

航空旅客群体移动行为特性分析

黄飞虎 彭舰 由明阳

Analyses of characteristics of air passenger group mobility behaviors

Huang Fei-Hu Peng Jian You Ming-Yang

引用信息 Citation: *Acta Physica Sinica*, 65, 228901 (2016) DOI: 10.7498/aps.65.228901

在线阅读 View online: <http://dx.doi.org/10.7498/aps.65.228901>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2016/V65/I22>

您可能感兴趣的其他文章

Articles you may be interested in

基于优化核极限学习机的风电功率时间序列预测

Wind power time series prediction using optimized kernel extreme learning machine method

物理学报.2016, 65(13): 130501 <http://dx.doi.org/10.7498/aps.65.130501>

自调整平滑区间粒子滤波平滑算法

Auto-adjust lag particle filter smoothing algorithm for non-linear state estimation

物理学报.2016, 65(4): 040502 <http://dx.doi.org/10.7498/aps.65.040502>

小管径气液两相流空隙率波传播的多尺度相关性

Multi-scale cross-correlation characteristics of void fraction wave propagation for gas-liquid two-phase flows in small diameter pipe

物理学报.2016, 65(1): 010501 <http://dx.doi.org/10.7498/aps.65.010501>

语音信号序列的 Volterra 预测模型

Volterra prediction model for speech signal series

物理学报.2015, 64(20): 200507 <http://dx.doi.org/10.7498/aps.64.200507>

基于改进教学优化算法的 Hermite 正交基神经网络混沌时间序列预测

Hermite orthogonal basis neural network based on improved teaching-learning-based optimization algorithm for chaotic time series prediction

物理学报.2015, 64(20): 200506 <http://dx.doi.org/10.7498/aps.64.200506>

航空旅客群体移动行为特性分析*

黄飞虎 彭舰[†] 由明阳

(四川大学计算机学院, 成都 610065)

(2016年7月3日收到; 2016年8月16日收到修改稿)

利用民航旅客订座数据分析了航空旅客群体移动的特性. 研究发现: 航空网络对于旅客的出行有一定的限制, 从而使得群体的出行距离更符合拉伸指数分布且没有无标度特性; 旅客的出行时间间隔满足截断幂律分布而不是幂律分布; 同时, 节假日等时间因素对旅客的出行有很大影响, 特别是春节、暑假、国庆这些时间节点使得旅客的出行量远大于其他时间节点, 而且在假节日出行的旅客, 更喜欢提前购票. 通过对吞吐量的分析, 发现吞吐量具有一定的周期性, 易受节假日的影响. 另外, 机场之间的吞吐量与其相互通航的城市数量有很大关系.

关键词: 航空旅客, 移动特性, 吞吐量

PACS: 89.40.Dd, 05.45.Tp

DOI: 10.7498/aps.65.228901

1 引言

随着航空运输业的发展以及国民经济的增长, 越来越多的人选择乘坐飞机出行. 对民航旅客出行数据进行挖掘和分析, 实证民航旅客的出行特性, 发现民航客运流的演化规律, 可以为航空公司和民航相关部门合理规划航空网络、提高管理效率提供决策支持, 具有实际应用价值.

2005年, Barabasi^[1]在《自然》发表论文提出人类行为的时间间隔分布具有幂率特性. 这使得很多学者开始对人类行为进行研究^[2-6]. 利用航空数据探索人类行为的时间特性和空间特性也得到很多学者的关注. 人类行为时间特性是指人多次从事某特定事件表现出来的时间上的统计规律^[7]. 文献^[1]统计了人们收发邮件的时间间隔分布; 文献^[8]提出阵发性的计算公式对时间间隔分布进行描述; 文献^[9]提出人类行为事件的时间序列也可以用记忆性进行刻画; Zhou等^[10]通过对用户观看

电影的时间间隔进行分析, 观察到了用户行为的周期性和波动性. 人类行为的空间特性是指人类活动在空间上表现出的特性, 最直观的就是出行距离分布. Brockmann等^[11]通过研究钞票的移动轨迹, 间接反映人的出行轨迹; 文献^[12]通过个体移动的回旋半径, 均方位移来描述个体的移动特性; Song等^[13]利用手机通讯记录, 探索了人类出行规律的可预测性.

吞吐量特性是航空旅客在各个机场聚集情况的总体反映. 张子栋和徐建红^[14]探索了影响机场吞吐量的主要因素; Zhang等^[15]通过分析2010年的民航数据, 发现除了吞吐量之外, 工业输出和机场度等因素也是影响航线之间连接的因素; 文献^[16]研究了香港的航空数据, 并利用时间序列模型对吞吐量进行预测, 得出香港的机场吞吐量会呈现稳步的整张态势; Letavkova等^[17]结合小波变换和时间序列模型提出一个新的预测吞吐量的方法. 总的来说, 国内外学者对于机场吞吐量的研究主要集中在预测和影响因素分析这两个方面.

* 国家自然科学基金 (批准号: U1333113, 61303204) 资助的课题.

[†] 通信作者. E-mail: penguest@163.com

本文利用航空数据探索航空旅客群体移动的特性,也主要从时间和空间两个方面进行分析.同时,机场吞吐量也是航空旅客在机场航线上的聚集反映,因此本文也对机场吞吐量进行了讨论.本文的安排如下:第2节介绍数据预处理方式;第3—5小节分别介绍航空旅客移动行为的空间特性、时间特性以及在吞吐量方面表现的特性;最后为文章的总结.

2 数据预处理

原始数据包含从2014年1月—2014年12月的民航旅客出行数据.航空公司提供的出行记录数据集中的每条记录即为旅客的一次出行记录.本文在Spark大数据分析的平台上对数据做了如下处理:

- 1) 除去记录中存在冲突和错误的的数据;
- 2) 本文考虑到国外机场吞吐量数据获取的不完整情况,因此除去了国外的航线记录,主要研究国内航空旅客的群体移动特性;
- 3) 为了消除临时航线对研究旅客出行特性的影响,统计了航线中航班的运行情况,最后在实验数据中删除了周航班飞行次数少于1次的不固定航线.

经过上述数据预处理步骤,获得了15616181位旅客的30069995条出行记录,共有119个机场.

3 时间特性分析

3.1 出行间隔天数分析

出行间隔天数反映了旅客的乘机频率,是刻画民航旅客出行行为的重要度量参数之一.本文统计了航空旅客的出行间隔天数分布(间隔天数为0的数据未显示).从图1中可以看出,与以往研究中发现的移动数据间隔时间服从幂律分布的结果不同^[10],航空旅客乘机间隔天数更符合指数截断幂律分布,其分布函数为 $P(d) = d^{-\alpha} e^{-\lambda d}$,其中 $\alpha = 1.08, \lambda = 0.08$.在人类移动模型的研究中也观察到了这种现象^[11,12].航空出行属于远距离出行,旅途花费时间 t 和费用 m 是用户需要考虑的因素或者称为出行所需代价.下面我们将从这方面对出行时间间隔分布进行讨论.文献^[18]利用最大熵原则得出了用户出行所需代价与旅途花费时间和费用之间的关系为

$$P(c) \propto \exp(c/\bar{c}), \tag{1}$$

其中, c 为代价, \bar{c} 为所有线路的平均代价, c 的表达式为

$$c \approx \eta t + \mu m, \tag{2}$$

其中,旅途花费时间 $t \approx \phi \ln r + \omega r + \psi$,费用 $m \approx \nu r$, r 为移动距离, $\eta, \mu, \phi, \omega, \psi$ 为系数.由于人类的出行具有高度的有界性,即经常会在短距离的范围活动,远距离活动的次数则相对很少.因此可以认为活动的时间间隔 d 与移动距离 r 呈正比关系,即 $d \propto r$.于是可以得到代价 c 与时间间隔 d 的关系为

$$c \approx \eta \phi \ln d + (\eta \omega + \mu \nu) d + \eta \psi. \tag{3}$$

因此,时间间隔的分布函数为

$$P(d) \propto d^{-\alpha} e^{-\lambda d}. \tag{4}$$

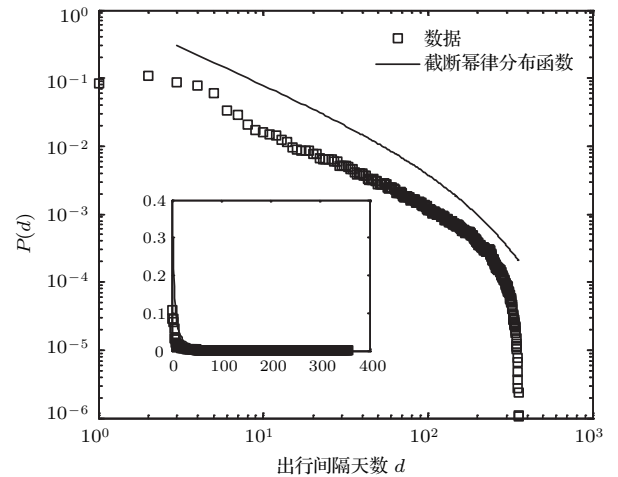


图1 出行间隔时间分布

Fig. 1. Distribution of travel interval time.

从上面的讨论可以看出,在航空出行这样的远距离移动中,在用户考虑旅途花费时间和费用的情况下,出行时间间隔分布满足指数截断幂律分布.这种现象表明航空旅客的出行时间间隔不是均匀分布的.下面将通过阵发性来衡量这种不均匀的程度.所谓阵发性,即是某些时候在很短的时间内会有很多事件发生,随后又会有较长时间的空档期,其计算公式为^[8]

$$B = \frac{\sigma_t - m_t}{\sigma_t + m_t}, \tag{5}$$

其中 σ_t 表示分布的标准差, m_t 为均值.从(5)式可以看出,对于泊松分布,其均值和标准差相等,因此 $B = 0$;对于严重的胖尾分布,标准差远大于均

值, B 靠近 1; 对于周期性规则信号, 标准差为 0, $B = -1$. 这里计算得到的 $B = 0.36$ ($B \in [-1, 1]$), 这说明航空旅客出行有较强的阵发性. 从前面出行间隔时间满足指数截断幂律分布可以看出, 分布具有明显的拖尾现象, 这就使得方差大于均值, 导致阵发性取值较大. 从旅客出行特点来看, 具有较强的阵发性这很符合实际情况. 当旅客间的隔时间为几天时, 乘坐飞机的概率很大, 这是因为旅客在出行之后, 一般会在几天内乘坐飞机返回出发地. 航空出行为远距离移动, 人们不可能像坐公交一样随时乘飞机出行, 因此会间隔较长的时间再次乘坐飞机.

3.2 提前购票天数分析

一般情况下, 航空旅客在选择乘机出行时大多会选择提前多天购买机票, 而提前购票天数可以反映出旅客的出行目的的可计划性和紧急程度, 同时机票价格与购票日期之间也有关联, 在一定程度上可以反映旅客类型. 图 2 为民航旅客的提前购票天数的分布. 分布拟合结果显示, 民航旅客的提前购票天数可以用拉伸指数分布拟合, 分布函数为 $e^{-\lambda x^\beta}$, 其中参数分别为 $\lambda = 0.2, \beta = 0.64$. 从图 2 可以看出大部分旅客提前购票天数都比较短. 一般情况下提前购票天数越短, 机票价格的折扣也相对越少, 说明这些对价格不敏感的旅客很有可能是公务出差, 因为大部分旅客对机票价格都很敏感 [19]. 如果说提前购票时间在一定程度上能反映旅客的类型, 那么这些提前购票时间短的旅客很可能是因为工作原因出行, 因此出行日期会在工作日期间. 同时提前购票日期较长的旅客则是非工作原因出行, 则出行日期会在节假日期间. 为了验证这一猜想, 本文提取了旅客数据中出行次数不少于 24 次的常旅客出行数据, 并分析了提前购票时间与出行日期类型的关系.

图 3 为常旅客出行日期类型与提前购票天数的对应关系. 从图 3 中可以看出, 随着提前购票天数的增加, 在假节日出行的旅客所占的比例逐渐增大, 说明以非工作为目的旅客, 或者可以说是与旅游为目的的旅客, 他们趋向更早地安排出行计划. 在图 3 中, 工作日出行的旅客比例也是最多的, 节假日出行的旅客比例最少, 这也正好反映了图 2 的结果, 即大部分旅客提前购票天数都比较短, 而提前购票时间长的旅客比例很少.

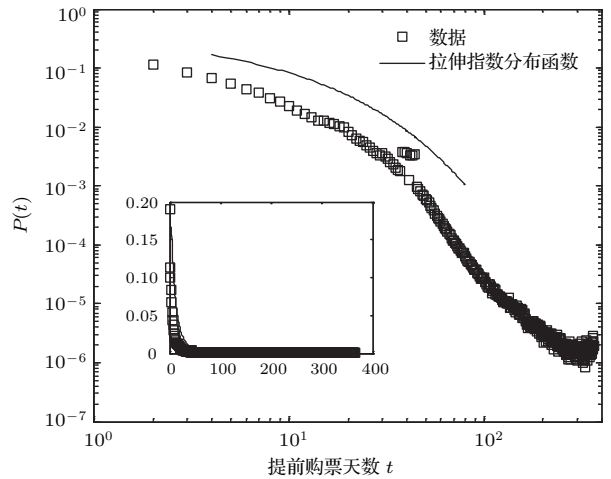


图 2 提前购票时间分布

Fig. 2. Distribution of advance booking time.

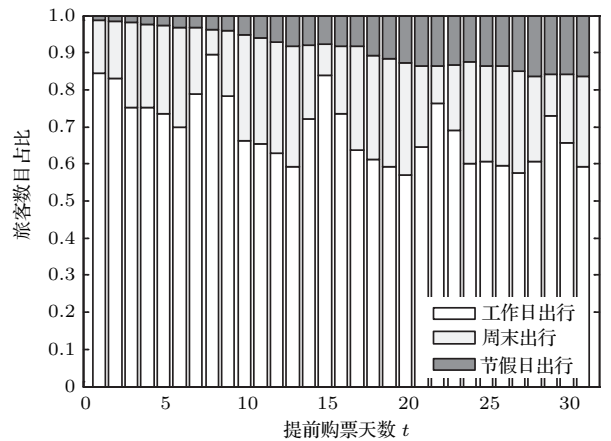


图 3 提前购票时间与出行日期类型

Fig. 3. Advance booking time and type of travel date.

4 空间特性分析

4.1 出行距离分布分析

出行距离是描述旅客出行的一个重要特性, 在公交、出租车等基础设施网络中已有广泛研究 [20–22]. 王明生等 [23] 对石家庄的公交数据进行研究发现, 人们的出行距离符合指数分布. 而 Camille 等 [24] 对伦敦市的地铁数据研究发现地铁乘客的出行距离满足负二项分布, 并指出这种现象是由于城市经济分布结构的影响. 本文统计了旅客出发机场和到达机场之间的距离, 发现国内旅客的出行距离分布并没有无标度特性, 如图 4 所示. 从图 4 (黑色) 中可以看出, 出行距离分布在 900—1000 km 出现了峰值, 表明这个范围是国内航空旅客的典型出行距离. 国内航空旅客的出行距

离总体上并不符合特定分布, 但是通过研究发现航空旅客出行距离分布与航空网络有密切联系. 本文统计了机场航线的距离分布(白色), 可以看出每个机场与相邻机场的距离分布也满足拉伸指数分布. 这表明航空网络对航空旅客出行有极大的影响. 航空旅客出行属于远距离移动, 必须依赖航空网络, 因此这也就限制了旅客的出行距离. 同样, 以公交、出租车数据为基础的人类移动特性研究, 也依赖于道路网络或者城市经济分布等因素, 从而使得人们出行距离分布没有无标度特性. 在 Brockmann 等^[11]的研究中, 钞票的流动并不受限于交通网络, 才使得钞票流动的距离分布满足幂律特性.

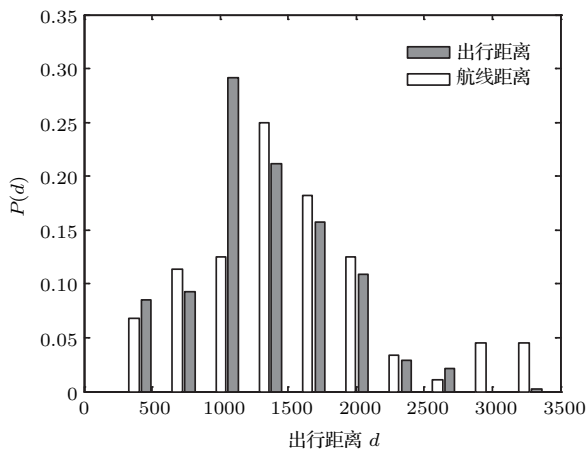


图 4 出行距离和航线距离分布

Fig. 4. Distribution of travel distance and route distance.

4.2 出行距离特性分析

人类出行特性研究中, 主要使用回旋半径和均方位移 (mean squared displacement, MSD) 随着时间的增长规律来表示人类的扩散特性.

图 5 为民航旅客回旋半径随出行次数的变化趋势, 图中的回旋半径由全体旅客出行次数为 n 时的回旋半径值按出现次数加权计算取得. 全体民航旅客经过 n 次出行后的回旋半径定义为^[12]

$$r_g(n) = \sqrt{\frac{1}{n} \sum_{k=1}^n (\mathbf{x}_k - \mathbf{x}_t)^2}, \quad (6)$$

上式中, \mathbf{x}_t 表示全体旅客在 n 次出行中最频繁访问的机场, \mathbf{x}_k 表示旅客每次出行的到达机场. 特别地对于出行总次数为 1 的旅客 \mathbf{x}_t 和 \mathbf{x}_k 分别表示出发机场和到达机场. 从图 5 可以看出, 不同于以往的研究结果, 民航旅客整体的回旋半径在出行次数

1 就已经达到最大值, 而出行次数 n 为 2 时回旋半径急剧下降并在 10 次至 21 次左右达到了平衡, 当出行次数大于 21 次后回旋半径又出现了下降趋势. 造成上述现象的原因主要是当 n 为 1 时包含了大量出行总次数为 1 的旅客, 这部分旅客普遍选择的航线距离较远, 他们的回旋半径即为实际的出行距离; 而当出行次数 n 大于 1 时民航旅客的往返率很高, 即旅客在到达目的机场后, 以很大概率返回上次出发的机场, 这就造成了旅客回旋半径的迅速减少. 出行次数在 10—21 次之间的旅客为常旅客的主体, 他们在民航网络中游走情况比较均衡. 而出行次数在 21—45 次的旅客所占比例很小, 这部分旅客有机会选择距离较近的航线, 并且出行频率很高, 因而回旋半径随出行次数的增加出现了下降的趋势. 出行次数大于 45 次的旅客所占比例极小, 出现了不同于其他类型旅客的特征.

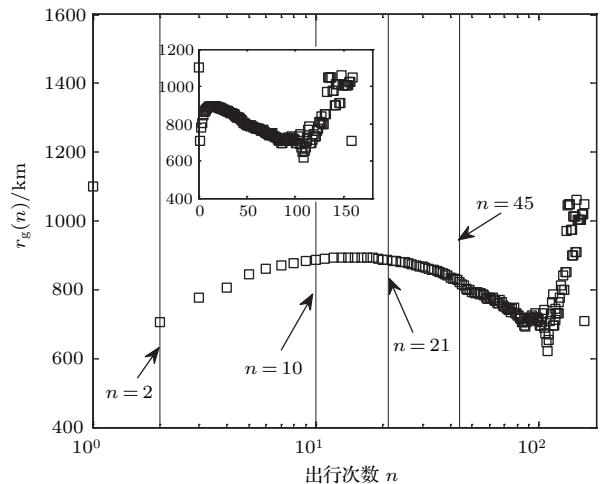


图 5 回旋半径分布

Fig. 5. Distribution of cyclotron radius.

图 6 为民航旅客的均方位移分布, 民航旅客经过 n 次出行后的 MSD 定义为^[12]:

$$MSD(n) = \langle (\mathbf{x}_n - \mathbf{x}_0)^2 \rangle, \quad (7)$$

其中 \mathbf{x}_0 为旅客出行记录中首个出发机场的位置, $\langle \cdot \rangle$ 运算符表示计算 n 次出行记录的平均值. 从图中可以看出, 均方位移在出行次数为 2 时出现了极值, 这是因为大量出行总次数为 2 的旅客在民航网络中的出行距离较远. 出现次数在 [2, 45] 范围内的旅客访问了民航不同距离的航线, 随着出现次数的增大, 相应的旅客数量变少, 因而出现了均方位移下降的趋势. 而出行次数大于 45 次的旅客访问航

线具有各向异性, 因此均方位移也出现了异质性变化.

从回旋半径和均方位移来看, 大部分旅客的移动距离并没有随出行次数的增加而无限增加, 而是在一定的范围内变化, 这说明航空旅客的出行具有高度的有界性特征.

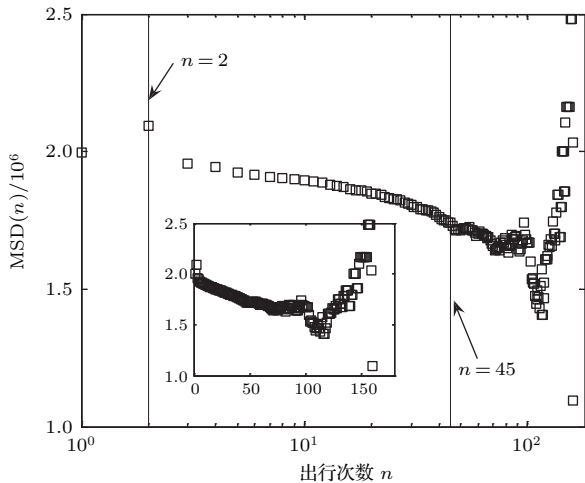


图6 均方位移分布

Fig. 6. Distribution of MSD.

5 吞吐量特性分析

5.1 周期与波动

周期性数据序列的自相关系数呈连续振荡波形, 对吞吐量的周期性检测对于吞吐量的预测有一定的指导意义^[25]. 图7(a)为某机场一年吞吐量序列的自相关图, 实验中设置了滞后阶数为0到50. 自相关系数呈连续振荡波形, 而且在第7, 14, 21, 28, 35, 42, 49阶有峰值. 在每个峰值之间也有起伏现象, 随着滞后阶数的增加自相关系数慢慢收敛. 图7(b)为人工设置的周期为7的数据序列, 从图中也能看出第7, 14, 21, 28, 35, 42, 49阶有峰值, 随着滞后阶数的增加, 自相关系数也慢慢收敛. 通过人工设置的周期为7的数据自相关系数来看, 该机场吞吐量自相关系数与其有相同的地方, 即机场吞吐量的变化是以7天为一个周期. 实验中对数据中119个机场的自相关系数图进行观察, 也得到了类似结论. Choi等^[26]在实验中也用该方法探索数据的周期性. 通过查询航空公司的航班安排, 我们发现航班是按周进行安排的, 这正好解释了为什么机场吞吐量得到变化周期是7天. 在4.1节中分析了旅客出行距离的分布情况, 发现旅客的出行受航空网络的影响. 航空网络是由航线组成的, 航线基本

不会有变化, 同时每条航线的乘客也是固定的, 因此机场的吞吐量在很大程度上受航班安排的影响, 所以机场航班的安排规律使得机场吞吐量以周为周期.

对于吞吐量的波动性, 本文主要研究假期对于机场吞吐量的影响. 在实验中, 分别计算了每个机场包含节假日数据的方差 v_1 以及除去节假日后的方差 v_2 . 为了方便显示, 实验中如果 v_1 大于 v_2 , 则取值为1, 否则取值为-1. 图8可以看出, 大部分机场吞吐量的方差 v_1 比 v_2 大. 主要是在节假日期间旅客的出行增多, 吞吐量增加, 使得包含节假日吞吐量的方差大于工作日吞吐量的方差. 这说明假期对于机场吞吐量也有一定影响.

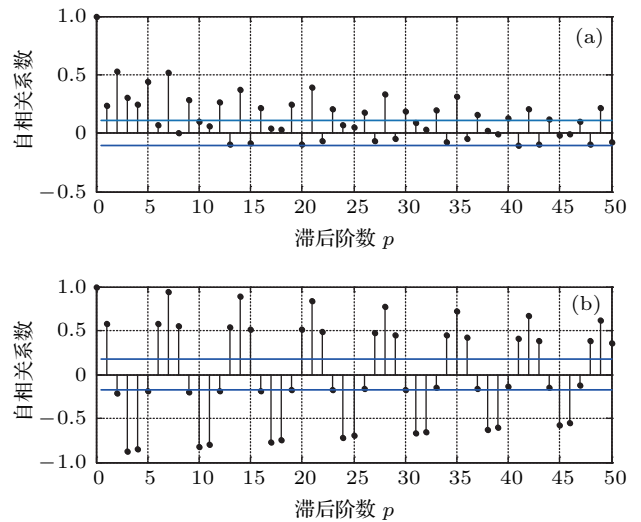


图7 自相关系数 (a) 某机场一年吞吐量序列的自相关系数; (b) 周期为7的数据序列自相关系数

Fig. 7. Autocorrelation coefficient: (a) Autocorrelation coefficient of an airport's throughput series in one year; (b) autocorrelation coefficient of data sequence with period of 7.

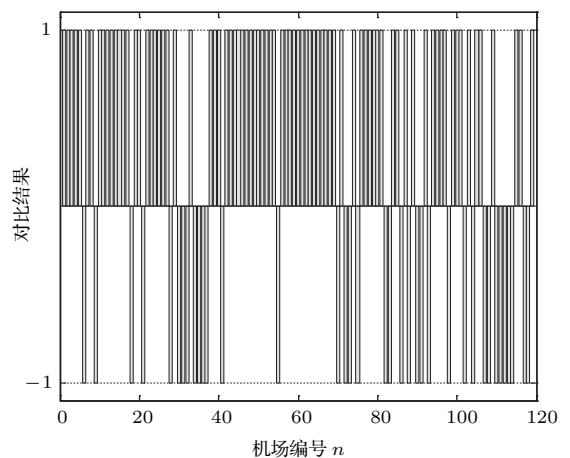


图8 方差对比结果

Fig. 8. Variance comparison results.

5.2 季节与趋势

为了探索吞吐量的趋势变化,统计了每个机场的月吞吐量,然后进行拟合.实验中利用matlab的fit函数进行二元混合高斯模型拟合, R^2 的平均值为0.81,表1为图9和图10的拟合结果.

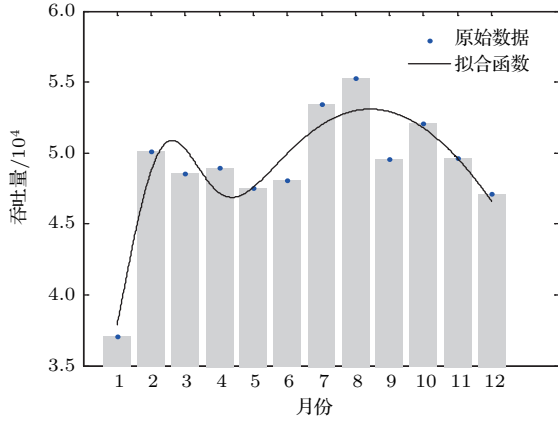


图9 所有机场月平均吞吐量

Fig. 9. Monthly average throughput of all airports.

图9和图10分别为所有机场月平均吞吐量和某机场的月吞吐量及拟合结果.从总体的平均吞吐

量来看,2,7,8,10月份的吞吐量比较大.在5.2节的讨论中,我们知道假期对旅客的出行有影响,这几个月正好包含国内的几个长假.2月份是寒假期间,同时也有国内旅客流动最大的春节长假;7月和8月份包含暑假;10月份有国庆长假.这就使得国内的机场吞吐量在这几个月较大,表现出每年都相同的季节性和变化趋势.

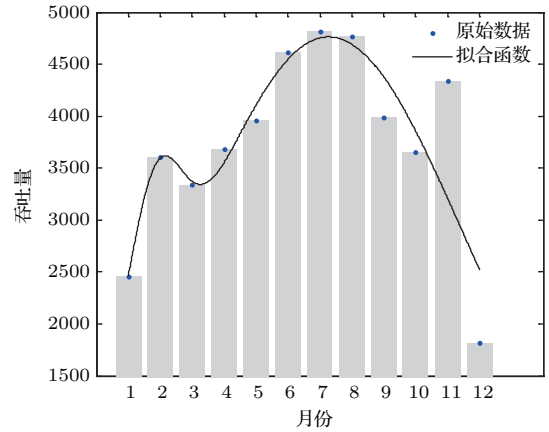


图10 某机场月吞吐量

Fig. 10. An airport's monthly throughput.

表1 拟合结果参数统计表

Table 1. The statistics results of parameter fitting.

	a_1	a_2	μ_1	μ_2	σ_1	σ_2	R^2
某机场月吞吐量	4763.2	1481.4	7.2819	1.8195	5.921	1.1871	0.77477
所有机场月平均吞吐量	53092	14463	8.4478	2.2113	9.8351	1.5673	0.85696

5.3 机场与吞吐量的关系

在出行距离分析中,航空旅客的出行受到机场网络的影响.那吞吐量与机场又有怎样的关系呢?在已有的研究中,机场吞吐量与机场所在城市的经济、人口等因素有关.本文将从另一个角度来探讨这个问题.

首先,给出机场差异度的定义

$$S_{ab} = d_a + d_b - d_{ab}, \quad (8)$$

d_a 和 d_b 分别为与机场A和机场B通航的城市数量, d_{ab} 为城市A和城市B共有的通航城市数量.根据(4)式, S_{ab} 的值越小,城市A和B的相似性越高.图11为差异度为1的两个城市A和B的吞吐量序

列.可以发现这两个城市之间的吞吐量序列非常相似.原因很明显,两个城市A和B只有一个通航城市不一样,所以吞吐量序列有很高的相似性.实验中利用自组织特征映射网络(self-organizing feature map, SOM)将119个机场的吞吐量序列分为10个类别[27],分类采用的方法是动态时间规整(dynamic time warping, DTW)[28].图12中统计了有相同差异度的城市之间吞吐量序列分类的数量.在机场差异度为10时,机场的吞吐量总类就达到分类的总数了.可以看出,随着差异度的增加有相同差异度的城市之间吞吐量序列也逐渐变得差异很大.这表明机场之间吞吐量的相似性与它们之间通航的城市负相关.

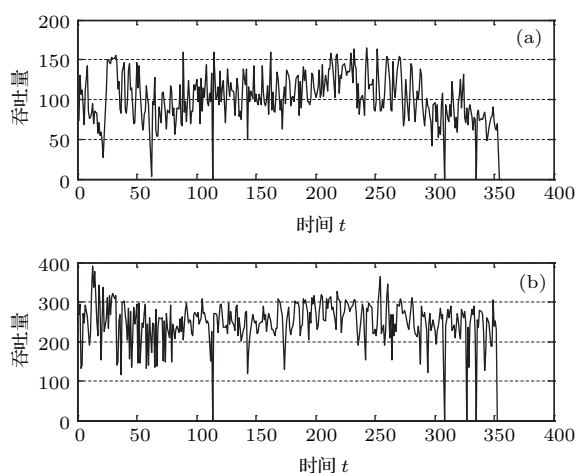


图 11 吞吐量序列 (a)大同机场吞吐量序列; (b)佳木斯机场吞吐量序列

Fig. 11. Throughput sequence: (a) Datong airport throughput sequence; (b) Jiamusi Airport throughput sequence.

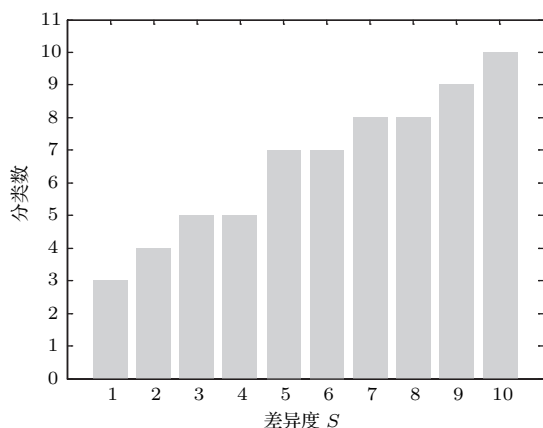


图 12 相同差异度的机场吞吐量序列分类的数量

Fig. 12. Classification number of airport throughput sequence with the same difference degree.

6 总 结

利用民航旅客的出行数据, 本文主要讨论了国内航空旅客群体移动行为在空间、时间以及吞吐量方面表现出来的特性. 由于航空旅客的出行依赖于航空网络, 这也使得航空旅客的出行特性与其他人类活动的出行特性具有一定的差异性.

从实验分析来看, 航空旅客出行距离的分布更符合拉伸指数分布却并不具有无标度特性. 而且从回旋半径和均方位移来看, 大部分旅客的移动距离并没有随出行次数的增加而无限增加, 而是在一定的范围内变化, 这说明航空旅客的出行具有高度的有界性特征, 即更趋向于在一个有限的范围活动.

航空旅客的出行间隔时间也不是幂律分布, 而

是满足截断幂律分布. 同时通过分析旅客提前购票的时间, 发现随着提前购票天数的增加, 在节假日出行的旅客所占的比例逐渐增大. 这也使得节假日对于吞吐量产生了影响.

在吞吐量特性分析的过程中, 计算了数据序列的自相关系数, 发现吞吐量序列呈连续振荡波形, 并每隔 7 天就会产生一个峰值, 这表明其具有一定的周期性. 通过对比了除去节假日与包含节假日之间的吞吐量方差, 讨论了节假日对于吞吐量的影响. 同时, 利用二元高斯混合模型拟合机场的月吞吐量, 得到了不错的拟合效果. 从拟合的结果来看, 春节、暑假、国庆三个时间节点的旅客吞吐量很大. 这很符合国内旅客的出行特点, 春节和国庆是旅客出行的高峰期, 而暑假则是很多学生选择出行的时间. 为了研究机场之间吞吐量的关系, 定义了机场差异度的计算, 结果发现机场之间的吞吐量与机场之间通航的城市有一定相关性.

参考文献

- [1] Barabasi A L 2005 *Nature* **435** 207
- [2] Guo J L 2010 *Acta Phys. Sin.* **59** 3851 (in Chinese) [郭进利 2010 物理学报 **59** 3851]
- [3] Zhao F, Liu J F, Cha Y L, Zhou T 2011 *Acta Phys. Sin.* **60** 118902 (in Chinese) [赵飞, 刘金虎, 查一龙, 周涛 2011 物理学报 **60** 118902]
- [4] Zhao Z D, Huang Z G, Huang L, Liu H, Lai Y C 2014 *Phys. Rev. E* **90** 050802
- [5] Pappalardo L, Simini F, Rinzivillo S, Pedreschi D, Giannotti F, Barabasi A L 2015 *Nat. Commun.* **6** 8166
- [6] Cui P, Liu H, Aggarwal C, Wang F 2016 *IEEE Intell. Syst.* **31** 77
- [7] Zhou T, Han X P, Yan X Y, Yang Z M, Zhao Z D, Wang B H 2013 *J. UESTC* **4** 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 **4** 481]
- [8] Goh K I, Barabasi A L 2006 *Euro. Phys. Lett.* **81** 48002
- [9] Cai S M, Fu Z Q, Zhou T, Gu J, Zhou P L 2009 *EPL* **87** 68001
- [10] Zhou T, Kiet H A T, Kim B J, Wang B H, Holme P 2008 *Euro. Phys. Lett.* **82** 28002
- [11] Brockmann D, Hufnagel L, Geisel T 2006 *Nature* **439** 462
- [12] González M C, Hidalgo C A, Barabási A L 2008 *Nature* **453** 779
- [13] Song C, Qu Z, Blumm N, Barabasi A L 2010 *Science* **327** 1018
- [14] Zhang Z D, Xu J H 2007 *Urban Transp. Chin.* **5** 54 (in Chinese) [张子栋, 徐建红 2007 城市交通 **5** 54]
- [15] Zhang Y, Peng T, Fu C Y, Cheng S W 2016 *J. Air Transp. Mgmt.* **50** 12

- [16] Tsui W H K, Balli H O, Gilbey A, Gow H 2014 *Tourism Mgmt.* **42** 62
- [17] Letavkova D, Matuskova S, Kebo V 2015 *16th International Carpathian Control Conference Miskolc, Hungary, May 27–30 2015*, pp295–299
- [18] Yan X Y, Han X P, Wang B H, Zhou T 2012 *Sci. Rep.* **3** 454
- [19] Ghobrial A, Kanafani A 1995 *J. Transp. Eng.* **121** 135
- [20] Jiang B, Yin J, Zhao S 2009 *Phys. Rev. E* **80** 1711
- [21] Liang X, Zheng X, Lü W, Zhu T, Xu K 2011 *Physica A* **391** 2315
- [22] Peng C, Jin X, Wong K C, Shi M, Liò P 2012 *PLoS ONE* **7** e34487
- [23] Wang M S, Huang L, Yan X Y 2012 *J. UESTC* **41** 2 (in Chinese) [王明生, 黄琳, 闫小勇 2012 电子科技大学学报 **41** 2]
- [24] Roth C, Kang S M, Batty M 2011 *PLoS One* **6** e15923
- [25] Zhang S J, Qi L X 2003 *Time Series Analysis Introductory Tutorial* (Beijing: Tsinghua University Press) p12–13 (in Chinese) [张树京, 齐立心 2003 时间序列分析简明教程 (北京: 清华大学出版社) 第12—13页]
- [26] Choi T M, Yu Y, Au K F 2011 *Decis. Support Syst.* **51** 130
- [27] Brugger D, Bogdan M, Rosenstiel W 2008 *IEEE Trans. Neural Netw.* **19** 442
- [28] Song C, Pei T 2012 *Prog. Geog.* **31** 1317 (in Chinese) [宋辞, 裴韬 2012 地理科学进展 **31** 1317]

Analyses of characteristics of air passenger group mobility behaviors*

Huang Fei-Hu Peng Jian[†] You Ming-Yang

(College of Computer Science, Sichuan University, Chengdu 610065, China)

(Received 3 July 2016; revised manuscript received 16 August 2016)

Abstract

In this paper, we use the civil aviation passenger reservation record data to analyze the characteristics of air passenger group mobility behaviors. We find that the airline network has restriction on passengers traveling, which causes the group travel length distribution to be more consistent with the stretched exponential distribution and to have no scale-free property. The passenger travel interval time satisfies a truncated power law distribution rather than the power law distribution. Meanwhile, holidays have a great influence on passengers traveling. In particular, during the Spring Festival, the summer vacation, and National Day, the number of traveling passengers is far more than at any other time. And passengers who will travel during the holidays prefer to book tickets in advance. By analyzing the throughput, we find that it has a periodicity and is easy to be affected by festival and holidays. Moreover, the throughput of the airport is greatly related to the number of cities which are navigable to each other.

Keywords: air passenger, mobility characteristics, throughput

PACS: 89.40.Dd, 05.45.Tp

DOI: 10.7498/aps.65.228901

* Project supported by the National Natural Science Foundation of China (Grant Nos. U1333113, 61303204).

† Corresponding author. E-mail: penguest@163.com