

空间交互网络研究进展

闫小勇

北京交通大学交通系统科学与工程研究院, 北京 100044

摘要 空间交互网络是人、商品和信息等在地点之间流动而形成的嵌入在空间中的有向流网络。典型的空间交互网络包括国际贸易网络、人口迁移网络、人群出行网络及电话通信网络等。理解和预测空间交互网络中的流量分布模式不仅是区域科学、交通科学、经济地理学等很多领域长期以来的一个重要研究主题,在城市和交通规划、疾病传播防控、商业服务等领域也具有广泛应用价值。本文在简要介绍引力模型、介入机会模型等经典空间交互模型的基础上,着重对近年来复杂系统研究领域在空间交互网络建模方面的研究成果进行介绍,包括辐射模型、人口权重机会模型及空间交互网络上的随机游走模型等,并且对空间交互网络研究中存在的挑战性问题进行探讨,包括个体多样性行为建模、群体空间交互决策行为实验、数据驱动的活动-出行行为研究等。

关键词 空间交互;复杂网络;人类移动模式;出行分布预测

空间交互是指人、商品和信息等在地点之间的流动^[1],这种流动自然形成了嵌入在地理空间中的有向流网络——空间交互网络^[2]。空间交互网络在日常生活中无处不在:身处不同城市的人们通过电话、社交 App 等进行沟通联络,形成了以城市为节点、电话或网络通信量为边权的通信网络^[3];不同国家之间进行商品贸易,形成了以国家为节点、贸易量或物流量为边权的贸易网络^[4]。更直接的例子则是人类自身的移动:在长时间尺度上,人口在城市或国家之间流动,形成了人口迁移网络^[5];而在短时间尺度上,每天进行的出行活动形成了人群出行网络^[6]。显然,研究这些空间交互网络上的流量模式形成机制、预测这些网络上的连边交互强度等都是具有重要价值的问题。这些问题的研究对于制订区域经济和人口发展政策、规划和管理交通、通信等基础设施网络、预防和控制流行病的传播、提供基于位置的商业服务等具有重要的实际意义^[7-8]。

空间交互网络研究中一个基本问题是:在已知各地点的人口(或经济产值、流出总量等反映地点体量差异的指标)和地点之间距离(或移动成本、出行时间等反映地点之间阻隔程度的指标)等数据的前提下,预测地点之间空间交互的强度(即流量)。解决这类问题的模型在区域科学和经济地理学中被称为空间交互模型^[1],而在交通科学中则被称为出行分布预测模型^[9-10]。早在 100 多年前,研究者们就发现在人群出行网络、人口迁移网络、商品贸易网络等空间交互网络中

普遍存在着引力定律^[11-16],即网络节点之间的交互强度近似正比于节点强度乘积与节点间距离幂函数的比值。而类比牛顿万有引力定律提出的空间交互引力模型^[17],也自然地成为解释和预测空间网络交互强度的经典模型,被广泛应用于包括交通规划^[9-10]、人口迁移^[5]、国际贸易^[4,18]、通信网络^[3]甚至在线商品推荐^[19]等在内的诸多领域中。进入 21 世纪后,随着手机、全球定位系统(GPS)、移动互联网等现代信息通信技术的飞速发展,大规模、高精度的人类空间交互数据的可获取程度不断提高,一些数据驱动的空间交互新模型不断涌现,极大地丰富和发展了空间交互网络研究的理论与方法。本文将简要介绍经典的空间交互模型引力模型,并重点对近年来统计物理和复杂系统研究领域在空间交互网络建模方面的一些研究成果进行介绍。

1 空间交互引力模型

1.1 引力模型及其扩展

引力模型是最早被提出的空间交互强度预测模型^[1,17],其基本假设是两地之间(人口、出行、贸易等)的交互量与两地的人口数成正比,与两地间距离的幂函数成反比

$$T_{ij} = \alpha \frac{m_i m_j}{d_{ij}^\beta} \quad (1)$$

式中, T_{ij} 为从起点*i*到终点*j*的流量; m_i 和 m_j 分别为起终点的人口数; d_{ij} 为两者间的距离; α, β 为 2 个可调参数,需要根据实际

收稿日期:2017-05-12;修回日期:2017-06-14

基金项目:国家自然科学基金项目(71671015,61304177)

作者简介:闫小勇,副教授,研究方向为复杂网络和出行行为复杂性,电子邮箱:yanxy@bjtu.edu.cn

引用格式:闫小勇. 空间交互网络研究进展[J]. 科技导报, 2017, 35(14): 15-22; doi: 10.3981/j.issn.1000-7857.2017.14.001

数据进行估计^[20]。

式(1)中原始的引力模型无法保证预测所得到的流量矩阵 T 满足起点流出总量 $O_i = \sum_j T_{ij}$ 的约束以及终点流入总量 $D_j = \sum_i T_{ij}$ 的约束,因此研究者又发展出了双约束引力模型^[10]

$$T_{ij} = A_i B_j O_i D_j f(d_{ij}) \quad (2)$$

式中, $f(d_{ij})$ 为阻抗函数,一般可取距离(或广义出行成本)的幂函数 $d_{ij}^{-\beta}$ 、负指数函数 $e^{-\beta d_{ij}}$ 、幂指复合型函数 $d_{ij}^{\alpha} e^{-\beta d_{ij}}$ 等形式^[9], $A_i = 1 / \sum_j B_j D_j f(d_{ij})$ 和 $B_j = 1 / \sum_i A_i O_i f(d_{ij})$ 为 2 个平衡因子,它们在迭代计算中产生,目的是使得模型预测结果同时满足流出和流入总量约束。

1.2 引力模型的统计物理解释

引力模型具有形式简单、概念易懂的优点,是目前被应用最为广泛的空间交互模型^[1]。但引力模型是类比物理中的万有引力定律提出的,缺乏严格的理论依据^[21]。Wilson^[22]以出行者目的地选择问题为背景,提出了一种从最大熵原理导出引力模型的方法:在给定出行总成本约束 $C = \sum_i \sum_j T_{ij} c_{ij}$ 、出行总量约束 $O_i = \sum_j T_{ij}$ 和 $D_j = \sum_i T_{ij}$ 的条件下,系统出行分布的微观状态数为

$$\Omega = T! / \prod_i T_{ij}! \quad (3)$$

根据统计物理中的最大熵原理^[22],系统最可能出现的出行分布是微观状态数最多的分布(最可几分布)。用拉格朗日乘数法求解式(3)中的约束极值问题,可以得到

$$T_{ij} = A_i B_j O_i D_j e^{-\beta c_{ij}} \quad (4)$$

恰好是个带有指数出行成本函数的双约束引力模型。研究^[23]发现,混合交通方式出行者的出行成本与距离具有近似对数关系 $c_{ij} \approx \eta \ln d_{ij} + \mu d_{ij} + \psi$,代入式(4)中得到 $T_{ij} \approx A_i B_j O_i D_j d_{ij}^{\alpha} e^{-\beta d_{ij}}$,能够覆盖大多数实证中观测到的引力模型距离函数^[17-8]。

Wilson 最大熵原理为引力模型建立了一个统计物理的理论基础。但是,统计物理中的最大熵模型并不是一个机制模型,它仅能给出系统最可能的宏观分布状态,无法对系统中个体的微观决策过程进行描述^[22],同时系统的总成本约束从何而来也并不清楚^[24]。从个体空间交互决策行为的角度,引力模型的低层机制仍未得到满意解答^[21,24]。

2 空间交互行为决策模型

2.1 介入机会模型

与类比物理定律得来的引力模型不同,介入机会模型^[25]完全是从个体对目的地选择的决策过程出发提出模型假设的。其基本假设是:对于每个从地点 i 出发的出行者,先将所有目的地按照距离地点 i 由近及远的顺序排序,并假设出行者会以固定的概率 P 被某个地点的机会所吸引,并停留在那里。如果假设地点 j 的机会数正比于人口数 m_j ,则出行者被地点 j 吸引的概率可假设为 αm_j 。令 q_i^j 表示从 i 点出发的出行

没有被第 j 个地点所吸引的概率,则可以得到关系式

$$q_i^j = q_i^{j-1} (1 - \alpha m_j) \quad (5)$$

它等价于

$$\frac{q_i^j - q_i^{j-1}}{q_i^{j-1}} = -\alpha m_j = -\alpha (S_{ij} - S_{i,j-1}) \quad (6)$$

式中, S_{ij} 为地点 i 到 j 之间(包含 i 和 j 点)的总人口数量。将 q 和 S 视为连续变量,则式(6)可写成微分方程 $\frac{dq}{q(S)} = -\alpha dS$ 。

解此方程可得到

$$q_i(S) = \frac{e^{-\alpha S}}{1 - e^{-\alpha M}} \quad (7)$$

式中, M 为总人口数量。由于从 i 点出发被 j 点吸引的总出行量为

$$T_{ij} = O_i [q_i(S_{i,j-1}) - q_i(S_{ij})] \quad (8)$$

综合式(7)、(8)可得

$$T_{ij} = O_i \frac{e^{-\alpha(S_{ij} - m_j)} - e^{-\alpha S_{ij}}}{1 - e^{-\alpha M}} \quad (9)$$

这即为介入机会模型。

与引力模型相比,介入机会模型的假设中体现了个体选择目的地时的行为决策过程,是一个微观机制类模型。但介入机会模型在形式上比引力模型更为复杂,且容易低估长距离出行的比例^[26],因此在实际中并未获得广泛应用,但介入机会模型中提出了一个非常重要的思想:用排序而不是真实的空间距离来度量各目的地距离起点的远近,这一思想直接启发了后续的大量研究工作。

2.2 辐射模型

在包括引力模型、介入机会模型在内的空间交互经典模型中,普遍存在的一个问题是包含待估参数,必须用历史的空间交互流量数据估计模型参数之后才能实施预测。但通常地点之间详实的流量数据难以获取,这在很大程度上限制了这些经典模型的适用范围。针对这一问题,Simini 等^[27]提出了一个名为辐射模型的空间交互模型,它的基本假设是个体在选择目的地时,会选择距起点最近且比起点收益(收益值按正比于地点人口数随机抽取)更多的一个地点移动。辐射模型的原始推导过程可见文献^[27]的补充材料,本文给出一种更简单直观的推导方式。

前已述及,介入机会模型假设出行者被地点 j 吸引的概率是 αm_j (即正比于 j 的人口数)。如果假设这一概率正比于目的地人口数与起终点之间人口总数的比值 $\frac{m_j}{S_{ij}}$,则可以得到从 i 点出发的出行没有被第 j 个地点所吸引的概率为

$$q_i^j = q_i^{j-1} \left(1 - \frac{m_j}{S_{ij}}\right) = \prod_j \frac{S_{i,j-1}}{S_{ij}} = \frac{m_j}{S_{ij}} \quad (10)$$

进一步可得从 i 点出发的出行被第 j 个地点所吸引的概率为

$$P_{ij} = q_i^{j-1} \frac{m_j}{S_{ij}} = \frac{m_j m_j}{S_{i,j-1} S_{ij}} \quad (11)$$

该模型中出行者的目的地选择过程非常类似物理中的粒子

辐射过程:粒子(出行者)从一个源头(出行起点)向外四散,可以被环境以一定概率吸收(选择了某个终点),而粒子行进的距离(出行距离的长度)则依赖于材料的厚度(人口密度)^[28]。因此Simini等将这一模型命名为辐射模型。

由式(11)可以看到,辐射模型需要的输入数据只有各地的人口数量,而不包含任何可调整的参数。换句话说,该模型仅需要输入人口分布数据就能实施空间交互强度预测,显示出了巨大的应用优势。文献[27]通过一系列实际数据验证了这一模型的预测效果,结果表明辐射模型能够相当准确地预测城市间的通勤出行量、人口迁移量和货物运输量等。

2.3 人口权重机会模型

辐射模型的突出特点是不需要任何可调参数、仅输入人

口分布数据就可以相当准确地预测空间交互网络流量,克服了引力模型等经典模型中存在的缺陷。但是,辐射模型假定个体在选择目的地时只选择距自己最近的高收益地点出行,这使得其预测结果在一些情况下与实际有较大的偏差。图1显示了用北京市出租车GPS数据进行的对比。图1(a)~(c)显示的是从一个位于市中心的地点出发、在不同地点下车的人数分布;图1(d)~(f)显示的是从一个位于南四环的地点出发、在不同地点下车的人数分布;最左列是辐射模型的预测结果,中间一列是实际的出租车乘客移动模式,最右列是人口权重机会(population-weighted opportunity, PWO)模型的预测结果。从图1中可以看出,辐射模型预测的人群移动范围明显小于真实情况。

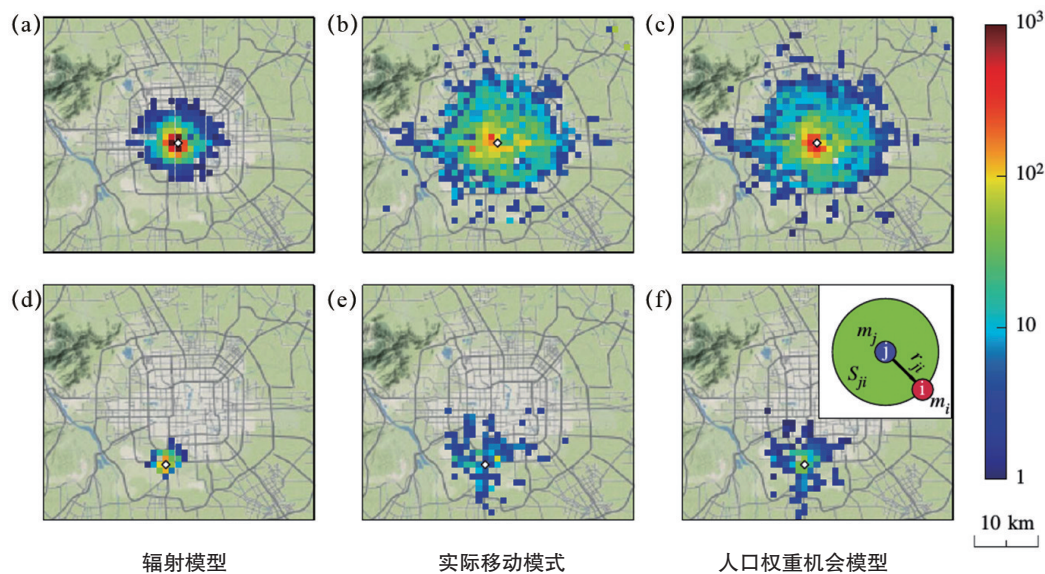


图1 北京市出租车乘客实际移动模式与辐射模型、人口权重机会模型预测结果的对比

Fig. 1 Comparing the observed mobility patterns of Beijing taxi passengers with the mobility patterns predicted

研究认为,出行者在选择目的地时不是仅考虑距离起点最近的高收益目的地,而是会综合衡量备选空间范围内所有潜在目的地的收益(或机会数)。一般而言,对于具有相同机会数的目的地来说,出行者更倾向于选择近的目的地;而对于距离相近的目的地来说,机会数多的地点对出行者的吸引力更大。换句话说,一个地点被出行者所感受的机会数是该地点的实际机会数按某种方式衰减后剩余的部分。引力模型中用一个阻抗函数描述这种衰减^[10],但这一函数中不可避免地要引入可调参数。是否存在更自然的、不依赖于具体参数的方式描述这种机会衰减现象,文献[29]提出了一个新的衰减机制,认为个体在选择目的地的时候会考虑起终点之间其他个体(见图1(f)中的插入图)对目的地机会数的竞争效应。为体现这种竞争效应,假设出行者选择一个目的地的概率正比于目的地的机会数(假设为正比于人口数),反比于出行者所在地点到目的地之间的人口总数

$$P_{ij} \propto \frac{m_j}{S_{ji}} \quad (12)$$

将这一模型命名为人口权重机会模型。与辐射模型类似,PWO模型中所需要的输入数据仅仅是各地点的人口分布,同样不存在可调参数。为验证这一形式异常简单的PWO模型的预测效果,收集了14个城市的人群移动量数据,既包括传统的居民出行调查数据(5个美国城市),也包括新型的人群移动轨迹数据如手机用户数据(1个非洲城市)、社交网络签到数据(6个欧洲城市)和出租车GPS数据(2个中国城市)。从图1中可以看出,PWO模型对北京出租车乘客群体出行模式的预测结果与实际数据显示出的结果十分接近。图2为在14个案例城市中用PWO模型和辐射模型进行移动量预测的准确率对比,从中可以看到PWO模型的预测准确率在所有案例中都高于辐射模型。尽管这些城市在人口数量和城市规模、社会文化背景、经济发展水平等方面是非常多样化的,但

PWO 模型却取得了十分稳定的预测效果(准确率在 70% 左右),说明这一模型或许捕捉到了与具体城市背景无关的人群移动模式形成的普遍机制。

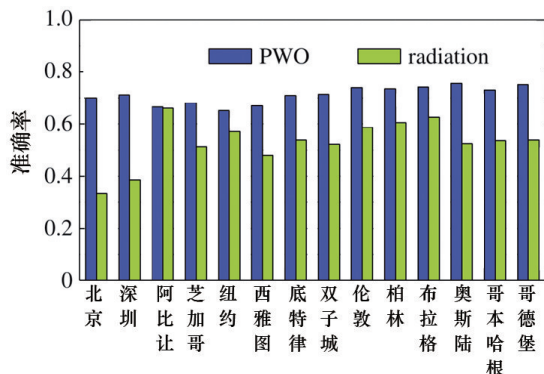


图2 人口权重机会模型与辐射模型整体预测精度对比
Fig. 2 Comparison between the prediction ability of the PWO and radiation models

2.4 几种模型的内在联系

前述分别介绍了引力模型、介入机会模型、辐射模型和 PWO 模型的基本原理。本节将进一步探讨这几种模型的内在联系,以深化对这些模型底层机制的理解。

首先考虑地点人口数均匀分布这样一种极端简化的情形,在这种情况下显然有

$$S_{ij} = S_{ij} = \rho \pi d_{ij}^2 \quad (13)$$

式中, ρ 为人口密度。将式(13)代入式(12),可以得出人口均匀分布情况下的简化 PWO 模型

$$T_{ij} \propto \frac{m_i m_j}{d_{ij}^2} \quad (14)$$

与式(1)进行对比,可以看出式(14)是一个带有平方阻抗函数的引力模型。有趣的是,简化的 PWO 模型与牛顿万有引力定律具有几乎一样的形式。由于实际空间中人口肯定是异质性分布的,因此在引力模型中不能直接使用平方阻抗函数,必须要用实际数据来估计引力模型的参数。但却可以直接使用 PWO 模型中的人口函数 $1/S_{ij}$ 作为阻抗函数,因为人口的异质性已经被反映在 S_{ij} 中^[29]。

在人口均匀分布的假设下,也可以将式(11)中的辐射模型改写为

$$T_{ij} \propto \frac{m_i m_j}{d_{ij}^4} \quad (15)$$

式(15)是一个带有 4 次方阻抗函数的引力模型。与简化的 PWO 模型相比,简化辐射模型的幂指数更大,说明个体选择远距离目的地出行的概率更低。这是导致辐射模型在预测城市人群移动模式时精度不高的原因之一,因为在城市人群中出行的选择范围显然比辐射模型预言的要大得多(图 1)。这一点也可以从文献[26]中对 8 个城市进行引力模型参数估

计的结果上得到印证——在这些城市中估计的引力模型幂指数范围介于 1.63~2.43 之间,更接近简化 PWO 模型中的幂指数 2,但与简化辐射模型中的幂指数 4 则相距甚远。

类似地,在人口均匀分布的假设下也可以把介入机会模型简化为

$$T_{ij} \propto m_i \left(e^{a m_j} - 1 \right) e^{-a \rho \pi d_{ij}^2} \quad (16)$$

式(16)同样是一个类引力模型的形式,其中阻抗函数是距离的二阶指数函数,说明长距离出行产生的概率极低,这可能正是介入机会模型低估长距离出行的主要原因。

通过上述对各种模型的分析可知,在人口均匀分布的假设下, PWO 模型、辐射模型、介入机会模型都可以转化为某种形式的引力模型。这说明,尽管这些模型基于不同的假说提出,但它们具有非常类似的底层机制。可以把这种机制归纳为:个体选择一个目的地的概率会随着某种阻挠因素的增加而下降,在引力模型中,这种阻挠因素是空间距离,而在介入机会模型、辐射模型和 PWO 模型中,这种阻挠因素是地点间的人口数量。不同的是,在引力模型和介入机会模型中,需要一个带参数的阻抗函数来描述目的地选择概率随阻挠因素增加而下降的过程,而在辐射模型和 PWO 模型中,目的地选择概率是随着地点间人口数的增加而自然衰减的。

3 空间交互网络上的随机游走模型

3.1 空间随机游走模型概述

第 2 节介绍的介入机会模型、辐射模型和 PWO 模型都是从个体目的地选择决策行为出发提出相关假设,通过输入人口分布、空间距离等数据预测地点间的空间交互流量矩阵。但这些模型本质上都是静态模型,缺乏对个体连续移动过程的刻画,因而无法再现个体移动轨迹的时空统计模式,例如地点访问频率的异质性分布和访问地点的异速增长现象^[30-32]等。在研究人口迁移、居住或就业地选址等长时间尺度的空间交互行为时,使用这类静态模型没有问题。但在研究出行者交通行为^[9-10]、疾病时空传播过程^[33-34]、点对点通信网络设计^[35-36]等问题时,需要详细刻画个体在空间中的连续移动过程,此时静态模型不再适用。因此,对个体的空间连续移动过程进行建模,也是空间交互网络研究中的一个重要问题。

传统上,处理粒子扩散过程的标准模型是二维随机游走模型^[37],而列维飞行模型(一种带有无尺度步长分布的随机游走模型)则被广泛应用于对动物移动行为的建模^[38]。近年来,随着人类移动轨迹数据可获取程度的提高,人们可以用更翔实的数据探究自身的空间移动规律。这方面的研究始于 Brockmann 等^[30]于 2006 年发表在《Nature》上的开创性工作,他们利用美元流通记录网站的数据间接对人类空间移动行为进行了实证统计,并用连续时间随机游走模型^[39]再现了美元流通的低速扩散特征。随后, Song 等^[31-32]又从对手机数据所进行的统计实证中发现人类的空间移动模式中存在一些无法用连续时间随机游走模型解释的标度异常现象,例如访

问地点数量 $S(t)$ 随着时间亚线性增长、地点访问频率分布服从齐普夫律等,并建立了一个探索和偏好返回模型^[32]解释这些现象。关于个体空间移动模式建模^[30-32,40-43]的更多详细介绍可见文献[8],本文重点介绍作者最近在空间网络随机游走建模方面的一些研究进展。

3.2 记忆性随机游走模型

之前的很多实证研究^[31-32,43]都发现,人类个体在空间移动过程中具有强烈的记忆性——个体总是倾向于很快就返回之前访问过的一个地点。Szell 等^[43]在他们的工作中已经意识到了记忆性对人类个体移动行为的重要影响,他们直接使用返回时间间隔分布 $P(\tau)$ 作为建模假设,即个体每次从该分布中抽取一个间隔 τ 来决定返回之前 l 步曾访问的地点。该模型能够生成和实际一致的访问地点增长速度 $S(t)$,但本文认为这是用一个现象去解释另一个现象,并没有触及个体移动行为的本质机制。而探索和偏好返回模型^[32]则假设个体在移动过程中会以一定的概率 $P \propto S^{-\gamma}$ 访问一个新地点(即随着访问地点数量 S 的增加,个体越来越不倾向于访问新地点,而更倾向于在已访问过的地点之间移动),同时以概率 $1-P$ 返回之前访问过的一个地点,选择某个已访问地点的概率正比于该地点之前已被访问过的频率。探索和偏好返回模型能够生成和实际数据相符合的地点访问频率分布和基本符合实际的地点增长速度曲线(因为该模型只能再现 $S(t)$ 增长曲线的尾部特征)。但是该模型中却无法解释为什么个体扩展新地点的概率 P 会随访问地点数 S 的增加而下降。

本文认为,造成这种现象的可能原因是个体在移动过程中会对已经访问过的地点形成记忆性偏好,这种记忆性偏好会随着个体对一个地点的访问次数增加而得到不断强化。基于这些思考,建立在有限空间中的记忆性偏好随机游走模型(图3)。在该模型中,位于点的个体在每步随机游走过程中,按照式(17)中的概率随机访问地点 β

$$p_{\alpha \rightarrow \beta} \propto 1 + \lambda k_{\beta} \quad (17)$$

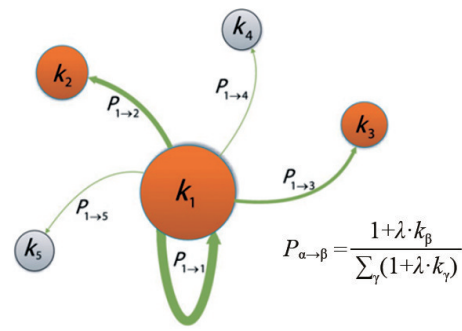


图3 记忆性随机游走模型示意

Fig. 3 Illustration of memory-preferential random walk model

式中, k_{β} 为地点 β 已被访问过的次数, λ 为记忆强度因子。当 $\lambda=0$ 时,式(17)退化为一般的无记忆随机游走模型;当 $\lambda > 0$ 时,个体会对已访问地点形成记忆性。随着个体访问地点数量的增多,已访问过地点的总吸引力就相对越来越强,因此个体访问新地点的可能性自然就降低了。

对记忆性随机游走模型进行定量解析分析和实际数据验证,结果表明它能很好地再现个体返回时间间隔的异质性分布、地点访问频率服从齐普夫律和访问地点的亚线性增长这些重要的时空统计模式。同时,记忆性偏好游走模型不仅能对真实地理空间中的个体移动行为进行建模,还能再现虚拟空间中个体跳转行为的时空统计特征^[45]。

3.3 个体和群体移动模式预测的统一模型

前已述及,包括记忆性随机游走模型在内的个体移动模型^[30-32,40-45],只能再现个体空间移动网络的各种时空统计特征,但无法直接用于对空间交互网络流量分布模式的预测;而引力模型、辐射模型等空间交互模型,又无法用于对个体连续移动轨迹的刻画。事实上,无论是个体空间移动网络,还是群体空间交互网络,它们都由同一群个体的移动行为所产生(图4),只是对网络时空特征的统计角度有所不同。因

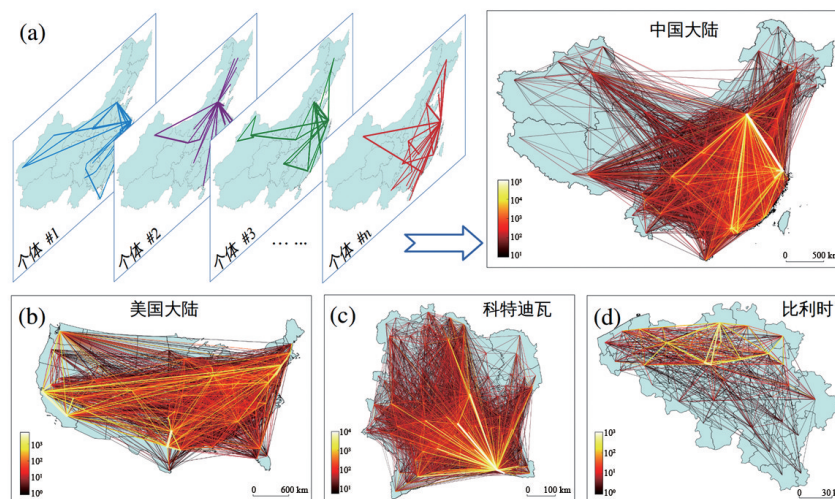


图4 个体空间移动网络和群体空间交互网络的关系

Fig. 4 Relationship between individual mobility networks and collective spatial interaction networks

此,本质上应该能用某种统一的模型来对个体和群体2个层面的空间移动模式同时进行描述和预测,但目前还仍然缺乏这样的模型。

针对这一问题,在PWO模型^[29]和记忆性随机游走模型^[44]的基础上,设计了一种能够同时预测个体和群体空间移动模式的统一模型^[46]。假设每个地点对于个体来说有一个固有的吸引力,这与个体是否访问过该地点无关,可以通过PWO模型计算这个固定吸引力;而地点的另一部分吸引力则是由个体的记忆性带来的,即个体访问过的地点相对于他没有访问过的地点会多出一部分附加吸引力。根据实际数据中观察到个体访问地点频率分布近似服从齐普夫律^[31-32],假设各地点的附加吸引力依齐普夫律递减,即被访问越多的地点其排序就越靠前,这与记忆性随机游走模型^[44]具有相似机制。根据以上假设,可以给出出行者在每步移动过程中选择一个地点的概率

$$P_{ij} \propto \frac{m_j}{S_j} \left(1 + \frac{\lambda}{r_j} \right) \quad (18)$$

式中, λ 为记忆强度参数, r_j 为地点 j 附加吸引力的排序。

分别用图4中的4个数据集(包括科特迪瓦全国的手机用户通信记录、新浪微博上中国大陆用户的签到记录、Four-square网站上美国大陆用户的签到记录和Gowalla网站上比利时用户的签到记录)验证了式(18)模型的预测效果。结果显示,无论是在个体的访问地点增长速度、返回时间间隔分布、地点访问频率这些个体移动模式上,还是在出行距离分布、地点间流量分布和地点间流量矩阵这些群体移动模式上,统一模型的预测结果都能与实际数据相当精确地符合^[46]。统一模型能够在文化背景、经济水平、空间尺度等方面差异非常显著的4个国家中都取得较好的预测结果,说明模型的基本机制是相当普适的。更为重要的是,由于统一模型能够同时对个体和群体的空间移动模式进行建模,在疾病传播等需要精细刻画个体群体移动模式的研究中更具应用价值^[33-34]。

4 结论

自19世纪中叶Desart用万有引力定律研究铁路网络客流分布问题^[11]以来,对空间交互网络建模的研究已有超过百年的历史^[47],以引力模型为代表的经典空间交互模型已在诸多领域取得了广泛而成功的应用^[1-19]。而在过去的10年,随着GPS、手机通信、移动互联网、物联网等社会感知技术的快速发展^[48],从大规模空间交互数据中发现普遍模式、提出个体行为决策假说并建立预测模型这种大数据驱动的新兴研究范式^[49],为空间交互网络研究又注入了新的活力^[50]。尽管这一方向目前已取得丰硕的研究成果,仍存在很多具有挑战性的问题亟待解决。

1) 现有的空间交互模型大多假定群体中的所有个体都具有相同的行为模式或决策机制,但实际中个体的行为模式

却是非常多样性的^[23]。是否可以通过对群体进行聚类分析的方法,挖掘个体的多样性行为决策机制并依此建立模型进一步提高预测精度,是一个值得关注的问题^[51]。但过于细致的个体行为分类与追求普适化建模之间又存在明显的矛盾^[32],如何在两者之间取得平衡是难点所在。

2) 以辐射模型为代表的新型空间交互模型^[27,29,52]虽然仅通过输入人口分布数据就可预测交互强度,但这类模型中缺乏对交互成本因素的直接考量(只是在地点排序时间间接反映了成本差异)。这使得这类模型对成本变化的反应不敏感,在分析诸如交通需求管理政策(如区域拥堵收费)对出行者目的地选择行为的影响等问题时存在局限性。如何在数据驱动模型中敏感地反应成本因素变动的影响,也是一个值得关注的问题。但由于从真实数据中往往难以直接或全面观测到各种影响因素变动对个体行为的影响^[53],实验室可控实验的方式或许是揭示个体和群体在各种影响因素复杂变动情况下选择决策机制的一种有效手段^[54]。

3) 在个体移动模式建模方面,目前统计物理和复杂系统领域所发展的大部分个体移动模型整体上还处于简化,能考虑个体活动和移动之间的有机联系、个体出行链的时空约束、不同交通方式对出行者的竞争关系等关键特征的研究工作还非常有限。而交通行为研究领域所发展的基于活动的分析方法^[55-57]则特别强调移动与活动之间的有机联系及时空制约,以活动-出行链的整体视角对个体日常活动-移动行为规律进行分析和建模,在理论上具有突出优势。目前,已有一些学者基于手机等新型数据源开展了个体活动-出行链时空结构特征分析和建模的研究工作^[58-60]。随着手机信令数据、移动社交App数据等富含个体活动和出行信息的数据可获得程度日益提高,将数据驱动的复杂性建模和基于活动的分析法深度结合,开展社会感知数据驱动的个体活动-出行决策机理分析与建模研究,将是未来一段时期内一个非常值得关注的研究方向。

致谢:感谢狄增如、樊瑛、高自友、韩筱璞、来颖诚、汪秉宏、王文旭、杨紫陌、赵琛、赵毅鸣、赵志丹、曾安、周涛等对本文中相关工作所做出的贡献。

参考文献(References)

- [1] Roy J R, Thill J C. Spatial interaction modelling[M]. Berlin: Springer, 2004.
- [2] Dejon B. Spatial interaction network flow models[M]. Heidelberg: Physica-Verlag, 1978, 1(1): 377-386.
- [3] Krings G, Calabrese F, Ratti C, et al. Urban gravity: A model for inter-city telecommunication flows[J]. Journal of Statistical Mechanics, 2009, 7: L07003.
- [4] Duenas M, Fagiolo G. Modeling the international-trade network: A gravity approach[J]. Journal of Economic Interaction and Coordination, 2013, 8(1): 155-178.
- [5] Davis K F, D'Odorico P, Laio F, et al. Global spatio-temporal patterns

- in human migration: A complex network perspective[J]. *PLoS One*, 2013, 8(1): e53723.
- [6] Kwan M P. Mobile communications, social networks, and urban travel: Hypertext as a new metaphor for conceptualizing spatial interaction[J]. *The Professional Geographer*, 2007, 59(4): 434–446.
- [7] Barthélemy M. Spatial networks[J]. *Physics Reports*, 2011, 499(1): 1–101.
- [8] 周涛, 韩筱璞, 闫小勇, 等. 人类行为时空特性的统计力学[J]. *电子科技大学学报*, 2013, 42(4): 482–540.
Zhou Tao, Han Xiaopu, Yan Xiaoyong, et al. Statistical mechanics on temporal and spatial activities of human[J]. *Journal of University of Electronic Science and Technology of China*, 2013, 42(4): 482–540.
- [9] 邵春福. 交通规划原理[M]. 北京: 中国铁道出版社, 2008.
Shao Chunfu. *Transportation planning theory*[M]. Beijing: China Railway Publishing House, 2008.
- [10] de Dios Ortúzar J, Willumsen L G. *Modelling transport*[M]. Chichester: Wiley, 2001.
- [11] Desart H G. *Théorie des mouvements de voyageurs*[M]. Bruxelles: E. Devroye, 1847.
- [12] Carey H. *Principles of social science*[M]. Philadelphia: Lippincott, 1858.
- [13] Ravenstein E G. The laws of migration[J]. *Journal of the Royal Statistical Society*, 1885, 48: 167–235.
- [14] Young E C. The movement of farm population[M]. Cornell University Agricultural Experiment Station, 1924.
- [15] Reilly W J. *The law of retail gravitation*[M]. New York: WJ Reilly, 1931.
- [16] Zipf G K. The P1P2/D hypothesis: On the intercity movement of persons[J]. *American Sociological Review*, 1946, 11(6): 677–686.
- [17] Stewart J Q. An inverse distance variation for certain social influences [J]. *Science*, 1941, 93(2404): 89–90.
- [18] Kaluza P, Kölsch A, Gastner M T, et al. The complex network of global cargo ship movements[J]. *Journal of the Royal Society Interface*, 2010, 7(48): 1093–1103.
- [19] Liu J H, Zhang Z K, Yang C, et al. Gravity effects on information filtering and network evolving[J]. *PLoS ONE*, 2014, 9: e91070.
- [20] 闫小勇. 一种改进的重力模型标定方法[J]. *交通与计算机*, 2003, 21(4): 93–95.
Yan Xiaoyong. A modified calibration method for gravity model[J]. *Transportation and Computer*, 2003, 21(4): 93–95.
- [21] Sheppard E S. Theoretical underpinnings of the gravity hypothesis[J]. *Geographical Analysis*, 1978, 10(4): 386–402.
- [22] Wilson A G. A statistical theory of spatial distribution models[J]. *Transportation Research*, 1967, 1(3): 253–269.
- [23] Yan X Y, Han X P, Wang B H, et al. Diversity of individual mobility patterns and emergence of aggregated scaling laws[J]. *Scientific Reports*, 2013, 3: 2678.
- [24] Hua C I, Porell F. A critical review of the development of the gravity model[J]. *International Regional Science Review*, 1979, 4(2): 97–126.
- [25] Stouffer S A. Intervening opportunities: A theory relating mobility and distance[J]. *American Sociological Review*, 1940, 5(6): 845–867.
- [26] 闫小勇. 人类移动模式分析与预测[D]. 北京: 北京师范大学系统科学学院, 2014.
Yan Xiaoyong. *Characterizing and predicting human mobility patterns* [D]. Beijing: School of Systems Science, Beijing Normal University, 2014.
- [27] Simini F, González M C, Maritan A, et al. A universal model for mobility and migration patterns[J]. *Nature*, 2012, 484(7392): 96–100.
- [28] Offenhuber D, Ratti C. *Decoding the city: Urbanism in the age of big data*[M]. Berlin: Birkhäuser, 2014.
- [29] Yan X Y, Zhao C, Fan Y, et al. Universal predictability of mobility patterns in cities[J]. *Journal of the Royal Society Interface*, 2014, 11(100): 20140834.
- [30] Brockmann D, Hufnagel L, Geisel T. The scaling laws of human travel [J]. *Nature*, 2006, 439(7075): 462–465.
- [31] González M C, Hidalgo C A, Barabási A L. Understanding individual human mobility patterns[J]. *Nature*, 2008, 453(7196): 779–782.
- [32] Song C, Koren T, Wang P, et al. Modelling the scaling properties of human mobility[J]. *Nature Physics*, 2010, 6(10): 818–823.
- [33] Belik V, Geisel T, Brockmann D. Natural human mobility patterns and spatial spread of infectious diseases[J]. *Physical Review X*, 2011, 1(1): 011001.
- [34] Balcan D, Vespignani A. Phase transitions in contagion processes mediated by recurrent mobility patterns[J]. *Nature Physics*, 2011, 7(7): 581–586.
- [35] Camp T, Boleng J, Davies V. A survey of mobility models for ad hoc network research[J]. *Wireless Communications and Mobile Computing*, 2002, 2(5): 483–502.
- [36] Bai F, Helmy A. A survey of mobility models[J]. *Wireless Adhoc Networks*, 2004, 206: 1–30.
- [37] *reactions in fractals and disordered systems*[M]. Cambridge: Cambridge University Press, 2000.
- [38] Viswanathan G M, Da Luz M G E, Raposo E P, et al. *The physics of foraging: An introduction to random searches and biological encounters* [M]. Cambridge: Cambridge University Press, 2011.
- [39] Montroll E W, Weiss G H. Random walks on lattices[J]. *Journal of Mathematical Physics*, 1965, 6(2): 167–181.
- [40] Hu Y, Zhang J, Huan D, et al. Toward a general understanding of the scaling laws in human and animal mobility[J]. *Europhysics Letters*, 2011, 96(3): 38006.
- [41] Han X P, Hao Q, Wang B H, et al. Origin of the scaling law in human mobility: Hierarchy of traffic systems[J]. *Physical Review E*, 2011, 83(3): 036117.
- [42] Yan X Y, Han X P, Zhou T, et al. Exact solution of the gyration radius of an individual's trajectory for a simplified human regular mobility model[J]. *Chinese Physics Letters*, 2011, 28(12): 120506.
- [43] Szell M, Sinatra R, Petri G, et al. Understanding mobility in a social petri dish[J]. *Scientific Reports*, 2012, 2: 457.
- [44] Zhao Y M, Zeng A, Yan X Y, et al. Unified underpinning of human mobility in the real world and cyberspace[J]. *New Journal of Physics*, 2016, 18: 053025.
- [45] Zhao Z D, Yang Z, Zhang Z, et al. Emergence of scaling in human-interest dynamics[J]. *Scientific Reports*, 2013, 3: 3472.
- [46] Yan X Y, Wang W X, Gao Z Y, et al. Universal model of individual and population mobility on diverse spatial scales[J]. *Working Paper*, 2016.
- [47] Odlyzko A. The forgotten discovery of gravity models and the inefficiency of early railway networks[J]. *Oeconomica: History, Methodology, Philosophy*, 2015, 5(1): 157–192.
- [48] Liu Y, Liu X, Gao S, et al. Social sensing: A new approach to understanding our socioeconomic environments[J]. *Annals of the Association of American Geographers*, 2015, 105(3): 512–530.

- [49] Lazer D, Pentland A S, Adamic L, et al. Life in the network: The coming age of computational social science[J]. *Science*, 2009, 323(5915): 721–723.
- [50] O’Sullivan D, Manson S M. Do physicists have geography envy? and what can geographers learn from it?[J]. *Annals of the Association of American Geographers*, 2015, 105(4): 704–722.
- [51] Pappalardo L, Simini F, Rinzivillo S, et al. Returners and explorers dichotomy in human mobility[J]. *Nature Communications*, 2015, 6: 8166.
- [52] Liang X, Zhao J, Xu K. A general law of human mobility[J]. *Science China Information Sciences*, 2015: 58(10): 1–14.
- [53] Hensher D A, Rose J, Greene B. *Applied choice analysis: A primer* [M]. Cambridge: Cambridge University Press, 2005.
- [54] Mahmassani H S. *Learning from interactive experiments: Travel behavior and complex system dynamics*[M]. Bingley: Emerald Publishing, 2009.
- [55] Timmermans H, Arentze T, Joh C-H. Analysing space-time behaviour: New approaches to old problems[J]. *Progress in Human Geography*, 2002, 26(2): 175–190.
- [56] Buliung R N, Kanaroglou P S. Activity-travel behaviour research: Conceptual issues, state of the art, and emerging perspectives on behavioural analysis and simulation modelling[J]. *Transport Reviews*, 2007, 27(2): 151–187.
- [57] 张文佳, 柴彦威. 时空制约下的城市居民活动-移动系统——活动分析法的理论和模型进展[J]. *国际城市规划*, 2009, 24(4): 60–68. Zhang Wenjia, Chai Yanwei. Urban activity-travel systems in the condition of space-time: A review of activity-based theories and models [J]. *Urban Planning International*, 2009, 24(4): 60–68.
- [58] Schneider C M, Belik V, Couronné T, et al. Unravelling daily human mobility motifs[J]. *Journal of the Royal Society Interface*, 2013, 10 (84): 20130246.
- [59] Widhalm P, Yang Y, Ulm M, et al. Discovering urban activity patterns in cell phone data[J]. *Transportation*, 2015, 42(4): 597–623.
- [60] Jiang S, Ferreira Jr J, González M C. Activity-based human mobility patterns inferred from mobile phone data: A case study of Singapore [J]. *IEEE Transactions on Big Data*, 2015, 12: 0163.

Advances in modeling spatial interaction network

YAN Xiaoyong

Institute of Transportation System Science and Engineering, Beijing Jiaotong University, Beijing 100044, China

Abstract The spatial interaction network is the space-embedded directed network with flows of people, goods or information among places. Examples of spatial interaction networks include the international trade network, the migration network, the transportation network and the inter-city telecommunication network. Understanding and predicting spatial interaction patterns of these networks are of importance in various disciplines, including the regional science, the transportation science and the economic geography, with many practical applications in the urban or transportation system planning, the epidemiology of infectious diseases, and the location-based services. This paper first introduces briefly two classic spatial interaction models: the gravity model and the intervening opportunity model, and then reviews some recent advances in data-driven spatial interaction models, including the radiation model, the population-weighted opportunity model and the random walk models on spatial interaction networks. This paper also discusses some challenging problems in modeling the spatial interaction networks, such as the individual mobility behavior diversity, the group choice decision behavior experiment and the data-driven modeling approach for the human activity-travel behavior.

Keywords spatial interaction; complex networks; human mobility patterns; trip distribution

(责任编辑 刘志远)