

人类行为时空规律初探

周涛

1. 引言

复杂系统在物理学研究对象中行为最难以捉摸，性质最难以刻画，覆盖面却最为广泛。对复杂系统的深入研究，不仅产生了新的理论体系，例如耗散结构理论、突变论、微循环论、协同学等，而且大大拓展了统计物理学研究视野并带来了“从还原论到系统论”这一研究理念和方法论上的革新。

大量典型的复杂系统，都直接或间接和人发生关系。经济金融系统和社会系统是由人的意志和行为驱动的，如计算机互联网这样的大规模技术网络是由人建造的，被很多学者认为最复杂的系统之一的大脑及神经系统，是人身体的一部分——从人的社会属性出发，我们研究社会经济系统中通过人的相互作用和策略博弈涌现出来的复杂性；从人的生物属性出发，我们研究人体内的神经系统、代谢系统、基因调控系统等。尽管这方面的研究已经是硕果累累，我们还是注意到前者把人看得太小，后者把人看得太大，对于与人相关的复杂系统的整体认识，尚缺失一个环节，就是人类自身行为在时间和空间上表现出来的复杂性。

追根溯源，对人类行为的思索古已有之。现代心理学和经济学热衷的关于群体压力和从众行为的研究，早在金代文学家蔡松年诗赋“槽床过竹春泉句，他日人云吾亦云”中便可觅得踪影，而近年传播学关于面对谣言和恐慌情绪时人类非理性行为和判断的报道研究，亦可视作三人成虎、曾子杀人这些典故的科学注脚。对人类行为进行科学而系统的研究，或始于美国心理学家华生（J. B. Watson）的行为主义流派——如此算来，也不过百年历史。尽管这一百年里，对人类行为的理解一直是社会学、心理学和经济学共同关注的焦点，但人类自身的复杂性和多样性，对一切科学的尝试来说都是巨大的挑战。最近统计物理学家在人类行为时空统计特性分析上面的一系列工作，迈出了定量理解人类行为统计规律坚实的一步。本文将介绍人类行为时间规

律和空间规律的实证和模型，以及这些新规律的发现对于我们理解若干复杂社会经济现象的价值。有兴趣的读者可以参考刘怡君和笔者在科学出版社出版的名为《社会动力学》的小册子，其中笔者与同事韩筱璞、闫小勇、杨紫陌、赵志丹共同撰写了6篇论文，详细介绍了人类行为时空规律研究的前沿进展。

2. 人类行为的时间规律

泊松过程被广泛应用于人类活动的量化模型中，描述单位时间内或指定范围内特定事件出现次数的统计规律性，例如模拟交通流量的模式或事故发生频率，商业话务中心的配置，库存管理，以及估计移动通信中占线的电话数量。该过程假设每个时间片段一件事情发生的概率都是恒定的，如果把事件发生的时间点都画在时间轴上，这些点的分布是非常均匀的，事实上，可以证明，两个相邻事件间隔时间的分布是指指数衰减的。如图1(a~c)所示，符合泊松过程的情况下，事件的到达比较均匀，间隔时间长短往往相差不大，其分布在单对数坐标（ y 轴取对数， x 轴线性坐标）下是一条负斜率的直线。

2005年，巴拉巴西（Barabási）在《自然》上报道了大学里，以3个月为周期的3188位用户收发的共129135封电子邮件，与泊松过程所刻画的行为大不一样，用户相继发送两封邮件的间隔时间服从的不是指数分布，而是一个幂函数分布。图1(d~f)所描述的就是这种情况。从图1(d)中可以看到，事件的到达非常不均匀，有时候很多事件犹如爆发般连续到达，而这些爆发之间往往是长时间的静默。图1(e)显示了差异非常大的间隔时间，其分布在具有双对数坐标（ y 轴和 x 轴都取对数）的图1(f)中呈现为一条负斜率的直线。类似图1(d~f)的在时间上非均匀的行为，在人类工作生活的方方面面都被观察到，包括电子邮件、水陆信件、手机短信、手机通话、图书借阅、金融交易、网页访问、即时通信、网络搜索、论坛活动、博客活动、社交行为、任务

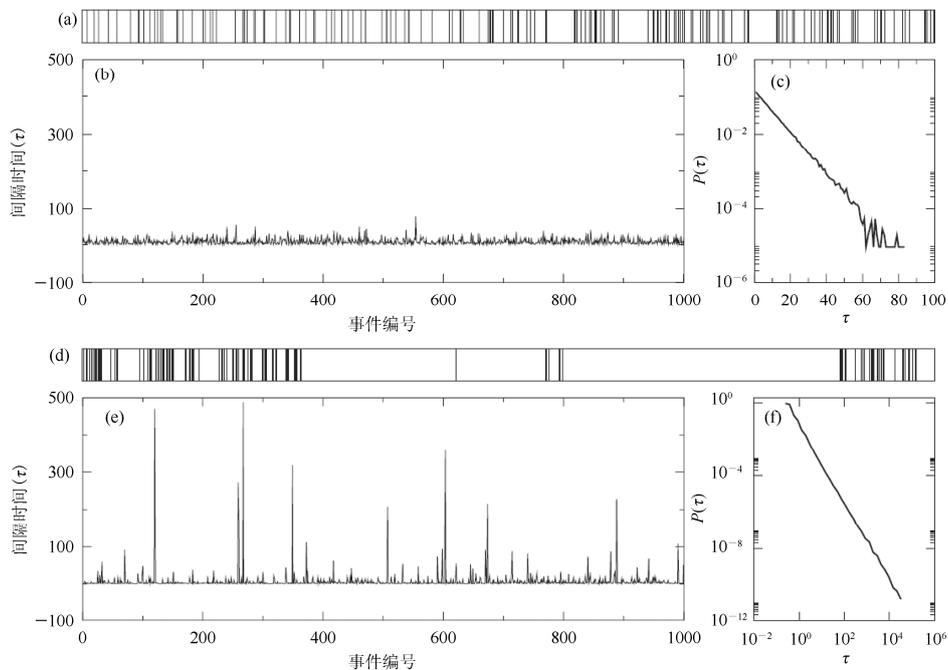


图1 (a) 泊松过程所预示的事件到达过程, (b) 相邻事件间隔时间, (c) 以及相邻事件间隔时间的分布; (d), (e), (f) 事件间隔时间符合幂函数律的对应情形

执行、在线服务、战争叛乱、文字写作、物理接触, 等等。

大部分已知的人类行为都不能用泊松过程来刻画这一事实, 向物理学家提出了一个问题: 什么样的机制导致了这种非泊松的行为? 2005 年以来, 研究人员对这个问题从各个方面进行了全方位的探索, 先后提出了数十种理论模型。这些模型大多数有着深刻的现实背景, 其基本假设往往植根于在日常生活中影响我们的行为的实际因素。从研究思路上划分, 这些典型模型大体可以分成三类: 基于任务队列理论的模型, 基于人类兴趣或记忆的模型, 以及基于社会交互影响的模型。

任务队列模型建立在这样一种生活经验上: 一个人在他的日常生活中往往需要面对很多工作, 如发送电子邮件或者水陆信件, 打电话, 阅读书报, 写文章等。一般来说, 在生活中人们是按照一定的顺序一件一件来做这些工作的。在针对人类行为的建模中, 可以把这样一些日常活动概括为需要处理的“任务”。个体对于这些陆续到来的任务, 有三种简化的处理方案: 第一种是先进先出方案, 这种方案是个体按照其获得任务的顺序执行任务; 第二种是随机执行方案, 即执行任务完全不受任务优先级和到达时间的影响; 第三种方案是按照任务优先级

执行, 在这种方案中, 优先级最高的任务最先执行, 即使它后加入任务列表。巴拉巴西和巴兹克斯 (Vázquez) 等人的研究显示, 真实人类行为更接近第三种方案所预示的结果。任务队列理论作为人类动力学研究领域的第一种重要模型, 到目前为止仍然是该领域最主流模型。针对早期模型所存在的各种问题, 近年来研究者也从各种现实情况出发, 更深入地挖掘了它的内在机制, 提出了多种改进型模型。例如, 研究者深入讨论了它的严格数学解析, 提出了其背后的等效渗流理论, 引入了任务本身的时间特性, 任务处理优化等各种现实因素。

韩筱璞和汪秉宏等人注意到很多表现出非泊松特性的人类行为不是由任务驱动的, 譬如在线游戏、电影观看等。2008 年, 他们在《新物理学》上提出了一个理论模型, 并指出自适应的兴趣变化可以导致人类行为的非泊松特性。举个例子, 如果一个人很久没有浏览网页了, 但是一个偶然的上网会立刻唤起他对上网的兴趣; 此后, 由于兴趣的持续, 上网的频率也不断增加, 但一段时间后就会出现兴趣的衰退, 从而就会减少上网的次数。韩筱璞和汪秉宏等认为在很长的时间范畴上进行观察, 人类从事某特定活动的兴趣具有准周期变化的特征, 他们进一步证明这种变化会驱动并导致间隔时间近似服从

幂函数律的行为序列。

以上所叙述的模型都是考虑孤立个体的行为。在现实生活中，每个人都是生活在一个与其他人互相交流影响的社会环境中，这些社会交互很大程度上影响了我们的行为，如电子邮件通信，电话通信和其他合作性工作。2007年，奥利维拉（Oliveira）和巴兹克斯在《物理 A》撰文提出了一种最简化的考虑社会相互作用模型。该模型依然基于任务队列模型的框架，并且只考虑了两个个体之间的相互作用——存在一种需要两人合作的任务，只有当这个任务同时在两个人的任务列表中都是优先级最高的任务时才能被执行。这一模型尽管极为简单，但它提供了一个研究社会交互行为中的非泊松特性的成功范例。随着大量针对社会交互行为的实证统计研究的涌现，研究者在奥利维拉和巴兹克斯模型的基础上，成功将其推广到了更为实际的条件之下，并取得了一系列成功。例如，吴晔和周昌松等人2010年在《美国科学院院刊》上提出了一种复杂的交互机制，可以很好解释在手机短信通信行为中观察到的统计特征。

除了上述具有明确社会心理机制的模型外，还有一些学者关注到外在因素对人类行为时间规律的影响，譬如我们进行活动的周期和节律——白天活动晚上睡觉也可能是产生这些非平凡统计特性的重要原因。赵志丹和杨紫陌等2012年在《欧洲物理快报》的研究，却显示这种周期节律对人类行为事件规律的影响非常有限。

3. 人类行为的空间规律

布罗克曼（Brockmann）等2006年在《自然》上撰文分析了美元流通记录网站的数据，该数据集包含美国境内100多万条美元流通数据，每条数据记录了钞票的编号，使用地点和使用日期，可以由此推算出每张钞票的空间移动轨迹。由于钞票是被人携带的，因此钞票的移动轨迹可以近似看作是人的空间运动轨迹集的一个样本。如图2（a）所示，在10~3500千米的范围内，钞票移动步长分布符合幂函数律。

需要指出的是，钞票的移动轨迹可能混合了多个钞票持有人的移动，因此仅通过观测钞票的移动并不能完全反映出人类个体的空间运动规律。尽管如此，布罗克曼等的工作是最早用统计物理方法对人类空间运动特性所进行的实证研究，对后续的人类空间运动行为的统计和建模研究具有重要的启发

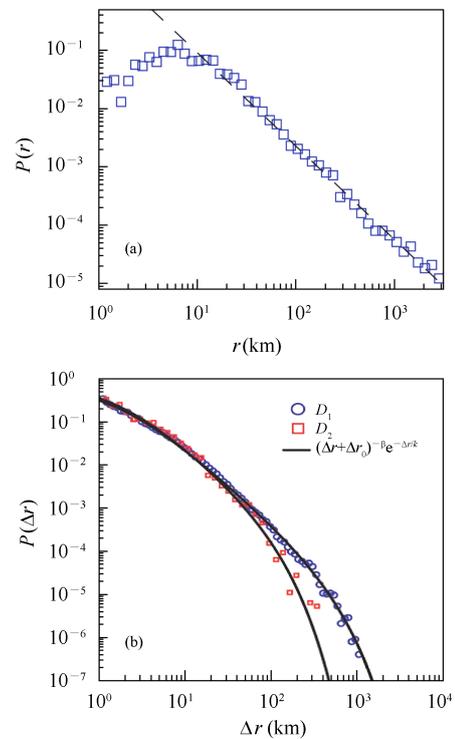


图2 钞票移动步长 r 和手机两次激活距离的概率密度分布

意义。对于布罗克曼等的观察结果也有不同的解读，他们自己认为这个现象可以用连续时间随机游走来解释，而韩筱璞等却认为是全球交通系统自身存在的层次结构导致了这个结果，而个体的特征和属性反而不一定是最重要的因素。

布罗克曼等的的数据主要刻画人类长距离旅行的特征，利用手机定位系统的数据，巴拉巴西小组2008年在《自然》上撰文分析了人类日常出行的距离分布，发现也具有类似的广阔分布区间。如图2（b）所示，手机两次收发短信或接打电话之间的距离分布也具有类似与钞票步长分布的下降规律。尽管巴拉巴西小组的实验结果是基于群体数据，而个体数据量还不够做出准确的判断，他们还是推断个体行为也具有类似的性质。2010年，该小组在《自然-物理》上发表了一个基于个体探索返回机制的动力学模型，该模型显示，如果探索新位置的概率负相关于已经访问过的位置数，而访问老位置的概率正相关于该位置被访问过的次数，则个体在不同位置间移动距离的分布符合手机数据的实证结论。然而，闫小勇等通过对数百名瑞士志愿者出行日志的分析发现，不同类型的个体（学生、工作人员、退休人员、待业者、家庭主妇……）有着截然不同的

出行模式，而且几乎都不符合巴拉巴西小组所强调的标度规律（图3）。事实上，群体数据中观察到的漂亮的标度律，只是迥异个体行为模式叠加的结果。闫小勇等还给出了一个基于麦克斯韦-玻尔兹曼统计的解释，这里就不细讲了。人类行为的空间规律适用范围如何？仅仅只适用于群体还是对个体也有效？这些空间规律背后的机制是什么？这些问题，目前都没有一个让人信服的答案。

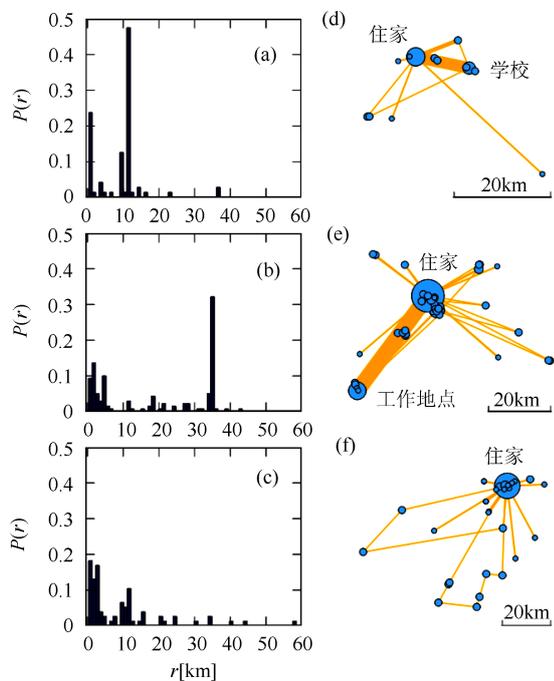


图3 志愿者出行日志所记录的个体距离分布
(a) (学生)、(b) (工作人员) 和 (c) (退休人员) 是三个典型用户出行距离的分布, (d)、(e)、(f) 是其对应的出行网络, 其中节点的大小代表个体在该位置访问的次数, 边的粗细表示个体在对应路径上出行的频数

4. 结语与讨论

对人类行为时空规律的讨论古已有之，但是真正量化的研究并推动形成一门受广泛关注的新兴交叉研究方向，主要归功于巴拉巴西和布罗克曼2005年及2006年在《自然》上发表的两篇论文。在这短短几年时间里，研究成果层出不穷，颠覆了很多我们以前对于人类行为统计特性“似乎正确”却未经严格检验的认识。本文作为一篇科普性质的论文，着重于介绍该领域的研究背景与研究思路，以及研究成果中最有代表性以及最容易描述的部分——人类时空行为的非均匀性。还有很多性质本文没有介绍，譬如人类行为在时间上的阵发性和记忆性，人类空间出行回转半径极端缓慢的增长，人类

空间出行的强可预测性，等等。

作为一门新兴的交叉学科方向，人类动力学的提出不过寥寥数年，但其重大的理论和应用意义已初露峥嵘。从理论层面来看，我们观察到很多由人的行为驱动的复杂系统的各类现象，其中人类动力学的研究提供了新视角和新解释。举例而言，对于部分计算机病毒以极低感染比例长期存在这个问题，最近巴兹克斯等提出了基于人处理电子邮件时间上的非泊松特性的解释，该解释完全独立于基于接触网络度分布异质性的解释。与此同时，人类动力学的研究还提出了很多新的理论问题，譬如对于到达时间间隔分布二阶矩发散的服务请求，以及有优先级限制的服务请求，如何用排队论的方法进行处理？又如具有向心性和局域化倾向的出行行为能否纳入某种扩散动力学加以解释？我们有理由相信，人类动力学的研究会推动随机过程和排队论经典理论的丰富和发展。

人类行为时空统计特性的挖掘，还具有相当的应用价值。对人出行，特别是长途旅行规律的理解，有助于理解、预测和控制全球性疾病的流行；对人出行和网络活动中表现出来的阵发性和记忆性的分析，可以应用于异常行为检测，例如疑似恐怖分子非正常的频繁跨越国境的旅行和处于非正当目的的网络舆情操控行为的检测等；对于人类群体出行行为时空图的分析，可以为交通路网设计和道路流量规划，包括公共服务设施和商业设施的选点提供借鉴；对人网上行为的分析可以为更好的开展互联网信息推荐服务提供借鉴；等等。

作为一篇科普文章，笔者希望本文能够带给普通读者对于传统物理学一些不同的认识，特别是展现具有交叉视野的物理研究所蕴含的新的观念和思路——物理学所关注的对象与你我的距离或许比读者想象的更近。对于从事统计物理学与复杂性科学研究的学者，笔者希望本文能够激起他们对人类行为研究的兴趣。随着互联网和其他科技手段的发展，越来越多人类日常行为被精确地数字化地记录下来，我们第一次真正有机会去探寻和揭示人类行为背后的驱动机制。本文所介绍的这一研究方向必将成为统计物理学未来一段时间的热点和焦点。

(电子科技大学互联网科学中心 610054)