的移动和传播。这就是地点影响力下的移动集合种群感染模型。

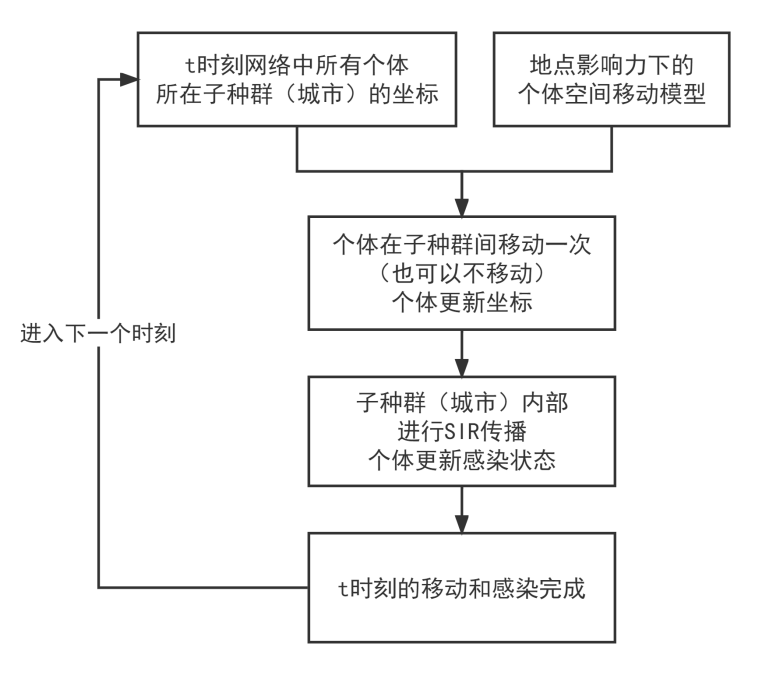


图3-3 地点影响力下的移动集合种群感染模型框架图

3.3 本章小结

SIR传播模型作为经典的流行病模型，被人们广泛地使用并改进，本章基于SIR模型建立了一种移动集合种群模型。区别于以往的集合种群模型，该模型运用强化学习的算法，使个体在各个子种群地点影响力的作用下进行子种群间的移动和传播流行病。地点影响力下的个体移动更符合人类出行轨迹，能更好的地帮助人类研究流行病的传播特点，找到合适的干预措施。

1. 仿真实验结果与分析

本章主要完成了两个实验：第一个实验是不同增长系数下的流行病传播情况，第二个实验是特定增长系数下的不同地点影响力分布对流行病传播过程的影响。主要从实验具体设置和实验结果两方面介绍实验并分析总结。

4.1 实验一：不同增长系数下的流行病传播情况。

在没有对流行病传播进行人为干涉的情况下，影响流行病传播情况最主要的因素就是流行病的感染率和恢复率。在第三章中已经提到，感染率和恢复率满足式子（3-5）和式子（3-6）：



（3-5）



（3-6））

当网络中所有城市的规模CS和所有城市拥有的医疗资源MR不变时，影响城市感染率β和恢复率μ大小的因素只有感染率增长系数γβ和恢复率增长系数γμ。因此本实验在已知城市规模、医疗资源和基本感染率、基本恢复率的前提下，分为两个部分，分别研究不同感染率增长系数对流行病传播的影响和研究不同恢复率增长系数对流行病传播的影响。本课题将从具体设置与结果分析两个方面分别对这两个部分进行阐述。

* 4.1.1 不同感染率增长系数对流行病传播的影响

1. **实验具体设置**

本实验假设二维网络中一共有16个城市，且网络中的总人数为100人。网络中没有医疗资源MR，即MR=0，此时所有城市的恢复率都满足μ=μ0。所有城市的规模CS都为1，即所有城市的感染率满足β = γβ+β0。基本感染率β0 和基本恢复率μ0是常数，分别等于0.005和0.002。实验模拟八次流行病的传播，初始情况时感染率增长系数 γβ=0，随后的每一次模拟，γβ 都增长0.007。将八次流行病传播的情况迭代60次画在一张图中，通过观察和分析，得出感染率增长系数 γβ 对流行病的感染规模、波峰到达时间和波峰人数的影响。

1. **实验结果与分析**

图4-1是在城市规模、医疗资源、基本感染率、基本恢复率和恢复率增长系数已知且不变的情况下，不同感染率增长系数的感染规模和I态波峰随时间演化的情况。

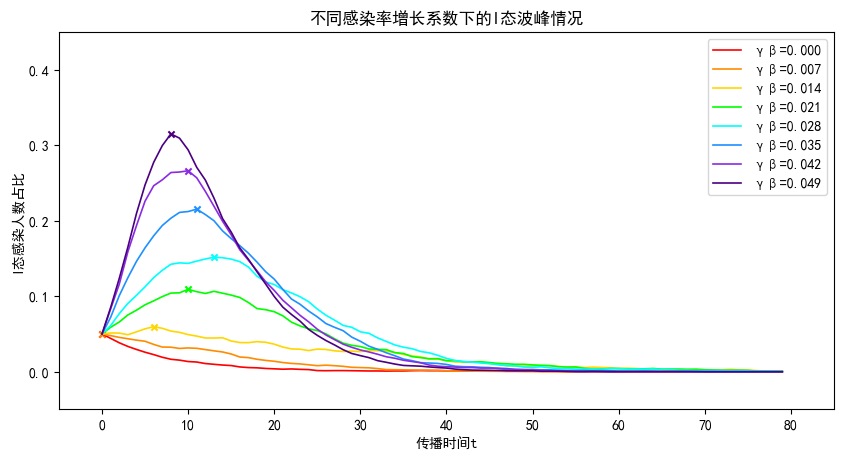
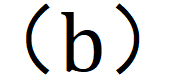
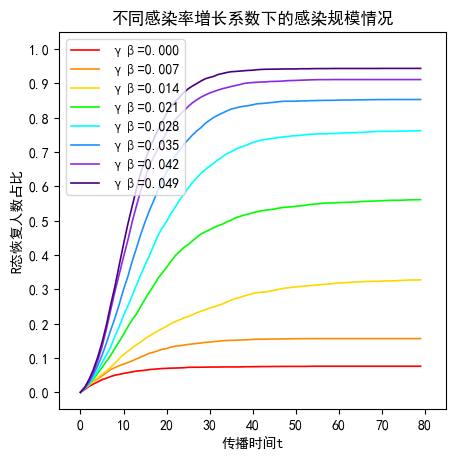


图4-1 不同感染率增长系数下的感染规模和I态波峰随时间演化情况

（a）R态恢复人数占比随时间演化情况；（b）I态感染人数占比随时间演化情况

图4-1（a）的曲线是不同时刻下R态免疫者的人数占比曲线，每一个颜色代表一个感染率增长系数 γβ 的取值，所有 γβ 的值成等差关系。当传播时间t≥60时，R态曲线处于稳态，此时流行病已经停止传播，R态对应的总人数占比就是感染过流行病的总人数占比，即流行病的感染规模。稳态时，通过比较不同感染率增长系数 γβ 下的R态曲线可知：感染规模随着 γβ 的增大而增大。

图4-1（b）的曲线是不同时刻下I态免疫者的人数占比曲线。曲线上标有“X”的代表波峰，I态曲线的波峰在很大程度上能反映流行病的传播情况。对于波峰人数占比，当0≤γβ≤0.007时，I态曲线的波峰到达时间为t=0，波峰人数等于初始的感染者人数，表明当 γβ 在这个区间时，流行病没有传播开来；当γβ＞0.007时，波峰人数随着感染率增长系数 γβ 的增大而增大。因此当恢复率μ=0.002时，γβ在0.007与0.014之间存在一个阈值，使流行病能传播开来。对于波峰到达时间，当0≤γβ≤0.028时，波峰到达时间随着感染率增长系数 γβ 的增大而增加，即流行病的传播得到控制的时间推后；当γβ≥0.028时，波峰到达时间随着感染率增长系数γβ的增大而减少，即流行病得到控制的时间提前。因此当恢复率μ=0.002时，γβ在0.021与0.035之间存在一个阈值，使得波峰到达时间最晚，流行病传播进程最长。

为了进一步观察随着感染率增长系数的增加感染规模和I态波峰的变化趋势，本文引进了增长量的概念：假设γβi（1≤i≤8）是八个感染率增长系数中第i个取值，则满足∆γβi = γβi+1 - γβi = 0.007（1≤i≤7），其中∆γβi是相邻两个 γβ 的增长量，如∆γβ1 = γβ2 - γβ1。另外∆Rmax/Nsum（0≤∆R≤1）和∆Ipeak/Nsum（0≤∆R≤1）分别表示感染规模增长量占总人数的比值和波峰人数增长量占总人数的比值。图4-2是随着感染率增长系数的增加，相邻 γβ 的感染规模占比增长量与波峰人数占比增长量。

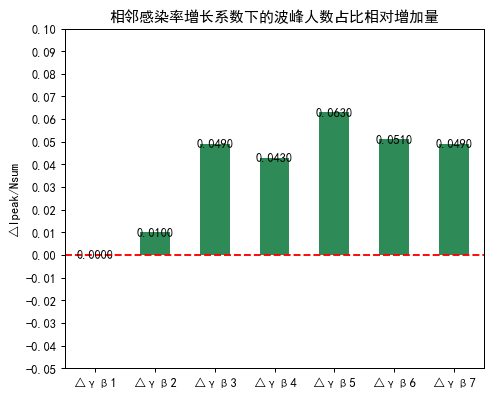
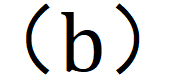
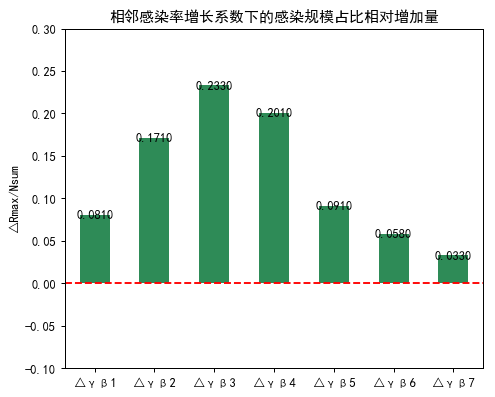


图4-2 相邻γβ的感染规模占比增长量与波峰人数占比增长量随着γβ增加的情况

（a）相邻感染规模占比增长量随γβ增加情况；（b）相邻波峰人数占比增长量随γβ增加情况

图4-2（a）中，纵坐标表示感染规模占比的增长量，横坐标表示相邻感染率增长系数的增长量。从柱状图可知，随着γβ的增加流行病的感染规模始终增大，且规模的扩大有先加速后减缓的趋势。当∆γβi3 = γβ4 - γβ3 = 0.021-0.014时，感染规模增长量最大，即当0.014≤γβ≤0.021时，存在流行病感染规模增长最快的时候。

图4-2（b）中，纵坐标表示波峰人数占比的增长量。已知流行病只有在当γβ＞0.007时才能开始传播，因此∆γβ1时波峰人数为0；当流行病开始传播时，波峰人数始终随着γβ的增加而增加，且波峰人数的增长量先急剧增加后趋于平稳。

* 4.1.2 不同恢复率增长系数对流行病传播的影响

**一、实验具体设置**

本实验假设二维网络中一共有16个城市，网络中总人数为100人。网络中所有城市的规模MR都相同，且为0，即所有城市的感染率满足 β = β0。医疗资源CS在网络中均匀分布，且为1，即所有城市的感染率满足 μ = γμ + μ0。基本感染率β0和基本恢复率μ0是常数，分别为0.1和0.001。实验模拟了8次流行病的传播，在初始情况时恢复率增长系数 γμ=0，随后每一次模拟 γμ 都增长0.0005。将8次流行病传播的情况迭代多次画在一张图中，通过观察和分析，得出恢复率增长系数 γμ 对流行病的感染规模、波峰到达时间和波峰人数的影响。

**二、实验结果与分析**

图4-3是CS、MR、β0、μ0和γβ已知且不变的情况下，不同恢复率增长系数γμ

下的感染规模和I态波峰随时间演化的情况。

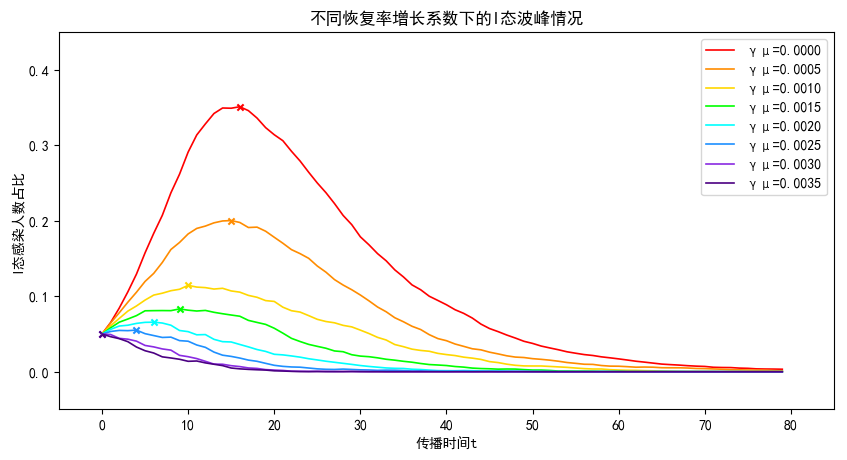
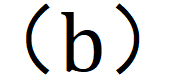
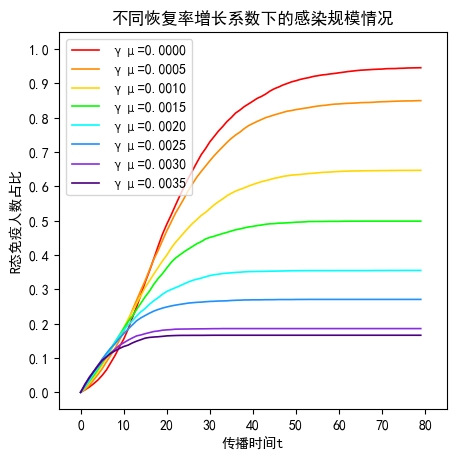


图4-3 不同恢复率增长系数下的感染规模和I态波峰随时间演化情况

（a）R态恢复人数占比随时间演化情况；（b）I态感染人数占比随时间演化情况

图4-3（a）是R态人数占比曲线，每个颜色代表了恢复率增长系数γμ的一个取值，γμ成等差关系，公差0.0005就是γμ的增长量。当t≥70时，所有R态曲线到达稳态，此时感染人数趋于0，免疫人数占比达到最高，流行病的传播结束。通过观察不同恢复率增长系数得：当γμ增大时，感染规模随着γμ的增大而减小。

图4-3（b）是不同恢复率增长系数下的I态感染人数占比变化趋势，通过比较曲线上标有“X”的波峰点，可以反映不同 γμ 取值下流行病的传播情况。对于波峰到达时间，当0≤γμ≤0.0030时，波峰到达时间随着 γμ 的增大逐渐减少，即流行病的传播进程逐渐缩短；当γμ≥0.0030时，I态波峰到达时间为t=0，此时流行病无法传播开来，因此流行病能扩散开的阈值在0.0025和0.0030之间。对于波峰人数，随着 γμ 的增加，流行病的I态波峰人数逐渐减少直到趋于最初的感染人数，即n=5人。

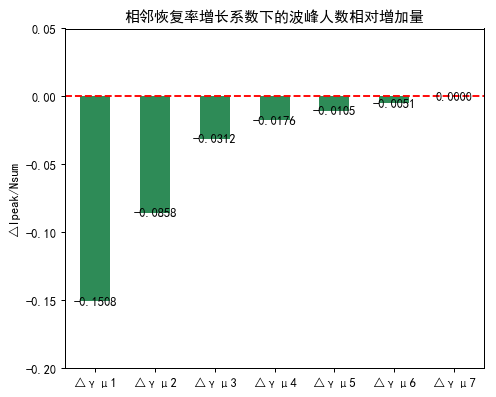
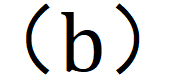
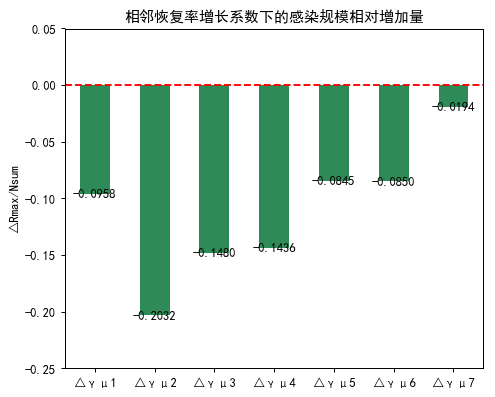


图4-4 相邻γμ的感染规模占比增长量和波峰人数占比增长量随γμ增加的情况

（a）相邻感染规模占比增长量随γμ增加情况；（b）相邻波峰人数占比增长量随γμ增加情况

图4-4是随着恢复率增长系数的增加，相邻 γμ 的感染规模占比增长量与波峰人数占比增长量的变化趋势。与感染率增长系数相同，假设γμi（1≤i≤8）是八个恢复率增长系数中第i个取值，则满足∆γμi = γμi+1 - γμi = 0.0005（1≤i≤7），其中∆γμi是相邻两个 γμ 的增长量，如∆γμ1 = γμ2 - γμ1。

图4-4（a）中，纵坐标是感染规模增长量与总人数的占比，横坐标表示相邻恢复率增长系数的增长量。随着 γμ 的增加流行病的感染规模始终减小，且规模的减小量先增大后减小。当∆γμ2 = γμ3 - γμ2 = 0.0010-0.0005时，感染规模减小量最大，即当0.0010≤γμ≤0.0005时，存在流行病感染规模减小最快的时候。

图4-4（b）的纵坐标是波峰人数增长量与总人数的占比。由柱状图可知，波峰人数随着 γμ 的增加而减少，且波峰人数的减少量逐渐减小直至趋于0。

4.2 实验二：特定增长系数下的流行病传播情况

4.1节研究了感染率增长系数 γβ 与恢复率增长系数 γμ 对流行病传播的影响，本节使 γβ 和 γμ 为常数，研究与地点影响力有关的两个因素：城市规模大小MS和医疗资源数量MR对流行病传播的影响。

* 4.2.1 实验具体设置

本实验假设了四种不同的地点影响力分布情况，通过城市规模CS和医疗资源MR不同的分布情况，研究地点影响力对流行病传播的影响。实验假设二维网络大小为10\*10，即一共有100个城市，网络中的总人数为10000人。另外感染率增长系数γβ=0.01，恢复率增长系数γμ=0.004，基本感染率β0=0.01，基本恢复率μ0=0.005。四种地点影响力分布情况如下：

**Case1：城市规模均匀分布，且网络中没有医疗资源**

此时城市规模CS相等且全为0，医疗资源MR也全为0。由公式（2-1）可知，这是一个地点影响力均匀分布的情况；由公式（2-2）可知，每个城市的人口密度相等。因此个体均匀分布在二维网络中，且在每个时刻都做等概率的随机移动。地点影响力分布情况如图4-5所示：在100个城市中，每个城市的地点影响力均匀分布且都为0，此时所有城市的β=β0，μ=μ0。



图4-5 Case1下的地点影响力分布情况

（L=10，网络中的城市地点影响力均匀分布，没有大城市、没有医疗资源）

**Case2：城市规模不均匀分布，且网络中没有医疗资源**

在100个城市中，有40%的城市是大城市，即40个城市的规模CS=1，60个城市的规模CS=0。网络中没有医疗资源，即MR=0。大城市的分布是随机的，初始撒人时，每个城市的人口密度 ρi 满足公式（2-2），因此大城市的人口密度比其他城市的人口密度大；由公式（2-1）和公式（3-5）可知，大城市的地点影响力也比其他城市的大、感染率比其他城市高。地点影响力分布如图4-6所示：在100个城市中，有40个城市的地点影响力为2，其余60个城市的地点影响力为0，此时城市的地点影响力不均匀分布，个体在城市间做强化学习下的空间移动。



图4-6 Case2下的地点影响力分布情况

（L=10，城市地点影响力不均匀分布，城市规模分布不均匀且没有医疗资源）

**Case3：城市规模不均匀分布，且医疗资源均匀分布**

Case3在Case2的基础上拥有了医疗资源：在网络中每个城市的医疗资源数量相等且MR=1。当一个城市被分配到医疗资源时，城市的地点影响力和恢复率相应增加，满足公式（2-1）和公式（3-6）。而城市规模的分布情况与Case2相同。地点影响力分布如图4-7所示：在100个城市中，有40个城市地点影响力为4，有60个城市地点影响力为2，地点影响力分布不均匀。

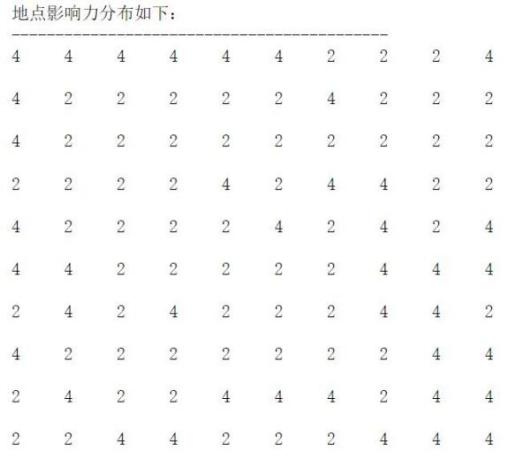


图4-7 Case3下的地点影响力分布情况

（L=10，城市地点影响力不均匀分布，城市规模不均匀分布且医疗资源均匀分布）

**Case4：城市规模不均匀分布，且医疗资源不均匀分布**

Case4的网络大小、网络总人数、城市规模分布、初始人口密度分布、网络中的医疗资源总量与Case3相同，不同的是医疗资源的分布不再是均匀的，而是随机的，即每个城市被分配到的医疗资源MR的数量不同，有的城市可能没有医疗资源，有的城市拥有很多医疗资源。但无论资源如何随机，整个网络的医疗资源总数100始终不变。城市的感染率、恢复率、地点影响力都满足相应的公式，具体地点影响力分布如图4-8所示：在100个城市中有1个城市地点影响力为10，3个城市地点影响力为8，9个城市地点影响力为6，29个城市地点影响力为4，38个城市地点影响力为2，其余城市影响力为0。地点影响力不均匀分布，个体做强化学习下的空间移动。

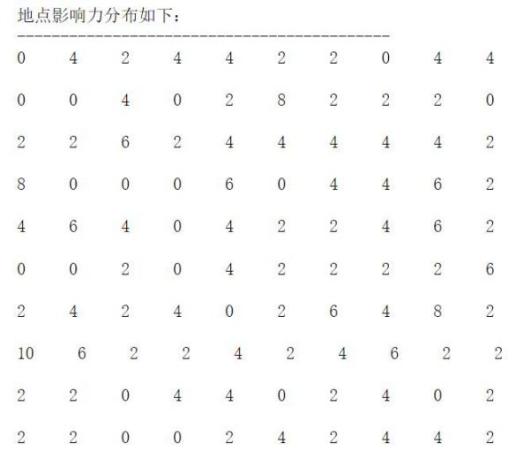


图4-8 Case4下的地点影响力分布情况

（L=10，城市地点影响力不均匀分布，城市规模不均匀分布且医疗资源不均匀分布）

4.2.2 实验结果及分析

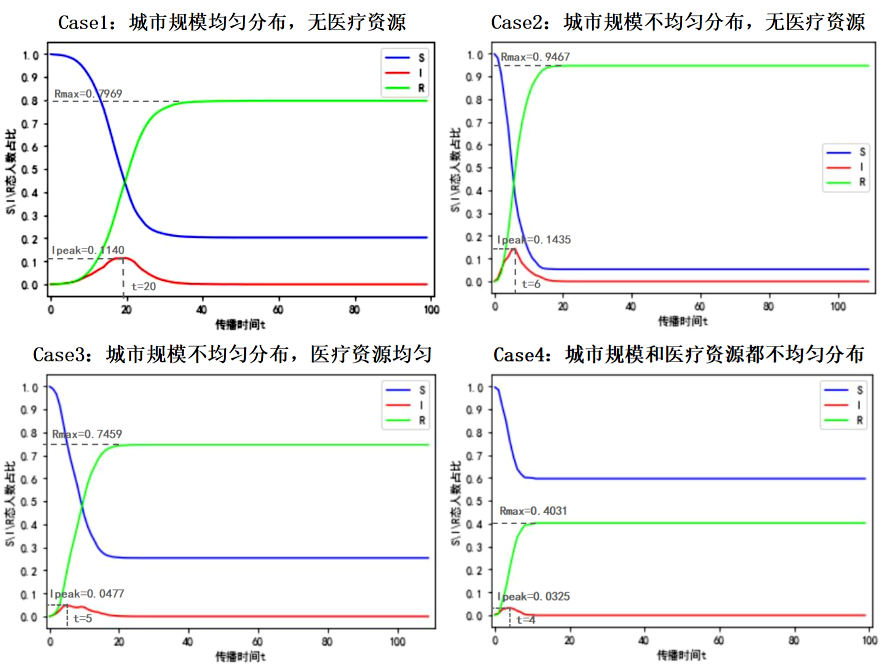


图4-9 四种情况下的流行病传播情况

（四种情况分别是4.2.1中的Case1、Case2、Case3、Case4）

图4-9分别是四种情况下的流行病传播情况，记录着每一个时刻t下易感者S、感染者I、免疫者R人数与总数的占比。通过观察四张图，我们可以了解到流行病传播的大致过程。随着时间的增加，S态易感者的人数首先会急剧减少，然后趋于平缓；R态免疫者的人数先急剧增加，后趋于平缓；I态感染者人数先增加后减少，最终趋于平缓。无论地点影响力怎样分布，在流行病传染过程中总会出现一个I态波峰，当波峰出现时，表明当前时刻是流行病传播最严重的时刻，超过了该时刻，流行病的传播就会得到遏制，并且逐渐趋于稳态。当S、I、R三态都达到稳态时，流行病传播结束，此时I态感染者总数为0，R态免疫者的总数就是此次流行病的感染规模。

感染规模、I态波峰到达时间、I态波峰人数是三个重要的指标，能反映流行病的传播情况。本文将通过比较四种情况下这三种指标的大小来分析地点影响力对流行病传播的影响，具体见图4-9。

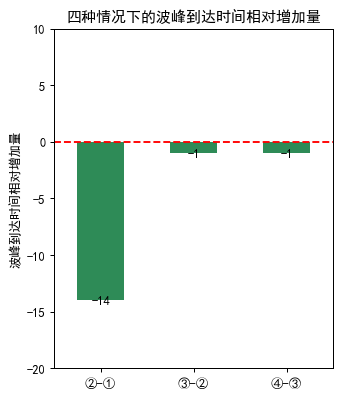
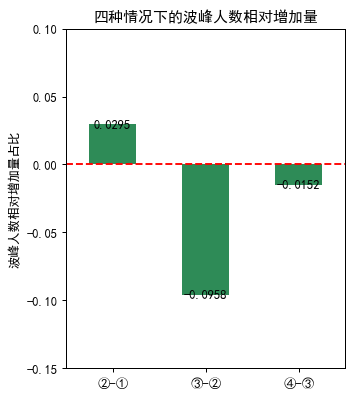
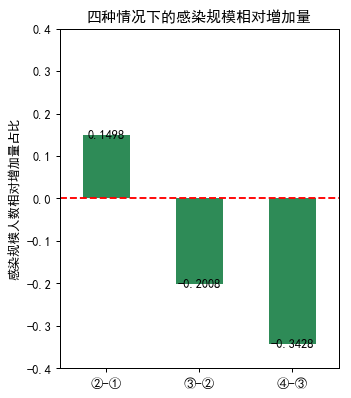


图4-10 四种情况下的感染规模、波峰人数和波峰到达时间增量的比较

（四种情况分别是4.2.1中的Case1、Case2、Case3和Case4）

1. 四种情况下的感染规模相对增加量；（b）四种情况下的波峰人数相对增加量；
2. 四种情况下的波峰到达时间相对增加量

图4-9分别是四个Case下感染规模、I态波峰人数和I态波峰到达时间的相对增量。横坐标表示Case2、Case3、Case4分别与前一种情况比较，即城市规模不均匀分布与城市规模均匀分布相比，城市有医疗资源与城市无医疗资源相比，城市医疗资源不均匀分布与城市医疗资源均匀分布相比。纵坐标表示每两种情况的增量与网络中的总人数之比，比值大于0时，表示该指标增加；比值小于0时，表示该指标减少。

通过比较这三个指标可知：城市规模分布的均匀性与医疗资源的分布确实对流行病传播有很大影响；与城市规模均匀分布相比，城市规模分布不均匀时，感染规模变大，I态波峰人数增加，波峰到达时间提前，整个流行病传播进程缩短；医疗资源的分配对减小感染规模和减少I态波峰人数有很大帮助。其中不均匀的医疗资源分配比均匀的医疗资源分配更能有效地控制感染规模和I态波峰人数。

4.3 本章小结

本章通过两个实验，研究了影响感染率和恢复率的四个变量，即分别研究感染率、恢复率的增长系数和城市规模、城市医疗资源的分布对流行病传播的影响。其中，γβ的增长对扩大感染规模、增加波峰人数有很大影响，而 γμ 的增长能减小感染规模、减少波峰人数，并且缩短流行病传播进程。另外，城市规模的均匀分布和医疗资源的分配能有效地控制流行病传播，其中医疗资源的不均匀分配对控制病毒传播的能力更强。因此如何分配医疗资源，将资源的效益最大化是未来值得研究的方向。

1. 仿真实验的可视化

本章主要实现个体空间移动与流行病传播过程两个方面的可视化，介绍了可视化的实现方法，并且展示了部分可视化效果。

5.1 个体空间移动轨迹的可视化

* 5.1.1 实现方法

个体空间移动的可视化，就是将个体的移动轨迹通过直观的方式展现出来。本文将个体空间移动可视化分为两个步骤：二维网络可视化与个体移动可视化。

1. **二维网络的可视化**

已知二维网络的边长为L，大小为L \* L，则网络中一共有L \* L个城市。每个城市都有一个编号i（0≤i≤L\*L-1），则城市对应的横坐标为Xi=i//L，纵坐标为Yi=i%L。将每个城市按照坐标排列，给城市边长CityLength赋值（这里CityLength=50）就能得到图5-1的二维网络（以L=4为例）。图中每个网格代表了一个城市，一共有16个城市。城市编号与城市坐标的对照如图5-2所示。

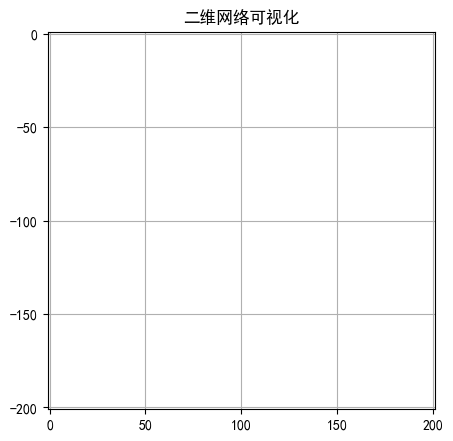
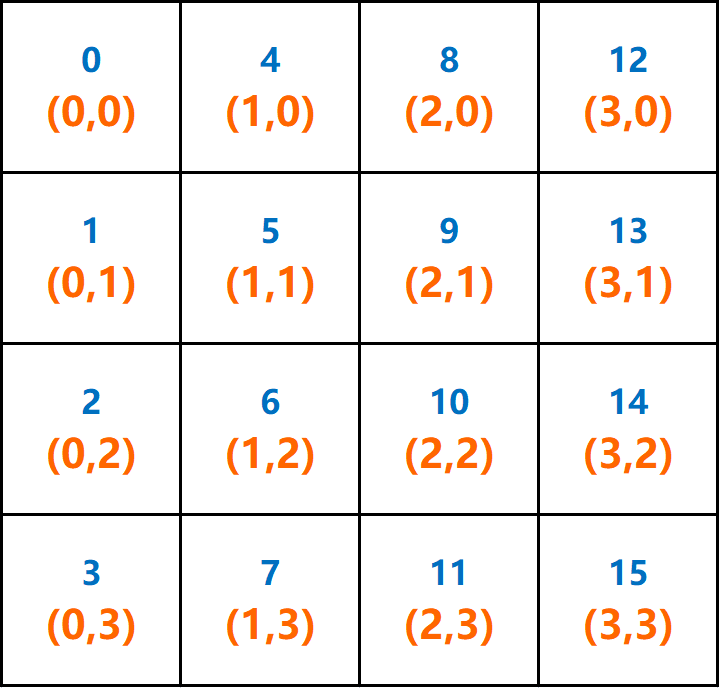
 

图5-1 二维网络可视化（L=4） 图5-2 城市编号与城市坐标对照图（L=4）

**二、个体移动可视化**

对于个体移动轨迹的可视化，本文是用折线图实现的。将每个时刻t下指定个体所在的城市坐标存入两个数组中，数组的长度都为个体移动的总次数t，其中X[ ]数组存的是个体所在城市的横坐标，Y[ ]数组存的是个体所在城市的纵坐标。将数组X\*CityLength作为横坐标，数组Y\*CityLength作为纵坐标画折线图，就能得到个体在城市之间的移动轨迹，但是当个体移动次数较多，或者需要画多个个体的移动时，轨迹就会因为重合而看不清。为了便于观察，在画图时每次都随机产生两个不大于城市边长CityLength的增量Xplus和Yplus，使得个体在每个城市中的位置随机，这样更符合SIR传播模型中种群内部均匀混合的假设。将X[ ]\*CityLength+Xplus作为折线图中各点的横坐标，Y[ ]\*CityLength+Yplus作为折线图中各点的纵坐标就能实现个体空间移动的可视化。

* 5.1.2 可视化展示

本文以L=4为例，随机生成了一个二维网络，网络中各城市的地点影响力分布如图5-3所示：在16个城市中有6个城市地点影响力为4，1其余10个城市地点影响力为2。指定一个个体，使之在网络中进行强化学习下的移动，初始坐标为（0,0）。



图5-3 个体空间移动可视化的地点影响力分布情况（L=4）

图5-4模拟了六次个体移动100步的轨迹。在地点影响力的作用下，个体每次空间移动的轨迹都不相同：轨迹a、b、c集中在一个地点影响力较大的城市；轨迹d、e在多个地点影响力较大的城市间移动；轨迹f在多个城市之间移动。但无论轨迹如何，个体总是倾向于向地点影响力高的城市靠近。

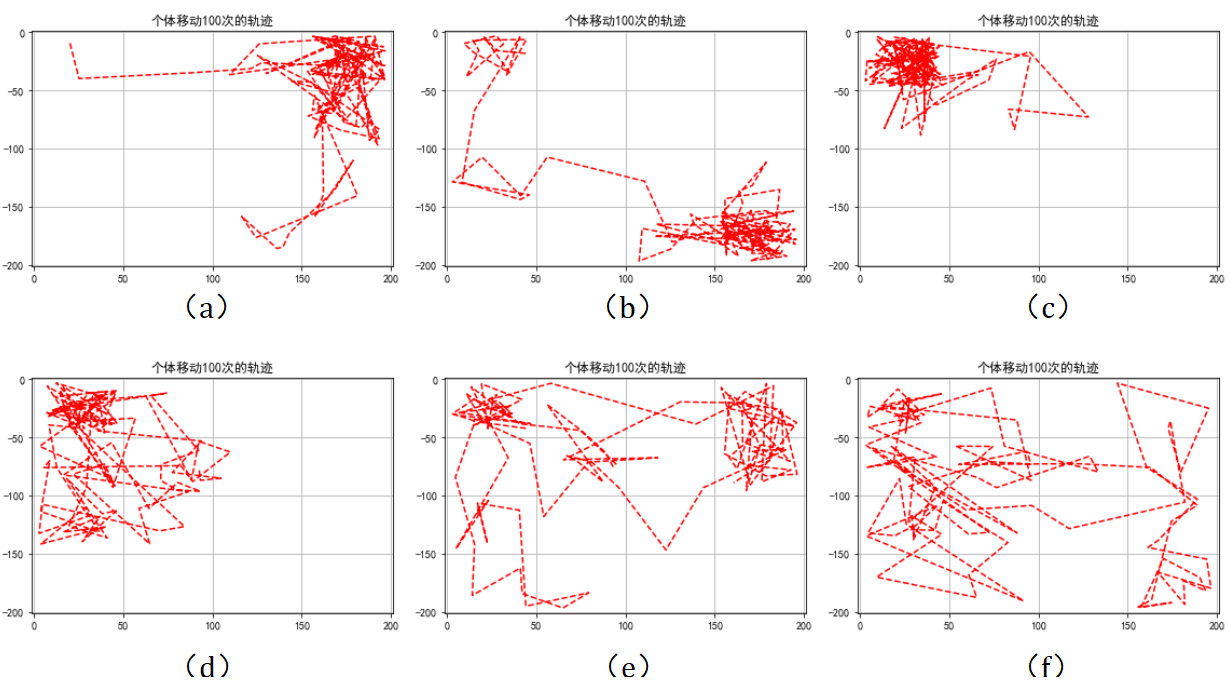


图5-4 个体空间移动轨迹的可视化例子

实现个体空间移动轨迹的可视化，能够直观形象地展现人类出行轨迹，帮助人类更好地理解地点影响力对个体空间移动的影响。

5.2 流行病传播过程可视化

* 5.2.1 实现方法

对于流行病传播过程的可视化，本文是用动态图实现的。在每个时刻t，都用一张散点图展现网络中所有个体的位置分布和个体的感染状态，用另一张SIR曲线图展现各个状态人数的占比，两张子图画在同一个画布上作为一张图。t个时刻的情况逐一显示，就形成了一个动态图。

散点图的画法与5.1节中折线图的画法基本一致，只是把图的类型从折线改为散点。散点图在折线图的基础上加上了感染状态的可视化，即用三个颜色的点分别表示三种感染状态：蓝色的点表示一个易感者S，红色的点表示一个感染者I，绿色的点表示一个免疫者R。

SIR曲线图的画法比较简单，本文不过多阐述。

* 5.2.2 可视化展示

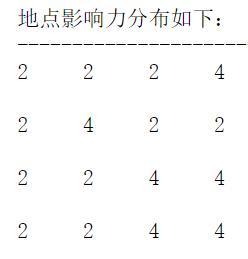
 

图5-5 传播过程可视化的地点影响力分布 图5-6 城市规模和医疗资源分布图

对于传播过程可视化，本文采用了4.2节中的Case3，即城市规模不均匀分布，医疗资源均匀分布。以L=4为例，图5-5为地点影响力的分布情况，图5-6是城市规模和医疗资源分布情况：在16个城市中，有6个城市的规模比较大，另外，每个城市都有医疗资源且数量相等。本文从动态图中截取了三个时刻。

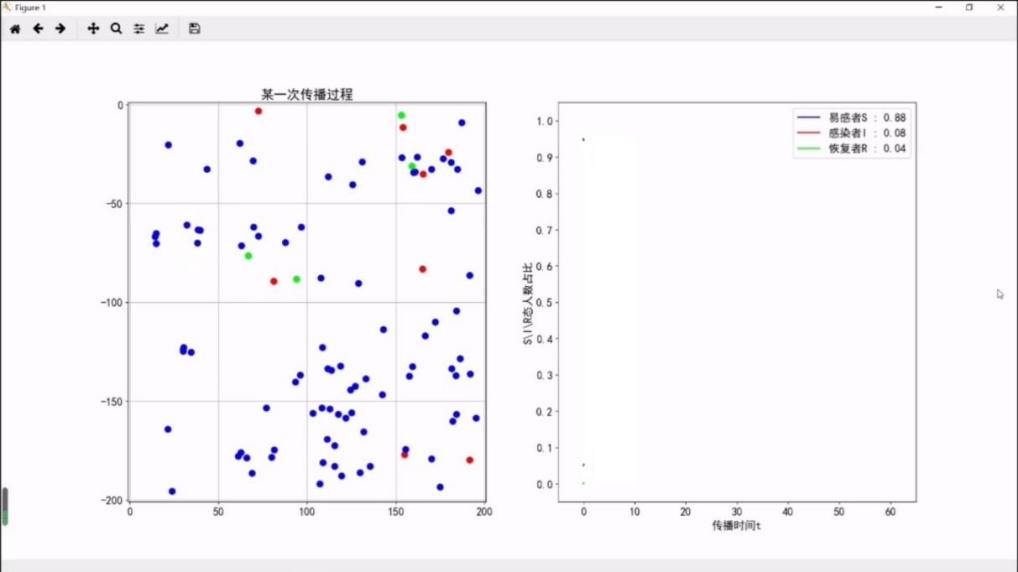


图5-7 传播过程可视化（t=0）

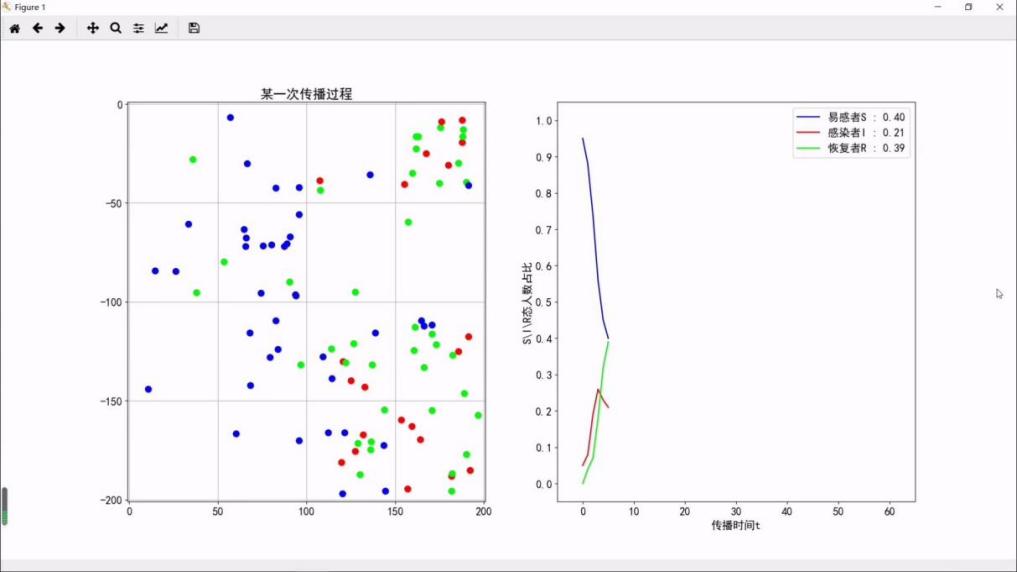


图5-8 传播过程可视化（t=5）

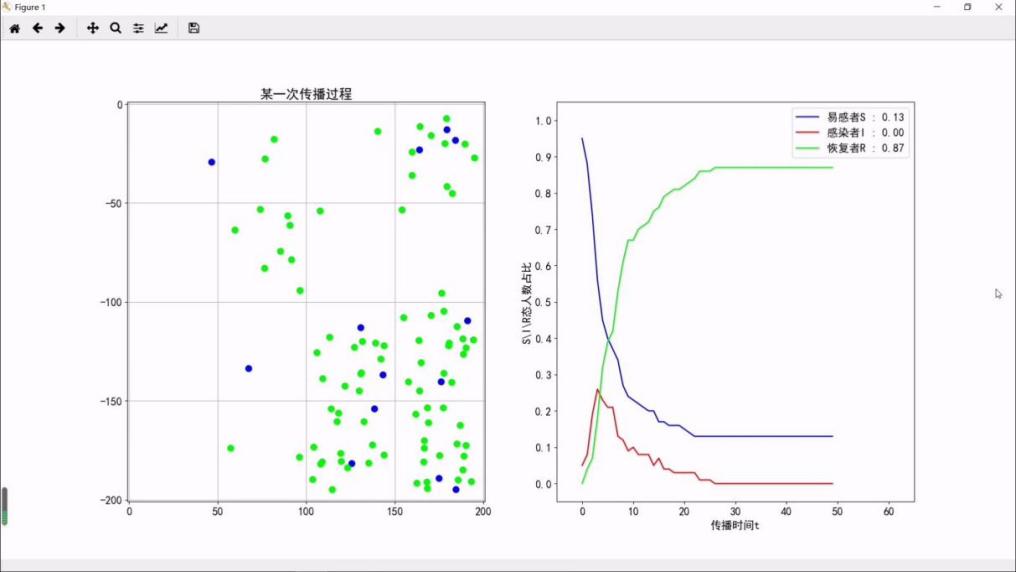


图5-9 传播过程可视化（t=49）

图5-7是流行病刚开始传播时的情况，在二维网络中一共有8个感染者和4个免疫者。城市的初始人口密度是由城市规模决定的，城市规模越大，人口密度越大，因此在散点图中，编号为5、10、11、12、14、15的城市人口密度明显比其他城市的大。由SIR曲线图可知，此时易感者人数占总人数的0.88，感染者占总人数的0.08，免疫者占总人数的0.04。

图5-8是流行病传播了一段时间后的情况，在16个城市中已有21个感染者和39个免疫者，且免疫者总人数逐渐增加，易感者总人数逐渐减少。个体都趋向于向地点影响力大的城市移动。

图5-9是流行病结束传播后的情况，此时在网络中已经没有感染者的存在，免疫者的总数就是此次流行病的感染规模。在此次流行病传播中，总人数的0.87曾经被感染，剩下的13%的人幸免于流行病的传播。

传播过程的可视化，为人类提供了一个直接观察流行病传播的机会，有助于人类理解地点影响力下的个体空间移动对流行病传播的影响。

5.3 本章小结

本章利用折线图和散点图生动、直观并且形象地展现了个体的空间移动和流行病的传播过程，有助于验证模型建立的正确性和理解人类移动与流行病传播的关系。

1. 总结与展望

本章总结了全文的主要内容，提出了课题的创新点，并且指明了未来可以进一步研究的方向。

6.1 本文总结

本文建立了一种运用强化学习的方法使个体在地点影响力作用下进行种群间的空间移动来传播流行病的模型。运用建立好的移动集合种群模型做了两个实验，研究了城市规模和医疗资源对流行病传播的影响，并且通过可视化方法，直观地展现个体空间移动和流行病传播的过程。

* 6.1.2 本文主要工作

本文研究地点影响力下个体空间移动对流行病传播的影响，主要工作有以下几点：

（1）地点影响力定义：定义了与地点影响力有关的两个因素，即城市规模大小和医疗资源数量，并且建立了这两个因素与地点影响力的数学公式。

（2）个体空间移动实现：使用Q-learning算法，强化学习地点影响力的分布情况，使个体在构建的二维网络中进行种群间的空间移动。

（3）感染模型建立：将强化学习、个体移动和基于SIR模型的移动集合种群模型相结合，建立了一个新的流行病传播模型。

（4）地点影响力与流行病传播研究：研究四种不同的地点影响力分布情况，分析城市规模与医疗资源的分配对流行病传播的影响。

（5）可视化实现：用动态图与折线图直观形象地展现病毒传播的整个过程和个体在地点影响力的作用下在城市间的移动行为。

* 6.1.3 本文主要创新点

在流行病传播模型上，本文主要有以下创新：

1. 在传统的个体空间移动流行病传播模型的基础上，引入地点影响力的概念，使得模型中的个体移动更符合实际生活中人类的移动轨迹。
2. 使用强化学习的算法，规划个体移动的路径。

6.2 展望

本文实现了基于地点影响力的个体空间移动与流行病传播模型的构造，并且成功将其可视化，但仍存在一些需要改进的地方：

1. 仿真模型的规模有待扩大。实验一中二维网络的大小是4\*4，总人数为100人，而实验二中的网络大小为10\*10，总人数为10000人，实验一的规模明显小于实验二的规模。小规模的仿真实验只能得出定性研究的结论，如果想要进行定量研究，就必须扩大规模。然而要将实验一模型扩大到10\*10，总人数为10000的大规模，使用目前的算法需要花费极大的时间成本。因此要扩大规模就必须先优化代码的实现和模型的算法。
2. 决定地点影响力大小的因素有待增加。本文定义了两个控制地点影响力大小的因素，然而在现实生活中，除了城市规模和医疗资源数量，经济活跃度、教育水平、气候环境、消费水平、就业水平等等其他因素都能影响一个城市的地点影响力。
3. 仿真模拟的数据有待改为真实数据。在本文实现的模型中，地点影响力、城市规模以及医疗资源的分布都是随机的，通过仿真实验可以研究地点影响力对个体空间移动与传播行为的影响。但是如果想要进一步预测传播趋势、研究干预方法，就需要使用真实的数据进行模拟。

**致谢**

做毕设的时间仿佛在弹指一挥间溜走了。从一月选题时的迷茫，到二月有想法后的激动；从三月修改BUG时的烦躁到四月看到成果时的喜悦；从五月优化代码、写论文的焦虑，到如今毕设渐进尾声、大学生涯即将结束的感慨与不舍……这五个月我过得很充实也很精彩，学习了很多从没接触过的知识，也获得了新的技能。毕设能顺利地完成，除了有我自己的努力外，还离不开一群关心、爱护、督促我的人。

首先，我要感谢我的导师王冰老师。感谢她一丝不苟的工作态度，与循循善诱的指导思路。如果不是她每周督促我完成任务，我的进展不会如此顺利；如果不是她要求我以答辩的形式交流进展，我写论文的思路不会如此清晰；因为有她每周都对我的肯定，我在写代码期间才能以平和的心态面对各种问题，才能有动力克服难题；因为有她将看起来复杂困难的课题拆分，一步一步循序渐进地提高要求，我才能在不知不觉中完成了毕设。王老师是我的灯塔，指引着我前进的方向，帮助我稳稳地迈出每一个脚印。

其次，我要感谢我的同学。感谢他们常常与我交流课题进展、倾听我的思路与困惑；感谢他们时不时地为我推荐相关资料，为我积攒爆发灵感的能量；感谢他们对我的关心与鼓励，让我意识到我不是一个人在奋斗。

再次，我要感谢我的家人。疫情期间居家两月，足不出户又粮食短缺。每个人住在不大的房子里，远离自然、远离伙伴，居家工作也没有止境。但我的父母没有在我面前展现出负面的情绪，每天变着花样安排一日三餐，保持平和的心态，为我提供良好的写代码环境。

最后，我要感谢两个学院与学校。感谢中欧学院与计算机学院的合作，让我有幸能够遇见这么好的课题、这么好的导师；感谢学校四年来对我的栽培。

毕业既是人生一个阶段的结束，也是新的阶段的开始。我知道前路必定困难重重，但我已做好准备，无畏无惧，勇往直前！

**参考文献**

1. 邰鑫玉. 公元2-4世纪罗马帝国瘟疫研究[D].上海师范大学,2021.DOI:10.27312/d.cnki.gshsu.2021.001151.
2. 张乐.黑死病：欧洲中世纪之殇[J].中国经济评论,2020(01):82-85.
3. 吕亚茹,王磊,常晓未.冷面杀手“天花”消亡史[J].中国医学人文,2022,8(04):66-68.
4. Barbanel Dove. No Remedy to Cholera in Haiti: How the UN Undermines Human Rights[J]. Journal of Human Rights Practice, 2022, 13(2) : 386-406.
5. Leon Piterman,Marika Vicziany,林楚玲,黄文静,杨辉.大流行的盘点：西班牙流感大流行和新型冠状病毒肺炎大流行的对比[J].中国全科医学,2021,24(13):1592-1596.
6. 金倩莹,李星明.流行病学方法在医学研究中的应用概述[J].北京医学,2020,42(05):444-451.DOI:10.15932/j.0253-9713.2020.05.023.
7. Kermack W,McKendrick A G. Contributions to the mathematical theory of epidemics-I [J]. Bulletin of Mathematical Biology,1991,53:33-55
8. 冯姗姗. 集合种群网络上传染病动力学建模分析[D].中北大学,2020.DOI:10.27470/d.cnki.ghbgc.2020.001049.
9. 于晨光. 基于人类移动性的集合种群传染病动力学模型研究[D].南京邮电大学,2020.DOI:10.27251/d.cnki.gnjdc.2020.001068.
10. 白云飞. 基于强化学习的AGV动态路径规划研究[D].四川大学,2021.DOI:10.27342/d.cnki.gscdu.2021.000279.
11. 梅俊,顾和平.基于自适应动态规划的非线性系统最优采样控制[J].中南民族大学学报(自然科学版),2022,41(02):251-256.
12. 刘思嘉. 基于强化学习的城市交通路径规划问题研究[D].烟台大学,2021.DOI:10.27437/d.cnki.gytdu.2021.000307.
13. 杨倩.中国人口迁移的空间格局变化及其影响因素研究[J].中小企业管理与科技(上旬刊),2020(06):125-127.d
14. 赵迪,陈鹏,李海成,苗红斌.2005—2018年北京市外来人口迁移特征与影响因素分析[J].地球信息科学学报,2022,24(04):698-710.
15. 王政忠. 基于强化学习的路径规划算法研究[D].桂林理工大学,2021.DOI:10.27050/d.cnki.gglgc.2021.000373.
16. 余小军,王昌晶,屈文建,左正康,罗海梅.COVID-19疫情传播建模分析[J].江西师范大学学报(自然科学版),2021,45(06):559-565.DOI:10.16357/j.cnki.issn1000-5862.2021.06.02.
17. Levins Richard. Some Demographic and Genetic Consequences of Environmental Heterogeneity for Biological Control[J]. Bulletin of the Entomological Society of America, 1969, 15(3) : 237-240.
18. 许冬焱.论复合种群的理论与模型在保护生物学中的应用[J].肇庆学院学报,2003(05):48-51.
19. 宋卫信,张锋,刘荣堂.集合种群理论研究的数学模型(专论)[J].甘肃农业大学学报,2009,44(03):133-139.
20. 李玉喜,强化学习应用场景,https://zhuanlan.zhihu.com/p/78191585,2019.
21. billlee,强化学习入门知识超全梳理,https://zhuanlan.zhihu.com/p/255111887,2020.

**附录：部分源程序清单**

*#定义个体结构，包括每个人的城市位置坐标和一个状态  
#S易感者；I感染者；R免疫者***class** Agent:  
 **def** \_\_init\_\_(self,a,b,status):  
 self.position = [a,b] *#城市位置坐标* self.status = status *#感染状态  
#定义城市结构，包括他的坐标、感染率和恢复率、影响力***class** City:  
 **def** \_\_init\_\_(self,a,b,infection,recovery,attract):  
 self.position = [a,b] *#城市坐标* self.infection = infection *#感染率* self.recovery = recovery *#恢复率* self.attract = attract *#影响力*

*#初始化撒人（在L\*L的网络中撒NUM个人，随机挑选n个设为感染者）  
#NUM、n、L已赋值  
#random.randint( , )是闭区间***def** initialize(NUM,n,L):  
 P\_population = [] *#存放每个城市的撒人概率（与城市规模成正比）* res = [] *#存放所有个体的信息* sum = 0   
 **for** i **in** range(L\*L): *#归一化，使所有撒人概率之和为1* P\_population.append((CITY[i].attract+1)/(L\*L+BigCity\*attractUP))  
 sum = sum+P\_population[i]  
 **for** i **in** range(L\*L): *#计算每个城市撒人概率* P\_population[i] = P\_population[i]/sum  
 **for** i **in** range(NUM): *#撒人* people\_num = np.random.choice(np.arange(0,L\*L),p=P\_population) *#按概率随机选择个体所在城市的编号* X\_people = people\_num//L *#个体所在城市的横坐标* Y\_people = people\_num%L *#个体所在城市的纵坐标* res.append(Agent(X\_people,Y\_people,**'S'**))  
 **for** i **in** range(0,n,1): *#将n个人状态改为感染态I* res[i].status = **'I'  
 return** res  
agents\_before = initialize(NUM,n,L)

*#agents\_before作为原始数据保存，用作画动态图。保存的是所有个体的城市坐标和状态*agents = copy.deepcopy(agents\_before) *#深拷贝 agents保存的是要处理的数据*

*#函数：建立一个行为价值表，Q表***def** build\_q\_table(world\_r,world\_c,actions):  
 k = 0  
 I = np.zeros([world\_r \* world\_c,2],int)  
 **for** i **in** range(world\_r):  
 **for** j **in** range(world\_c):  
 I[k,0] = i  
 I[k,1] = j  
 k+=1  
 I = np.transpose(I).tolist() *#把矩阵转置一次* table = pd.DataFrame(np.zeros((world\_r \* world\_c,len(actions))),index=I ,

columns=actions) *#table是一个二维表* **return** table

*#计算每个城市被选为终点的概率*Attract\_all = 0 *#所有城市地点影响力之和，用于计算每个城市被选为终点的概率***for** i **in** range(0,L\*L,1):  
 Attract\_all = Attract\_all + CITY[i].attract  
P\_end = [] *#每个城市被选为终点的概率***for** i **in** range(0,L\*L,1):  
 **if** Attract\_all==0: *#等概率选择终点* P\_end.append(1/(L\*L))  
 **else**: *#有概率选择终点* P\_end.append(CITY[i].attract/Attract\_all)

*#函数：选择终点***def** END\_num(): *#根据地点影响力大小，有概率地随机选择Q-Learning的终点* END = np.random.choice(np.arange(0,L\*L),p=P\_end)  
 **return** END

*#函数：选择下一步行为***def** choose\_action(pos\_x,pos\_y, q\_table): *#根据当前所在位置，选择下一步动作* pos\_actions = q\_table.loc[(pos\_x,pos\_y), :]

*#[pos\_x,pos\_y, :]表示pos\_x,pos\_y那一行的所有可执行动作都列出来* **if** (np.random.rand() > EPSILON) **or** (max(pos\_actions) == 0):

*#大于贪婪度或者全0，则随机选择行为* action\_name = np.random.choice(ACTIONS)  
 **else**:

*#小于贪婪度，则选择最大Q值，得到它的索引（可能有多个最大值，随机选其中一个）* action\_name =np.random.choice(pos\_actions[(pos\_actions==max(pos\_actions))].index)  
 **return** action\_name  
*#函数：反馈行为（下一步坐标，奖励）***def** get\_env\_feedback(pos\_x,pos\_y,Action):  
 **if** Action == **'up'**: *#行为：上* **if** pos\_y == 0: *#到顶了，无法继续up* next\_pos\_x = pos\_x  
 next\_pos\_y = pos\_y  
 Reward = CITY[next\_pos\_x\*L+next\_pos\_y].attract

*#奖励值=到达的城市的地点影响力* **else**:  
 next\_pos\_y = pos\_y - 1  
 next\_pos\_x = pos\_x  
 Reward = CITY[next\_pos\_x\*L+next\_pos\_y].attract  
 **elif** Action == **'down'**: *#行为：下* **if** pos\_y == WORLD\_R - 1: *#到底了，无法继续down* next\_pos\_x = pos\_x  
 next\_pos\_y = pos\_y  
 Reward = CITY[next\_pos\_x\*L+next\_pos\_y].attract  
 **else**:  
 next\_pos\_y = pos\_y + 1  
 next\_pos\_x = pos\_x  
 Reward = CITY[next\_pos\_x\*L+next\_pos\_y].attract  
 **elif** Action == **'left'**: *#行为：左* **if** pos\_x == 0: *#到最左了，无法继续left* next\_pos\_x = pos\_x  
 next\_pos\_y = pos\_y  
 Reward = CITY[next\_pos\_x\*L+next\_pos\_y].attract  
 **else**:  
 next\_pos\_x = pos\_x - 1  
 next\_pos\_y = pos\_y  
 Reward = CITY[next\_pos\_x\*L+next\_pos\_y].attract  
 **elif** Action == **'right'**: *#行为：右* **if** pos\_x == WORLD\_C - 1: *#到最右了，无法继续right* next\_pos\_x = pos\_x  
 next\_pos\_y = pos\_y  
 Reward = CITY[next\_pos\_x\*L+next\_pos\_y].attract  
 **else**:  
 next\_pos\_x = pos\_x + 1  
 next\_pos\_y = pos\_y  
 Reward = CITY[next\_pos\_x\*L+next\_pos\_y].attract  
 **else**: *#行为：不动* next\_pos\_x = pos\_x  
 next\_pos\_y = pos\_y  
 Reward = CITY[next\_pos\_x\*L+next\_pos\_y].attract  
 **return** next\_pos\_x, next\_pos\_y, Reward

*#主函数：强化学习 Q-learning 获得Q表***def** RL():  
 q\_table = build\_q\_table(WORLD\_R,WORLD\_C,ACTIONS) *#实例化Q表* **for** episode **in** range(MAX\_EPISODES): *#在总回合数以内，每一回合* pos\_x = random.randint(0,L-1) *#起始位置随机* pos\_y = random.randint(0,L-1)  
 is\_terminated = **False** *#是否结束* END = END\_num() *#随机选择一个终点* end\_pos\_x = END//L *#终点的横纵坐标* end\_pos\_y = END%L  
 **while not** is\_terminated: *#没有结束时，每一步，每次结束返回for循环* Action = choose\_action(pos\_x,pos\_y, q\_table) *#选择动作* next\_pos\_x,next\_pos\_y,Reward = get\_env\_feedback(pos\_x,pos\_y,Action)q\_predict = q\_table.loc[(pos\_x,pos\_y),Action]

*#获取当前位置当前动作的Q值,loc通过标签检索，是估计值* **if** next\_pos\_x != end\_pos\_x **and** next\_pos\_y != end\_pos\_y:

*#如果下一步没有结束* q\_target = Reward + GAMMA \* q\_table.loc[(next\_pos\_x,next\_pos\_y), :].max() *#计算真实Q值* **else**: *#下一步就结束* q\_target = Reward + GAMMA \* q\_table.loc[(next\_pos\_x,next\_pos\_y), :].max()  
 is\_terminated = **True** *#结束，不再进入循环* q\_table.loc[(pos\_x,pos\_y), Action] += ALPHA \* (q\_target - q\_predict)

*#新的Q值=老的Q值+学习率\*目标与实际获得Q值的差值* pos\_x = next\_pos\_x *#更新位置* pos\_y = next\_pos\_y  
**return** q\_table *#更新Q表*

*#函数：城市坐标（a,b）的人的下一步到达的城市坐标（按照强化学习Q表走）*SMALL = q\_table.min().min() *#Q表中的最小Q值***def** RANKWALK(a,b):  
 pos\_actions = q\_table.loc[(a,b), :]

*#[a,b, :]表示a,b那一行的所有可执行动作都列出来* POS\_ACTIONS = pos\_actions.sort\_values(ascending=**False**) *#按降序排列* POS\_ACTIONS\_average = 0**for** i **in** range(0,5,1): *#算指定坐标对应所有动作的平均值* POS\_ACTIONS\_average = POS\_ACTIONS\_average+POS\_ACTIONS[i]  
 POS\_ACTIONS\_average = POS\_ACTIONS\_average/5  
 POS\_ACTIONS\_small = POS\_ACTIONS[4] *#指定坐标对应动作的最小Q值* **for** i **in** range(0,5,1): *#对数据进行处理* POS\_ACTIONS[i] = POS\_ACTIONS[i]-0.985\*POS\_ACTIONS\_small ACTIONS\_ = POS\_ACTIONS.index *#提取所有动作名称* **if** max(POS\_ACTIONS) != 0:  
 ALL = POS\_ACTIONS[0]+POS\_ACTIONS[1]+POS\_ACTIONS[2]+POS\_ACTIONS[3]+POS\_ACTIONS[4]  
 P0 = POS\_ACTIONS[0]/ALL *#选择第一个动作的可能性* P1 = POS\_ACTIONS[1]/ALL *#选择第二个动作的可能性* P2 = POS\_ACTIONS[2]/ALL *#选择第三个动作的可能性* P3 = POS\_ACTIONS[3]/ALL *#选择第四个动作的可能性* P4 = POS\_ACTIONS[4]/ALL *#选择第五个动作的可能性* action\_num = np.random.choice(np.arange(0,5),p=[P0,P1,P2,P3,P4])

*#按照概率随机选择动作* Action = ACTIONS\_[action\_num]  
 **else**: *#Q值全为0时，等概率随机选择行为* Action = np.random.choice(ACTIONS\_)  
 x,y,Reward = get\_env\_feedback(a,b,Action) *#获得行动反馈* **return** [x,y] *#返回新城市的坐标*

*#SIR模型***def** SWEEP(agents,L):  
 *#先所有个体移动一次* **for** i **in** range(NUM):  
 agents[i].position = RANKWALK(agents[i].position[0],agents[i].position[1])  
 *#再每个城市SIR感染* **for** i **in** range(NUM):  
 **for** j **in** range(i+1,NUM,1):  
 **if** agents[i].position==agents[j].position **and** bernoulli(CITY[agents[i].position[0]\*L+agents[i].position[1]].infection):  
 *#在同一个城市，且不是同一个人，且在感染率以内* **if** agents[i].status == **'S' and** agents[j].status == **'I'**:  
 agents[i].status = **'I'  
 elif** agents[i].status == **'I' and** agents[j].status == **'S'**:  
 agents[j].status = **'I'  
 elif** agents[i].position==agents[j].position **and** bernoulli(CITY[agents[i].position[0]\*L+agents[i].position[1]].recovery):  
 *#在同一个城市，且在恢复率以内* **if** agents[i].status == **'I'**:  
 agents[i].status = **'R'  
 if** agents[j].status == **'I'**:  
 agents[j].status = **'R'  
 return** agents

*#函数：可视化（散点图）***def** visualize(agents,L,title):  
 **for** i **in** range(0,NUM,1):  
 **try**:  
 x = agents[i].position[0]\*CitySize+random.uniform(3,CitySize-3)  
 y = -agents[i].position[1]\*CitySize-random.uniform(3,CitySize-3)  
 plt.scatter(x,y,s=70,c=colour(agents[i].status),marker=**'o'**)  
 **except**: *#排除重合的可能* **continue** plt.title(title)  
 plt.xlim(-1, L\*CitySize+1) *#横纵坐标范围* plt.ylim(-L\*CitySize-1,1)  
 plt.xticks(fontsize=15) *#横纵坐标字体大小* plt.yticks(fontsize=15)  
 x\_major\_locator=MultipleLocator(CitySize) *#设置x轴坐标的刻度间隔* y\_major\_locator=MultipleLocator(CitySize) *#设置y轴坐标的刻度间隔* ax=plt.gca()  
 ax.xaxis.set\_major\_locator(x\_major\_locator)  
 ax.yaxis.set\_major\_locator(y\_major\_locator)  
 plt.grid() *#画主刻度网格* plt.show