

空间交互作用中的距离影响及定量分析

刘瑜[†] 龚俐 童庆禧

北京大学遥感与地理信息系统研究所, 北京 100871; [†] E-mail: liuyu@urban.pku.edu.cn

摘要 对地理学研究中空间交互问题的距离衰减现象进行描述和分析, 并总结重力模型的求解方法。首先, 大部分空间交互都受到距离衰减效应的影响, 距离衰减函数定量表示交互中距离的作用。对比幂律、指数等多种形式的距离衰减函数, 认为幂律形式能够更好地刻画距离影响的本质特征, 并且可以在不同交互中进行比较, 其对应的交互模型即为在地理学和区域经济学中广泛应用的重力模型。然后, 在重力模型的基础上, 总结解析法和模拟法两大类基于给定交互矩阵和距离矩阵定量计算距离衰减函数及衰减系数的方法。由于难以先验确定交互公式中节点规模的数值, 因此拟合重力模型, 同时估计距离衰减系数, 是应用重力模型的恰当方式。

关键词 空间交互; 距离衰减; 重力模型; 解析法; 模拟法

中图分类号 P208

Quantifying the Distance Effect in Spatial Interactions

LIU Yu[†], GONG Li, TONG Qingxi

Institute of Remote Sensing and Geographical Information Systems, Peking University, Beijing 100871; [†] E-mail: liuyu@urban.pku.edu.cn

Abstract Distance decay issue of spatial interactions in geography was described and analyzed, and a number of approaches were summarized to solve the gravity model. Firstly, many spatial interactions are governed by the distance decay effect, and the distance decay function quantitatively represents the distance effect. Among kinds of distance decay functions such as power law, exponential and other equations, the power law function reveals the inherent distance impacts behind spatial interactions so that the interaction formula becomes the gravity model, which is widely used in geography and regional economics. Secondly, the authors summarize some methods to quantify the distance decay function according to an interaction matrix and a distance matrix. The simulation results show that fitting the gravity model to estimate nodal attractions as well as the distance decay parameter is an appropriate way to apply the model.

Key words spatial interactions; distance decay; gravity model; analytic approach; simulation approach

在地理空间分析与建模中, 距离是一个重要因素。Tobler^[1]描述的地理学第一定律——“任何事物都相关, 只是相近的事物关联更紧密”阐述了距离对于地理现象的影响。Miller^[2]指出, 第一定律可以从两个方面解读, 即空间上相近的区域具有更高的属性相似性和交互强度。

在空间交互中, 随着距离的增长, 交互强度相应减弱, 这种现象称为距离衰减(distance decay)。通

常, 距离衰减有宏观和微观两个层面的表现形式。在宏观层面, 由于距离衰减, 两个地理实体(或地理对象、地物、节点)之间的交互总量(如两个城市之间的客流量)在其他变量(如城市人口等)相对稳定的前提下, 与距离表现为负相关关系。而在微观层面, 距离衰减表现为在不考虑其他影响因素下, 单个物体、动物或人移动到不同目的地的概率与距离负相关。近年来, 基于轨迹数据的移动模式研究中, 发

现每次移动步长 Δd 的概率分布通常可以定量表示为幂律形式 $p(\Delta d) \sim \Delta d^{-\beta}$ 、指数形式 $p(\Delta d) \sim \exp(-\alpha \Delta d)$ ，或者指数截断幂律形式 $p(\Delta d) \sim \exp(-\alpha \Delta d) \cdot \Delta d^{-\beta}$ [3-7]，这体现了微观层面的距离衰减。

距离衰减形成可以归结为两个原因。首先，由于空间移动的成本与距离正相关，为了最小化成本，需要降低长距离交互的强度。从该角度出发，Wilson^[8]和 Colwell^[9]分别基于不同假设进行数学推导，得到指数和幂律形式的距离衰减函数。其次，在交互目的地的选择过程中，由于中介机会(intervening opportunities)的存在^[10]，降低了长距离交互的概率。在 Liben-Nowell 等^[11]、Noulas 等^[12]、Simini 等^[13]以及其他人的研究中，为了解释空间交互，采用中介机会原理，并认为潜在交互目的地之间的相对距离比绝对距离更为重要。

近年来，随着 IT 技术的发展，人类进入大数据(big data)时代，使得我们可以从微观和宏观两个层面获取具有较高时空分辨率的空间交互信息。例如，基于手机和 GPS，可以得到海量个体的移动轨迹数据^[14]；而通过带有位置标记的社交媒体签到(check-in)数据，可以提取大空间尺度的移动模式，并通过好友关系估计城市间的联系强度。表 1 列举近年来采用大数据开展关于空间交互的主要研究。

大数据所包含的空间交互信息为地理学家量化和比较不同类型空间交互中距离的作用提供了便

利，从而能够更好地理解空间在诸多社会经济现象中所发挥的作用。本文在既有工作的基础上，总结不同形式的距离衰减函数，梳理常用距离衰减估算方法，以期对空间交互的定量研究提供指导。

1 空间距离衰减函数和距离系数

在地理学研究中，通常采用 $G_{ij} = CP_i P_j f(d_{ij})$ 来量化距离对于空间交互的影响。其中 C 为常量系数； G_{ij} 表示地物 i 和 j 之间的交互强度，在宏观层面可以用客流量、货流量以及通讯量等表示，在微观层面可以用概率表示； P_i 和 P_j 反映这两个地理实体的规模；距离衰减函数 $f(d_{ij})$ 是以距离 d 为自变量的函数，用来定量刻画距离因素的作用。在实践中，距离 d 可以采用不同度量，如绝对地理距离、旅行成本、旅行时间等。尽管由于假设不同，距离衰减函数的形式不尽相同，但是函数 $f(d)$ 的值通常是随着 d 的增加而减小的。目前最常用的距离衰减函数包括指数函数、幂函数及高斯函数(图 1)，这 3 种距离衰减函数的定义^[21-23]如下。

指数型距离衰减函数：

$$f(d) = Ce^{-\alpha d} \quad (\alpha > 0)。 \quad (1)$$

幂律型距离衰减函数：

$$f(d) = Cd^{-\beta} \quad (\beta > 0)。 \quad (2)$$

高斯型距离衰减函数：

$$f(d) = Cd^{-\alpha d^2} \quad (\alpha > 0)。 \quad (3)$$

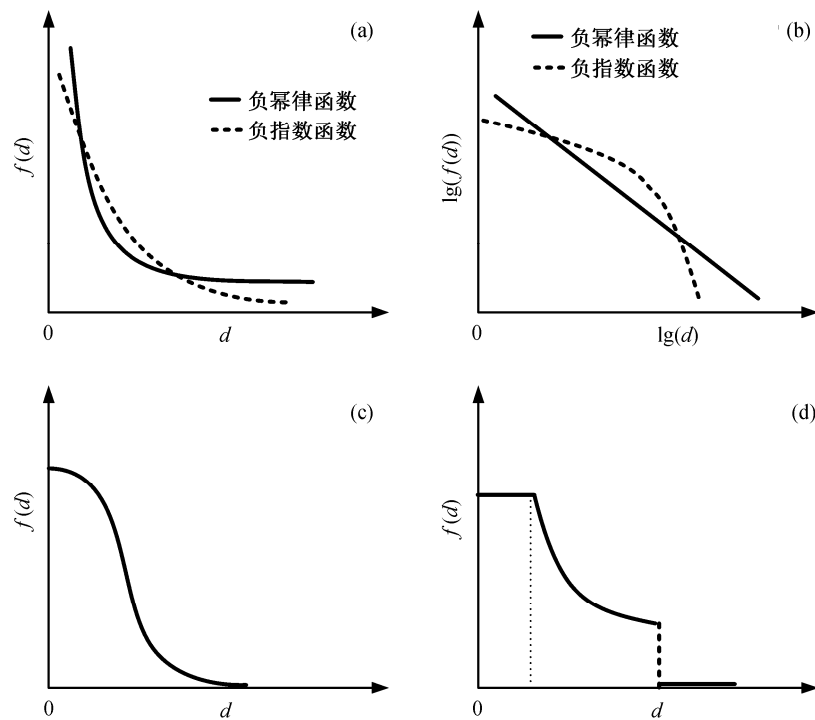
此外，在不同研究领域和应用环境中，距离衰减函数也会产生一定的变形。例如使用指数截断的幂律函数 $f(d) = Ce^{-\alpha d} d^{-\beta}$ ($\alpha, \beta > 0$)；或者采用分段函数控制不同距离区间内的距离衰减程度^[23]。

在地理学中最基本的两种交互模型是重力模型和 Wilson 最大熵模型^[8](表 2)，它们分别使用幂律和指数形式的距离衰减函数。Carey^[24]最早应用重力模型研究空间交互作用，其中幂律形式的距离衰减源于牛顿的引力模型，但是由于缺乏严格理论证明，尤其由于量纲问题^[25]，受到一些批评。而最大熵模型因其严格的数学证明而被更多地接受。但是，最大熵模型的距离衰减函数是负指数函数，对应的空间相互作用是局部性的，这点与 Tobler 的地理学第一定律不一致^[26]。另外，Colwell^[9]基于中心地理论得到幂律形式衰减，陈彦光^[25]通过修改最大熵模型推导过程中的距离成本假设，也得到相似结果。然

表 1 部分基于大数据的空间交互研究

Table 1 Some spatial interaction research based on big data

作者	数据源	空间交互形式
Brockmann 等(2006) ^[3]	可追踪物体(钱币)	空间移动
González 等(2008) ^[4]	手机数据	空间移动
Cheng 等(2011) ^[15]	社交网络数据 (Foursquare)	空间移动
Kang 等(2012) ^[16]	手机数据	空间移动
Liu 等(2012) ^[7]	出租车轨迹数据	空间移动
Noulas 等(2012) ^[12]	社交网络数据 (Foursquare)	空间移动
Liang 等(2012) ^[6]	出租车轨迹数据	空间移动
Kang 等(2013) ^[17]	手机数据	城市和区域联系
Ratti 等(2010) ^[18]	手机数据	城市和区域联系
Liu 等(2014) ^[19]	网页数据	城市和区域联系
Liben-Nowell 等(2004) ^[11]	社交网络数据 (LiveJournal)	好友关系
Onnela 等(2011) ^[20]	手机数据	好友关系



(a) 负幂律函数和负指数函数; (b) 双对数坐标下负幂律函数和负指数函数; (c) 高斯函数; (d) 分段函数。 d 为距离, $f(d)$ 为距离衰减函数

图 1 几种不同的距离衰减曲线

Fig. 1 Different distance decay functions

表 2 重力模型和 Wilson 最大熵模型对比

Table 2 Comparison of gravity model and Wilson's spatial interaction model

交互模型	通用公式	参数解释	距离函数
重力模型	$G_{ij} = kP_i^\gamma P_j^\delta d_{ij}^{-\beta}$	G_{ij} : 实体 i 和 j 之间的作用强度 P_i : 实体 i 的规模 d_{ij} : 距离 γ, δ : 规模参数 β : 距离衰减系数	$f(d)=d^{-\beta}$
Wilson 最大熵模型	$T_{ij} = A_i O_i B_j D_j e^{-\alpha d_{ij}}$	T_{ij} : 实体 i 和 j 之间的作用强度 O_i, D_j : 实体 i 和 j 的输出和输入量 A_i, B_j : 归一化因子 d_{ij} : 距离 α : 距离衰减系数	$f(d)=e^{-\alpha d}$

而,上述工作未得到广泛认可,这从另一个侧面说明地理现象的复杂性以及建立普适模型的难度。

对比幂律和指数两种形式的距离衰减,负指数函数随着距离增加,衰减速度比负幂律函数快得多,因此负指数函数在空间上的作用域有限,长程的交互作用偏少,而负幂律函数适用于大尺度的复杂空间交互系统。此外幂函数具有无标度(scale-free)的特征,暗示空间作用距离没有特征长度,从而使构造的空间过程模型与地理学第一定律没有矛盾,还能够反映并比较不同类型空间活动的本质特征。当

采用幂律形式的距离衰减函数时,指数 β 是距离衰减系数,也称为距离摩擦系数或阻尼系数,用于表达距离对流量的影响。 β 越大,距离 d 对 G_{ij} 的阻碍作用也越大,反之亦然^[27]。尤其是当 β 小于1.0时,距离衰减作用就变得不那么明显了。根据我们的研究,对于个体在城市尺度的移动,距离衰减系数在1.0~2.0之间^[7,16,28],而在城市间的移动,距离衰减系数略低^[29]。通过非空间移动所表现的空间交互,如手机通信,往往具有更低的距离衰减系数,即距离对于这类空间交互的影响并不明显^[17,19]。由于幂

律函数所具有的优势, 我们认为负幂函数形式的重力模型能够更好地刻画大尺度的空间范围内地理实体之间的相互作用。下面主要讨论在重力模型下的距离衰减系数的估算问题。

2 距离衰减系数定量估算方法

在很多实际应用中, 给定一组地理对象, 任意两个对象之间的交互强度和距离已知, 我们希望定量得到这种交互行为受距离的影响, 如果采用幂律形式的距离衰减函数, 那么这个影响可以简单地用指数 β 的值表示。直观上, 距离衰减作用的存在, 使得长距离地理对象间交互强度变弱。因此, 可以制作交互强度与距离的散点图, 从而反映它们之间的关系。利用签到数据, 可以得到海量个体在中国城市间的移动信息, 进而统计每两个城市间的移动次数, 该次数表示城市间的交互强度。图2是基于签到数据获取的中国城市间交互强度与欧氏距离在 log-linear 坐标系中的散点图, 可以看出交互强度与距离之间的关系可以较好地用指数函数拟合。然而, 考虑到两个区域之间的交互强度除了受到距离影响外, 还与这两个区域的规模有关, 因此, 不能简单用图2所表示的关系来量化相应空间交互行为中的距离衰减特征。基于此, 我们探讨了如何依据实际观测的交互强度来估计距离衰减影响的方法, 基本上可以分为基于重力模型的解析方法和模拟方法两大类。其中, 解析法主要通过公式推导或问题转换的方法, 提出计算距离衰减系数 β 的解析值; 而模拟的方法可以处理任意形式的距离衰减函数, 通过给定的假设条件, 由计算机模拟大量的重复实验,

得到符合条件的距离衰减函数或 β 值。

在实际应用中, 为了简化计算, 重力模型的公式通常表示为式(4)的形式, 即令 $P_i = P_i'$ 和 $P_j = P_j'$ (参见表2):

$$G_{ij} = kP_iP_jd_{ij}^{-\beta} \quad (i \neq j), \quad (4)$$

其中 G_{ij} 是实体 i 与实体 j 之间的作用强度, P_i 和 P_j 分别是实体 i 和 j 的规模(引力)大小, d_{ij} 是 i 和 j 之间的距离, k 是常量系数。

2.1 解析法

一些研究者使用正向求解重力模型的方法估算距离衰减系数, 具体方法如下。

首先, 对式(4)进行等式变形, 得到

$$G_{ij} / P_iP_j = kd_{ij}^{-\beta} \quad (i \neq j)。 \quad (5)$$

对式(5)做对数变换, 公式两边取对数, 得到

$$\ln(G_{ij} / P_iP_j) = \ln k - \beta \ln d_{ij} \quad (i \neq j)。 \quad (6)$$

令 $A_{ij} = G_{ij} / P_iP_j$, 如果在双对数坐标系中绘制 A_{ij} 和 d_{ij} 的散点图, 可以观察到一条直线, 且该直线的斜率刚好等于距离衰减系数 β 。在交通网络中, 可以用节点的度近似代表节点的规模大小; 在地理学领域, 一般用节点的总人口数作为 P 的取值。但是在很多空间相互作用的网络中, 节点规模并不与节点的人口规模和节点的度线性相关, 因此, 该方法具有一定的局限性。

假设观测的流量分布符合重力模型, 从真实流量分布反推节点规模大小的过程称为逆重力模型。很多研究者提出逆重力模型的求解方法, 如线性规划法、代数方法和粒子群优化算法等, 这些方法除了可以估计距离衰减系数, 还可以用于得到每个地理对象对应于实际交互强度的理论大小(theoretical size), 在文献[30]中称为吸引力(attraction), 反映地理对象在交互网络中的地位, 通常与其规模正相关。如 Xiao 等^[29]利用航空客流量, 用重力模型估计中国主要城市的吸引力, 结果基本与城市人口规模相符, 但存在部分不一致, 如海口和昆明的吸引力排名高于其人口规模排名, 反映这两个城市航空交通相对重要的事实。下面介绍两种基于逆重力模型估算距离衰减系数及节点规模大小的解析方法。

2.1.1 线性模型算法

线性回归法(linear regression method)和线性规划法(linear programming method)是目前求解逆重力模型的常规算法, 基本思路都是对重力模型进行函数变换, 将其转化为求解线性系统的优化问题, 可

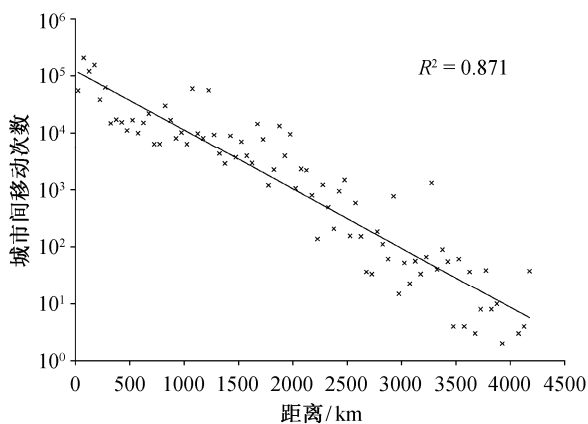


图2 基于签到数据的空间交互强度与距离关系散点图
Fig. 2 Relationships between spatial interaction intensity and displacements

以统称为线性模型算法。

将式(4)进行对数变换,可以得到

$$\ln G_{ij} = \ln k + \ln P_i + \ln P_j - \beta \ln d_{ij} \quad (i \neq j)。(7)$$

令 $X_i = \ln P_i$, $X_j = \ln P_j$, $\alpha_{ij} = \ln d_{ij}$ 和 $b_{ij} = \ln G_{ij} - \ln k$, 对于任意两个有空间交互作用的地理对象都可以得到线性方程:

$$X_i + X_j - \beta \alpha_{ij} = b_{ij} \quad (i \neq j)。(8)$$

在实际应用中,观察到的实际流量并不完全等于由重力模型得到的理论值,因此几乎不存在一组解使得由式(8)组成的线性系统成立,但是我们可以一定的允许条件下,寻找到 X_i , X_j 和 β 的最优解。

传统的线性回归方法将一系列 0-1 变量 O_i 和 D_i 引入式(8)得到回归方程:

$$\Sigma X_i O_i + \Sigma X_j D_j - \beta \alpha_{ij} = b_{ij} \quad (i \neq j)。(9)$$

其中 O_i , D_j 和 α_{ij} 是自变量, b_{ij} 是因变量,那么 X_i 和 X_j 成为变量 O_i 和 D_i 的回归系数。对每对有流量联系的地理实体 i 和 j 而言,式(9)中 $O_i = D_j = 1$,剩下的变量 O 和变量 D 均为零。可以使用最小二乘估计法(OLS)计算回归系数。

线性规划方法引入误差项 e_{ij} 来表示理论流量与实际流量的偏差,将误差项作为优化对象,并构建线性规划模型。例如,可以使用最小化最大误差(MINIMAX)和最小化误差和(MAD)等优化准则建立目标函数^[31]。但是线性规划模型中所有约束条件都必须满足的约束过于严格,O'Kelly 等^[30]提出用目标优化方法(goal programming, GP),在考虑约束条件的优先级情况下,重新构建线性规划模型:

$$e_{ij} = X_i + X_j - \beta \alpha_{ij} - b_{ij} \quad (i \neq j)。(10)$$

这一类方法在处理节点和连接均为海量的大规模网络交互问题时,对估算结果的稳定性和求解时间的长短都是一个很大的挑战,需要进一步改进和优化。

2.1.2 代数法及简化算法

Shen^[32]基于代数的方法(algebraic method, AM),根据式(4)推导出地理对象引力 P 的解析解,提供了一种新的求解逆重力模型的方法。对于由 n 个地理对象组成的交互网络,根据式(4)可以得到 $n(n-1)/2$ 个等式,将它们左右分别相乘,得到:

$$P_1 P_2 \cdots P_n = \sqrt[n]{\prod_{i=1, j=i+1}^n (G_{ij} d_{ij}^\beta / k)} \quad (i \neq j)。(11)$$

再将等式组中包含 P_1 的等式左右相乘,得到

$$P_1^{n-2} (P_1 P_2 \cdots P_n) = \prod_{j=2}^n (G_{1j} d_{1j}^\beta / k) \quad (i \neq j)。(12)$$

将式(11)和(12)相除,就能得到 P_1 的解析解,同

理可以得到 P_2, P_3, \dots, P_n 的解析式。简化的代数法 SAM^[33]与 AM 方法相似,但是仅使用整个网络的部分流量数据,减少了参与运算的等式,从而降低了运算的复杂度:

$$P_i = \sqrt[n-2]{\prod_{j=2}^n (G_{1j} d_{1j}^\beta / k)^{1/n-1}} \sqrt{\prod_{i=1, j=i+1}^n (G_{ij} d_{ij}^\beta / k)} \quad (i \neq j)。(13)$$

该类方法需要对观测流量为 0 的交互进行特殊处理,一般将 G_{ij} 的值设置为 1 或者任意常数。此外距离衰减系数 β 也需要经验预先设定,通过尝试对 β 取不同的数值,最终选择使得估算流量和观测流量拟合度最好的 β 值。代数法寻找最优距离衰减系数的方法十分繁琐,在处理稀疏交互网络(即部分节点间的交互流量缺失)时,难以得到较好的拟合结果。

2.2 模拟法

2.2.1 粒子群优化算法

粒子群优化算法(particle swarm optimization, PSO)是 Kennedy 等^[34]受鸟群捕食行为的启发,提出的一种随机搜索算法。PSO 算法将优化问题的每个解看做一个粒子,初始每个粒子都获得一个随机位置 $X_i \in [X_{\min}, X_{\max}]$ 和初始速度 V_i , X_i^t 是 t 时刻粒子的位置,粒子 $t+1$ 时刻的位置 X_i^{t+1} 是由它的速度 V_i^{t+1} 决定的,计算公式见式(14)和(15)。粒子通过不停迭代直到找到满足给定的拟合优度(goodness of fit)的最佳位置,即最优解。

$$V_i^{t+1} = w V_i^t + c_1 r_1 (P_{\text{best}}^t - X_i^t) + c_2 r_2 (G_{\text{best}}^t - X_i^t), \quad (14)$$

$$X_i^{t+1} = X_i^t + V_i^{t+1}。 \quad (15)$$

Xiao 等^[29]2013 年首次将粒子群算法引入到重力模型求解问题中,把估算节点规模的问题转化为优化问题。对于包含 n 个节点的空间交互网络,粒子 X_i 被定义为 $X_i = \{P_1, P_2, \dots, P_n\}$, P 是每个节点的规模,可以选择估算流量与实际流量的偏差和作为适应度函数。距离衰减函数 β 可以看做待求变量,作为粒子 X_i 的一个维度带入求解,也可以给定不同的 β 得到相应的估算结果,并选取估算结果最优的 β 值。当问题规模较大时,粒子群优化算法搜索效率高,搜索速度快,并且能够较好地处理稀疏交互网络以及非幂律形式的距离衰减函数。

2.2.2 空模型法

对一组地理对象,根据两两之间的交互强度,可以构造一个带权的交互网络,其中每个节点对应

一个地理对象, 每条边的权重表示节点间的交互强度。基于该网络, 可以采用空模型(null model)法, 即构建一个与距离无关的交互网络模型, 并将空模型中流量分布与观测到的流量分布进行比较, 两者的差异能够反映空间距离的作用。在空模型中, 首先根据每个节点的流量计算该节点的总流量:

$$W_i = \sum_j G_{ij} \quad (16)$$

令 $F = \sum \sum G_{ij}$ 并且 $N = \sum \sum W_i W_j$, 则在空模型中, 节点 i 和 j 之间的流量为

$$G_{ij}^{\text{null}} = W_i W_j F / N \quad (17)$$

我们用实际流量与空模型流量的比值 $R_{ij} = G_{ij} / G_{ij}^{\text{null}}$ 作为衡量距离衰减影响的指标, 通过证明可以得到 $R_{ij} \sim f(d_{ij})$, 因此可以通过回归的方法得到距离衰减函数。

2.2.3 蒙特卡洛法

在观测到区域间交互强度统计特征的基础上, 可以采用蒙特卡洛(Monte Carlo)模拟法来估算距离衰减的影响。其基本思路是, 根据重力模型, 事件 m (即区域 i 和 j 之间产生一次交互, 通常是一次出行)的概率与 i, j 以及它们之间的距离有关, 可以表示为

$$\text{Prob}(m) \propto P_i P_j f(d_{ij}) \quad (18)$$

基于式(10), 通过设置不同的距离衰减函数, 可以在 n 个区域间随机产生大量的出行, 进而计算这些出行的统计特征, 并与观测到交互的统计特征进行比较。如果能够得到匹配的分布, 则表明该距离衰减函数设置较为合理。在 Liu 等^[7]的研究中, 根据上海市出租车获取的出发到达点对数据, 计算其步长和角度分布特征, 然后, 将研究区划分为 1 km^2 的栅格单元, 并以 LandScan 数据获取的人口密度数据作为每个栅格单元规模的估计, 最后根据式(18)并采用幂律衰减函数, 生成模拟的出行数据。实验表明, 当设置幂律衰减系数为 1.08 时, 模拟产生出行的距离分布以及角度分布与实际分布较为吻合, 这说明上海市出租车出行这个行为所展示的距离衰减函数可以用 $d^{-1.08}$ 来近似, 该数值与采用重力模型拟合得到的结果基本一致。蒙特卡洛法能够从微观角度, 根据每个个体的行为特征, 探究群体的宏观模式。

3 讨论

基于幂律距离衰减形式的重力模型, 以及其更

一般的形式——距离衰减函数的空间交互模型, 可以用于估计两个空间对象之间的交互强度。但是, Simini 等^[13]指出重力模型的不足。Thiemann 等^[35]试图采用重力模型对基于钱币追踪轨迹所反映的空间交互进行建模, 却发现重力模型不能很好地解释所观察到的交互强度特征。我们认为在其拟合中采用人口数作为地理对象的估计存在缺陷, 因为区域间交互除了与人口数有关, 还受其他因素影响, 如经济发展程度。正确应用重力模型应该采用本文总结的方法, 根据实际交互强度和距离估算每个地理对象的理论大小或者引力, 同时得到最适合的距离衰减系数。

在一些应用中, 将每个地理对象的总流量作为其规模的近似。然而需要指出的是, 如果空间交互完全遵循重力模型, 每个节点的交互总量与其规模存在差异。图 3 通过一组模拟数据展示了这一点。图 3(a)为 500 个随机生成的点, 其用于计算交互强度的规模通过点的大小表现。在生成过程中, 控制点的位置为随机分布, 但点的规模表现出正的空间自相关, 其基于倒数权重的 Moran's I 值为 0.960。设距离衰减函数为幂律形式, 并且 $\beta = 1.5$, 可以计算每两个点之间的交互强度, 进而得到每个点的交互总量(图 3(b))。图 3(c)为节点规模和交互总量的散点图, 可以看出它们存在一定相关性, 但并不很强。将每个节点的规模和交互总量进行归一化处理并计算其差值, 得到图 3(d), 其中三角形符号表示交互总量相对低于规模的节点, 方形符号表示情况相反的节点。从图 3(d)可以看出, 在节点, 尤其是规模较大的节点, 分布密集的区域, 相对交互总量大于相对规模, 而在研究区边缘以及节点分布较为稀疏的区域, 相对交互总量小于相对规模。以上差异反映了区位的作用, 即具有区位优势节点往往会拥有比其规模地位更高的交互总量。图 3 也展示了重力模型与最大熵模型的差异, 最大熵模型本质上是在节点总流量已知的前提下在节点间分配流量, 是一个边界确定的封闭系统; 而重力模型根据规模和距离衰减函数计算交互量, 是一个开放系统。

另外, 比较图 3(a)和(b), 可以看出从交互总量上看, 其高值节点分布较节点规模分布更为集中。这种现象的形成与两个因素有关: 首先, 由于距离衰减作用的存在, 使得短距离交互的频率增加; 其次, 由于空间自相关(spatial autocorrelation), 造成规模较大的节点往往在空间上集中分布(参见图 3

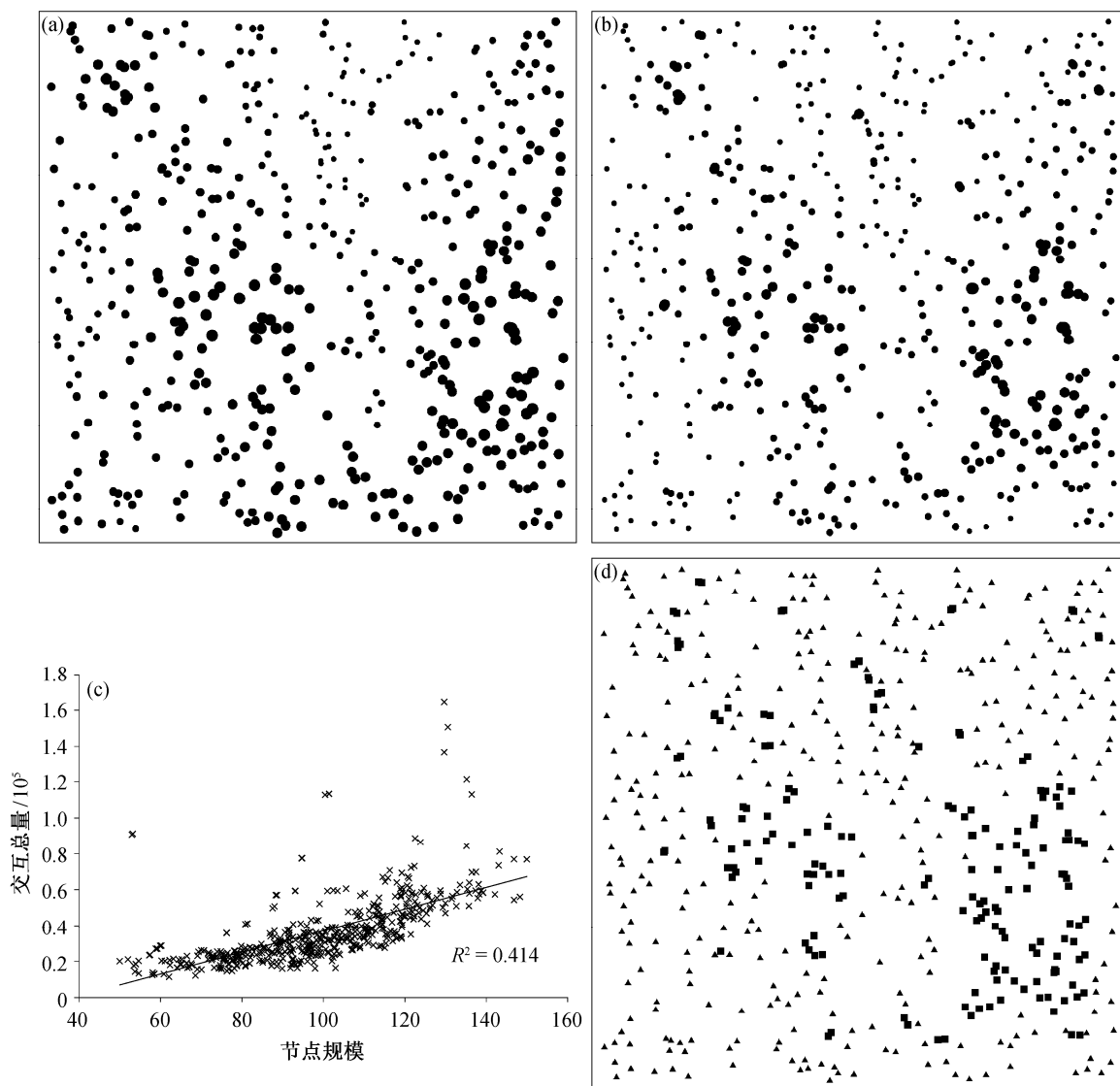


图3 500个随机生成的空间节点(a), 根据重力模型生成的节点交互总量分布(b), 交互总量和节点规模的相关性(c), 以及交互总量和节点规模的相对差异空间分布(d)

Fig. 3 Positions and sizes of 500 random nodes (a), total interaction distribution of each pair of nodes using the gravity model (b), relationship between total interaction flows and sizes for 500 random nodes (c), and spatial distribution of relative difference between total interaction flows and node sizes (d)

(a)), 这进一步提高了短距离交互的数量。Liu 等^[7]指出基于出租车数据获取的上海市居民出行距离分布呈现指数截断的幂律分布, 就是由于幂律距离衰减函数以及城市人口密度的空间分布共同作用, 降低了长程移动频率所造成的。而图 2 展示的指数步长分布, 也可以用 $\beta=0.8$ 的重力模型加以解释, 该衰减系数量化了城市间移动受距离的影响, 与 Xiao 等^[29]基于城市间航空客流量得到的结果接近。

如前所述, 距离衰减作用在宏观聚集层面以及微观个体层面均有影响。在微观个体层面, 由于距离衰减作用, 降低了长距离移动的概率。在 Lévy 飞

行模型中, 个体移动步长为幂律分布。假定空间是均值的, 其反映的距离衰减与重力模型一致。然而, 现有研究已经揭示人以及动物的移动受到地理环境的影响, 造成不同的分布特征^[36], 并且人的移动存在较多的重访点(如家、工作地点)^[37], 这也影响最终观察到的步长分布。在实践中, 由于难以采集个体较长时间的轨迹, 个体间差异也较大, 难以直接对个体移动受距离的影响进行定量分析。因此往往利用大量个体的移动轨迹分析其模式。然而, 由于个体差异的存在^[4,38], 使得我们即使在宏观聚集层面得到距离衰减函数, 也不能简单地推测在个体层

面的移动简单受该函数影响。例如,在城市尺度的移动中,每个人的通勤活动占据了较大比重。对于每个个体而言,家和工作地点的距离分布频率较高,未必展现距离衰减效应。而在群体层面,由于长距离通勤的个体比例相对较低,这使得群体的移动步长会表现出距离衰减。因此,要基于大数据研究个体层面的移动模式以及聚集层面的空间交互,建立微观和宏观之间的桥梁,需要深刻理解人的空间行为特征。

4 结论

本文对空间相互作用模型的相关研究工作进行梳理,讨论了距离衰减函数的多种形式,并基于重力模型介绍了估算距离衰减系数的不同方法及其特点。地理学第一定律定性地描述了距离对空间交互的影响,随着信息通讯技术的快速发展和广泛应用,我们能从更多的数据源对距离因素进行定量的刻画。空间交互模型已广泛应用于交通预测、疾病传播等众多领域,距离因素在其中如何发挥作用,是地理学者最为关注的问题。不同学者对指数和幂律两种形式的距离衰减函数持不同的看法,但是负幂律函数具有的重尾分布和无标度等特征,使其更适用于大尺度的复杂空间交互系统,通常我们的研究对象均具有这种特性,此外距离衰减系数也能够表征不同类型的交互活动,因此我们认为基于负幂律衰减的重力模型能够更好地刻画大尺度的空间范围内地理实体之间的相互作用。在此基础上,本文总结了一些距离衰减系数的估算方法:线性模型法和代数法是最常见的模型解析估算方法,蒙特卡洛模拟法和空模型法是经典的模拟估算方法,而粒子群优化算法为逆重力模型的求解提供了新的思路。每种距离衰减系数的估算方法都有其适用的条件,在实际应用中需要根据具体问题选择相应的方法。本文同时比较了重力模型与最大熵模型,指出应用重力模型中应该注意的问题,探讨了距离衰减在微观个体和宏观聚集两个层面产生影响的差异。

参考文献

- [1] Tobler W. A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 1970, 46(2): 234–240
- [2] Miller H J. Tobler's First Law and spatial analysis. *Annals of the Association of American Geographers*, 2004, 94(2): 284–289
- [3] Brockmann D, Hufnagel L, Geisel T. The scaling laws of human travel. *Nature*, 2006, 439: 462–465
- [4] González M C, Hidalgo C A, Barabási A L. Understanding individual human mobility patterns. *Nature*, 2008, 453: 779–782
- [5] Jiang B, Yin J, Zhao S. Characterizing the human mobility pattern in a large street network. *Physical Review E*, 2009, 80(2): 021136
- [6] Liang X, Zheng X, Lv W, et al. The scaling of human mobility by taxis is exponential. *Physica A: Statistical Mechanics and Its Applications*, 2012, 391(5): 2135–2144
- [7] Liu Y, Kang C, Gao S, et al. Understanding intra-urban trip patterns from taxi trajectory data. *Journal of Geographical Systems*, 2012, 14(4): 463–483
- [8] Wilson A G. The use of entropy maximizing models in the theory of trip distribution, mode split and route split. *Journal of Transport Economics and Policy*, 1969, 3: 108–126
- [9] Colwell P F. Central place theory and the simple economic foundations of the gravity model. *Journal of Regional Science*, 1982, 22(4): 541–546
- [10] Stouffer S A. Intervening opportunities: a theory relating to mobility and distance. *American Sociological Review*, 1940, 5(6): 845–867
- [11] Liben-Nowell D, Novak J, Kumar R, et al. Geographic routing in social networks. *Proceedings of the National Academy of Sciences of the USA*, 2005, 102: 11623–11628
- [12] Noulas A, Scellato S, Lambiotte R, et al. A tale of many cities: universal patterns in human urban mobility. *PLoS ONE*, 2012, 7(5): e37027
- [13] Simini F, González M C, Maritan A, et al. A universal model for mobility and migration patterns. *Nature*, 2012, 484: 96–100
- [14] 刘瑜, 肖昱, 高松, 等. 基于位置感知设备的人类移动性研究综述. *地理与地理信息科学*, 2011, 27(4): 8–13
- [15] Cheng Z Y, Caverlee J, Lee K, et al. Exploring millions of footprints in location sharing services // *ICWSM*. Barcelona, 2011: 81–88
- [16] Kang C, Ma X, Tong D, et al. Intra-urban human mobility patterns: an urban morphology perspective. *Physica A: Statistical Mechanics and its Applications*, 2012, 391(5): 2135–2144

- 2012, 391(4): 1702–1717
- [17] Kang C, Zhang Y, Ma X, et al. Inferring properties and revealing geographical impacts of intercity mobile communication network of China using a subnet data set. *International Journal of Geographical Information Science*, 2013, 27(3): 431–448
- [18] Ratti C, Sobolevsky S, Calabrese F, et al. Redrawing the map of Great Britain from a network of human interactions. *PloS ONE*, 2010, 5(12): e14248
- [19] Liu Y, Wang F, Kang C, et al. Analyzing relatedness by toponym co-occurrences on web pages. *Transactions in GIS*, 2014, 18(1): 89–107
- [20] Onnela J P, Arbesman S, González M C, et al. Geographic constraints on social network groups. *PLoS ONE*, 2011, 6(4): e16939
- [21] 周一星. 城市地理学. 北京: 商务印书馆, 1995: 360
- [22] Haggett P, Cliff A D, Frey A. *Locational analysis in human geography*. London: Edward Arnold, 1977: 339
- [23] Wang F. Measurement, optimization and impact of health care accessibility: a methodological review. *Annals of the Association of American Geographers*, 2012, 102(5): 1104–1112
- [24] Carey H C. *Principles of social sciences* (Vol. 3). Philadelphia: JB Lippincott & Company, 1867
- [25] 陈彦光. 空间相互作用模型的形式、量纲和局域性问题探讨. 北京大学学报: 自然科学版, 2009, 45(2): 333–338
- [26] Chen Y. A wave-spectrum analysis of urban population density: entropy, fractal, and spatial localization. *Discrete Dynamics in Nature and Society*, 2008: doi: 10.1155/2008/728420
- [27] Fotheringham S. Spatial structure and distance-decay parameters. *Annals of the Association of American Geographers*, 1981, 71(3): 425–436
- [28] Gao S, Wang Y, Gao Y, et al. Understanding urban traffic flow characteristics: a rethinking of betweenness centrality. *Environment and Planning B: Planning and Design*, 2013, 40(1): 135–153
- [29] Xiao Y, Wang F, Liu Y, et al. Reconstructing gravitational attractions of major cities in china from air passenger flow data, 2001–2008: a particle swarm optimization approach. *The Professional Geographer*, 2013, 65(2): 265–282
- [30] O’Kelly M E, Song W, Shen G. New estimates of gravitational attraction by linear programming. *Geographical Analysis*, 1995, 27(4): 271–285
- [31] Schrage L. *Linear, integer, and quadratic programming with LINDO*. 3rd ed. Palo Alto: The Scientific Press, 1986
- [32] Shen G. Estimating nodal attractions with exogenous spatial interaction and impedance data using the gravity model. *Papers in Regional Science*, 1999, 78(2): 213–220
- [33] Shen G. Reverse-fitting the gravity model to inter-city airline passenger flows by an algebraic simplification. *Journal of Transport Geography*, 2004, 12(3): 219–234
- [34] Kennedy J, Eberhart R C. Particle swarm optimization. *Proceedings of the IEEE International Conference on Neural Networks*, 1995, 4: 1942–1948
- [35] Thiemann C, Theis F, Grady D, et al. The structure of borders in a small world. *PLoS ONE*, 2010, 5(11): e15422
- [36] Humphries N E, Queiroz N, Dyer J R, et al. Environmental context explains Lévy and Brownian movement patterns of marine predators. *Nature*, 2010, 456: 1066–1069
- [37] Song C, Koren T, Wang P, et al. Modelling the scaling properties of human mobility. *Nature Physics*, 2010, 6(10): 818–823
- [38] Yuan Y, Raubal M, Liu Y. Correlating mobile phone usage and travel behavior — a case study of Harbin, China. *Computers, Environment and Urban Systems*, 2012, 36(2): 118–130