专题：统计物理和复杂系统

社会引力定律追根溯源*

闫小勇 1)2)†

1) (北京交通大学, 交通系统科学与工程研究院, 北京 100044)

2) (电子科技大学, 复杂性实验室, 成都 611731)

(2019 年 11 月 4 日收到; 2019 年 11 月 19 日收到修改稿)

在交通出行、人口迁移、商品贸易、信息流通、社会交往、科研合作等大量人、物、信息的空间流动现象中, 都存在类似万有引力定律的规律, 即两地之间的某种流动量与两地活力的乘积成正比、与两地距离的幂成反比. 类比万有引力定律建立的引力模型也在交通出行分布预测、人口迁移量预测、地区间贸易量预测等诸多方面获得了广泛应用. 但复杂的社会系统中为何会有这样简单的引力定律存在? 这是个非常有趣也有价值的问题. 本文对从统计物理学、微观经济学和博弈论等不同视角探索社会引力定律根源的研究进行了综述.

关键词：社会引力定律, 最大熵原理, 离散选择模型, 拥挤博弈

PACS：89.75.–k, 89.65.–s, 89.70.Cf, 02.50.Le **DOI:** 10.7498/aps.69.20191686

1 引 言

预测地点间人、物、信息的流动是社会学、经济学、人口学、交通科学、网络科学等诸多学科长期以来的一个重要研究主题[1,2]. 一百多年来, 研究者们陆续提出了多种预测地点间流动量的模型 (在经济学中被称为空间交互模型[3]), 其中最有影响力的是引力模型[4] (gravity model, 也译为“重力模型”). 引力模型在许多方面都获得了广泛应用. 例如在人口学中, 引力模型被用来预测地区间的人口迁移量[5]; 在经济学中, 引力模型被用来预测国家间的商品贸易量[6]; 在交通科学中, 引力模型被用来预测地点间的交通出行量[7]; 在网络科学中, 引力模型被用来评估网络节点的相似性[8]、影响力[9,10]等. 引力模型之所以受到如此青睐, 是由于在许多空间流动现象中都存在类似万有引力定律[11]的规律, 即两地之间的某种流动量正比于两地“活力”(多用地点人口数量、经济发展水平、进出流动量等表示) 的乘积, 反比于两地之间空间距离的幂函数. 这种规律被称为社会引力定律[12].

早在 1846 年, Desart[13]就在比利时铁路客运数据中发现, 两个车站间的客运量正比于两个车站所在地的人口数乘积, 反比于两车站间距离的2.25次幂. 这可能是关于发现社会引力定律的最早记载[14], 其发现时间比 Carey[15]在人口迁移现象中发现社会引力定律的时间 (1858 年) 还要早 12 年. 后来, Ravenstein[16]从 19 世纪 70—80 年代英国人口普查数据中也发现两地人口迁移量与两地人口乘积成正比、与两地距离成反比的现象. 而Reilly[17]在 1929 年的著作中研究零售市场问题时也发现, 零售中心从其周围某个城镇吸引到的顾客数量与该零售中心的规模成正比, 与两地间的距离平方成反比. 此后的几十年中, 研究者在交通[18]、人口[19]、经济[20]等领域又陆续发现了许多服从社会引力定律的现象.

近年来, 随着现代电子与信息技术的不断发展, 有越来越多的手段 (如 IC 卡、GPS、手机、社

* 国家自然科学基金 (批准号: 71822102, 71621001, 71671015, 61304177) 资助的课题.

† 通信作者. E-mail: yanxy@bjtu.edu.cn

交网站等)可长期记录人、物、信息在空间中的流动数据[21,22]. 通过对这些大数据进行统计分析, 统计物理与复杂系统领域的学者在很多复杂系统中都发现了符合社会引力定律的现象. 例如, Viboud 等[23]通过分析美国人口普查数据, 发现两郡之间的通勤出行量与郡人口、郡间距离的关系符合社会引力定律; Jung 等[24]通过分析韩国高速公路收费数据, 发现两城市间的高速公路交通量与城市人口、城市间距离的关系符合社会引力定律; Krings 等[25]分析了比利时 250 万名匿名手机用户在 571 个城镇之间的手机通讯数据, 发现城镇间通讯量与城镇人口、城镇间距离的关系符合社会引力定律; Balcan 等[26]通过分析国际航空运输协会的航空客运数据, 发现全球 29 个国家主要机场之间的航空运输量与机场服务区域人口、机场间距离的关系符合社会引力定律; Kaluza 等[27]分析了全球船舶港口自动识别系统数据, 发现全球 951 个港口间的海运量与港口进出量、港口间距离的关系符合社会引力定律; Pan 等[28]通过分析美国科学信息研究所 2003—2010 年的论文数据, 发现全球 18199 个城市之间的科研合作强度与城市整体科研水平、城市间距离的关系符合社会引力定律; Goh 等[29,30]分析了韩国首尔市地铁网络和常规公交网络中的刷卡数据, 发现车站间客流与车站上下车客流、车站间距离的关系符合社会引力定律; Levy 和 Goldenberg[31]分析了 Facebook 等四类社交网络数据, 发现两人交往概率与两人距离之间的关系符合社会引力定律. 诸如此类的研究还有很多. 可以预见, 社会引力定律还会在更多的复杂系统中被发现.

从以上研究回顾中可以看出, 交通出行、人口迁移、商品贸易、信息流通、社会交往、科研合作等大量空间交互现象都符合社会引力定律. 但在诸多复杂的社会系统中为何会有如此简单的社会引力定律存在? 这是一个困扰人类上百年的问题, 也吸引了很多不同领域的学者对其追根溯源[32,33]. 从 20 世纪 60 年代起, 统计物理领域的学者用最大熵原理[34]、经济学领域的学者用效用理论[35,36]相继对社会引力定律进行了解释. 最近, 复杂系统领域的研究人员又从博弈论角度探索了社会引力定律的根源[37]. 本文将对这些从不同视角探索社会引力定律根源的研究进行综述.

2 社会引力定律的统计物理解释

2.1 最大熵原理导出引力模型

Wilson[34]以交通系统中地点间的出行分布问题为背景, 用统计物理学中的最大熵原理[38]对社会引力定律给出了最早的理论解释. 在交通系统的出行需求预测工作中, 常采用的模型框架是“四阶段法”, 即依次进行出行生成预测、出行分布预测、交通方式划分和交通分配[7]. 其中, 在出行生成预测阶段可得到地点 i 的出行发生量 O_i 和出行吸引量 D_i. 而在出行分布预测阶段, 则要求预测出的地点间出行分布量 T_{ij} 既满足发生量约束 $\sum_j T_{ij}=O_i$, 又满足吸引量约束 $\sum_i T_{ij}=D_j$. Wilson用最大熵原理导出了满足以上两类约束的引力模型: 在交通系统总出行量为 T 的情况下, 系统的宏观状态是任意两个地点 i, j 之间的出行量为 T_{ij}, 则系统的微观状态数为

$$\varOmega=\frac{T!}{\prod_i\prod_j T_{ij}!}. \tag{1}$$

Wilson 为地点间出行量 T_{ij} 又增加了一个出行成本约束

$$\sum_i\sum_j T_{ij}c_{ij}=C, \tag{2}$$

其中 c_{ij} 为从地点 i 到 j 的出行成本, C 为系统可支配的总成本. 那么, 系统最可能出现的出行分布就是以下最大熵模型的解:

$$\begin{aligned}\max \ln\varOmega = {}& \ln\frac{T!}{\prod_i\prod_j T_{ij}!},\\ \text{s.t.}\quad & \sum_j T_{ij}=O_i,\\ & \sum_i T_{ij}=D_i,\\ & \sum_i\sum_j T_{ij}c_{ij}=C.\end{aligned} \tag{3}$$

用拉格朗日乘子法可求得该模型的解为

$$T_{ij}=a_iO_ib_jD_j\mathrm{e}^{-\gamma c_{ij}}, \tag{4}$$

其中 γ 是模型参数, $a_i=1\Big/\sum_j b_jD_j\mathrm{e}^{-c_{ij}}$ 和 $b_j=1\Big/\sum_i a_iO_i\mathrm{e}^{-\gamma c_{ij}}$ 是两组相互依赖的迭代因子. 这

就是带有负指数成本函数的双约束引力模型[7].

2.2 成本与距离的对数关系

经典引力模型中的距离函数是幂律函数[4]. 但 (4) 式中用最大熵原理导出的引力模型, 其成本函数是负指数函数 $\mathrm{e}^{-\gamma c_{ij}}$, 这与经典引力模型中使用的距离幂函数不同. 简单地看, 直接将 $\gamma c_{ij}=\beta \ln d_{ij}$ 这种成本与距离之间的对数关系代入 (4) 式, 即可得到经典的双约束引力模型 $T_{ij}=a_i O_i b_j D_j d_{ij}^{-\beta}$. 但为什么成本与距离之间具有这种对数关系, 之前并未得到很好的解答[39].

Yan 等[40] 为社会引力定律中距离的幂函数形式提供了一种解释. 他们认为, 交通系统中的出行成本 c 主要由出行时间 t 和货币费用 m 两部分组成, 可以表示为二者的加权和形式 $c\approx \eta t+\mu m$[41]. 其中, 货币费用通常与出行距离 d 具有近似线性关系 $m\approx \nu d$[41]. 他们通过分析出行日志数据[42] 中出行时间与距离的关系, 发现出行时间与距离也具有近似对数关系 $t\approx \phi \ln d+\psi$[43]. 这种对数关系源于人们在不同出行距离所采用交通方式的速度差异[44], 例如人们在进行几百米的出行时往往是步行或骑自行车, 而几十千米时就要使用公交、小汽车等交通方式, 当几百千米时就要乘坐更快速的火车或飞机了, 这使得出行时间和距离之间呈现非线性关系. 综上, 可得到出行成本与距离之间的关系为 $c=\eta\phi \ln d+\eta\psi+\mu\nu d$. 进而结合最大熵原理, 还可导出群体出行距离的截尾幂律分布 $P(d)\sim d^{-\beta}\mathrm{e}^{-d/\kappa}$. 这一研究不仅能解释社会引力定律中距离函数的幂律形式, 还解释了一大批出行距离分布的实证统计结果[45–50].

3 社会引力定律的微观经济学解释

Wilson 用最大熵原理为社会引力定律提供了一个非常合理的宏观解释. 但是, 最大熵原理仅能给出系统最可能的宏观分布状态, 并不考虑系统中个体选择出行目的地的微观决策过程. 从出行目的地选择决策行为的角度来看, 社会引力定律的微观底层机制仍未得到满意解答[32,33]. 而经济学家则很早就开始用效用理论来研究个体选择出行目的地的微观决策行为[35,36]. 早期的研究使用确定效用理论来解释社会引力定律[35], 而影响更为广泛的研究则是基于随机效用理论的离散选择模型[36]. 离散选择模型是微观经济学中非常重要的描述个体选择行为的方法, 在交通系统出行选择问题上已获得了非常广泛的应用[51,52]. 本节将介绍用离散选择模型对个体出行目的地选择行为进行建模以及导出引力模型的方法.

3.1 目的地离散选择模型

离散选择模型被用来描述、解释和预测决策者对多个离散的选项进行选择的行为. 一次离散选择行为通常包含以下要素[51].

1) 决策者, 即做出选择行为的主体. 在出行目的地选择行为中, 决策者就是指位于某起点选择目的地的一个出行者.

2) 选项, 即供决策者选择的多项事物. 在出行目的地选择行为中, 选项是指供出行者选择的目的地.

3) 选项属性. 决策者选择选项时会考虑诸如价格、质量等因素, 每一种因素称为一个属性. 在出行目的地选择行为中, 选项属性一般包括各地点的活力、到各地点的距离或出行成本等.

4) 决策准则, 即决策者在做出选择时的行为准则. 最常用的是效用最大化准则[53], 即决策者在所有选项中选择效用最高的选项. 此处的效用是指决策者选择某个选项所能获得的满意程度. 在出行目的地选择问题中, 可用下式来表达这一准则:

$$U_{ij}>U_{ik},\ \forall\, k\in K-\{j\}, \tag{5}$$

其中 U_{ij} 是位于地点 i 的出行者所选择的目的地 j 的效用, U_{ik} 则是其他任一目的地的效用, K 是所有目的地的集合.

(5) 式的效用值 U_{ik} 本质上是指目的地 k 所具有的真实效用. 根据效用最大化准则, 决策者会选择所有选项中真实效用最大的一个选项. 但实际选择问题中, 决策者有可能并不是总选择一个固定的选项, 而是以不同的概率选择不同的选项. 这可能是决策者受到一些内部或外部因素的影响, 对选项效用的认知发生了变化. 但研究者并不能直接观测真实效用变化, 只能观测选项的确定效用值[54]. 在离散选择模型中, 这种概率选择的情况是用随机效用理论来处理的. 此时目的地 k 的真实效用表示为

$$U_{ik}=V_{ik}+\varepsilon_{ik}, \tag{6}$$

其中 V_{ik} 是研究者直接观测到的选项 k 的确定效用, ε_{ik} 是描述 V_{ik} 与真实效用 U_{ik} 偏差程度的随机项.

3.2 离散选择模型导出引力模型

根据 (5) 式和 (6) 式可知, 在出行目的地选择问题中, 所处 i 地点的决策者选择目的地 j 能获得的效用是 $U_{ij}=V_{ij}+\varepsilon_{ij}$, 则目的地 j 被 i 点决策者选择的概率是

$$P_{ij}=\Pr(V_{ij}+\varepsilon_{ij}>V_{ik}+\varepsilon_{ik},\ \forall k\in K-\{j\}),\quad (7)$$

其中 K 是所有备选目的地集合 (不包括起点 i).

如果 (7) 式中所有随机项 ε 都服从相互独立同分布的标准 Gumbel 分布 $F(\varepsilon)=\mathrm{e}^{-\mathrm{e}^{-\varepsilon}}$, 那么决策者选择目的地 j 的概率就是

$$\begin{aligned}P_{ij}&=\Pr(\varepsilon_{ik}<V_{ij}-V_{ik}+\varepsilon_{ij},\forall k\in K-\{j\})\\&=\Pr(\varepsilon_{ij}=x)\cdot\prod_{k\neq j}\Pr(\varepsilon_{ik}<V_{ij}-V_{ik}+x)\\&=\int_{-\infty}^{\infty}\frac{\mathrm{d}F(x)}{\mathrm{d}x}F(V_{ij}-V_{ik}+x)\mathrm{d}x\\&=\int_{-\infty}^{\infty}\mathrm{e}^{-x}\mathrm{e}^{-\mathrm{e}^{-x}}\prod_{k\neq j}\mathrm{e}^{-\mathrm{e}^{-(V_{ij}-V_{ik}+x)}}\mathrm{d}x\\&=\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{e}^{-x}\sum_k\mathrm{e}^{-(V_{ij}-V_{ik})}}\mathrm{e}^{-x}\mathrm{d}x\\&=\frac{\mathrm{e}^{V_{ij}}}{\sum_k\mathrm{e}^{V_{ik}}}.\end{aligned}\quad (8)$$

这就是 Logit 模型[54], 是微观经济学中应用非常广泛的离散选择模型.

由 Logit 模型可以很容易地导出引力模型: 假设目的地 j 对 i 点出行者的确定效用主要受目的地活力 A_j 和地点 i 到 j 的距离 d_{ij} 影响, 即

$$V_{ij}=\ln A_j-\beta\ln d_{ij}+\gamma.\quad (9)$$

如果所有出行者之间的选择决策行为无差异, 那么从点 i 到 j 的出行量 T_{ij} 就是从 i 出发的总出行量 O_i 乘以出行者们选择 j 点的概率, 即

$$T_{ij}=O_iP_{ij}=O_i\frac{\mathrm{e}^{V_{ij}}}{\sum_k\mathrm{e}^{V_{ik}}}=O_i\frac{A_j\cdot d_{ij}^{-\beta}}{\sum_k A_k\cdot d_{ik}^{-\beta}},\quad (10)$$

这就是经典的单约束引力模型[7].

3.3 收益与活力的对数关系

从 (9) 式中可以看出, 目的地的效用 V_{ij} 实际上是目的地活力为出行者带来的收益 $\ln A_j$ 与出行者到目的地的成本 $\beta\ln d_{ij}$ 的差值. 2.2 节已对出行成本与距离的对数关系进行了解释, 本节将解释为何使用地点活力的对数来表达地点会带给出行者收益.

根据心理物理学中著名的 Weber-Fechner 定律[55], 人的直观感觉差异 $\mathrm{d}p$ 正比于某种物理刺激强度 W 的相对变化量 $\mathrm{d}W/W$, 即 $\mathrm{d}p=\kappa\mathrm{d}W/W$, 其中 κ 是个常数. 据此可导出人的直观感觉程度 p 与物理刺激强度之间的对数关系 $p=\kappa\ln(W/W_0)$, 其中 W_0 可以解释为刺激阈值. Weber-Fechner 定律在行为经济学中被广泛应用于收益函数[55], 即认为选项带给决策者的收益是决策者对选项属性的实际度量指标 (如商品的价格、反映地点活力的人口数等) 的直观感觉. 在本文即将介绍的目的地选择博弈研究工作中, Weber-Fechner 定律导出的收益活力对数关系也会被再次使用.

4 社会引力定律的博弈论解释

基于随机效用理论的离散选择模型从个体选择决策角度为社会引力定律提供了新的解释, 但这类模型并未考虑实际出行目的地选择过程中个体之间的相互作用. 在包括交通系统在内的社会经济复杂系统中, 个体间的相互作用是一个非常普遍的现象, 而博弈论则是社会学、经济学以至军事、政治等社会科学学科研究个体相互作用的一个重要科学工具[56]. 最近, Yan 和 Zhou[37] 从博弈论的角度解释了社会引力定律的可能根源. 他们将出行者选择目的地的过程刻画为一种拥挤博弈[57,58], 体现了出行者个体之间的相互作用. 本节将介绍这种目的地选择博弈模型的基本框架及其导出引力模型的方法.

4.1 目的地选择博弈模型

出行目的地选择问题可以看作是多个参与者进行的博弈. 每个参与者在面对多个可供选择的目的地时, 总会选择带给自己收益 (或称效用) 最大的目的地. 位于地点 i 的个体选择目的地 j 获得的收益 U_{ij} 由两部分组成.

1) 出行成本 $C_{ij}+g(T_{ij})$, 即负收益, 与 i 到 j 的固定出行成本 C_{ij} 直接相关, 并且随两地间出行量 T_{ij} 的增加而增加. 其中 $g(T_{ij})$ 是一个增函数, 体现了路途上的拥挤效应.

2) 目的地带给选择者的收益 $h(A_j)-f(D_j)$, 取决于目的地的活力 A_j 和选择该目的地的人数

$D_j=\sum_i T_{ij}$. 其中$f(D_j)$是一个增函数, 体现了目的地的拥挤效应, 即目的地的收益会随着选择人数的增加而下降. 这种拥挤效应具有现实依据: 例如在城市出行中, 对于非通勤出行 (例如购物、娱乐等) 来说, 选择某一目的地的人数增加后可能导致环境不舒适或可获取资源的减少, 从而降低选择者的收益.

综上, i地点的出行者选择目的地j的收益可表示为

$$U_{ij}=h(A_j)-f(D_j)-C_{ij}-g(T_{ij}). \tag{11}$$

这种模型被命名为目的地选择博弈 (destination choice game, DCG) 模型[37]. 在 DCG 模型中, 当个体信息完备并总是选择使自己收益最大化的目的地时, 相同起点出行的所有个体都具有相等的收益, 没有人能通过单方面改变选择而增加自己的收益.

实际应用中, 需要先把 (11) 式中的各项成本和收益函数具体化. 根据出行成本与距离之间的对数关系 (见 2.2 节), 可令$C_{ij}\propto\ln d_{ij}$. 根据 Weber-Fechner 定律[55] (见 3.3 节), 则可将目的地活力收益函数$h(A_j)$、目的地拥挤成本函数$f(D_j)$和路途拥挤函数$g(T_{ij})$均表示为对数函数. 此时 (11) 式可写为

$$U_{ij}=\alpha\ln A_j-\beta\ln d_{ij}-\gamma\ln D_j-\ln T_{ij}, \tag{12}$$

其中α, β和γ是三个非负参数. 通过使用多种真实人类移动数据进行的模型测试结果[37]显示, DCG 模型相对于传统的引力模型[4]、介入机会模型[59]和新型辐射模型[60]、人口权重机会模型[61,62], 具有更高的预测精度.

4.2 拥挤博弈模型导出引力模型

若忽略 DCG 模型收益函数中的目的地拥挤成本, 则可得到一个简化的目的地选择博弈 (degenerated destination choice game, DDCG) 模型, 其收益函数如下:

$$U_{ij}=\alpha\ln A_j-\beta\ln d_{ij}-\ln T_{ij}, \tag{13}$$

这是一个典型的拥挤博弈模型[57,58]. 拥挤博弈可表示为一个四元组: $(N,R,(\varPsi_k)_{k\in N},(w_j)_{j\in R})$[58,63], 其中$N=\{1,2,\cdots,n\}$是所有参与博弈的个体 (在 DDCG 中就是某个出发地点i上数量为O_i的所有出行者), $R=\{1,2,\cdots,m\}$是备选资源集合 (在 DDCG 中就是备选目的地), $(\varPsi_k)\subseteq 2^R$是参与者k的策略空间, w_j是备选资源j的收益函数 (在 DDCG 中就是目的地j对i起点出行者的收益函数U_{ij}). 所有参与者的策略集合$S=S_1,\cdots,S_n$就是拥挤博弈的一个状态, 其中参与者k的策略$S_k\in\varPsi_k$. 资源j的拥挤程度$n_j(S)$(在 DDCG 中就是目的地j的路途拥挤程度) 则表示在状态S下选择该资源的参与者数量 (在 DDCG 中就是从起点i到目的地j的出行量T_{ij}). 据此可知, DDCG 中每个出行者获取的收益不仅取决于其所选择的目的地的活力与距离, 还受选择同样目的地的其他出行者数量的影响, 即个体之间的相互作用. 根据拥挤博弈理论中的势函数[58,63]定义, 可知 DDCG 的势函数为

$$\begin{aligned}\phi&=\sum_{j\in R}\sum_{k=1}^{T_{ij}(S)}U_{ij}(k)\\&=\sum_{j\in R}\sum_{k=1}^{T_{ij}(S)}(\alpha\ln A_j-\beta\ln d_{ij}-\ln k).\end{aligned} \tag{14}$$

最大化此势函数的策略集合, 就是拥挤博弈的纳什均衡解[57,63].

为求解 DDCG 的纳什均衡解, 可将T_{ij}视为一个连续变量, 并令势函数最大化, 即可得到一个最优化模型

$$\begin{aligned}&\max\phi(S)=\sum_j\int_0^{T_{ij}}(\alpha\ln A_j-\beta\ln d_{ij}-\ln x)\mathrm{d}x,\\&\text{s.t.}\quad\sum_j T_{ij}=O_i.\end{aligned} \tag{15}$$

用拉格朗日乘子法可求得该模型的解为

$$T_{ij}=O_i\frac{A_j^{\alpha}d_{ij}^{-\beta}}{\sum_j A_j^{\alpha}d_{ij}^{-\beta}}, \tag{16}$$

是一个双参数的单约束引力模型. 若令参数$\alpha=1$, 则可转换为标准的单约束引力模型, 见 (10) 式.

4.3 DDCG 模型与 Logit 模型的对比

DDCG 模型与 3.2 节中介绍的离散选择 Logit 模型都体现了个体选择目的地的微观决策行为, 且都假设个体总是追求效用最大. 二者导出的引力模型在形式上也没有本质区别, 见 (10) 式和 (16) 式. 但二者的底层机制并不相同: Logit 模型假设个体对效用的认知具有随机性, 但并未考虑复杂系统中广泛存在的个体相互作用[56]; 而 DDCG 模型则考

虑了群体中个体间的相互作用.

著名的红蓝巴士问题[64]可以用来说明这两个模型的机制差异. 在红蓝巴士问题中, 出行者可选择的交通方式有小汽车与蓝巴士两种. 简单起见, 假设这两种交通方式的确定效用是相等的, 根据 Logit 模型可知二者的被选概率为$P_{\rm car}=P_{\rm blue}=1/2$. 现在把一半蓝巴士的颜色改为红色, 并把红蓝巴士看成是两种交通方式, 那么两者的确定效用也必然相等. 按照 Logit 模型的结果, 此时出行者选择这三种交通方式的概率为$P_{\rm car}=P_{\rm blue}=P_{\rm red}=1/3$. 然而, 从常识推断, 将半数巴士涂红仅会影响原先选择蓝巴士的人, 而小汽车的被选概率仍为$P_{\rm car}=1/2$, 红、蓝巴士则各为$P_{\rm red}=P_{\rm blue}=1/4$. 但 Logit 模型却高估了巴士被选择的概率, 低估了小汽车被选择的概率, 形成了典型的悖论[64].

在考虑了个体相互作用的 DDCG 模型中则不存在这种悖论: 假设总共有b辆巴士, 每辆巴士的固定收益值为$A_{\rm bus}$, 拥挤成本 (体现了个体之间相互作用的结果) 为$\ln(ax)$, 其中x代表乘坐这辆巴士的人数. 那么当把所有巴士看成同一交通方式时, 其收益值为$U_{\rm bus}=A_{\rm bus}-\ln(ax/b)$; 而把红蓝巴士看作两种交通方式时, 二者的收益值则均为$U_{\rm red}=U_{\rm blue}=A_{\rm bus}-\ln(2ax/b)$. 根据 DDCG 模型的均衡条件, 如果在$m$个人中选择巴士和小汽车的比例各为 1/2, 说明$U_{\rm car}(m/2)=U_{\rm bus}(m/2)=A_{\rm bus}-\ln[am/(2b)]$. 那么把红蓝巴士看作两种方式时, 一定有$U_{\rm car}(m/2)=U_{\rm red}(m/4)=U_{\rm blue}(m/4)$, 即选择红蓝巴士的比例各为 1/4, 与实际相符.

当然, 实际中的出行者对目的地效用 (即收益) 的感知可能会有一定的随机性. 此时可对 DDCG 模型进行进一步的扩展: 假设出行者对目的地收益的理解与可观测的目的地固定收益之间存在服从独立同 Gumbel 分布的随机性偏差, 根据 Fisk[65] 的证明, 此时的均衡解等价于如下最优化模型的解

$$\begin{aligned}\min Z(x)=&\sum_j\int_0^{T_{ij}}(-\alpha\ln A_j+\beta\ln d_{ij}+\ln x)\mathrm{d}x\\&+\frac{1}{\theta}\sum_j T_{ij}\ln T_{ij},\\ \text{s.t.}\quad&\sum_j T_{ij}=O_i,\end{aligned}\tag{17}$$

其解为

$$T_{ij}=O_i\frac{A_j^{\frac{\theta\alpha}{1+\theta}}d_{ij}^{-\frac{\theta\beta}{1+\theta}}}{\sum\limits_j A_j^{\frac{\theta\alpha}{1+\theta}}d_{ij}^{-\frac{\theta\beta}{1+\theta}}},\tag{18}$$

式中的参数θ体现了随机性程度对目的地选择结果的影响: 当$\theta\to\infty$时该模型退化为 DDCG 模型, 均衡结果与 (16) 式一致; 当$\theta=0$时, 个体将完全随机地选择目的地, 各目的地被选择的概率相等. 这一结果说明, DDCG 模型可以扩展到个体对目的地收益感知有随机性的情形. 但是, 基于随机效用理论的离散选择模型却无法退化到目的地收益固定的情形: 如果离散选择模型不假设个体对收益感知具有随机性, 那么所有个体都会选择固定收益最高的同一目的地. 这是因为离散选择模型并未考虑实际社会系统中普遍存在的个体相互作用, 即群体拥挤问题.

5 结 论

社会引力定律是诸多社会复杂系统中广泛存在的普适规律, 受到社会学、经济学、地理学、交通科学、统计物理与复杂系统等学科学者持久的关注[1,4,6,7,32,33,39]. 本文综述了解释社会引力定律存在根源的三类主要理论, 既包括传统的最大熵原理和效用理论, 也包括最近提出的目的地选择拥挤博弈. 其中, 最大熵原理为社会引力定律建立了一个统计物理的理论基础, 但其仅能给出系统最可能的宏观分布状态, 却无法反映系统中个体的微观决策行为. 从个体决策行为角度建立的随机效用离散选择模型为社会引力定律提供了微观经济学解释, 但其并未考虑在实际社会系统中普遍存在的个体相互作用. 而目的地选择博弈则为社会引力定律建立了一个更简洁的理论框架: 它不需像最大熵模型一样事先指定出行总成本作为额外的约束条件, 也不需像随机效用理论一样作出个体对效用感知具有随机性的假设, 除"个体追求收益最大化"这条经济学公理之外, 目的地选择博弈的核心假设就只有"群体拥挤影响收益"这一体现个体相互作用的基本事实. 与现有对社会引力定律的其他理论解释相比, 从博弈论角度给出的解释更符合"奥卡姆剃刀"原理[66], 有助于我们理解很多复杂社会系统中由个体相互作用而涌现出的群体空间交互模式.

本文介绍的这些对社会引力定律的理论解释都是以交通出行问题为背景的, 离散选择模型和目

的地选择博弈模型分析的核心都是出行者如何选择目的地. 但这种个体选择交互对象的行为并不仅仅存在于交通系统中, 在空间交互模式符合社会引力定律的系统中几乎都存在. 例如, 人口迁移是选地点做居住地, 社会交往是选人做交流对象, 科研合作则是选研究者做合作伙伴等. 在这些系统中, 个体多会倾向于选择活力相对高、距离相对近的对象. 但往往选择同一对象的人越多, 该对象能带来的收益就会越低: 在交通系统中拥挤会导致出行成本上升, 在商品贸易中竞争对手增加会导致商品价格下降, 在科研合作中选择某人的合作者增多会导致双方的合作强度降低, 等等. 而这些不同系统中的空间交互对象选择行为, 都能被本文所介绍的目的地选择博弈等模型来刻画. 总之, 对社会引力定律追根溯源, 不仅有助于深入理解交通出行、人口迁移、商品贸易、信息流通、社会交往、科研合作等空间交互现象的底层机制, 对于更好地预测、引导甚至控制各种复杂社会系统中人、物、信息的流动, 都具有广阔的应用前景.

感谢合作者周涛、汪秉宏、韩筱璞对本文中相关研究成果所做出的贡献.

参考文献

[1] Yan X Y 2019 *Beyond Gravity Law* (Beijing: China Science Press) piii (in Chinese) [闫小勇 2019 超越引力定律 (北京: 科学出版社) 第iii页]

[2] Ding Y M, Yang C P 2012 *Acta Phys. Sin.* **61** 238901 (in Chinese) [丁益民, 杨昌平 2012 物理学报 **61** 238901]

[3] Fotheringham A S, O'Kelly M E 1989 *Spatial Interaction Models: Formulations and Applications* (Dordrecht: Kluwer Academic Publishers) p1

[4] Sen A, Smith T 1995 *Gravity Models of Spatial Interaction Behavior* (Berlin: Springer) p1

[5] Karemera D, Oguledo V I, Davis B 2000 *Appl. Econ.* **32** 1745

[6] Porojan A 2001 *Open Econ. Rev.* **12** 265

[7] de Dios Ortuzar J, Willumsen L G 2011 *Modelling Transport* (West Sussex: John Wiley & Sons) p182

[8] Liu J H, Zhang Z K, Chen L, Liu C, Yang C, Wang X 2014 *PloS ONE* **9** e91070

[9] Ma L L, Ma C, Zhang H F, Wang B H 2015 *Physica A* **451** 205

[10] Li Z, Ren T, Ma X Q, Liu S M, Zhang Y X, Zhou T 2019 *Sci. Rep.* **9** 8387

[11] Newton I 1729 *Mathematical Principles of Natural Philosophy* (London: Benjamin Motte Publisher) p5

[12] Dodd S C 1950 *Am. Sociol. Rev.* **15** 245

[13] Desart H G 1846 *Chemin de fer Direct de Bruxelles vers Gand, par Alost, en Communication avec les Stations Diverses* (Bruxelles: Devroye) p16

[14] Odlyzko A 2015 *Economia Hist. Meth. Philos.* **5** 157

[15] Carey H C 1858 *Principles of Social Science* (Philadelphia: Lippincott) p28

[16] Ravenstein E G 1889 *J. R. Stat. Soc.* **48** 167

[17] Reilly W J 1929 *Methods for the Study of Retail Relationships* (Texas: University of Texas Bulletin) p2944

[18] Zipf G K 1946 *Am. Sociol. Rev.* **11** 677

[19] Tobler W 1995 *Urban Geogr.* **16** 327

[20] Fagiolo G 2010 *J. Econ. Interact. Coor.* **5** 1

[21] Zhou T, Han X P, Yan X Y, Yang Z M, Zhao Z D, Wang B H 2013 *J. Uiv. Electron. Sci. Technol. China* **42** 481 (in Chinese) [周涛, 韩筱璞, 闫小勇, 杨紫陌, 赵志丹, 汪秉宏 2013 电子科技大学学报 **42** 481]

[22] Xu Z X, Wang Y, Si H B, Feng Z M 2011 *Acta Phys. Sin.* **60** 040501 (in Chinese) [徐赞新, 王钺, 司洪波, 冯振明 2011 物理学报 **60** 040501]

[23] Viboud C, Bjørnstad O N, Smith D L, Simonsen L, Miller M A, Grenfell B T 2006 *Science* **312** 447

[24] Jung W S, Wang F, Stanley H E 2008 *Europhys. Lett.* **81** 48005

[25] Krings G, Calabrese F, Ratti C, Blondel V D 2009 *J. Stat. Mech.* **2009** L07003

[26] Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco J J, Vespignani A 2009 *Proc. Natl. Acad. Sci. U.S.A.* **106** 21484

[27] Kaluza P, Kölzsch A, Gastner M T, Blasius B 2010 *J. R. Soc. Interface* **7** 1093

[28] Pan R K, Kaski K, Fortunato S 2012 *Sci. Rep.* **2** 902

[29] Goh S, Lee K, Park J S, Choi M Y 2012 *Phys. Rev. E* **86** 026102

[30] Goh S, Lee K, Choi M Y, Fortin J Y 2014 *PloS ONE* **9** e89980

[31] Levy M, Goldenberg J 2014 *Physica A* **393** 418

[32] Sheppard E S 1978 *Geogr. Anal.* **10** 386

[33] Hua C I, Porell F 1979 *Int. Regional Sci. Rev.* **4** 97

[34] Wilson A G 1967 *Transp. Res.* **1** 253

[35] Niedercorn J H, Bechdolt Jr B V 1969 *J. Regional Sci.* **9** 273

[36] Domencich T A, Mcfadden D 1975 *Urban Travel Demand: A Behavioral Analysis* (Amsterdam: North-Holland) p1

[37] Yan X Y, Zhou T 2019 *Sci. Rep.* **9** 9466

[38] Pressé S, Ghosh K, Lee J, Dill K A 2013 *Rev. Mod. Phys.* **85** 1115

[39] Barthélemy M 2011 *Phys. Rep.* **499** 1

[40] Yan X Y, Han X P, Wang B H, Zhou T 2013 *Sci. Rep.* **3** 2678

[41] Hensher D A, Button K J 2000 *Handbook of Transport Modelling* (Oxford: Pergamon) p20

[42] Chalasani V S, Engebretsen Ø, Denstadli J M, Axhausen K W 2005 *J. Transp. Stat.* **2** 1

[43] Rietveld P, Zwart B, Van Wee B, van den Hoorn T 1999 *Ann. Regional Sci.* **33** 269

[44] Brakman S, Garretsen H 2005 *Location and Competition* (London: Routledge) p205

[45] Brockmann D, Hufnagel L, Geisel T 2006 *Nature* **439** 462

[46] González M C, Hidalgo C A, Barabási A L 2008 *Nature* **453** 779

[47] Han X P, Hao Q, Wang B H, Zhou T 2011 *Phys. Rev. E* **83** 036117

[48] Liang X, Zheng X, Lü W, Zhu T, Xu K 2012 *Physica A* **391** 2153

[49] Gallotti R, Bazzani A, Rambaldi S 2012 *Int. J Mod. Phys. C* **23** 1250061

[50] Huang F H, Peng J, You M Y 2016 *Acta Phys. Sin.* **65** 228901 (in Chinese) [黄飞虎, 彭舰, 由明阳 2016 物理学报 **65** 228901]

[51] Ben-Akiva M E, Lerman S R 1985 *Discrete Choice Analysis: Theory and Application to Travel Demand* (Cambridge: MIT Press) p1
[52] McFadden D 2007 *Transp. Policy* **14** 269
[53] Herrnstein R J, Loewenstein G F, Prelec D, Vaughan Jr W 1993 *Journal of Behav. Decis. Making* **6** 149
[54] Train K E 2009 *Discrete Choice Methods with Simulation* (Cambridge: Cambridge University Press) p38
[55] Takemura K 2014 *Behavioral Decision Theory: Psychological and Mathematical Descriptions of Human Choice Behavior* (Berlin: Springer) p52
[56] He D R, Liu Z H, Wang B H 2009 *Complex Systems and Complex Networks* (Beijing: Higher Education Press) p59 (in Chinese) [何大韧, 刘宗华, 汪秉宏 2009 复杂系统与复杂网络 (北京: 高等教育出版社) 第59页]
[57] Rosenthal R W 1973 *Int. J Game Theory* **2** 65
[58] Monderer D, Shapley L S 1996 *Games Econ. Behav.* **14** 124
[59] Stouffer S A 1940 *Am. Sociol. Rev.* **5** 845
[60] Simini F, González M C, Maritan A, Barabási A L 2012 *Nature* **484** 96
[61] Yan X Y, Zhao C, Fan Y, Di Z R, Wang W X 2014 *J R Soc. Interface* **11** 20140834
[62] Yan X Y, Wang W X, Gao Z Y, Lai Y C 2017 *Nat. Commun.* **8** 1639
[63] Voorneveld M, Borm P, Van Megen F, Tijs S, Facchini G 1999 *Int. Game Theory Rev.* **1** 283
[64] McFadden D 1980 *J. Bus.* **1980** S13
[65] Fisk C 1980 *Transp. Res. B* **14** 243
[66] Burgess J 1998 *Occam's Razor and Scientific Method* (Oxford: Clarendon Press) p195

SPECIAL TOPIC—Statistical physics and complex systems

Exploring the roots of social gravity law*

Yan Xiao-Yong 1)2)†

1) (*Institute of Transportation System Science and Engineering, Beijing Jiaotong University, Beijing* 100044, *China*)

2) (*Complex Labortary, University of Electronic Science and Technology of China, Chengdu* 611731, *China*)

(Received 4 November 2019; revised manuscript received 19 November 2019)

Abstract

Many spatial mobility of people, goods and information, such as human travel, population migration, commodity trade, information communication, social interaction and scientific cooperation, follow a law similar to Newton's law of universal gravitation. This law, named social gravity law, is that the flow between two locations is directly proportional to the product of the vitality of these two locations, and inversely proportional to a power function of their distance. The gravity model established by analogy with the gravity law has also been widely used to predict trip distribution, population migration, interregional trade flows, etc. But why do many complex social systems have such a simple law? It is an interesting and valuable issue. This paper reviews the research on exploring the roots of the social gravity law from various perspectives, including statistical physics, microeconomics, and game theory.

Keywords: social gravity law, principle of maximum entropy, discrete choice model, congestion game

PACS: 89.75.–k, 89.65.–s, 89.70.Cf, 02.50.Le **DOI:** 10.7498/aps.69.20191686

* Project supported by the National Natural Science Foundation of China (Grant Nos. 71822102, 71621001, 71671015, 61304177).

† Corresponding author. E-mail: yanxy@bjtu.edu.cn