



北京大學

博士研究生开题报告

题目：介观尺度的城市交互问题

姓 名：修格致

学 号：1801110566

院 系：地球与空间科学学院

专 业：地图学与地理信息系统

研究方向：数字地球与智慧城市

指导教师：刘瑜 教授

2021 年 4 月 6 日

摘要

关键词：布偶, 英短, 莫布利

Abstract

Keywords: World Cup, Lancet, Blues

目录

第一章 选题背景	1
1.1 选题背景	1
第二章 研究现状与问题归纳	3
2.1 研究现状	3
2.1.1 介观尺度的城市内部交互问题	3
2.2 问题归纳	6
2.2.1 自适应空间单元的提取	6
第三章 研究内容	7
3.1 框架	7
3.2 研究内容	8
3.2.1 利用扩散映射提取城市隐藏特征	8
3.2.2 稳定性：衡量城市交互系统对抗冲击的能力	9
3.2.3 从众心理驱动下的介观模式形成问题	12
3.3 预期创新点	12
第四章 论文组织结构与时间进度安排	13
4.1 论文组织结构	13
4.2 时间进度安排	13

第一章 选题背景

1.1 选题背景

世界各国在快速城市化过程中，集中涌现出很多有共性的模式与问题。其中包括空气污染、城市热岛现象、能源消耗问题、社会空间不平等以及很多与可持续发展有关的问题。为了向决策者提供可靠的理论和新范式来缓解这些城市问题，对城市结构与演变规律进行建模成为尤为重要的一环。基于简单规则而可解析的统计物理模型可以在不同的角度给出城市发展机制的阐释和预判。近年来，可供城市研究参考的数据越来越多，建模的置信程度也在逐渐提高，这使得此类简洁有力的物理模型能够提供更好的连接理论和经验结果的工具和概念，从而在城市研究发挥重要作用。因此，构建可定量研究、能解释尽量多的城市现象、并有预测性的城市科学是城市研究中极具潜力的发展方向。

介观尺度的概念由 Van Kampen 于 1981 年所创，指的是介乎于微观和宏观之间的尺度。针对城市科学，介观尺度是指城市内部几公里范围内的空间尺度，这是城市研究和城市规划中最常用的空间尺度。处于介观尺度的地学现象，尽管包含很多更小的研究单元，但其系统尺度小于相干尺度，介观尺度单元的同一样本中的空间单元保持较强的相关性，但各个空间单元的差异极大，城市系统的全局平均性质不再有效的刻画系统中所有微观空间单元的性质。理论角度看，在介观尺度下，城市的人口和社会经济密度保持了足够的异质性（如用地功能分化等）；而对于该尺度的研究问题来讲，个体行为特征的聚合又可以体现出足够的规律性。实际角度看，介观尺度对于政策制定、感知城市空间、场所性等非具体、弱化实际度量的方面又有着很好的反馈效果。这使得基于街区、人口普查单元等介观尺度空间单元的研究体现出重要的意义。

城市是由环境和其中的人组成的。个体行为的总和构成了城市的动态。基于

个体行为模式的挖掘，科学家在人类移动模式、职住平衡分析等方面建树颇多。然而，从数据来讲，由于个体数据采集成本较高、隐私保护等问题，基于个体的研究受数据制约比较大；从研究效果来讲，基于个体行为模式的研究结论整体可复现性不高，研究结论受随机误差影响比较大，同时空间一致性不强，很难得到较为通用的、有意义的模式。在此背景下，选取合适的研究尺度，从城市的内蕴框架下探究城市模式的涌现，就成了一个更好的选择。

已有研究从介观尺度结构出发，已经挖掘出来很多有意义的模式。可以概括为如下三类：

1. 从空间分布（一阶量）来看，介观尺度模型可以反映空间上连续变化的事物的分布。在刻画空气污染 [Mijling and Van Der A, 2012]、城市热岛效应、城市形态演化 [Raimbault, 2018]、人口模式识别方面都是建模的核心。
2. 从空间交互（二阶量）来看，介观尺度对应着聚合尺度的空间交互，在流行病传播、对政策制定等角度有着核心意义。
3. 从介观尺度模式的形成来看，介观尺度观测很好地体现了城市微观-宏观之间的跨尺度特征，在交通拥堵、交互尺度提取等方面方兴未艾。

社会物理（social physics）与地理大数据（big geodata）的结合有助于我们建立统一的框架来理解介观尺度规律。然而，这个框架中已有的方法和工具暂时不足以完全解决城市科学家的观测需要。其中几个主要方面为：

1. 数据空间与真实空间的非线性对应关系。
2. 介观尺度结构空间交互的高耦合性。
3. 统计涨落对城市介观尺度结构稳定性的影响。

第二章 研究现状与问题归纳

2.1 研究现状

2.1.1 介观尺度的城市内部交互问题

城市中的空间交互问题在不同的聚合尺度上通常有着不同的研究结论。以人类移动性研究为例:个体层次的移动性常用列维飞行来近似 [Brockmann et al., 2006, Gonzalez et al., 2008, Mantegna and Stanley, 1994, Metzler et al., 2007], 其中体现了高度的随机性;而在群体层次,人类移动性则在出行目的、社群性质、频率分布等方面体现了很高的规律性 [Peng et al., 2012, Jiang et al., 2017, Dong et al., 2020, Wachowicz and Liu, 2016]。随着交通、手机信令、社交媒体、细粒度普查等元数据采集精度的提高,从随机性的个体行为中提取群体行为并进行研究也有了足够的可行性。

在更高的交互频率下,城市生活提供了更多机会,例如更好的就业机会、更广泛的文化活动以及更多样化的人口。然而,这种交互增益也有一些缺点,如犯罪问题更大,拥挤,低空气质量和较高的生活费用。许多研究表明,健康的社会结构在强调城市生活的积极方面起着关键作用。研究还表明,社会凝聚力可能受到设计决策的影响,如街道布局和交通限制 [Appleyard et al., 1976, Appleyard, 1980]。我们可以区分那些被精确定义和中央管制的**硬属性**,例如物质基础设施、交通规划和分区;和那些由居民的社会交往,例如社会凝聚力、城市精神和氛围构成的**软属性**。后者对于城市的居民幸福感、应对灾害能力、可持续发展、城市吸引力等方面都发挥着重要作用。因此,在发展基础设施建设的同时,考虑优化空间结构,提升交互性质,实现城市软属性的提升是一个有价值的目标。

通过数学模型,很多学者阐释了城市的软属性如何有序组织排列城市的硬

属性，这些组织性质可以帮助决策者理解城市的韧性和健康程度 [Batty, 1971, Louf and Barthélemy, 2013]。经济学家使用大量的多变量模型来预测扰动对城市系统的影响，比如土地利用与交通的相互作用、产业结构调整对城市单-多中心模式转换的影响等 [Fujita and Ogawa, 1982, Acheampong and Silva, 2015]。这些预测可能非常准确，但依赖于对大量参数的拟合。而城市经济学家和经济物理学家更倾向于使用简单的规则和变量，基于城市居民的个体选择的群体效应来复现城市现象。这些模型体现了比较强的解释能力，但缺点是无法完全重现城市的多样性，以及难以证伪。因此简单规则导出的模型预测通常不适用于为具体的干预措施或政策提供建议。

另一个思路是从物理学借鉴方法论，考虑一小组经过仔细选择的变量建立动态模型，然后根据对少数变量的分析制定定性或定量预测。这种方法的优点在于结果不是从黑箱中产生的，而是通过理解过程的动态而获得的。对软城市特性的出现进行建模的一个核心挑战是识别一组能够捕捉这些特性的变量。通过将城市的文化和经济方面进行量化，对城市建模面临着无数的可能性：可以选取很多变量，如家庭收入、家庭规模、就业状况等，这些变量之间都有很强的相互依赖性。因此，挑战在于，在不产生不必要的冗余的情况下，选择一个理想的小变量来表示一个区域的状态。

海量城市数据中找到关键变量

随着近年来机器学习技术的普及与发展，以及社会感知等认知框架的建立。地理学界在地理数据挖掘领域有了广泛的应用。’城市数据挖掘’描述了一种方法论，以揭示一组地理空间数据内部的模式和规律的逻辑或数学和部分复杂的描述 [Behnisch and Ultsch, 2009, Bendimerad et al., 2016]。基本的线性回归之外，诸如涌现自组织映射（Emergent Self Organizing Maps）、局部线性插值（Local Linear Embedding）等方法的引入逐渐使人意识到，在城市数据的参考系下，城市并不存在一个全局适定的度量。多维时空数据不均等地反映在不同区位优势与劣势。进而使得以下成了很重要的问题：如何选择合适的统计指标来反映城市某个侧面的性质？

第一个尝试是手动选择关键指标。[Bosetti et al., 2020] 中对流行因素和人口统计学指标进行了回归分析，证明了免疫力、人口分布、和人口流动的异质性的影响下，麻疹爆发的概率会显著被低估。

第二种尝试是利用机器学习方法进行降维。其中代表性的方法为主成分分析。[Deng et al., 2008] 利用多元卫星数据, 结合了 PCA 方法进行了用地功能分类研究; [Nagendra and Khare, 2003] 用 PCA 方法分析了高耦合性的交通、排放和气象数据, 结果显示, 24 小时平均的四个主成分对交通和排放变量的载荷最高, 且它们之间有很强的相关性。1 小时和 8 小时数据的 PC 载荷表明它们之间的变化最小; [Palmason et al., 2005] 分析了城市形态特征, 得到了比较好的抽象表述。此类方法的优点是可以简单而直观地找到关键的数据法向量。但其有一个不容忽视的缺点: 其无法很好地反馈城市中非全局、非线性的地理耦合关系。

空间交互视角下城市系统韧性与稳定性

高速增长的城市规模使得城市核心结构的人口和经济规模的密度到达了一个非常高的水平。如何在这种非均衡的系统中找到结构稳定性就成为了此背景下最重要的议题。一些城市系统的研究显示, 城市现有交互模式被证明与贫民窟的形成有着很高的关系 [Brelsford et al., 2018]。亦有研究表明, 长期发展下, 旧城市将不可挽回地被新城市取代 [Fujita and Mori, 1997, Cottineau et al., 2017]。

城市内部组分之间的流量构成是异常复杂的, 随着不同的观测尺度有着非常大的异质性 [Masucci et al., 2013]。

个体交互作用下介观尺度模式的涌现

学界通常认为, 城市空间结构是自组织的: 即城市作为一个混沌系统, 会自发形成平衡的耗散结构。在社会理论中, 尼古拉斯·卢曼引入了“自我指涉” Self-referentiality 的概念作为自组织理论的社会学应用 [Luhmann, 1984]。对于卢曼而言, 社会系统的“交流”是自我复制的, 即交流产生进一步的交流。因此, 只要存在进化的交流, 社会系统就可以自我复制。对于卢曼而言, 人类就是系统环境中的传感器。这与我们熟知的社会感知框架也不谋而合。卢曼通过功能分析和系统理论发展出一套关于社会及其子系统的进化论。自组织现象意味着特定特征尺度的现象涌现。[Friesen et al., 2018, Pelz et al., 2019] 验证了全球主要大城市的贫民窟规模都是类似的, 并会形成图灵不稳定性的图案; [Louf and Barthélemy, 2013, Courtat et al., 2011] 等工作通过交通与机会模式的平衡解释了城市多中心现象的出现; [Rosvall et al., 2005] 根据人在路网中导航所需要的信息量证明, 城市网络并不是为了沟通而优化的, 而局部地理环境比网络拓扑有着更重要的意义。这

些结论告诉我们，寻找驱动介观地理模式形成的原因有助于我们对城市组分交互的“属性”信息得到更好的先验知识。

“回音室效应” [Wang et al., 2020, Liu et al., 2020] 作为信息传播领域的一个研究热点，其中的研究方法和对人群的理解亦可以应用到理解介观模式形成问题中。由于人的从众心理和地理相关性的存在，地理空间单元容易产生极化和同质的意见集群，通常被称为回音室。回音室效应可能阻止人接触与其现有信仰相反的信息或观点，并进一步激化具有极端信仰的个人。在高度信息化的社会中，回音室效应在各种时空尺度加速出现 [Wang et al., 2020]。已有研究中，社会影响模型 [Noah et al., 2006, Friedkin and Johnsen, 2011, Parsegov et al., 2016]，阿克塞尔罗德的文化传播模型 [Axelrod, 1997]，有界信心模型 [Deffuant et al., 2000] 等理论工作确立了解释微观尺度偏差积累到宏观尺度的基础。这些结论与人类移动性模型的探索返回、偏好重访等机制又有着异曲同工之处 [Song et al., 2010, Gonzalez et al., 2008, Brockmann et al., 2006]。在这些人类移动的行为模式涌现的过程中，如何防止城市动态固化、局部同质化将是规划者保持城市活力的重要考虑方面。

为此，经济地理学家基于同质化人口给出了一些模型来定量解释城市模式固化的条件。

2.2 问题归纳

介观尺度聚合地理大数据和理论模型方法可以挖掘个体行为模式的确定性、城市的空间结构、以及城市作为一个系统的性质，有助于理解城市面对外来冲击、政策变化等因素的反应。不过，自适应空间单元的提取、城市交互系统稳定性指标、交互模式中的涌现现象等问题还欠缺足够好的处理方法。

2.2.1 自适应空间单元的提取

要深入地了解城市的社会结构，虽然现有丰富的数据源提供了很大的便利，但是这些数据本身的复杂性也是一个挑战。以英国的人口普查数据为例：人口普查主要统计数据 and 快速统计报告为每个人口普查产出地区提供 1450 种不同的统计特征。如何找到其中的主要变量，并摒弃选择的偏见，得到一个相对客观的结果，是一个不易解决的问题。

第三章 研究内容

3.1 框架

针对城市中的各种交互问题，找到合适的研究尺度是给出解答的重点。在两个极端的尺度（即个体-微观，与整个城市的特征-宏观）之间，所有的尺度问题都可以用介观尺度的方法论加以解决。

第一个问题是针对问题的空间区域的提取。我们在利用地理大数据进行研究时，采样精度往往与需要探究的问题不完全一致，需要我们聚合成针对问题的合适尺度。比如，对于社会不平等的剥夺指数问题，研究数据主要是普查数据（普查范围是 ~ 8000 人的无重合空间范围），而与之对应的政策调整范围则是更大的（区、县等数万人到数十万人的）区域。我们利用**扩散映射**技术，在多元数据中找到合适的参量，以针对需要的问题得到合适的空间范围。该方法不会从多元空间数据中选择特定的**列**，而是从完整的数据集中构造一个针对问题的指标。由于其结构，扩散映射所提取的索引不会引入源数据偏见之外的偏见，并且对于数据操纵的企图具有很高的弹性。

第二个问题是介观尺度下的信息传播与共识达成问题。随着城市规模的增大，城市中各个区域的沟通也变得更加紧密。根据 [?], 更频繁的微观尺度交互会使得政策贯彻和城市面对外来冲击的弹性降低。为此，我们借鉴 Hubbell 模型 [?], 考虑不同社区中有若干个体，而每个个体针对一件时事（比如，疫情期间政府对于戴口罩的倡议）都有一个意见，在随机初始状态下意见统一的期望时间。我们发现，加入社区/集合种群等介观结构更容易达成意见统一。其他细致结构也对于我们理解介观结构对于形成城市意见共同体的影响。

第三个问题是介观尺度城市稳定性。城市各组分沟通得紧密也伴随着城市面临外在冲击时，冲击在城市内部快速的传导。城市面对流行病、战争等外来冲击

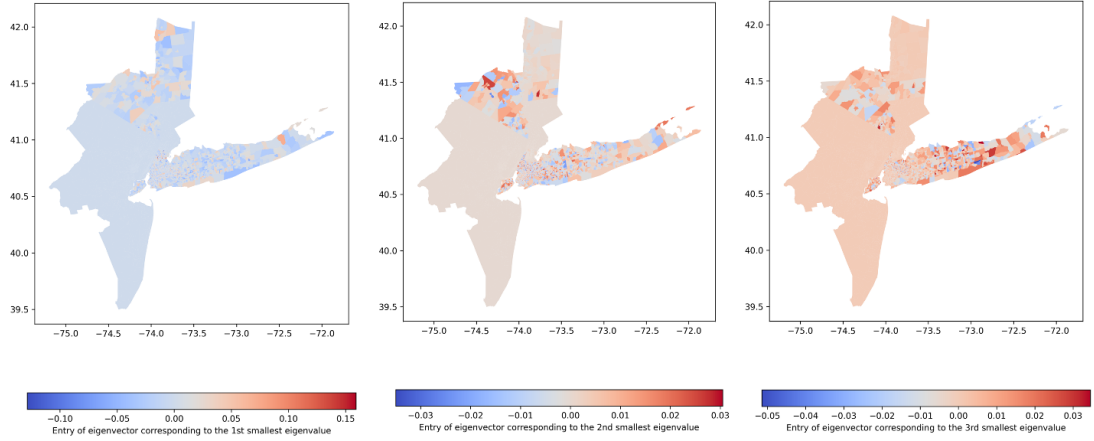


图 3.1: 扩散映射方法下, 人口统计相关矩阵的前三小的正特征值对应的特征向量的空间分布。分别对应着纽约州的教育优势区域、贫困分布、和游客景点。

时的抵抗能力通常被概括为城市韧性/稳定性。城市稳定性也对可持续发展的诸多方面（多样性，连通性，去中心化和自给自足）有解释作用。我们试图基于矩阵理论，利用城市中小区块的交互强度提出一种稳定性度量。在此理论下，可以通过求解矩阵的实部为负的特征值所占全部特征值的比例来衡量外来冲击在城市内部爆发的倾向性。

根据我们的理论和实证结果，介观空间单元的交互视角下，城市的空间连续性、一致性、稳定性等特征的结果都与微观尺度交互的结果不同。这说明城市中介观尺度观测结构不容忽视，是城市建模的重要一环。

3.2 研究内容

3.2.1 利用扩散映射提取城市隐藏特征

从认知、规划等角度来讲，城市空间都不是一个典型的欧氏空间：同样的欧氏距离下，穿不穿越河流、人流密集还是稀疏，都会产生不同的时间消耗和其他感知差异。这个事实说明城市空间的刻画是复杂的，全局的空间度量很难刻画城市空间的本质。为了更好的描述城市空间，我们需要更复杂的城市认知框架。数学上基于局部距离定义的流形提供了一个可能的方案。

Diffusion map 是一种流形学习方法。其出发点在于：很多多元数据的数据

元只能反映客体的一部分特征，只有将多个条目的 survey 合起来才能得到一个比较完整的刻画。提取多个条目的 survey 的“权重”并不是一件简单的事，因为很多时候数据的特征是非线性变化的。这种特征往往很难用一种全局度量来衡量。流形强调的是局部性质：流形 (Manifold) 是局部具有欧式空间性质的空间，包括各种纬度的曲线曲面，例如球体、弯曲的平面等。流形的局部和欧式空间是同构的。流形学习假设所处理的数据点分布在嵌入于外维欧式空间的一个潜在的流形体上，或者说这些数据点可以构成这样一个潜在的流形体。流形是线性子空间的一种非线性推广。以英国的普查数据为例：英国在每个区域统计两类人口统计学指标：关键指标和快速指标。总共有 1450 个特征。[Barter and Gross, 2019]，作者用布里斯托及其周边的 3490 个人口统计学单元作为主要研究对象。其中的分析找到了解释统计反馈的主要变量：大学生密度和贫困程度。

3.2.2 稳定性：衡量城市交互系统对抗冲击的能力

本研究借鉴了生态学的“稳定性”概念 [May, 1972]，拟提出一种衡量城市系统对抗外来冲击的指标，以判别流行病、谣言等外来因素在连通度日益提高的城市系统中是否会扩散开来。该框架的建立可以衡量城市系统的各个子系统的保守性，进而识别出城市中易于受攻击的子区域。该工作可以极大地取代多主体建模的复杂性，而得到同样鲁棒性的结果。

从概念上来讲，稳定性衡量一个系统各组成部分的平衡量由其相互作用形成的敏感程度，一直是大多数网络科学家关注的要点。由于缺乏足够的稳定性，城市流动网络这样的人类系统中经常无法幸免传染病和流言的爆发。这种性质通常是由于城市系统中过多的“节点”处于正反馈的网络路径上，进而难以规避。冲击扩散所留下的痕迹在缺乏稳定性的流动网络中积累。考虑到人类系统自发的稳定性缺失，我们研究：人类系统如何吸收外来的“冲击”？我们使用一个分析框架来量化人类系统自消化的稳定性，并给出了一个达到完全稳定的充分条件。最后，我们将城市内疫情爆发作为人类系统对外部冲击的不稳定诱导反应的案例进行研究，控制组件与自身的相互作用对于稳定疫情传播具有突出作用。因此，局部的保守性保证了人类系统的弹性，对它的监测增强了我们应对外部冲击的能力。

该工作的数学基础如下：介观尺度上，城市内部的人类移动性可以理解为场所 $i = 1, 2, \dots, n$ 之间的交互流，进而可以抽象成一个交互矩阵 W ，其中矩阵 W

的元素 w_{ij} 代表场所 i 到场所 j 的流量。外来的冲击改变系统是由冲击系统的局部开始的，即各个场所受影响的人口 $Y_1 > 0$ ，而 $Y_n = 0, n = 1, 2, \dots, n$ 。以流行病传播为例：某个场所出现了病例，病例进而随着交互流和局部扩散传播。疫情是否会扩散开，取决于患病子人口 $(Y_1, \dots, Y_n)^T$ 在矩阵 W 作用下的稳定性：如果稳定性高，则疫情流行度会维持在 0 附近，即不会爆发；如果稳定性低，则疫情会以很高的概率爆发。鉴于不同的城市特征和流行病等冲击的不同，不同城市的交互流对外部冲击的反应非常异质。然而，从统计角度来看上看，城市系统的交互模式有几个共同的特征：首先，人们经常去人口密集或热门的地方，使得这类地方成为控制冲击的关键位置；其次，停留时间加权下，局部交互占总交互的比例对于不同的城市系统都保持在 10-20% 左右，但是，城市在面对疫情等冲击时，短距离的交互会因为补偿效应而占有更高的比例（如图3.2f）。这些人类流动的特征都会影响社会在面对外部冲击时的稳定性。基于以上的观察，我们建立了一个模型来解释外来冲击是如何随着人类流动而积累的。每个场所受冲击的人口数随时间的变化可以由下述方程确定：

$$\begin{aligned} \frac{dY_n}{dt} &= \alpha X_n (w_{nn} \theta Y_n / N_n + (1 - \delta) \Delta F_n^I) - \beta Y_n, \\ \text{其中 } \Delta F_n^I &= \sum_{m \neq n} (w_{mn} Y_m / N_m - w_{nm} Y_n / N_n) \end{aligned} \quad (3.1)$$

是场所 n 受其他场所冲击的人口数， θ 和 δ 分别代表了外在冲击影响下人类移动性在局部和全局的扩散系数。传播速度和遗忘速度则分别是 α 和 β 。利用一个疫情传播的易感-患病-移除（SIR）场景下的符号系统，我们将受冲击子人口的交互矩阵近似为

$$A_{mn}^t = \begin{cases} \alpha(1 - \delta) [w_{mn} I_m / I_n N_m - w_{nm} / N_n] & m \neq n, \\ \theta \alpha S_n - \beta & m = n. \end{cases} \quad (3.2)$$

最后，我们定义：受冲击人口 Y 的交互矩阵 A 的负特征值的比例为外来冲击下，城市交互系统的稳定性。

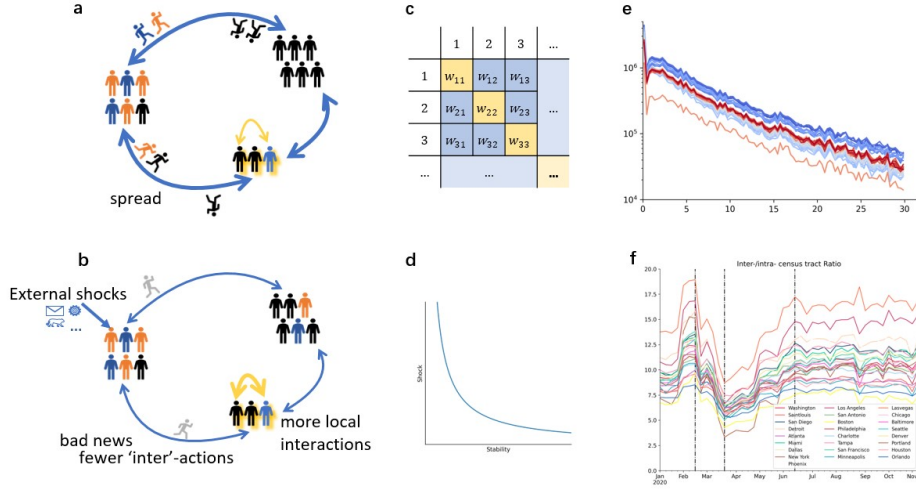


图 3.2: 城市稳定性的框架图。a-b, 人类系统（城市）面对外部冲击时的反应。冲击从一些元人口渗透到城市中，并通过经常性的流动网络和地方互动传播。c, 城市的抽象流动矩阵：在冲击期间，社区内部的访问更为频繁，对应 $\theta > 1$ 。d, 系统的稳定性与其抵抗冲击转为爆发的能力之间的负向关系。e, 华盛顿特区个体流动性的半径分布。最冷的颜色代表 2020 年的第一周，而最暖的颜色代表最新的。考虑到 COVID-19 的冲击，跨社区个体流动性的规模下降，恢复模式不变。f, 2020 年 1 月至 6 月 1 日美国 25 个城市人口普查区间和人口普查区内流动意向的周比值。趋势代表了城市地区个别的运动半径，在 2 月中旬达到顶点，即左边的虚线，直到 3 月下旬才下降，即中间的虚线。这一结果支持了局部相互作用的作用在检疫期增加的说法。

3.2.3 从众心理驱动下的介观模式形成问题

3.3 预期创新点

本文通过引入和构建较多的数学工具，对城市介观尺度下空间模式的提取、交互模式的性质、局部模式的涌现进行研究。本文的几个主要结果改进了地理大数据采集的设计，提出了定量衡量城市政策效果的方法，证明了介观尺度模式涌现的必然性。预期创新点如下：

1. 改进扩散映射方法，引入负相似性的概念，建立更广义的城市组件关联关系，挖掘了更多的诸如对偶结构、收缩城市结构等城市内蕴结构；
2. 建立了城市稳定性的量化框架，以极快的运算效率衡量出城市面临疫情等外来冲击时的抵抗能力，以及各种时空尺度的政策在控制冲击。同时该框架亦可发现冲击扩散的关键节点。
3. 通过考察一类较为通用的复杂网络动态及其

第四章 论文组织结构与时间进度安排

4.1 论文组织结构

第一章为绪论，

4.2 时间进度安排

参考文献

- [Acheampong and Silva, 2015] Acheampong, R. A. and Silva, E. A. (2015). Land use–transport interaction modeling: A review of the literature and future research directions. *Journal of Transport and Land use*, 8(3):11–38.
- [Appleyard, 1980] Appleyard, D. (1980). Livable streets: protected neighborhoods? *The ANNALS of the American Academy of Political and Social Science*, 451(1):106–117.
- [Appleyard et al., 1976] Appleyard, D., Gerson, M. S., and Lintell, M. (1976). *Liveable Urban Streets: Managing Auto Traffic in Neighborhoods: Final Report: Prepared for Federal Highway Administration, Department of Transportation...* US Department of Transportation, Federal Highway Administration.
- [Axelrod, 1997] Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of conflict resolution*, 41(2):203–226.
- [Barter and Gross, 2019] Barter, E. and Gross, T. (2019). Manifold cities: social variables of urban areas in the uk. *Proceedings of the Royal Society A*, 475(2221):20180615.
- [Batty, 1971] Batty, M. (1971). Modelling cities as dynamic systems. *Nature*, 231(5303):425–428.
- [Behnisch and Ultsch, 2009] Behnisch, M. and Ultsch, A. (2009). Urban data-mining: spatiotemporal exploration of multidimensional data. *Building Research & Information*, 37(5-6):520–532.

- [Bendimerad et al., 2016] Bendimerad, A. A., Plantevit, M., and Robardet, C. (2016). Unsupervised exceptional attributed sub-graph mining in urban data. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pages 21–30. IEEE.
- [Bosetti et al., 2020] Bosetti, P., Poletti, P., Stella, M., Lepri, B., Merler, S., and De Domenico, M. (2020). Heterogeneity in social and epidemiological factors determines the risk of measles outbreaks. *Proceedings of the National Academy of Sciences*, 117(48):30118–30125.
- [Brelsford et al., 2018] Brelsford, C., Martin, T., Hand, J., and Bettencourt, L. M. (2018). Toward cities without slums: Topology and the spatial evolution of neighborhoods. *Science advances*, 4(8):eaar4644.
- [Brockmann et al., 2006] Brockmann, D., Hufnagel, L., and Geisel, T. (2006). The scaling laws of human travel. *Nature*, 439(7075):462–465.
- [Cottineau et al., 2017] Cottineau, C., Hatna, E., Arcaute, E., and Batty, M. (2017). Diverse cities or the systematic paradox of urban scaling laws. *Computers, environment and urban systems*, 63:80–94.
- [Courtat et al., 2011] Courtat, T., Gloaguen, C., and Douady, S. (2011). Mathematics and morphogenesis of cities: A geometrical approach. *Physical Review E*, 83(3):036106.
- [Deffuant et al., 2000] Deffuant, G., Neau, D., Amblard, F., and Weisbuch, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(01n04):87–98.
- [Deng et al., 2008] Deng, J., Wang, K., Deng, Y., and Qi, G. (2008). Pca-based land-use change detection and analysis using multitemporal and multisensor satellite data. *International Journal of Remote Sensing*, 29(16):4823–4838.
- [Dong et al., 2020] Dong, L., O’Keeffe, K., Santi, P., Vazifeh, M., Anklesaria, S., Schlöpfer, M., West, G., and Ratti, C. (2020). The spectral dimension of human mobility. *arXiv preprint arXiv:2002.06740*.

- [Friedkin and Johnsen, 2011] Friedkin, N. E. and Johnsen, E. C. (2011). *Social influence network theory: A sociological examination of small group dynamics*, volume 33. Cambridge University Press.
- [Friesen et al., 2018] Friesen, J., Taubenböck, H., Wurm, M., and Pelz, P. F. (2018). The similar size of slums. *Habitat International*, 73:79–88.
- [Fujita and Mori, 1997] Fujita, M. and Mori, T. (1997). Structural stability and evolution of urban systems. *Regional science and urban economics*, 27(4-5):399–442.
- [Fujita and Ogawa, 1982] Fujita, M. and Ogawa, H. (1982). Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional science and urban economics*, 12(2):161–196.
- [Gonzalez et al., 2008] Gonzalez, M. C., Hidalgo, C. A., and Barabasi, A.-L. (2008). Understanding individual human mobility patterns. *nature*, 453(7196):779–782.
- [Jiang et al., 2017] Jiang, H., Li, Q., Zhou, X., Chen, Y., Yi, S., Wang, H., and Lu, Z. (2017). A collective human mobility analysis method based on data usage detail records. *International Journal of Geographical Information Science*, 31(12):2359–2381.
- [Liu et al., 2020] Liu, L., Wang, X., Tang, S., and Zheng, Z. (2020). Modeling confirmation bias and peer pressure in opinion dynamics. *arXiv preprint arXiv:2012.10838*.
- [Louf and Barthelemy, 2013] Louf, R. and Barthelemy, M. (2013). Modeling the polycentric transition of cities. *Physical review letters*, 111(19):198702.
- [Luhmann, 1984] Luhmann, N. (1984). *Soziale systeme: grundriss einer allgemeinen theorie*. Suhrkamp.

- [Mantegna and Stanley, 1994] Mantegna, R. N. and Stanley, H. E. (1994). Stochastic process with ultraslow convergence to a gaussian: the truncated lévy flight. *Physical Review Letters*, 73(22):2946.
- [Masucci et al., 2013] Masucci, A. P., Serras, J., Johansson, A., and Batty, M. (2013). Gravity versus radiation models: On the importance of scale and heterogeneity in commuting flows. *Physical Review E*, 88(2):022812.
- [May, 1972] May, R. M. (1972). Will a large complex system be stable? *Nature*, 238(5364):413–414.
- [Metzler et al., 2007] Metzler, R., Chechkin, A. V., Gonchar, V. Y., and Klafter, J. (2007). Some fundamental aspects of lévy flights. *Chaos, Solitons & Fractals*, 34(1):129–142.
- [Mijling and Van Der A, 2012] Mijling, B. and Van Der A, R. (2012). Using daily satellite observations to estimate emissions of short-lived air pollutants on a mesoscopic scale. *Journal of Geophysical Research: Atmospheres*, 117(D17).
- [Nagendra and Khare, 2003] Nagendra, S. S. and Khare, M. (2003). Principal component analysis of urban traffic characteristics and meteorological data. *Transportation Research Part D: Transport and Environment*, 8(4):285–297.
- [Noah et al., 2006] Noah, F. et al. (2006). *A Structural Theory of Social Influence*. Cambridge University Press, New York.
- [Palmason et al., 2005] Palmason, J. A., Benediktsson, J. A., Sveinsson, J. R., and Chanussot, J. (2005). Classification of hyperspectral data from urban areas using morphological preprocessing and independent component analysis. In *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS'05.*, volume 1, pages 4–pp. IEEE.
- [Parsegov et al., 2016] Parsegov, S. E., Proskurnikov, A. V., Tempo, R., and Friedkin, N. E. (2016). Novel multidimensional models of opinion dynamics in social networks. *IEEE Transactions on Automatic Control*, 62(5):2270–2285.

- [Pelz et al., 2019] Pelz, P. F., Friesen, J., and Hartig, J. (2019). Similar size of slums caused by a turing instability of migration behavior. *Physical Review E*, 99(2):022302.
- [Peng et al., 2012] Peng, C., Jin, X., Wong, K.-C., Shi, M., and Liò, P. (2012). Collective human mobility pattern from taxi trips in urban area. *PloS one*, 7(4):e34487.
- [Raimbault, 2018] Raimbault, J. (2018). Calibration of a density-based model of urban morphogenesis. *PloS one*, 13(9):e0203516.
- [Rosvall et al., 2005] Rosvall, M., Trusina, A., Minnhagen, P., and Sneppen, K. (2005). Networks and cities: An information perspective. *Physical Review Letters*, 94(2):028701.
- [Song et al., 2010] Song, C., Koren, T., Wang, P., and Barabási, A.-L. (2010). Modelling the scaling properties of human mobility. *Nature physics*, 6(10):818–823.
- [Wachowicz and Liu, 2016] Wachowicz, M. and Liu, T. (2016). Finding spatial outliers in collective mobility patterns coupled with social ties. *International Journal of Geographical Information Science*, 30(9):1806–1831.
- [Wang et al., 2020] Wang, X., Sirianni, A. D., Tang, S., Zheng, Z., and Fu, F. (2020). Public discourse and social network echo chambers driven by socio-cognitive biases. *Physical Review X*, 10(4):041042.

附录

扩散映射的数学基础

扩散映射的方法来源于数学中网络上的扩散过程。我们考虑一个复杂网络 $G(V, E)$ ，其上有足够多的人在进行随机游走。每个节点 i 上分配一个随机过程 $x_i(t)$ ，它可以代表 t 时刻恰好处于节点 i 上的人数。每一个人都以同样的概率 $1/k_i$ 访问所处节点 i 的邻居，其中 k_i 是节点 i 的度。于是节点上人数随时间的变化可以写成一个微分方程

$$\dot{x}_i = -k_i x_i + \sum_{j \rightarrow i} x_j = -k_i x_i + \sum_{j \neq i} A_{ij} x_j, \quad (4.1)$$

其中 A 是复杂网络的邻接矩阵。我们试图找到一个矩阵 L ，使得网络上的动态可以用方程

$$\dot{X} = -LX \quad (4.2)$$

来表示，其中 $X(t) = (x_1(t), x_2(t), \dots, x_n(t))$ 。于是容易得到 $L_{ij} = \delta_{ij} k_i - A_{ij}$ 。我们称这个矩阵 L 为 Laplace 矩阵，并记它的特征值为 λ_n ，对应的特征向量为 v_n ， $n = 1, 2, 3, \dots$ 。根据基本的线性代数知识，我们有

$$Lv_n = \lambda_n v_n. \quad (4.3)$$

这一组特征向量可以作为一组基，唯一表示所有的 n 维向量。特别的，可以表示 $X(t) = (x_1(t), x_2(t), \dots, x_n(t))$ 为

$$x_n = \sum_n c_n v_n, \quad (4.4)$$

其中 c_n 为系数。将 (4.3), (4.4) 代入 (4.2), 可得

$$\partial_t \sum c_n(t) v_n = -L \sum c_n(t) v_n \quad (4.5)$$

$$= - \sum c_n(t) \lambda_n v_n. \quad (4.6)$$

根据唯一表示定理, 每个特征向量 v_n 在等式两边的系数也是相等的, 即

$$\dot{c}_n = -c_n \lambda_n, \quad \forall n. \quad (4.7)$$

可以容易地解出每一个 $c_n(t) = c_n(0) e^{-\lambda_n t}$. 进而系统可以写成:

$$x(t) = \sum_n c_n(t) v_n = \sum_n c_n(0) e^{-\lambda_n t} v_n. \quad (4.8)$$

真实的扩散过程转移概率都是非负的, 所以特征值 λ_n 都非负。而行和又都是 0, 所以 Laplace 矩阵 L 都存在一个 0 特征值为 0, 其对应的特征向量为 $(1, 1, \dots, 1)^T$. 该特征向量只保证了矩阵的归一性, 而不包含其他信息。公式 (4.8) 告诉我们, 对于一个 Laplace 动力系统来说, 最小的特征值和其对应的特征向量是最重要的。因为系统沿着该特征向量方向收敛的速度是最慢的, 可以在更长的时间尺度上保留这个方向上的性质。