

基于习惯的人类动力学建模

焦玉^{①②}, 刘衍珩^{①②*}, 王健^{①②}, 王静^③

① 吉林大学计算机科学与技术学院, 长春 130012;

② 吉林大学符号计算与知识工程教育部重点实验室, 长春 130012;

③ 吉林大学档案馆, 长春 130012

* 联系人, E-mail: lyh_lb_lk@yahoo.com.cn

2009-08-03 收稿, 2009-11-11 接受

国家自然科学基金(批准号: 60973136)、科技部国际科技合作与交流专项(批准号: 2008DFA12140)、吉林大学研究生“985工程”创新计划(批准号: 20080235)和教育部高等学校博士学科点基金(批准号: 20060183043)资助项目

摘要 定量地理解人类行为是现代科学的中心议题之一,但由于人类行为的复杂性,其规律是难以发掘的.当前使用的人类动力学模型都假设人类行为发生的时间是随机分布的.为更好地解释人类行为中的阵发性与重尾特征,基于习惯的人类动力学模型通过结合考虑排队模型与人类行为特有的习惯特征,根据事件已发生的次数与稳定程度调整分布参数,使用正态分布模拟连续事件发生的间隔时间,并利用随机参数模拟现实中随机出现的打断人类习惯行为的突发事件,进一步通过考虑事件发生的持续时间,以模拟人类对事件持续关注现象.

关键词

人类动力学

习惯

正态分布

建模理论是一种重要的科学方法,在很多领域都有广泛的应用.如时翔等人^[1]对目标辐射截面的建模用于制导与探测等军事领域,张曙光等人^[2]建立了考虑线路、列车、弓网和气流耦合的高速轮轨系统动力学模型,拓展了现有车辆系统动力学研究的范围和作用.利用建模结果不仅能够大大减少系统的设计时间,还会提高系统的设计质量.如王友渔等人^[3]借助子结构综合思想,提出了一种考虑复杂机架柔性的并联构型装备整机静刚度半解析建模方法.通过统计建模对结果集进行训练与预测是建模领域中一项重要的应用方法,如王仁华等人^[4]提出的一种基于统计建模的可训练单元挑选语音合成方法.但对人类行为建模的研究仍处于初始阶段,因为人类行为不但由人类复杂的决策模式决定,还受到外界资源等各种因素的影响,因此人类行为的规律往往是难以发掘的.然而随着科学技术的进步,人类的各种信息都可以通过数字化的方式方便地统计与获取,如李安安等人^[5]就将人体解剖结构数字化和可视化,建立了具有组织形态、物理功能和生理功能的人体信息系统.于是对人类行为的研究也引起了广泛的

关注,通过人类活动的时间选择机制来定量研究人类行为成为一种有效地理解社会经济系统复杂性来源的方法.在以往的研究中,常常把单个人的行为简化为稳态的,或者完全随机的、可以使用泊松过程描述的^[6].泊松过程已经被广泛地应用于量化人类的活动,例如交通流量的模式及移动通讯流量的预测等.泊松假设必然导致的推论是人的行为的时间统计特征是较为均匀的,行为事件之间存在极大的时间间隔的概率很小.但是,最近的研究通过对电子邮件和信件的发送^[7]、在线音乐点播^[8]、网页浏览^[9]和移动电话通信^[10]等大量人类行为的时间间隔的统计,揭示出人类活动具有明显的非泊松特点,常常在短时间内频繁发生,然后在很长的一段时间内销声匿迹.数据分析结果表明,这种人类活动模式的事件间隔时间分布由幂律分布刻画,并带有衰减缓慢的重尾特征^[11].泊松与重尾分布的区别非常显著,泊松分布按指数下降,致使每两个行为之间的时间间隔长度几乎是均匀的,任务不会进行长时间的等待;而重尾过程下降缓慢,在巨大的空闲间隔时间后仍有活动出现,而那些长长的空闲时间之间被阵发的密集

的活动所填充. 这种非泊松的阵发行为可能反映了某些人类动力学基本和潜在的普适特征, 因此对产生这种特征的机制进行研究是非常必要的.

Barabási^[11]认为人类活动的阵发现象是基于排队过程的决策结果: 若个体依据某种感知的轻重缓急来执行任务, 那么任务等待完成的时间分布就会出现重尾特征——其中大部分工作很快被处理, 而少数工作需要忍受漫长的等待时间. 相反, 如果不考虑轻重缓急盲目地执行, 则工作等待时间是近似均匀的分布. Vázquez 等^[12]进一步将人类行为分为幂指数为-1 和-1.5 的两大普适类并给出了基于排队论模型的解释. 但是否存在其他的普适类, 及产生不同普适类的深层原因仍有待研究与证明. 之后, 很多学者基于排队论模型做了改进^[13-15]. Blanchard 等^[14]提出了带有“老龄化机制”的动力学模型, 每个等待被处理的任務有一个截止时间并且随着时间的流动平稳地递减, 在每个给定时刻, 任务的排程遵循“最早-截止时间-最先服务”的准则.

但是, 人类作为一种高等的智能生物, 并不会像机器一般一直被动地等待任务, 再以某种次序完成任务. 按照排队论模型, 一个人可能要一直面对无穷无尽的待完成的“任务”. 为此, 一些研究人员尝试通过考虑人类的心理特性研究人类行为动力学. Vázquez 等人^[16]认为, 导致幂律分布的另一个重要原因是人类的记忆特性, 人类对他们过去的活动率有一个直观感觉, 并根据以前的活动率确定加速或减速活动, 并认为这种与人类以往的活动经验相联系的加速或减速趋势是导致幂律的关键因素. Han 等人^[17]认为, 人类的很多活动是被个人的兴趣或欲望所驱动, 而一般并不能被视作“任务”, 并基于人类对自身行为的兴趣或活性的自适应调节现象, 提出了一个非排队论的人类动力学模型.

本文通过综合考虑排队模型与人类的心理特征, 提出了一个基于习惯的人类动力学模型. 习惯是人类行为中的一项重要特征, 人类多数行为都是由习惯决定或受习惯影响的, 从每天吃饭睡觉到每天处理邮件, 每天浏览网页, 每周末看电影玩游戏等等. 这些行为依照每个人的习惯不同而按照一定的周期定时发生, 这种习惯性行为导致的周期性事件的发生有可能是导致人类行为的重尾特征的一项重要原因. 而对于同类事件中的不同任务, 人类则按照一定

的方式排序完成. 例如在处理邮件时先处理重要的, 或者先处理较早接收的; 浏览新闻时先浏览自己更为关注的. 为解释这类行为中出现的非泊松特性, 本文通过事件已发生的次数与已发生事件的平均间隔时间调整分布函数, 模拟人类习惯行为的养成过程; 加入随机参数以模拟现实中可能出现的打断习惯行为的突发事件, 如资源短缺、不定期的休假等; 并考虑了每项队列在服务台停留一段时间的情况, 以模拟人类对某一事项的持续关注; 最后将模型得到的数值结果与实际数据进行拟合, 并得到了较好的拟合结果.

1 模型规则

心理学研究认为, 一个人一天的行为中, 大约只有 5%是非习惯性的, 而 95%是习惯性的. 即使是打破常规的创新, 最终也会演变成习惯性的创新. 行为心理学研究结果显示: 3 周以上的重复会形成习惯, 3 个月以上的重复会形成稳定的习惯. 且人类的习惯行为通常表现出有规律的周期性, 例如每天处理邮件、浏览新闻, 则处理邮件、浏览新闻这些事件以一天左右的时间为周期发生; 又如每个周末进行看电影, 玩游戏等娱乐活动, 则这些娱乐活动的事件以一周左右的时间为周期发生. 其中每个事件中又可能包含多项任务, 例如多封等待处理的邮件, 或多条等待浏览的新闻等. 因此本文提出的模型中有若干项队列, 每个队列代表一项习惯行为, 每个队列包含若干等待的任务. 这里模型中队列的数目是被忽略的, 而着重讨论同一队列所代表的事件发生的间隔时间及队列内部的任务执行情况. 队列长度为不定长, 到达的任务一直在队列中等待.

习惯行为导致人类每间隔一定的时间后会重复做同样的事情, 同类事件连续发生的间隔时间在某一固定周期时间左右波动, 该习惯越稳定, 波动就越小, 反则越大. 但即使是非常稳定的习惯, 事件发生的间隔时间也不会严格地在指定的周期时间, 对于每日发生的事件来说, 当间隔时间在 22~26 h 左右即可看做是稳定的习惯, 而不稳定的习惯会导致事件的间隔时间在几小时到上百甚至几百小时波动. 对于每星期发生的事件来说, 不稳定的习惯会导致事件的间隔时间在几天到上百天波动. 并且事件发生的时间间隔不是服从一个固定不变的分布函数,

其分布情况是随时间的变化根据事件已经发生的情况而动态变化的. 当该习惯行为被重复的次数越多, 发生的间隔时间越接近标准周期时(例如 24 h), 该习惯就越稳定, 事件发生的时间间隔也越稳定. 对于已经养成的稳定的习惯, 仍然有可能被临时出现的事情打断, 例如资源的短缺, 不定期的休假等. 而当临时事件结束时, 例如在结束假期后, 人们通常回到原来的生活状态, 继续保持原有习惯.

在以前的研究中, 通常假设任务以均匀的间隔时间到达, 本文也仍然使用该假设. 任务到达后, 按照到达的先后顺序在队列中等待. 当一个队列被选择执行时, 假设任务仍然以到达的先后顺序被选择执行. 在实际中, 有的任务例如电子邮件可能并不会被回复, 而直接被删除, 但这种删除操作也可以看做是该任务被执行的一种方式. 在队列被选择后, 如果当前队列中的所有任务都被执行完, 该队列所代表的事件也并不会一定会结束. 例如对于邮件处理, 众多邮件客户端软件的出现, 使得用户可以在打开客户端处理完收到的邮件后, 再继续保持客户端打开一段时间, 以处理随时收到的邮件. 类似地, 在新闻浏览的事件中, 新出现的 rss 订阅等技术使得用户可以随时关注新发生的新闻事件. 这里出现的事件持续时间, 同样受到人类习惯行为的影响. 而对于由于人类自身的兴趣或欲望引发的事件, 例如看电影, 玩游戏等非任务型事件的娱乐活动, 事件的持续时间完全由人类自身的习惯行为决定, 例如人们通常在周末的两天或三天内进行娱乐活动, 即事件的持续关注时间为两天至三天.

按照以上的思想, 本文模型的规则具体如下:

(1) 为使习惯行为的周期在多数情况下表现出较稳定的状态, 模型假设每项队列连续被选择的时间间隔, 即该类事件连续发生的时间间隔的分布函数的概率密度函数服从正态分布:

$$p(\tau, t) = \frac{1}{(\sqrt{2\pi} \times \exp(1/\xi)\sigma(r(t), f(t)))} \times \exp\left(-\frac{(\tau - T)^2}{2(\exp(1/\xi)\sigma(r(t), f(t)))^2}\right), \quad (1)$$

其中, $p(\tau, t)$ 表示在时刻 t 事件发生的时间间隔分布, 正态分布的参数 σ 是关于 $r(t)$ 和 $f(t)$ 的函数, 其中 $r(t)$ 表示到时刻 t 为止事件已经发生的次数, $f(t)$ 表示已经发生的事件的平均间隔时间, 使得该分布随着事件

已发生的次数和发生的平均间隔时间而变化.

$$(2) \quad \sigma(r(t), f(t)) = a \times \exp(1/r(t)) + b \times \ln(|f(t) - T| + 1). \quad (2)$$

其中 a, b 为参数. 该式使用相加的两部分分别控制事件的重复次数与已发生事件的平均间隔时间对时间分布的影响. 事件已发生次数在初始状态下为 1, $r(t)$ 值为 1, $\exp(1/r(t))$ 值为 e , 为使初始时连续事件发生的间隔时间波动的幅度较大, 使用参数 a 与之相乘以对 $\exp(1/r(t))$ 放大; 随着事件重复次数的增多, $r(t)$ 增大, $\exp(1/r(t))$ 减小, 参数 σ 也随之减小, 时间间隔分布趋于稳定, 人类的习惯行为也越来越稳定. 在极限状态下, $r(t)$ 趋于无穷, $\exp(1/r(t))$ 的值减至极小值为 1, 此时的时间间隔分布也最稳定. 该式后半部分 $b \times \ln(|f(t) - T| + 1)$ 用以控制已发生事件的平均间隔时间 $f(t)$ 对时间分布的影响, 当 $f(t)$ 与指定的标准周期 T 相差越小时, $\ln(|f(t) - T| + 1)$ 越小, 参数 σ 也随之减小, 时间间隔分布趋于稳定. 在理想情况下, $f(t)$ 的值与 T 相等, 即人类行为完全按照指定的标准周期定时进行, $\ln(|f(t) - T| + 1)$ 为 0, 时间间隔分布最为稳定. 在极限情况下, $f(t)$ 的值趋于无穷大, $\ln(|f(t) - T| + 1)$ 的值也趋近无穷大, 则参数 σ 同样趋近无穷大, 此时分布情况波动巨大, 几乎难以存在习惯行为. 参数 b 用于对 $\ln(|f(t) - T| + 1)$ 的值适当放大或缩小, 控制平均间隔时间对分布造成的影响程度及参数 σ 的大小. 由此可以看出, 习惯的稳定程度决定了连续事件发生的时间间隔波动的幅度.

(3) 参数 ξ 是连续区间 $[0, 0.1]$ 上的随机数. 该参数设置的作用是为了模拟现实中随机出现的打断人的习惯安排的事件发生. 将该参数的范围设置为 $[0, 0.1]$ 这个较小的范围, 使得习惯行为在较低的概率情况下出现较大的波动.

(4) 在排队论中经常假设新任务相互随机独立, 并以稳定的速率到达队列^[12], 因此本文模型假设队列所代表的事件类型的任务到达的时间间隔服从正态分布, 使得任务以理想的稳定速度到达:

$$p(t_{\text{arv}}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(t_{\text{arv}} - T_{\text{arv}})^2}{2}\right). \quad (3)$$

(5) 模型假设每项队列被选择后, 服务台将在该队列上停留一段时间 t_{last} , 该时间长度服从正态分布:

$$p(t_{\text{last}}) = \frac{1}{(\sqrt{2\pi} \times \sigma_{\text{last}})} \exp\left(-\frac{(t_{\text{last}} - T_{\text{last}})^2}{2\sigma_{\text{last}}^2}\right), \quad (4)$$

该项的设置用于模拟人类对某项事件的持续关注。

(6) 在排队模型中, 服务过程通常被模拟为泊松过程, 任务选择协议有先进先出, 随机选择, 或者选择具有最高优先级的任务^[12]. 本文模型简单地假设每项队列被选择后, 顺序执行当前队列中等待的任务, 这也相当于每个到达的任务被按照到达的先后顺序赋予优先级, 具有高优先级的任务被优先处理, 任务的处理时间同样假设为服从泊松分布:

$$P(t_{\text{proc}} = k) = \frac{\lambda_{\text{proc}}^k}{k!} \mu e^{-\lambda_{\text{proc}}}, \quad (5)$$

其中参数 μ 的计算方法如下: $\mu = \begin{cases} 1, & \varpi \geq \Omega \\ \frac{1}{\varpi}, & \varpi \leq \Omega \end{cases}$, ϖ

是(0, 1)之间的随机变量, Ω 是(0, 1)之间指定的门限值, 用于模拟实际中偶尔出现的需要长时间处理的任务。

2 实验分析与讨论

为验证本文模型的有效性, 本文分析 Kiel 大学的电子邮件收发记录^[18]与 Netflix 网站(<http://www.netflixprize.com>)提供的在线电影点播记录数据。

2.1 与电子邮件发送数据的比较

本节所使用的数据来自 Kiel 大学的电子邮件收发记录, 包含发送者和接收者以 3 个月为周期的 9665 位用户发送的每封电子邮件的发送时间. 本实验中参数设置与说明如下:

(1) 每项队列连续被选择的时间间隔分布参数 $T=1440$ min, 即模拟人类的每日习惯;

(2) $\sigma(r(t), f(t))$ 的计算中, 参数 $a=4, b=60$, 使得 $\sigma(r(t), f(t))$ 的取值大约在 (1, 400) 之间. 过大的 $\sigma(r(t), f(t))$ 会使间隔时间分布过于不稳定, 波动过大; 过小的 $\sigma(r(t), f(t))$ 则使间隔时间分布过于稳定, 过于密集地集中在 T 左右. 而对于人类来讲, 即使是已经形成的非常稳定的习惯, 也不会按照指定的周期准确地定时执行, 也就是说, 即使用户养成了每天都处理邮件的习惯, 也未必会在每天的同一时间打开

邮箱, 从而使得在稳定习惯下, 事件的间隔时间仍可能有 5~10 h 的波动. 因此, 此时参数设置使用户每两次收发邮件的平均间隔时间对时间分布的影响更大;

(3) 任务到达的时间间隔分布参数 $T_{\text{arv}} = 60$ min. 这里所指的到达的任务, 并不一定只代表实际现象中信件的到达, 还可以认为隐含了用户的一些主观需求, 当用户需要主动发起一封信件时, 该需求也可以被视作一项需要完成的任务;

(4) 服务台在队列上停留的时间分布参数 $T_{\text{last}} = 120$ min, $\sigma_{\text{last}} = 15$, 使得用户每天在该项事件上持续关注的时间在几个小时之间波动;

(5) 每个任务的处理时间分布参数 $\lambda_{\text{proc}} = 5$ min, $\Omega = 0.4$. 使得用户处理每项任务的时间在几分钟左右波动, 但是会有较低的概率出现处理时间较长的任务。

按照本文提出的人类习惯模型和上文的参数设置, 队列被选择的间隔时间变化情况如图 1 所示. 从该图可以看出, 随着事件发生次数的增多与稳定, 人类的习惯变得越来越稳定, 但即使该事件的发生已经成为稳定的习惯, 也会由于资源短缺等各种突发事件的发生而打破, 如图 1 尾部的相对大幅度的波动; 但当突发事件结束后, 原有习惯也仍易于继续保持, 事件重复发生大约两个月后形成相对稳定的习惯。

图 2 则分别描绘了真实数据与模型生成数据的模拟情况. 相对于真实数据, 由模型生成的数据有较多的部分集中在 60 min 左右处, 而真实数据相对模型生成数据有较多的部分集中在十几分钟处, 这是由于模型中假设任务以平均一个小时候左右的时间间隔到达, 并且用户必须执行每个任务, 即对于邮件发送的事件来说, 平均大约一个小时候就会有一封邮件到达或者是一次用户的主动需求, 且用户需要立即回复收到的每封邮件. 而在实际行为中, 用户并不一定要回复收到的每封邮件, 也不一定会立即处理收到的邮件. 然而更为重要的是, 模型通过对习惯行为的模拟, 使得队列连续两次被选择的时间间隔相对队列被选择后任务的连续执行时间间隔较大, 而在实际中, 人们通常会在工作时处理邮件, 工作结束后便关注其他的事情, 直到第二天工作的开始再处理邮件, 因此基于习惯的人类动力学模型较好地模拟了人类真实行为的习惯特性导致的连续任务之间较长的时间间隔. 在图 2(a)中可以看到, 信件连续发送

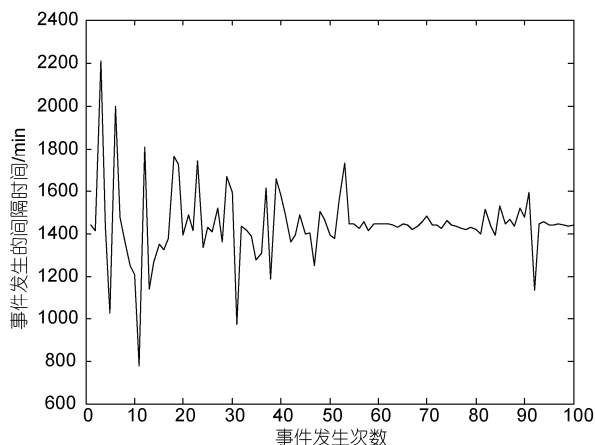


图1 模型生成的事件发生的时间间隔变化情况

事件发生的间隔时间在 1440 min 左右波动, 在初始时波动幅度较大, 随着事件发生次数的增多, 发生的间隔时间越来越趋近稳定. 图像尾部突然出现的大波动可被看做是突发事件的发生导致习惯被暂时破坏, 之后事件发生的间隔时间仍恢复稳定

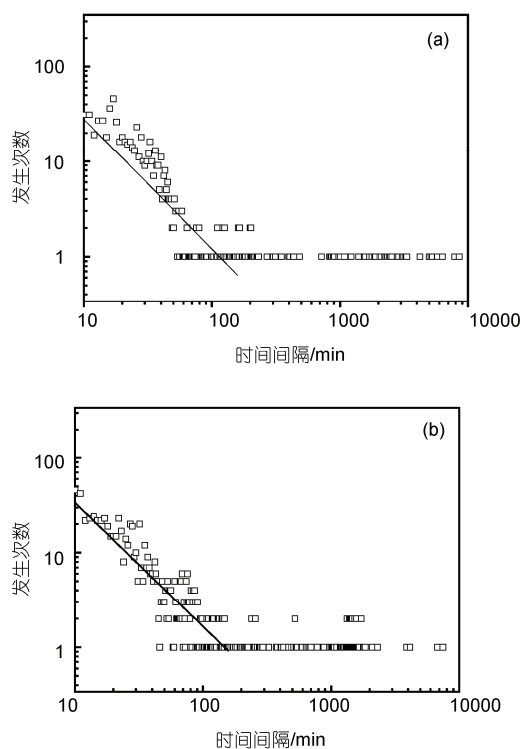


图2 模型与真实数据中观察到的时间间隔分布比较

(a) 展示了按照来自 Kiel 大学电子邮件真实数据描绘的单个用户 3 个月内连续发送电子邮件的时间间隔分布, 该分布在双对数坐标系下与幂指数约为 -0.89 的幂律分布可以较好地拟合, 并且分布情况带有明显的重尾特征; (b) 描绘了按照本文提出的基于习惯的人类动力学模型生成的人类活动模式数据分布, 该模式分布在双对数坐标系下的直线斜率为 -0.8936, 并带有明显的重尾特征

的时间间隔以接近幂指数的速度下降, 并在这个幂指数过程结束后, 呈现一条长长的“尾巴”, 即所谓的“重尾特征”. 这条尾巴使得两次连续发送信件之间存在很长的空闲时间. 图 2(b)中也存在着类似的幂指数下降及较长的重尾过程. 拟合结果的幂指数 0.89 与已有的研究结果认为, 电子通讯模式的拟合结果服从幂指数为 1 的分布相近. 模型对于随机干扰事件的模拟使得习惯行为偶尔会中断, 并产生相当大的时间间隔, 这是图中重尾尾部在 10000 左右的时间间隔仍然出现数据的原因, 这一现象在真实数据中同样存在.

2.2 与在线电影点播数据的比较

本节所使用的数据来自 Netflix 网站公开的数据, Netflix 是全球最大的在线电影共享系统之一, 其公开数据库(只包括整个 Netflix 中的部分数据)包含 17770 部电影, 447139 个用户和将近 1 亿条用户点播电影的记录. 该记录中点播时间的单位为天, 因此本文在处理数据时将一天内点播多次电影视为事件的一次发生, 忽略一天内的间隔时间. 通过跟踪单个用户的点播记录, 可以获取每两天之间点播的间隔天数.

本实验中参数设置与说明如下:

(1) 每项队列连续被选择的时间间隔分布参数 $T=7$ d, 即模拟人类的每周习惯;

(2) $\sigma(r(t), f(t))$ 的计算中, 参数 $a=3$, $b=0.3$, 使得取值大约在 (1, 10) 之间. 对于人类每周进行的习惯行为, 时间间隔波动多数情况下只在几天到几十天的范围内, 偶尔会出现上百天甚至几百天的波动, 因此该参数设置使得已发生时间的平均间隔时间对时间分布的影响相对较小;

(3) 任务到达的时间间隔分布参数 $T_{\text{arr}}=1$ d. 对于诸如看电影等由人类兴趣自发产生的事件, 该参数并没有什么实际的意义, 因为此时人类并不把事件的发生当做“任务”来对待, 在排队模型中可以认为用户并不需要以任何排序方式完成队列中的所有任务, 因此该参数将任务到达时间间隔设为一个较小的值, 使得队列中永远有足够多的任务, 而每次事件发生时, 用户完成的任务数由用户在该事件上持续关注的时间决定;

(4) 服务台在队列上停留的时间分布参数 $T_{\text{an}}=2$ d, $\sigma_{\text{last}}=1$, 使得用户每天在该项事件上持续关注的时间通常在 2 d 左右, 即一个周末;

(5) 尽管看一部电影的时间通常只有几个小时,但由于本实验以天为单位,忽略 1 d 内发生多次事件的时间间隔,因此每个任务的处理时间设置为 1 d.

图 3 分别描绘了真实数据与模型生成数据的模拟情况. 由于 Netflix 网站的点播时间数据以天为单位,无法对 1 d 内多次点播的时间间隔进行分析,因此产生的数据图中数据量较小,但仍然可以明显看到时间间隔的发生次数在双对数坐标系下以接近直线的速度下降,并产生了重尾现象. 在模型产生的数据中,队列被选择后,短暂的持续关注时间使得间隔时间在 2~3 d 的事件发生较多. 队伍以周期为 7 d 左右的时间被连续选择,并在十几天左右内波动,导致间隔时间发生次数在双对数坐标系下出现了接近直线下降的趋势. 习惯被扰动时产生的较大波动在

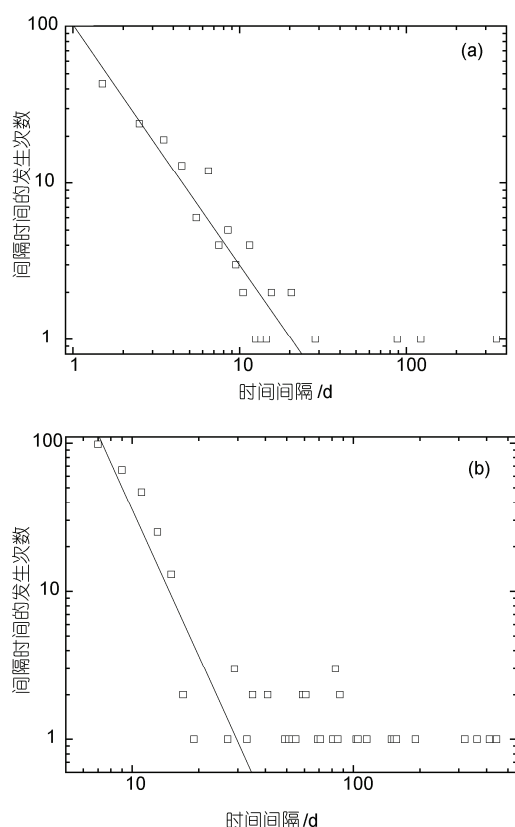


图 3 模型与真实数据中观察到的时间间隔分布比较

(a) 展示了来自 Netflix 网站真实数据中单个用户 3 个月内连续在线点播电影的时间间隔分布,该分布在双对数坐标系下的直线斜率约为 -0.9255 ,并且分布情况带有明显的重尾特征;(b) 描绘了按照本文提出的基于习惯的人类动力学模型生成的人类活动模式,该模式分布在双对数坐标系下的直线斜率为 -0.9061 ,并带有明显的重尾特征

数据图的尾部产生明显的重尾现象.

由于真实数据量较小,本文按照前文的模型参数设置,并指定任务执行次数为 2000 次,进行了多次实验,图 4 展示了其中两次实验结果. 由于模型设置队列被选择的周期为 7 d 左右,因此可能在时间间隔为 7 d 的次数产生一个较大的值,并在之后出现直线下降的趋势. 事实上,由于以天为单位,事件发生的时间间隔只能在最小一两天至几十天或几百天的范围内波动,因此表现出的幂律性质并不十分令人满意,但数据图中出现的接近直线下降的趋势与明显的重尾特征仍然可以作为探究人类动力学阵发与重尾现象产生的潜在原因的参考.

3 结论和展望

本文给出了一种基于习惯的人类动力学模型,通过结合考虑排队模型和人类具有的习惯行为特点,

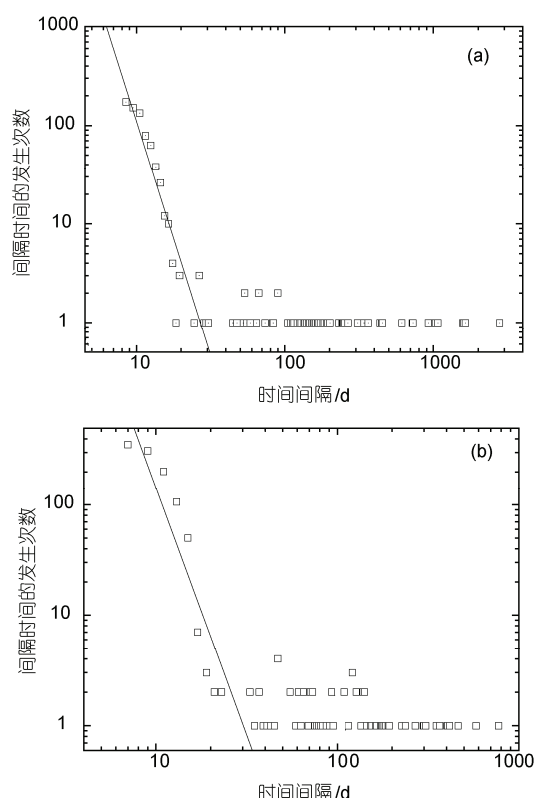


图 4 基于习惯的人类动力学模型预测的人类活动模式

描绘了按照基于习惯的人类动力学模型生成的人类活动模式,该模式分布在双对数坐标系下的直线斜率分别为 -0.94431 和 -0.92853 ,并带有明显的重尾特征

使用正态分布模拟人类习惯行为连续发生的间隔时间,并根据行为发生的次数与稳定程度动态调整分布函数的参数以控制间隔时间的波动幅度,使得习惯行为在多次重复后逐渐趋于稳定,以模拟人类的习惯养成过程.模型进一步为习惯行为选择的每项事件设置一定的持续时间,以模拟人类对事件持续关注现象.

事实上,之前关于人类动力学的研究多数以Barabasi使用的排队模型为基础,按照该模型的思想,人类动力学中出现的任务密集阵发与长时间静默的现象是由于低优先级的任务在队列中等待较长时间所导致的.而本文的研究认为,这种长时间的静默是由于人类的行为繁多而复杂,同一事件只能按照人

类习惯以相对较长的间隔时间周期性执行,每次执行只保持短暂的关注时间,实验结果表明,模型的生成结果与实际数据拟合结果的幂指数相近,并带有显著的重尾特征.

模拟人类行为对于病毒传播研究,资源分配,网络服务设置等都具有重大的意义与价值,具有非泊松分布特征的人类动力学自提出就引起了巨大的关注,其研究结果具有重要的理论意义与商业价值,但人类动力学在实证和建模方面尚存在巨大的研究空间.本文提出的模型旨在为挖掘人类动力学中潜在的机制提供一个可参考的框架.未来的工作将集中于通过对模型参数的调整以进一步探讨其他类型人类行为中所表现出的幂律分布指数.

参考文献

- 1 时翔, 娄国伟, 李兴国. 地面目标辐射截面的建模与分析. 科学通报, 2007, 52: 2206—2210
- 2 张曙光, 金学松. 高速轮轨系统动力学研究体系及其系统建模. 科学通报, 2007, 52: 855—860
- 3 王友渔, 黄田, 梅江平, 等. 考虑复杂机架柔性的并联构型装备刚度半解析建模方法. 科学通报, 2008, 53: 364—372
- 4 王仁华, 戴礼荣, 凌震华, 等. 基于统计建模的可训练单元挑选语音合成方法. 科学通报, 2009, 54: 1133—1138
- 5 李安安, 刘谦, 曾绍群, 等. 高分辨数字人体三维结构数据集的构建与可视化. 科学通报, 2008, 53: 1189—1195
- 6 Haight F A. Handbook of the Poisson Distribution. New York: Wiley, 1967
- 7 Li N N, Zhang N, Zhou T. Empirical analysis on temporal statistics of human correspondence patterns. Phys A-Stat Mech Appl, 2008, 387: 6391—6394
- 8 Hu H B, Han D Y. Empirical analysis of individual popularity and activity on an online music service system. Phys A-Stat Mech Appl, 2008, 387: 5916—5921
- 9 Goncalves B, Ramasco J J. Human dynamics revealed through web analytics. Phys Rev E, 2008, 78: 26123
- 10 Candia J, González Marta C, Wang P, et al. Uncovering individual and collective human dynamics from mobile phone records. J Phys A-Math Theo, 2008, 41: 224015
- 11 Barabási A L. The origin of bursts and heavy tails in human dynamics. Nature, 2005, 435: 207—211
- 12 Vázquez A, Dezsó O J. Modeling bursts and heavy tails in human dynamics. Phys Rev E, 2006, 73: 36127
- 13 Bedogne C, Rodgers G J. A continuous model of human dynamics. Phys A, 2007, 385: 356—362
- 14 Blanchard Ph, Hongler M O. Modeling human activity in the spirit of Barabási's queueing systems. Phys Rev E, 2007, 75: 26102
- 15 Cajueiro D O, Maldonado W L. Role of optimization in the human dynamics of task execution. Phys Rev E, 2008, 77: 35101
- 16 Vázquez A. Impact of memory on human dynamics. Phys A-Stat Mech Appl, 2007, 373: 747—752
- 17 Han X P, Zhou T, Wang B H. Modeling human dynamics with adaptive interest. New J Phys, 2008, 10: 073010
- 18 Ebel H, Mielsch L I, Bornholdt S. Scale-free topology of e-mail networks. Phys Rev E, 2002, 66: 35103