

# 基于手机信令和导航数据的出行方式识别方法\*

杜亚朋<sup>1,2</sup>, 雒江涛<sup>1,2</sup>, 程克非<sup>1,2</sup>, 唐刚<sup>1,2</sup>, 徐正<sup>1,2</sup>, 罗克韧<sup>3</sup>, 余疆<sup>3</sup>

(1. “新一代信息网络与终端”重庆市协同创新中心, 重庆 400065; 2. 重庆邮电大学 电子信息与网络工程研究院, 重庆 400065; 3. 中国移动通信集团重庆有限公司, 重庆 400065)

**摘要:** 基于手机信令识别居民出行方式对于智慧交通规划具有重要意义。通过结合信令和导航地图数据, 利用聚类算法以及时间关联性算法, 实现步行、驾车、公共交通等出行方式的识别。结果表明, 结合导航地图数据后识别正确率得到明显提高, 与只利用手机信令的识别方法相比, 整体准确率提升超过 15%, 具有较高的识别准确率, 同时算法执行时间为 187 s, 效率较高。整体而言, 该识别算法适合在实际工程环境中使用。

**关键词:** 城市交通; 出行方式识别; 聚类分析; 时间关联; 手机信令; 导航数据

**中图分类号:** U491.1      **文献标志码:** A      **文章编号:** 1001-3695(2018)08-2311-04

doi:10.3969/j.issn.1001-3695.2018.08.018

## Recognition of urban travel method based on cell phone signaling and navigation map data

Du Yapeng<sup>1,2</sup>, Luo Jiangtao<sup>1,2</sup>, Cheng Kefei<sup>1,2</sup>, Tang Gang<sup>1,2</sup>, Xu Zheng<sup>1,2</sup>, Luo Keren<sup>3</sup>, Yu Jiang<sup>3</sup>

(1. Chongqing Collaborative Innovation Center for New Generation Information Network & Terminal, Chongqing 400065, China; 2. Electronic Information & Networking Engineering Institute, Chongqing University of Posts & Telecommunications, Chongqing 400065, China; 3. China Mobile Communications Group Chongqing Company Limited, Chongqing 400065, China)

**Abstract:** Based on the cell phone signaling data to identify the residents' travel method plays an important role in intelligent transportation planning. This paper combined cell phone signaling data with navigation map data, used clustering algorithm and time correlation algorithm to identify the travel mode such as walking, driving and public transportation. The experimental results indicate that the recognition accuracy is significantly improved when the cell phone signaling and navigation map data are combined. And compare with the recognition scheme that only uses cell phone signaling, the comprehensive recognition accuracy of this model increases by more than 15% and it is a high recognition accuracy. At the same time, the algorithm execution time is 187 seconds and it is a high efficiency. In general, the recognition algorithm is suitable for use in real engineering environments.

**Key words:** urban traffic; recognition of travel method; cluster analysis; time correlation; mobile phone signaling; navigation data

## 0 引言

如何方便居民的出行是城市交通规划的重要使命。及时掌握城市居民出行方式及其发展变化, 不仅可以及时了解城市当前交通结构状态, 同时对于日后涉及城市规划、交通管理等方面的发展也有着重要的参考价值<sup>[1]</sup>。

传统居民出行方式数据获取的方法大多是基于调查问卷以及电话问询等, 此类方法成本高、效率低、样本量小, 并且由于出行的隐私性导致的心理顾虑, 所以可信度不高。随着信息化时代的到来以及智能终端的普及, 利用手机构建用户轨迹进而判别用户出行方式逐渐成为一种重要手段, 利用手机中的相关传感器是最直接、精度最高的一种方法。部分学者利用手机中的 GPS 或加速度计等模块, 提取定位数据, 计算平均速度、加速度、速度中值等特征参数, 利用支持向量机、神经网络等机器学习算法<sup>[2-4]</sup>和模糊推理<sup>[5]</sup>等算法识别出行方式。但由于密闭空间的干扰以及需实时开启 GPS 模块, 该方法无法对大规模人群长期的出行方式进行有效识别<sup>[6]</sup>。利用手机信令数据是另外一种重要方法。手机信令数据具有样本量大、覆盖面

广和时间周期长等优点, 但其定位精度通常依赖于通信蜂窝小区的大小, 较 GPS 等定位系统较差。Luo 等人<sup>[7]</sup>利用手机信令数据, 建立一个快速估计区域驻留人数和驻留时长的监测模型, 并且在实际中进行了部署。赖见辉等人<sup>[8]</sup>利用手机信令定位信息, 结合轨道交通线路信息识别地铁乘客出行路径。Bloch 等人<sup>[9]</sup>利用信令数据中位置数据和传感器数据, 建立机器学习模型识别居民出行方式。导航数据库可以根据出行起点与终点提供不同出行方式的出行道路轨迹与出行时长信息, 因此可以将信令分析得到的出行轨迹和出行时长与导航提供的相应信息进行比对, 从而分析出行方式。现已有学者进行了相关的研究, Wang 等人<sup>[10]</sup>利用基于划分的 K-means 聚类算法对出行起止点相同人群的出行时长进行聚类, 获取时长不同的出行聚类簇, 利用谷歌地图提供的不同出行方式的时长分布对每个聚类簇样本的出行方式进行划分。但是这种方法缺少了对起止点相同人群不同出行轨迹的关联分析, 仅对人群出行时长与谷歌地图提供的时长进行了关联分析。

为进一步提高基于手机信令的出行方式判断的准确度, 本文提出采用手机信令数据, 同时结合导航数据的识别方法。将

**收稿日期:** 2017-04-11; **修回日期:** 2017-05-25      **基金项目:** 重庆市基础科学与前沿技术研究重点项目(cstc2015jcyjBX0009); 重庆市科技创新领军人才计划支持项目(CSTCKJCXJRC20)

**作者简介:** 杜亚朋(1991-), 男, 河南郑州人, 硕士, 主要研究方向为移动数据挖掘、交通大数据(dyppmail@foxmail.com); 雒江涛(1971-), 男, 河南郑州人, 教授, 博导, 主要研究方向为新一代网络技术、移动数据挖掘; 程克非(1974-), 男, 重庆人, 教授, 主要研究方向为网络安全、移动大数据; 唐刚(1991-), 男, 重庆人, 硕士, 主要研究方向为移动数据挖掘、交通大数据; 徐正(1989-), 男, 湖北孝感人, 博士研究生, 主要研究方向为移动数据挖掘、交通大数据; 罗克韧(1984-), 女, 重庆人, 数据产品经理, 主要研究方向为移动数据应用发展规划与设计; 余疆(1979-), 男, 重庆人, 大数据室经理, 主要研究方向为移动数据应用产品发展。

居民出行轨迹数据与导航数据聚类来进行轨迹相关性的分析,并且同时分析两类数据的时间关联性,进而对居民出行方式进行更准确的判别。

## 1 原理

手机用户在使用通信业务时会通过信令数据在当前基站小区进行注册与信息交互。手机信令数据中包括的字段有用户ID、位置区码(location area code, LAC)、小区识别码(cell identify, CID)、信令发生时间戳(timestamp)、基站网络类型(base station type, BS type, 包括2G、3G和4G)、基站经度(longitude, Lon)、基站纬度(latitude, Lat)、事件类型(event ID)等。

用户信令数据中的位置信息是用户在出行过程中接入周边基站产生的位置数据。若用户实际出行是沿某条道路,则较多的信令数据会与该道路轨迹点数据聚类成簇。基于此原理,本文通过利用导航API请求不同出行方式的道路轨迹点数据,然后与信令中位置数据融合进行聚类分析,统计含有导航轨迹聚类簇中信令数据的个数,将其与全部信令数据的比值作为轨迹匹配度,根据设定的轨迹匹配度阈值,记录符合阈值的出行方式集合以及请求的导航线路信息,包括线路坐标数据、线路距离以及花费时间等,同时结合实际出行时间与导航线路时间,计算每条导航线路与实际出行的时间匹配度,从轨迹匹配度与时间匹配度两方面综合进行出行方式的识别。

### 1.1 有效出行段识别

出行方式的定义是人或车辆等从出发地到目的地的一次有效出行段中使用的移动方式。因此首先需要识别驻留区域,进而提取信令数据中的有效出行段。其中,有效出行段包含了本次出行的开始位置、开始时间、结束位置、结束时间和中间轨迹点等信息。

驻留区域是指人们的出行轨迹中发生长时间驻留的区域,也即有效出行段的起点区域或终点区域。驻留区域的识别对于提取用户一次有效出行非常关键。由于用户在驻留时接收到周边基站信号的强度可能会随时间发生变化,所以会在多个基站间进行切换,这样的多个基站位置点构成了驻留区域。识别步骤如下:

a) 设定驻留范围距离阈值  $S_{MAX}$  与驻留范围最短时间阈值  $T_{MIN}$ 。

b) 将第一条信令放入比较序列 seq 中。若第二条信令的位置与第一条信令的位置在设定的距离阈值  $S_{MAX}$  之内,将第二条信令也放入比较序列 seq 中,按照信令发生时间顺序将余下信令与 seq 中的每个信令依次进行判断;若某条信令的位置与 seq 中任意一条信令的位置距离超过  $S_{MAX}$ ,中断此次判断。

c) 计算 seq 中最后一条信令与第一条信令的时间差。如果时间差值大于设定的最短时间阈值  $T_{MIN}$ ,则认为 seq 中信令位置区域为驻留区域,同样根据时间比例计算这些位置的驻留核心点坐标,用核心点坐标代表驻留区域;如果时间差值小于设定的最短时间阈值  $T_{MIN}$ ,则这些信令位置区域不是驻留区域,释放 seq 中的信令数据,继续判断剩下的信令。

在进行驻留区域识别之后,根据结果分别将每两个驻留区域之间信令数据作为一次有效出行产生的信令数据,然后利用DBSCAN聚类算法结合导航数据,对一次出行的出行方式进行识别。

### 1.2 出行方式识别

#### 1.2.1 DBSCAN 算法

DBSCAN 算法于1996年由 Ester 等人<sup>[11]</sup>提出,是一种基于密度的聚类算法。该算法利用样本对象的密度连通性,根据设定的半径(Eps)和最小包含点数(MinPts)这两个参数,将低密度区域进行过滤去除,从而可以把具有足够密度的区域划分为簇,并且可以在有噪声的空间数据中发现任意形状的簇。

#### 1.2.2 基于 DBSCAN 算法的轨迹匹配方法

根据基站的辐射范围以及与道路的映射关系,本文中设定DBSCAN算法的Eps邻域参数为400,表示扫描半径为400 m。同时在本模型中,DBSCAN聚类算法的目的是将基站位置与道路位置数据聚类成尽量多的簇,所以设定最小包含点数MinPts为2。

基于DBSCAN算法的轨迹匹配算法流程如图1所示。具体步骤如下:

a) 获取用户的某个有效出行段的信令数据集合  $signal_{set}$ ,包括起点驻留区域的核心点  $signal_{start}$  的信令数据、出行段信令数据以及终点驻留区域核心点  $signal_{end}$  的信令数据,设共有  $N_{set}$  条信令数据。

b) 使用导航API请求从  $signal_{start}$  到  $signal_{end}$  的各种交通方式(包括步行、驾车、公共交通等,其中公共交通包括公交车、轻轨和地铁等交通方式)的导航路线信息数据。假设获得的导航路线信息数据集合为  $NR = \{nr_1, nr_2, \dots, nr_i, \dots, nr_n\}$ ,其中  $nr_i$  是指某种交通方式的导航线路数据,包括从  $stay_{start}$  到  $stay_{end}$  的线路经纬度坐标点集合  $NR_{loc} = \{(lon_1, lat_1), (lon_2, lat_2), \dots, (lon_i, lat_i), \dots, (lon_n, lat_n)\}$ ,以及距离长度  $NRS_i$  和时间长度  $NRT_i$  等。

c) 将用户信令数据集合  $signal_{set}$  中的位置数据  $signal_{loc}$  与导航线路中坐标点集合  $NR_{loc}$  作为DBSCAN算法的输入数据集合  $data = \{set_{loc}, NR_{loc}\}$ ,然后根据设定的Eps和MinPts两个参数进行聚类分析。设聚类结果共获取  $p$  个聚类簇集  $C = \{c_1, c_2, \dots, c_i, \dots, c_p\}$ ,分析获取聚类簇  $C$  中含有导航轨迹数据的子簇  $C_{nr} = \{c_{nr1}, c_{nr2}, \dots, c_{nrj}, \dots, c_{nrq}\}$ 。统计子簇集合  $C_{nr}$  中信令数据的总数  $S_{nr}$ ,计算  $S_{nr}$  与  $N_{set}$  的比值  $Rtra_i$  作为当前导航线路与信令轨迹的轨迹匹配比值,其中  $Rtra_i = S_{nr}/N_{set}$ 。

d) 对导航路线信息数据集合  $NR$  中每个导航线路重复步骤c),直到所有线路被分析完毕,得到各导航线路与此用户有效出行段的轨迹匹配度集合  $Rtra = \{Rtra_1, Rtra_2, \dots, Rtra_i, \dots, Rtra_n\}$ 。

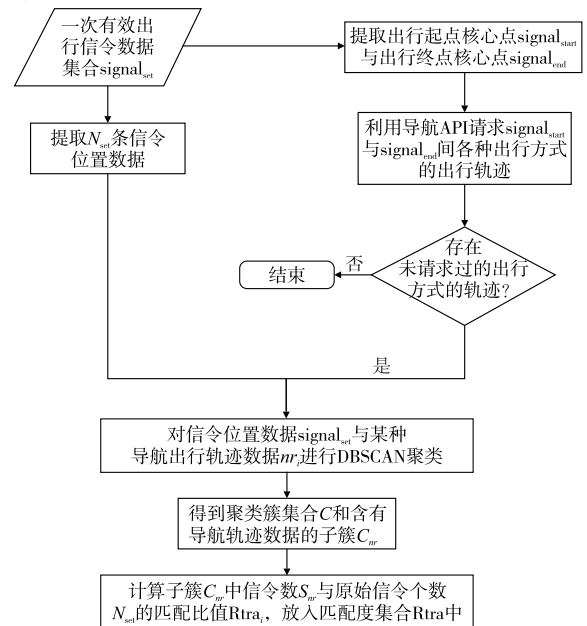


图1 基于DBSCAN算法的轨迹匹配方法流程

#### 1.2.3 基于轨迹匹配度集合与时间关联的出行方式识别方法

在获取到某次有效出行的轨迹匹配度集合后,同时结合该集合每条导航线路的时间长度与该有效出行的实际花费时长进行时间关联性分析,算法流程如图2所示。具体步骤如下:

a) 对于轨迹匹配度集合  $Rtra$  中某个轨迹匹配比值  $R_i$ ,获取其导航线路  $nr_i$  的时间长度  $NRT_i$  与该次有效出行的花费时间  $signal_{time}$ 。若导航线路时间长度与实际有效出行花费时间越接近,则该条导航线路与实际出行的时间匹配度越大。基于

以上原因,计算导航线路与实际出行线路的时间匹配度  $Rtm_i$  公式如下:

$$Rtm_i = \begin{cases} 1 - \frac{|signal_{time} - NRT_i|}{signal_{time}} & signal_{time} > NRT_i/2 \\ 0 & signal_{time} \leq NRT_i/2 \end{cases} \quad (1)$$

b) 对轨迹匹配度集合  $Rtra$  中的每条轨迹重复步骤 a), 计算每条轨迹的时间匹配度  $Rtm_i$ , 得到时间匹配度集合  $Rtm = \{Rtm_1, Rtm_2, \dots, Rtm_i, \dots, Rtm_n\}$ 。

c) 结合轨迹匹配度集合  $Rtra$  与时间匹配度集合  $Rtm$ , 根据集合属性设定相应的权值  $K$ , 利用公式计算出每种导航 API 请求的出行方式是实际出行方式的概率集合  $P = \{P_1, P_2, \dots, P_i, \dots, P_n\}$ 。其中  $P_i$  的计算公式如下:

$$P_i = K \times Rtra_i + (1 - K) \times Rtm_i \quad (2)$$

d) 选取集合  $P$  中最大概率值  $P_{max}$  对应的导航出行方式作为该有效出行段的出行方式。

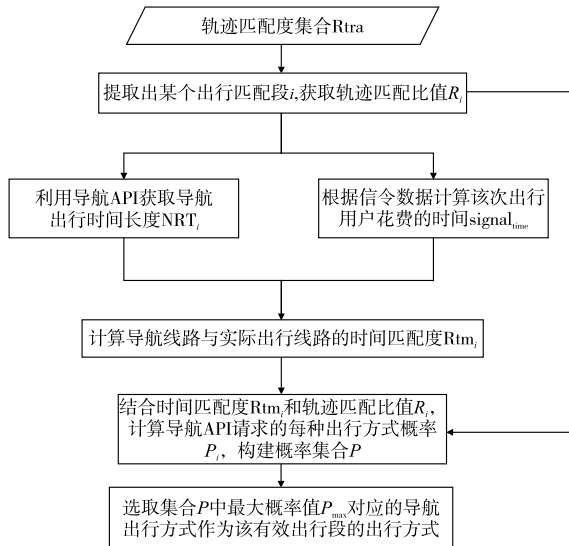


图2 基于轨迹匹配度集合与时间关联的出行方式识别流程

## 2 结果及分析

### 2.1 数据格式

#### 2.1.1 手机信令数据

本文中选用的数据包括重庆市某运营商提供的匿名脱敏后的 502 个居民某日某时段的手机信令数据和该时段的主要出行方式数据。部分原始手机信令数据如表 1 所示。

表1 部分原始手机信令数据

ID	LAC	CID	timestamp	BS type	Lon	Lat
U00024	1a3a **	C9694 **	20151023071007	4G	105.83188	30.19351
U00024	1a3a **	C9694 **	20151023071047	4G	105.83188	30.19351
U00321	133a **	C8849 **	20151022075604	4G	108.66034	31.95228

#### 2.1.2 导航数据

本文中利用百度地图提供的地图导航 API 接口, 实现获取地点信息、起止点出行方式查询等功能。本模型中主要利用一次出行中起始地与目的地的经纬度坐标, 通过 API 接口请求起止点间不同出行方式路径坐标、出行时长和出行距离等数据。

### 2.2 数据清洗

对信令数据分析发现, 连续时间内同一基站下会产生多条重复定位信令数据。另外由于受到移动通信系统自身机制的影响, 信令数据中存在乒乓效应数据以及错误定位数据, 所以本文首先对信令数据进行压缩和去噪等数据预处理工作, 减轻了后期存储与分析识别的压力。

### 2.3 用户实例

图 3 所示的是数据清洗后某用户的一次有效出行轨迹, 该出行段的相关信息如表 2 所示。



图3 某用户一次有效出行段轨迹

表2 该次有效出行段相关信息

属性	一次有效出行段信息
起点核心点坐标	(106.531829, 29.505902)
出发时刻	20151027070315
终点核心点坐标	(106.519985, 29.620844)
到达时刻	20151027080220
出行时长	59 min 5s
信令总数	187

对该用户的此次有效出行通过进行轨迹匹配以及时间关联匹配, 共获取 1 条步行线路、3 条公共交通线路、2 条驾车线路。对于时间匹配度权值, 由于理论时长与实际出行时间长度受路况等因素影响, 匹配准确度稍低于轨迹匹配度, 所以设定轨迹匹配度和时间匹配度的权值分别为 0.65 和 0.35, 计算结果如表 3 所示。

表3 该次有效出行相关匹配度计算结果

计算结果	步行	公共交通 A 线路	公共交通 B 路线	公共交通 C 路线	驾车 A 路线	驾车 B 路线
匹配信令个数	153	154	178	135	118	178
轨迹匹配度	81.82%	82.35%	95.19%	72.2%	63.1%	95.19%
导航线路时长	3 h 40 min	1 h 50 min	1 h 40 min	2 h	30 min	35 min
时间匹配度	0%	13.56%	30.51%	0%	50.85%	59.32%
交通方式概率	53.18%	58.27%	72.55%	46.93%	58.81%	82.64%

分析可得, 该次出行方式为驾车出行, 与提供的真实出行方式结果一致。另外从表 3 可看出, 公共交通 B 线路的出行方式概率也较高, 分析原因是该线路可能与实际的驾车出行线路较为一致, 导致轨迹匹配度较高。不过由于公共交通与驾车出行时间差异性较大, 时间匹配度较低, 所以时间匹配度对综合识别结果进行了修正, 从而准确地识别了该出行方式。

### 2.4 结果分析

为了对算法的有效性与合理性进行判断, 以原始数据中 502 个样本为研究对象, 运用上述识别方法进行出行方式判断, 并且采用查全率、查准率、 $F_1$ -measure 以及正确率对模型识别精度进行评价。

查全率: 即召回率 (recall), 本模型中表示对于某种出行方式, 判断结果中正确的个数与实际数据中该出行方式的比例。

查准率: 即准确率 (precision), 本模型中表示对于某种出行方式, 判断结果中正确的个数与判断结果总数的比例。

$F_1$ -measure: 即  $F_1$  值, 是查全率与查准率的加权调和平均值。本模型中用于对查全率与查准率两个指标进行综合评估。

正确率: 本模型中表示出行方式判断正确的样本个数与所有样本个数的比例。

相应的识别结果如表 4 所示。

表4 样本识别结果

实际的出行方式	识别的出行方式			查全率	$F_1$ 值
	步行	驾车	公共交通		
步行	116	11	9	85.3%	85.6%
驾车	12	167	18	84.8%	83.1%
公共交通	7	27	135	79.9%	81.6%
查准率	85.9%	81.5%	83.3%	正确率	83.3%

由表4可知,在步行、驾车以及公共交通这三种出行方式中,模型对于步行的识别率最高,查全率、查准率与 $F$ 值均在85%以上,驾车与公共交通出行方式识别率达到了80%以上,整体判断准确率达到了83.3%。

目前,部分学者利用单一的信令数据进行出行识别方面的研究,文献[12,13]中根据出行先验知识与模糊识别算法,利用不同出行方式构建出行平均速度、出行时长和出行距离等隶属度函数,从而进行出行方式的识别,但并未提供相应的识别结果。本文利用相同数据集,根据上述文献提供的识别方案与相关参数进行了实验并得出了结果:基于先验知识的出行方式模糊识别算法整体判断准确率为68%,本文提出的基于手机信令和导航数据的居民出行方式识别方法精度提升超过15%。两种方案的步行、驾车和公共交通出行识别具体结果对比如图4所示。由图4中可知,本文提出的识别方法比基于先验知识的出行方式模糊识别算法从查全率与查准率上均有所提升。

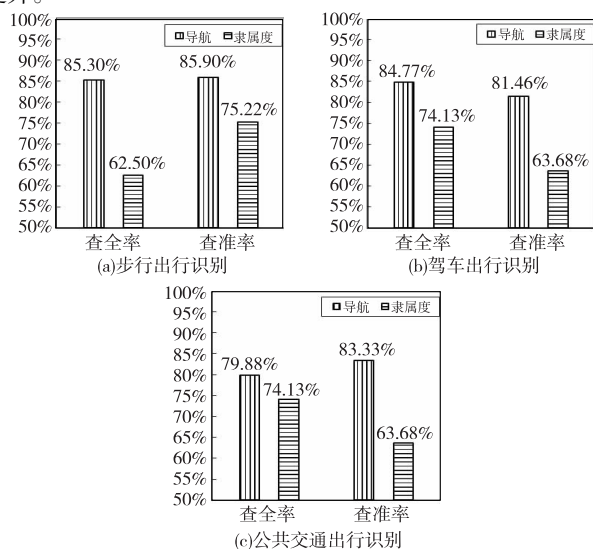


图4 两种方案识别结果的查全率、查准率对比

此外,算法模型执行效率是该算法模型能够在实际生产环境中部署的重要因素之一。本文算法模型的主要开发平台为Eclipse,版本为4.4;网络带宽为4 Mbps,无其他程序占用网络。经过对相同数据集的10次测算,该模型的平均运行时间为187 s,除去程序初始化以及读取文件等时间开销,平均运行时间为173 s,识别速度较快。但该模型由于需要使用百度地图API,所以需要一定的网络带宽来保障识别速度。

整体结果表明,本文提出的出行方式识别方法可以从轨迹相似性以及时间相似性等多角度对出行方式进行更加准确的识别,并且算法执行效率高,可用于工程环境中对居民出行方式进行识别。

### 3 结束语

居民的出行方式是城市交通规划的一个重要依据。目前,

随着智能终端设备的普及,手机信令数据也成为交通研究中一个重要的数据源。本文基于手机信令数据与导航数据,利用DBSCAN聚类算法进行轨迹聚类,同时进行时间关联,设计了一种识别居民出行方式的新方法。在基于502个居民某天某时段的手机信令数据和导航数据的基础上,利用该方法得到的结果整体正确率达到了83.3%,高于单一使用信令数据进行出行方式识别的方案,表明该方法可以更为准确地对居民的出行方式进行识别。此外,在本文的基础上,可以更加关注于多源数据的融合,如人物画像、公交IC卡和实时路况指数等数据,进一步提升居民出行方式的识别正确率。

### 参考文献:

- [1] 冉斌. 手机数据在交通调查和交通规划中的应用[J]. 城市交通, 2013, 11(1): 72-81.
- [2] Jahangiri A, Rakha H A. Applying machine learning techniques to transportation mode recognition using mobile phone sensor data [J]. IEEE Trans on Intelligent Transportation Systems, 2015, 16(5): 2406-2417.
- [3] 李喆,孙健,倪训友. 基于智能手机大数据的交通出行方式识别研究[J]. 计算机应用研究, 2016, 33(12): 3527-3529, 3558.
- [4] Shafique M A, Hato E. Use of acceleration data for transportation mode prediction [J]. Transportation, 2015, 42(1): 163-188.
- [5] Xu Chao, Ji Minhe, Chen Wen, et al. Identifying travel mode from GPS trajectories through fuzzy pattern recognition [C]//Proc of the 7th International Conference on Fuzzy Systems and Knowledge Discovery. Piscataway, NJ: IEEE Press, 2010: 889-893.
- [6] Abdelaziz A M, Youssef M. The diversity and scale matter: ubiquitous transportation mode detection using single cell tower information [C]//Proc of the 81st IEEE Vehicular Technology Conference. Piscataway, NJ: IEEE Press, 2015: 1-5.
- [7] Luo Jiangtao, Shu Zhonglin, Zhou Yunfeng, et al. Monitoring system of urban population traffic based on mobile network signaling [C]//Proc of IEEE/CIC International Conference on Communications in China. Piscataway, NJ: IEEE Press, 2014: 339-343.
- [8] 赖见辉,陈艳艳,钟园,等. 基于手机定位信息的地铁乘客出行路径辨识方法[J]. 计算机应用, 2013, 33(2): 583-586.
- [9] Bloch A, Erdin R, Meyer S, et al. Battery-efficient transportation mode detection on mobile devices [C]//Proc of the 16th IEEE International Conference on Mobile Data Management. Piscataway, NJ: IEEE Press, 2015: 185-190.
- [10] Wang Huayong, Calabrese F, Lorenzo G D, et al. Transportation mode inference from anonymized and aggregated mobile phone call detail records [C]//Proc of the 13th IEEE International Conference on Intelligent Transportation Systems. Piscataway, NJ: IEEE Press, 2010: 318-323.
- [11] Ester M, Kriegel, H P, Sander J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise [C]//Proc of International Conference on Knowledge Discovery and Data Mining. Palo Alto, CA: AAAI Press, 1996: 226-231.
- [12] 张博. 基于手机网络定位的OD调查的出行方式划分研究[D]. 北京: 北京交通大学, 2010.
- [13] 冯冲. 基于移动定位数据的用户出行模式识别[D]. 昆明: 昆明理工大学, 2011.
- [14] 郭敏,贾修一,商琳. 基于模糊化的决策粗糙集属性约简和分类[J]. 模式识别与人工智能, 2014, 27(8): 701-707.
- [15] Pawlak Z, Slowinski R. Rough set approach to multi-attribute decision analysis [J]. European Journal of Operational Research, 1994, 72(3): 443-459.
- [16] Liu Yong, Huang Wenliang, Jiang Yunliang, et al. Quick attribute reduct algorithm for neighborhood rough set model [J]. Information Sciences, 2014, 271(7): 65-81.
- [17] 刘遵仁,吴耿锋. 基于邻域粗糙模型的高维数据集快速约简算法[J]. 计算机科学, 2012, 39(10): 268-271.
- [18] 于洪,王国胤,姚一豫. 决策粗糙集理论研究现状与展望[J]. 计算机学报, 2015, 38(8): 1628-1639.

(上接第2259页)

- [8] Ziarko W. Variable precision rough set model [J]. Journal of Computer & System Sciences, 1993, 46(1): 39-59.
- [9] Zak D. Rough sets and Bayes factor [M]. Berlin: Springer-Verlag, 2005: 53-63.
- [10] Lezak D, Ziarko W. The investigation of the Bayesian rough set model [J]. International Journal of Approximate Reasoning, 2005, 40(1): 81-91.
- [11] Yao Yiyu. Decision-theoretic rough set models [C]//Lecture Notes in Computer Science, vol 4481. Berlin: Springer, 2007: 1-12.
- [12] Yao Y Y, Wong S K M. A decision theoretic framework for approximating concepts [J]. International Journal of Man-Machine Studies, 1992, 37(6): 793-809.
- [13] Yao Yiyu, Zhou Bing. Naive Bayesian rough sets [C]//Proc of Inter-