

DOI:10.3969/j.issn.1674-0696.2020.08.02

公交通勤个体出行特征图谱构建及相似性判别

梁 泉¹, 翁剑成¹, 周 伟², 荣 建¹

(1. 交通运输部管理干部学院 道路教研部, 北京 101601; 2. 北京工业大学 城市交通学院, 北京 100124;
3. 中华人民共和国交通运输部, 北京 100736)

摘要:从个体角度获取公共交通通勤乘客出行行为特征,有助于更准确地把握公交总体出行规律,更好地满足乘客的出行需求。基于公共交通刷卡数据与线站数据,针对公交通勤出行个体,借助图谱可视化表达优势,研究形成了以横轴为时间序列轴、纵轴为位置轴、节点大小为方向角度的公共交通通勤乘客出行特征图谱构建方法。在此基础上,选用出行方式、方向、时间和线路等属性指标,采用结构相似度与最长公共子序列相结合的分析方法,实现了图谱相似性判定。以个体乘客一周的公共交通出行数据为例,验证了研究结果的合理性和有效性。研究结果能够直观反映个体出行过程,为出行特征表达与提取提供了新思路。

关键词:交通工程;通勤乘客;特征图谱;相似性判别;个体特征

中图分类号:U491.1

文献标志码:A

文章编号:1674-0696(2020)08-006-08

Specific Chromatogram Construction and Similarity Identification of Individual Travel of Public Transport Commuters

LIANG Quan¹, WENG Jiancheng², ZHOU Wei³, RONG Jian²

(1. Road Teaching and Research Department, Transport Management Institute Ministry of Transport of the People's Republic of China; Beijing 101601, China; 2. College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China; 3. Ministry of Transport of the People's Republic of China, Beijing 100736, China)

Abstract: It is helpful to grasp the overall travel rules of public transport more accurately and better meet the travel needs of passengers to obtain the travel behavior characteristics of public transport commuters from the perspective of individuals. Based on the public transport card swiping data and line station data, by use of the representation advantage of graph visualization of public transport commuters, the travel graph construction method of public transport commuting trip was studied, which took the horizontal axis as the time series axis, the vertical axis as the position axis, and the node size as the direction angle. On this basis, the method of structure similarity combined with the longest common subsequence was used to realize the similarity identification of graph, which selected attribute indicators such as travel mode, direction, time and route. Taking the one-week public transport travel data of individual passengers as an example, the rationality and validity of the research results were verified. The research results can intuitively reflect the individual travel process, which provides a new idea for the expression and extraction of trip characteristics.

Key words: traffic engineering; commute passenger; travel graph; similarity identification; individual characteristics

0 引言

公共交通作为受众最广、影响最大的出行方式,

已成为通勤乘客出行的主体^[1]。目前北京市公共交通通勤乘客日均客流 300 万人,日均出行时间 54 min,日均出行距离 19.4 km,通勤交通在城市交

收稿日期:2018-07-22;修订日期:2019-01-03

基金项目:国家自然科学基金项目(51578028);北京市“科技新星”计划项目(Z171100001117100)

第一作者:梁 泉(1989—),女,山东潍坊人,博士研究生,主要从事智能交通、交通运行监测方面的研究。E-mail: lquan0730@163.com

通信作者:翁剑成(1981—),男,浙江金华人,副教授,博士,主要从事智能交通、交通运行监测方面的研究。E-mail: youthweng@bjut.edu.cn

通系统中举足轻重^[2]。

为进一步提高公共交通服务水平,满足公共交通乘客出行需求,精确掌握乘客出行特征尤为重要。以往的公交出行特征提取主要通过抽样问卷、走访调研、网络调查等手段完成。周钱等^[3]应用结构方程模型等提炼出行特征参数;V.S.WALLE 等^[4]结合调查数据分析了公交出行的时空特征;王珂^[5]通过调查轨道沿线家庭居住和通勤情况,分析了居民的通勤方式、时耗和费用等特征;杨敏等^[6]通过出行者空间特征分析,提出了出行者家庭和工作因素是出行模式选取的重要因素。但由于调查成本高,调查样本往往偏少且覆盖面窄,故小样本居民出行调查对出行总体特征描述精确度不够,无法实现公共交通出行特征的精细化提取^[7]。同时,以往研究集中反映宏观交通运行状态,缺乏对个体出行特征的描述^[8-9]。事实上,个人属性存在较大差异,从个体交通出行特征入手则能更科学地获取乘客出行规律。

伴随云计算、互联网+等新兴技术快速发展,公交系统智能化水平极大提升,可实现每位公共交通出行者出行交易过程和车辆运行状态信息记录,综合网络通信、地理信息、GPS 定位及电子控制等,形成了个体交通出行多源数据,为更加准确和精细化的公交出行特征提取奠定了数据支撑。戴霄等^[10]分别提出了基于线路与公交 IC 卡数据的出行特征分析方法;J.J. BARRY 等^[11]融合公交 IC 卡与车辆 GPS 数据,得到公交车行驶轨迹,并利用数据挖掘算法分析了乘客上下车站点特征;ZHAO Jinhua 等^[12]结合 AFC (automatic fare collection) 及 AVC (automatic vehicle collection) 系统,对轨道交通出行的 OD 矩阵进行估算;N.KARASMAA^[13]通过将公交 IC 卡刷卡数据与 GPS 定位数据联合分析,实现了车辆运行指标计算及乘客出行起终点、换乘点判别。更重要的是,基于智能公交大数据时代的多源数据为从个体角度分析公共交通出行特征提供了可能。

图谱作为特征可视化表达方式,以符号形式描述对象间的概念及相互关系,构成联结的知识结构,

能实现特征的直观表达^[14]。通过医学数据提取实体、关系、属性等知识图谱的组成元素,对医学知识进行表示、抽取、融合、推理及质量评估,医疗图谱形成了最佳知识表征框架;一方面便于医生通过图谱动态推理制定合适的诊疗方案;另一方面图谱可视化表达可帮助病人了解诊疗过程^[15]。利用大样本学术信息(如期刊论文、专利报告等)对科学知识结构、关系与演化过程进行可视化,图书情报领域的引证图谱能自动展示文献与作者关系,并进一步分析科学知识演化过程与结果^[16]。同样,基于行车过程中驾驶员操作行为和车辆运行状态的时序变化特征和编码组合,驾驶行为图谱能形象化地描述驾驶操作过程,辨别不同驾驶员的驾驶习惯^[17]。总体而言,通过对原始数据进行加工和处理,图谱能对数据特征进行有效地组织与表达,刻画其过程变化特征。

因此,笔者基于公共交通多源数据,借助图谱对复杂信息可视化表达的优势,从个体角度构建了表征乘客时空轨迹特性的特征图谱,准确反映个体乘客时间分布与空间位置波动;并选取多子区间相似性判别方法,度量个体特征图谱相似性;为个体出行特征提取提供新思路,为不同类型公共交通通勤乘客出行行为准确辨识提供借鉴,为精细化公共交通出行服务提供支撑。

1 数据基础

笔者依托北京城市交通协同创新中心和综合交通协同运行与超级计算应用技术协同创新平台,以公共交通刷卡与线站数据为数据基础进行分析。公共交通刷卡数据包括地面公交 IC 卡刷卡和轨道 AFC 系统刷卡数据;公共交通线站数据包括地面公交和轨道交通线站数据。结合前期研究成果,已完成了基于多源数据的预处理及公交出行过程匹配,进而提取获得出行阶段数据^[18],为出行特征图谱构建奠定数据支撑。

笔者以某乘客 4 d 的出行数据为例,获得该乘客出行阶段数据示意,如表 1。

表 1 某乘客连续 4 d 出行阶段数据
Table 1 Travel stage data of passenger for sequential 4 days

出行参数		出行数据		
一卡通卡号		21283200		
方式		地铁		
上车时间	2014-09-23T07:43	2014-09-23T17:43	...	2014-09-26T16:28

(续表 1)

出行参数	出行数据			
下车时间	2014-09-23T08:10	2014-09-23T18:12	...	2014-09-26T16:45
上车线路	10	10	...	2
下车线路	10	10	...	10
路径距离/km	13.13	13.13	...	11.37
上车站点	双井	天安门	...	朝阳门
下车站点	安贞门	双井	...	双井
上车经纬度/(°)	39.8935N, 116.4615E	39.9770N, 116.4060E	...	39.9248N, 116.4346E
下车经纬度/(°)	39.9770N, 116.4060E	39.8935N, 116.4615E	...	39.8935N, 116.4615E

2 图谱构建与相似性判别

笔者选用北京市全体公交乘客连续一周的出行数据,针对公共交通通勤出行者,通过个体出行特征图谱构建及图谱相似性判别,提取公共交通通勤乘客时空特性,并实现相同个体出行行为分类与不同个体出行行为相似性判别。

定义公交常乘客为一周工作日中至少有 4 d 采用公共交通出行的乘客。针对通勤乘客刷卡时间间隔及往返情况的阈值分析,研究规定公交常乘客在一周中,刷卡最大时间间隔大于 7 h 且往返地点为闭合的情况出现次数不少于 3 次,则判定其为公共交通通勤出行者。

2.1 个体出行特征图谱绘制

考虑周一相比其它工作日客流量激增,随机因素较多,出行特征可能存在较大差异,故暂不考虑。笔者以个体乘客为分析单元,同一乘客相邻出行阶段数据为分析基础,制定出行特征图谱绘制步骤。

2.1.1 步骤 1

公共交通出行阶段数据排序。选取乘客连续 4 d 的出行阶段数据,并按刷卡时间进行排序。

2.1.2 步骤 2

绘制图谱位置坐标点。以时间序列(min)为横坐标,以位置点(km)为纵坐标。

1) 横坐标设置

假定某乘客第 1 天出行的第 1 个刷卡时间点为 1(即横坐标),相邻位置点横坐标差表示前后刷卡的时间点分钟差值;同时,假定第 2 天首个刷卡横坐标比前一天的末横坐标大 100 min,即连续 4 d 横坐标值持续增大。

2) 纵坐标计算

选取北京天安门为出行位置对比点,假定每天第 1 个纵坐标值 P_1 为当天的首个出行阶段上车站点与天安门的曼哈顿距离,如式(1):

$$P_1 = (N_1 - N_0) \times 10^5 + \frac{T_1 - T_0}{7.5} \times 10^6 \quad (1)$$

式中: P_1 为每天首个纵坐标; N_1 为每天首个刷卡站点的经度; N_0 为天安门经度(116.388°); T_1 为每天首个刷卡站点的纬度; T_0 为天安门纬度(39.9075°)。

纵坐标依次为前一位置点的纵坐标值加或减两点间路径距离,差值体现相邻位置点实际路径距离。定义规则如式(2):

$$P_{i+1} = \begin{cases} P_i + D_{i+1}, & N_{i+1} - N_i \geq 0 \\ P_i - D_{i+1}, & N_{i+1} - N_i < 0 \end{cases} \quad (2)$$

式中: D_{i+1} 为从 i 个站点到第 $i+1$ 个站点的实际路径距离,km。

2.1.3 步骤 3

确定图谱元素。由节点大小表示出行方向,线型和线宽体现出行方式。

出行方向由出行阶段上下车站点经纬度通过反正切函数计算得到,为实际角度值。为了区分出行方向唯一性,利用实心节点表示角度值为 $0^\circ \sim 180^\circ$,空心节点为 $-180^\circ \sim 0^\circ$;每日首个出行方向以天安门为基准。对出行方式,粗实线代表轨道交通,细实线表示地面公交,虚线表示无公共交通出行。

2.2 通勤个体出行特征图谱相似性分析

基于个体出行特征图谱,通过相似性分析,可以实现相同乘客出行行为分类与不同乘客行为相似性识别。

2.2.1 相似性判定指标确定

图谱相似性判定从表征个体乘客出行行为特征的角度选取判定指标,根据文献[19],乘客出行特

征主要体现为出行路径、出行时间、出行位置等方面。故笔者提出了图谱相似性判别方法,分别从出行路径、出行时间、出行位置这3个层面选取判定指标。其中,采用出行方式和出行方向共同表征出行路径,采用出行时间表征乘客时间特性,采用乘坐线路表征位置波动特性。个体出行特征指标与图谱元素属性关联如表2。

表2 个体出行特征指标与图谱元素属性关联

Table 2 Relationship between individual travel characteristics indexes and element attributes

特性	特征指标	图谱属性
相似性	出行方式	线型、线宽
	出行方向	形状、角度
	出行时间	每日首个横坐标,横轴跨度
	乘坐线路	节点波动程度

1) 出行方式

由线型、线宽体现不同出行方式。

2) 出行方向

采用节点大小表示出行方向。

3) 出行时间

考虑到通勤者下班后可能存在娱乐、购物等活动,晚上到家时间跨度较大。故采用乘客每日首个上车时间,及当日出行阶段平均耗时体现出行时间。

4) 乘坐线路

定义波动系数体现所乘线路的动态性,波动系数越小,表示各天出行位置点的变化程度越小。计算方法如式(3)、(4)。

$$M = [E(l_1), E(l_2), \dots, E(l_n)] \quad (3)$$

$$T = \frac{\sqrt{D(M)}}{E(M)} \quad (4)$$

式中: M 为各天的位置点均值; $E(l_n)$ 为第 l_n 天位置点均值; T 为波动系数; $D(M)$ 为 M 方差; $E(M)$ 为 M 均值。

2.2.2 图谱相似性判定

在个体出行图谱的相似性判别中,笔者采用非连续性与连续性相结合进行综合判别。首先,由非连续性角度度量图谱各组成元素间的相似程度,通过计算结构相似度进行判别^[20-21];其次,选用公共子序列度量方法寻找不重叠的多个公共子区间,从时间先后角度判别子区间连续相似性,采用最长公共子序列衡量连续相似性,即两个或多个序列中存在最长的共同子序列^[21];最后,实现相同个体出行行为分类与不同个体图谱相似性判定。

1) 非连续性判定

为获取图谱间的结构相似度,研究首先计算结构距离(S_D)。 S_D 包括不同个体或相同个体不同天的特征属性比较。图谱结构距离计算如式(5):

$$S_D(C_i, C_j) = M_D \times W_M + D_D \times W_S + T_D \times W_T + L_D \times W_L \quad (5)$$

式中: $S_D(C_i, C_j)$ 为 i 与 j 天的结构距离; M_D 为方式距离; D_D 为方向距离; T_D 为时间距离; L_D 为线路距离; $W = \{W_M, W_S, W_T, W_L\}$ 为特征权重向量,分别表示轨迹的方式权重、方向权重、时间权重和线路权重,其中:各权重应满足权重取值非负且 $W_M + W_S + W_T + W_L = 1$,由于笔者未对各权重的重要程度进行研究,因此假定各权重不分主次,取 $W_M = W_S = W_T = W_L = 0.25$ 。

(1) 方式距离

方式距离主要表征乘客采用不同出行方式的差异性,不考虑出行方式的乘坐顺序,为各方式占比的公共差值绝对值。例如,某乘客 i 天的出行方式为4次轨道, j 天为3次轨道1次公交,则 i 天与 j 天的方式距离计为 $1/4$ 。

(2) 方向距离

方向距离从最大、平均和最小角度的层面表示结构距离。计算方法如式(6):

$$D_D(C_i, C_j) = \frac{1}{3} [D_{\max}(L_i, L_j) + D_{\text{avg}}(L_i, L_j) + D_{\min}(L_i, L_j)] \quad (6)$$

式中: $D_{\max}(L_i, L_j)$ 即 $|D_{\max}(L_i) - D_{\max}(L_j)|$,反映了 i 天与 j 天最大斜率差异绝对程度;类似地, D_{avg} 为平均斜率差异绝对值; D_{\min} 为最小斜率差异绝对值。

(3) 时间距离

时间距离从最长、平均和最短出行时间的层面表示结构距离。计算如式(7):

$$T_D(C_i, C_j) = \frac{1}{3} [T_{\max}(L_i, L_j) + T_{\text{avg}}(L_i, L_j) + T_{\min}(L_i, L_j)] \quad (7)$$

式中: $T_{\max}(L_i, L_j)$ 即 $|T_{\max}(L_i) - T_{\max}(L_j)|$,反映了 i 与 j 天最长出行时间差异绝对程度; T_{avg} 为平均出行时间绝对差异程度; T_{\min} 为最小出行时间绝对差异程度。

(4) 线路距离

采用波动系数表征 i 与 j 天的线路距离 $L_D(C_i,$

C_j),其波动系数计算如式(4)。

综上,由于不同特征值域不同,需将每个特征距离值做归一化处理。引入结构距离的归一化函数 $N(S_D)$ 。则结构相似度(G_S)计算如式(8):

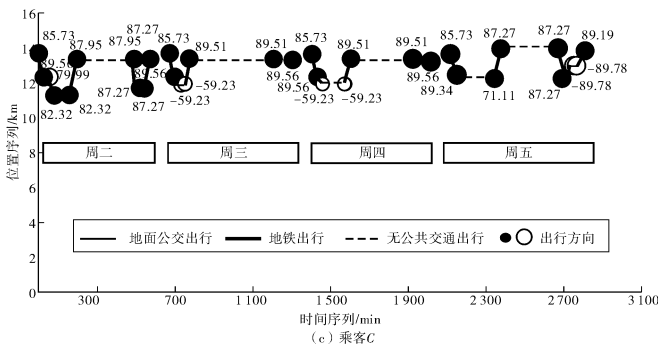
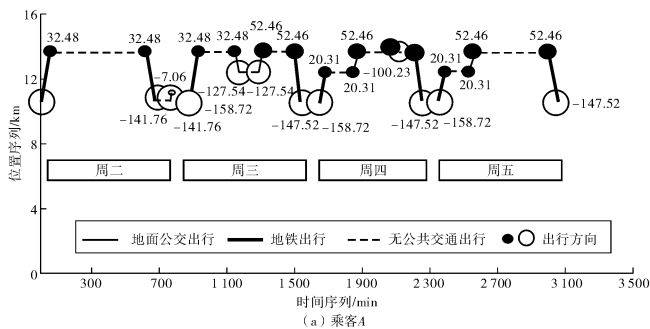
$$G_S(C_i, C_j) = 1 - N[S_D(C_i, C_j)] \quad (8)$$

$G_S(C_i, C_j)$ 体现了不同个体或相同个体不同天的特征图谱结构相似程度;其值越大,则相似性程度

$$L_C(i, j) = \begin{cases} 0, & m=n=0 \\ L_C(R(i), R(j)) + 1, & |i_{1,x} - j_{1,x}| \leq \delta, |i_{1,y} - j_{1,y}| \leq \varepsilon, |D_r(i) - D_r(j)| \leq \beta \\ \max\{L_C(R(j), i), L_C(i, \text{Rest}(j))\}, & \text{otherwise} \end{cases} \quad (9)$$

式中: $L_C(i, j)$ 为图谱*i*与*j*之间的LCSS长度; m 为图谱*i*的出行阶段数量; n 为图谱*j*的出行阶段数量; δ 为*x*轴相似阈值; ε 为*y*轴相似阈值; β 为方向相似阈值; $R(i)$ 、 $R(j)$ 分别为轨迹*i*、*j*去掉首个记录点后的轨迹间隔。

当轨迹记录点数*j*均为0时, $L_C(i, j) = 0$;若记录点个数不为0,则用递归方式判断共有子序列长度最大值。当同时满足横坐标差小于 δ 、纵坐标差小于 ε 、方向角度小于 β 时,该对记录点相似,LCSS距离值加1。考虑到北京公共交通通勤乘客从家出发时刻常集中在07:00—09:00,且不同天可能存在娱乐、餐饮、购物等活动的差异,取 $\delta = 1$ h; ε 取相邻两个轨道站点间距,按均值1.7 km计算;参照利用乘客轨迹行程统计居民出行方向分布的方向划分间隔^[23],设定 $\beta = 30^\circ$ 。



高。若结构相似度超过一半,需进一步判定连续相似性,否则,直接判定为不相似。

2) 连续相似性判定

笔者采用最长公共子序列(longest common subsequence, LCSS)衡量连续相似性。通过计算最长公共子序列并转化为LCSS距离衡量轨迹间连续性相似程度^[22],如式(9):

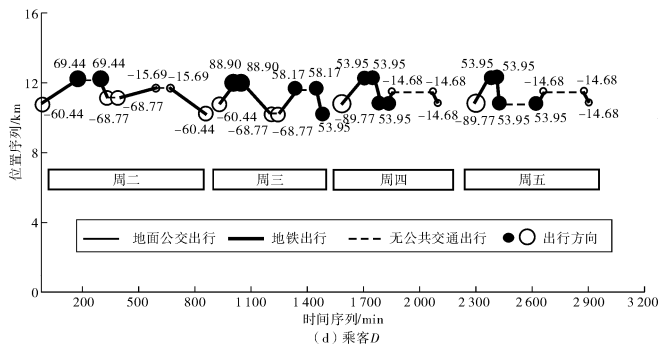
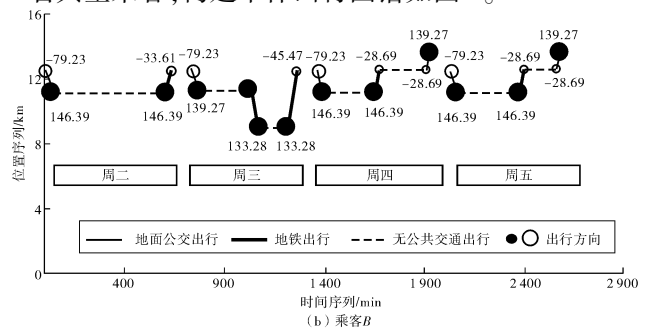
通过归一化处理,将LCSS距离转换为 $[0, 1]$ 的距离,如式(10):

$$D_{LCSS}(i, j) = 1 - \frac{L_C(i, j)}{\min(m, n)} \quad (10)$$

式中: $D_{LCSS}(i, j)$ 为特征图谱*i*与*j*间的LCSS距离; $\min(m, n)$ 为*i*与*j*记录点个数的较小值。

3 案例分析

为方便数据验证与核对,笔者选取北京工业大学附近居民区为研究区域。经计算统计,90%以上公共交通通勤乘客一天内出行时间跨度在14 h内,故*x*轴值域最大不超过3 500($14 \times 60 \times 4 = 3\ 360$ min);85%以上通勤乘客的位置波动在16 km内,故*y*轴采用统一性值域范围,即 $[0, 16]$ 。随机选取5名典型乘客,构建个体出行图谱如图1。



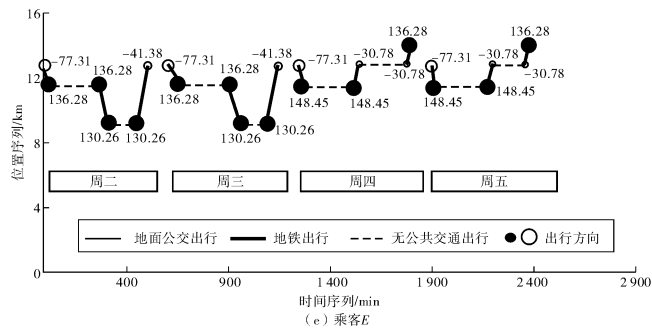


图 1 5 名乘客的个体出行特征图谱

Fig. 1 Individual travel specific chromatogram of 5 passengers

由图 1 可获得公共交通个体乘客的出行特征包括:① 通勤乘客的主要出行方式。由图谱线型可知乘客 A 在连续 4 d 的公共交通出行中以地铁为主,尤其高峰期间常乘坐地铁,平峰期间较多乘坐公交;乘客 B 则以公交出行为主,早高峰主要乘坐公交,平峰乘坐地铁。② 乘客路径距离分布。由图谱纵轴可知乘客 A 路径距离在 4 km 内,而乘客 D 则集中在 2 km 左右。③ 乘客出行目的地。乘客 B 日均有 3 个公共交通目的地,除去工作地与家,平峰出行 1

次;而乘客 C 日均 4~5 个目的地,无明显出行规律。④ 不同天的出行相似性。由图 1 可明显看出:乘客 A 后两天出行存在明显相似性,乘客 E 前两天与后两天出行也显著相似。

笔者以乘客 A 为例,计算相似性指标值如表 3。计算乘客 A 不同天的结构距离与 LCSS 距离,得到特征图谱相似性判别结果如表 4,其中: r_{ij} 指该乘客不同日期对比。由表 4 可知:乘客 A 后两天的出行相似性较大,而第 1、3 天出行几乎不存在相似性。

表 3 乘客 A 相似性指标值

Table 3 Similarity index value of passenger A

相似性指标	出行时间			
	周 2	周 3	周 4	周 5
出行方式	轨道 3 次	轨道 4 次	轨道 3 次,公交 1 次	轨道 3 次
最大角度/(°)	32.48	52.46	79.77	52.46
平均角度/(°)	-64.06	-29.97	-17.94	-24.92
最小角度/(°)	-158.72	-158.72	-158.72	-158.72
最长出行时间/min	56	58	50	52
平均出行时间/min	36	40	35	36
最短出行时间/min	9	26	24	27
波动系数	0.11	0.10	0.10	0.10

表 4 乘客 A 个体出行图谱相似性判别结果

Table 4 Individual travel specific chromatogram similarity identification results of passenger A

日期对比	非连续相似性	连续相似性	综合相似性
r_{12}	0.54	0.33	0.43
r_{13}	0.45	—	不相似
r_{14}	0.60	0.33	0.46
r_{23}	0.68	0.38	0.53
r_{24}	0.86	0.50	0.68
r_{34}	0.86	1.00	0.93

对于不同个体,通过非连续性与连续相似性判别可以得到乘客 A 和 B 的出行相似性,如表 5。其

中: r_i 指两名乘客在同日期对比。由表 5 知:乘客 A 和 B 在该周的出行行为不存在明显的相似性。

表 5 乘客 A 和 B 出行特征图谱相似性判别结果

Table 5 Travel specific chromatogram similarity identification results of passenger A and B

日期对比	非连续相似性	连续相似性	综合相似性
r_1	0.45	—	不相似
r_2	0.63	0	0.31
r_3	0.93	0	0.47
r_4	0.68	0	0.34

根据上述方法,分析这 5 名乘客出行特征图谱综合相似性判别结果如表 6。由表 6 可知:乘客 B

和 E 在该周出行行为明显相似,其他乘客则不存在明显相似性。

表 6 5 名乘客出行特征图谱综合相似性判别

Table 6 Travel specific chromatogram comprehensive similarity identification for 5 passengers

乘客	乘客				
	A	B	C	D	E
A	1.00	0.28	0.34	0.32	0.39
B	0.28	1.00	0.34	0.37	0.80
C	0.34	0.34	1.00	0.38	0.30
D	0.32	0.37	0.38	1.00	0.39
E	0.39	0.80	0.30	0.39	1.00

4 结 论

笔者针对公共交通通勤乘客,构建了个体出行特征图谱,并提出了图谱相似性判别方法,为分析公共交通出行特征提供了新思路。

1)通过设定公交常乘客刷卡最大间隔与往返地点闭合次数等规则鉴别通勤乘客,进而基于图谱可视化优势,以时间序列为横坐标,位置点为纵坐标,节点大小为出行方向,从个体的角度构建了通勤出行特征图谱,直观化地展示乘客乘坐公共交通的出行过程。

2)从非连续性与连续性判别的角度,通过结构相似度与最长公共子序列计算相结合,提出图谱相似性综合判别方法,实现相同乘客出行行为分类与不同乘客行为相似性判别。

3)笔者以北京 5 名公交常乘客一周的公共交通数据为例进行示例分析,为利用图谱进行乘客出行行为特征分析进行了探索。未来将采用计算机语言进行大样本研究与测试,进一步验证该方法的可靠性与稳定性,对相同个体通勤乘客,可实现其出行稳定性分类,比如将出行者划分为高稳定、中稳定、低稳定通勤乘客,为更加准确地掌握不同稳定性乘客的出行需求奠定支撑。同时,对于不同个体,出行者相似度判别可分析不同个体乘客间出行行为的差异性和一致性,进而实现公共交通总体出行人群细化分类。

参考文献(References):

[1] 张乐典.公共交通通勤出行时间影响模型研究[D].北京:北京工业大学,2015.
ZHANG Ledian. Research on Time Impact Model of Public Transportation Commuting Trip [D]. Beijing: Beijing University of

Technology, 2015.

- [2] 翁剑成,张乐典,陈智宏,等.基于多模式公交数据匹配的公共交通出行特征提取方法;CN 201510068077.7[P].2015-07-08.
WENG Jiancheng, ZHANG Ledian, CHEN Zhihong, et al. Public Transport Travel Feature Extraction Method Based on Multi-mode Public Transport Data Matching; CN 201510068077.7[P]. 2015-07-08.
- [3] 周钱,李一,孟超,等.基于结构方程模型的交通需求分析[J].清华大学学报(自然科学版),2008,48(5):879-882.
ZHOU Qian, LI Yi, MENG Chao, et al. Analysis of travel demand based on a structural equation model [J]. Journal of Tsinghua University (Science and Technology), 2008, 48(5): 879-882.
- [4] WALLE V S, STEENBERGHEN T. Space and time related determinants of public transport use in trip chains [J]. Transportation Research Part A: Policy and Practice, 2006, 40(2): 151-162.
- [5] 王珂.北京市轨道交通对居民职住地选择及通勤出行的影响研究[D].北京:北京交通大学,2010.
WANG Ke. Research on the Influence of Urban Rail Transit on Residents' Choice of Residence, Workplace and Commuting Travel Mode [D]. Beijing: Beijing Jiaotong University, 2010.
- [6] 杨敏,陈学武,王伟,等.通勤出行简单链和复杂链的选择行为研究[J].武汉理工大学学报(交通科学与工程版),2008,32(2): 191-194.
YANG Min, CHEN Xuewu, WANG Wei, et al. Behavioranalysis of simple/complex trip chaining [J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2008, 32(2): 191-194.
- [7] JANG W. Traveltime and transfer analysis using transit smart card data [J]. Journal of the Transportation Research Board, 2010, 2144(1): 1-2.
- [8] BOWMAN J L, BEN-AKIVA M E. Activity based disaggregate travel demand model system with activity schedules [J]. Transportation Research Part A: Policy and Practice, 2001, 35(1): 1-28.
- [9] SHEN Yue, KWAN M P, CHAI Yanwei. Investigating commuting flexibility with GPS data and 3D geo-visualization: A case study of Beijing, China [J]. Journal of Transport Geography, 2013, 32: 1-11.
- [10] 戴霄,陈学武,李文勇.公交 IC 卡信息处理的数据挖掘技术研究[J].交通与计算机,20016,24(1):40-42.
DAI Xiao, CHEN Xuewu, Li Wenyong. Study of data mining technique for bus intelligent card data processing [J]. Computer and Communications, 2006, 24(1): 40-42.
- [11] BARRY J J, FREIMER R, SLAVIN H L. Use of entry-only automatic fare collection data to estimate linked transit trips in New York city [J]. Journal of the Transportation Research Board, 2009, 2112: 53-61.
- [12] ZHAO Jinhua, RAHBEE A, WILSON N H M. Estimating a rail passenger trip origin-destination matrix using automatic data collection systems [J]. Computer-Aided Civil and Infrastructure Engineering, 2007, 22(1): 376-387.
- [13] KARASMAA N. Evaluation of transfer methods for spatial travel demand models [J]. Transportation Research Part A: Policy and

- Practice, 2007, 41(5): 411-427.
- [14] 刘峤,李杨,段宏,等.知识图谱构建技术综述[J].计算机研究与发展,2016,53(3):582-600.
- LIU Qiao, LI Yang, DUAN Hong, et al. Knowledge graph construction techniques [J]. *Journal of Computer Research and Development*, 2016, 53(3): 582-600.
- [15] 袁凯琦,邓扬,陈道源,等.医学知识图谱构建技术与研究进展[J].计算机应用研究,2017,35(7):1-12.
- YUAN Kaiqi, DENG Yang, CHEN Daoyuan, et al. Construction techniques and research development of medical knowledge graph [J]. *Application Research of Computers*, 2017, 35(7): 1-12.
- [16] 焦晓静,王兰成,韩锋.知识图谱在科技情报研究中的应用模型构建[J].图书情报知识,2017(3):118-128.
- JIAO Xiaojing, WANG Lancheng, HAN Feng. Application model construction of mapping knowledge domains in the science and technology intelligence research [J]. *Knowledge, Learning & Management*, 2017(3): 118-128.
- [17] CHEN Sei-wang, FANG Chiung-yao, TIEN Chih-ting. Driving behavior modelling system based on graph construction [J]. *Transportation Research Part C: Emerging Technologies*, 2013, 26: 314-330.
- [18] 王月玥.基于多源数据的公共交通通勤出行特征提取方法研究[D].北京:北京工业大学,2014.
- WANG Yueyue. *Research on Methods of Extracting Commuting Trip Characteristic Based on Public Transportation Multi-Source Data* [D]. Beijing: Beijing University of Technology, 2014.
- [19] 张弘毅.基于活动方法的个体出行行为分析与出行需求预测模型系统研究[D].南京:南京师范大学,2011.
- ZHANG Hongtao. *Research on Individual Travel Behavior and Travel Demand Forecasting Model* [D]. Nanjing: Nanjing Normal University, 2011.
- [20] CHANG Cheng, ZHOU Baoyao. Multi-granularity visualization of trajectory clusters using sub-trajectory clustering [C]// *Proceedings of the 7th IEEE International Conference on Data Mining Workshops*. Miami, Florida, USA, 2009: 577-582.
- [21] 陈红,章渺,王龙飞,等.高等级公路路段交通状态融合识别模型[J].重庆交通大学学报(自然科学版),2010,29(5):772-775.
- CHEN Hong, ZHANG Miao, WANG Longfei, et al. Fusion and recognition model of traffic state on high-grade highway segment [J]. *Journal of Chongqing Jiaotong University (Natural Science)*, 2010, 29(5): 772-775.
- [22] TAN P N, STEINBACH M, KUMAR V. *Introduction to Data Mining* [M]. Boston: Pearson Addison-Wesley, 2006: 60-65.
- [23] 吕振华.移动对象轨迹数据挖掘研究——以上海市出租车轨迹数据为例[D].上海:华东师范大学,2016.
- LYU Zhenhua. *A Study on Data Mining of Moving Object Trajectories-A Case of Shanghai Taxi Trajectories* [D]. Shanghai: East China Normal University, 2016.

(责任编辑:刘 韬)

(上接第5页)

- REN Qiling, XIE Xiaosong. Identification method of the road traffic conditions based on genetic-dynamic fuzzy clustering algorithm [J]. *Journal of Transportation Engineering and Information*, 2007, 5(3): 12-15.
- [5] 张亮亮,贾元华,牛忠海,等.交通状态划分的参数权重聚类方法研究[J].交通运输系统工程与信息,2014,14(6):147-151.
- ZHANG Lingling, JIA Yuanhua, NIU Zhonghai, et al. Traffic state classification based on parameter weighting and clustering method [J]. *Journal of Transportation Systems Engineering and Information Technology*, 2014, 14(6): 147-151.
- [6] 黄艳国,许伦辉,邝先验.基于模糊C均值聚类的城市道路交通状态判别[J].重庆交通大学学报(自然科学版),2015,34(2):102-107.
- HUANG Yanguo, XU Lunhui, KUANG Xianyan. Urban road traffic state identification based on fuzzy c-mean clustering [J]. *Journal of Chongqing Jiaotong University (Natural Science)*, 2015, 34(2): 102-107.
- [7] 巫威眺,靳文舟,林培群.基于BP神经网络的道路交通状态判别方法研究[J].交通信息与安全,2011,29(4):71-74.
- WU Weizhao, JIN Wenzhou, LIN Peiqun. The method of traffic state identification based on bp neural network [J]. *Journal of Transport Information and Safety*, 2011, 29(4): 71-74.
- [8] CORTES C, VAPNIK V. *Support-Vector Networks* [M]. Amsterdam: Kluwer Academic Publishers, 1995: 273-297.
- [9] 于荣,王国祥,郑继媛,等.基于支持向量机的城市道路交通状态模式识别研究[J].交通运输系统工程与信息,2013,13(1):130-136.
- YU Rong, WANG Guoxiang, ZHENG Jiyan, et al. Urban road traffic condition pattern recognition based on support vector machine [J]. *Journal of Transportation Systems Engineering and Information Technology*, 2013, 13(1): 130-136.
- [10] 廖瑞辉,周晶.基于云-支持向量机的道路交通拥挤预警模型[J].系统工程,2015(4):149-153.
- LIAO Ruihui, ZHOU Jing. Traffic congestion warning model based on cloud-support vector machine [J]. *Systems Engineering*, 2015(4): 149-153.
- [11] 吐松江·卡日,高文胜,张紫薇,等.基于支持向量机和遗传算法的变压器故障诊断[J].清华大学学报(自然科学版),2018,58(7):623-629.
- KARI TUO Songjiang, GAO Wensheng, ZHANG Ziwei, et al. Power transformer fault diagnosis based on a support vector machine and a genetic algorithm [J]. *Journal of Tsinghua University (Science and Technology)*, 2018, 58(7): 623-629.
- [12] 皮晓亮,王正,韩皓,等.基于环形线圈检测器采集信息的交通状态分类方法应用研究[J].公路交通科技,2006,23(4):115-119.
- PI Xiaoliang, WANG Zheng, HAN Hao, et al. Application research of traffic state classification method based on collected information from loop detector [J]. *Journal of Highway and Transportation Research and Development*, 2006, 23(4): 115-119.

(责任编辑:朱汉容)