

移动数据的交通出行方式识别方法

肖艳丽,张振宇,杨文忠

(新疆大学 信息科学与工程学院,新疆 乌鲁木齐 830046)

摘要:识别用户出行的交通方式,对理解用户移动性、交通状况的分析和预测、社会活动模式挖掘等方面起着非常关键的作用。随着无线网络技术的快速发展,越来越多的传感器被用于收集移动数据,如何通过收集的信息准确地识别用户不同的交通出行方式,近年来得到了广泛的研究。针对已有的从不同角度识别交通方式的方法,首先介绍了每种方法的具体内容及应用,然后对不同方法进行分类研究,并重点分析了各类方法的特点,分析几种不同方法在不同条件下的识别精确度,最后,给出了交通方式识别方法的进一步研究方向。

关键词:交通出行方式识别;用户行为;移动数据;无线网络技术;传感器

中图分类号: TP18 **文献标志码:** A **文章编号:** 1673-4785(2014)05-0536-08

中文引用格式:肖艳丽,张振宇,杨文忠. 移动数据的交通出行方式识别方法[J]. 智能系统学报, 2014, 9(5): 536-543.

英文引用格式:XIAO Yanli, ZHANG Zhenyu, YANG Wenzhong. Research of the identification methods for transportation modes based on mobile data[J]. CAAI Transactions on Intelligent Systems, 2014, 9(5): 536-543.

Research of the identification methods for transportation modes based on mobile data

XIAO Yanli, ZHANG Zhenyu, YANG Wenzhong

(School of Information Science and Engineering, Xinjiang University, Urumqi 830046, China)

Abstract: Identification of different transportation modes in the process of user travel plays an important role in understanding individuals' mobility, analyzing and forecasting traffic conditions and mining social activity pattern. With the rapid development of wireless network technology, more and more sensors are used to collect mobile data. Specially, how to accurately identify user's different transportation modes from the collected data has been extensively researched in recent years. In addition, the methods of identification proposed from different points of view to solve the problem were studied in this paper. Each method and its application was introduced in detail and then classified and researched, respectively. The focus of analysis is put on the characteristics of each method. The levels of the recognition accuracy of different methods under different conditions were analyzed in table form. Finally, the research directions of the identification methods for transportation modes were further discussed.

Keywords: identification of transportation modes; user behavior; mobile data; wireless network technology; sensors

交通出行方式是理解用户移动行为的基本属性^[1]。用户的移动状态总是表现出一定的规律性,而对交通出行方式的识别能够体现出这种规律性,并且可以为用户提供特定的服务,比如用户常常开

车经过某个地方,可以在该地方提供相应的业务推送。交通出行方式的识别还可以为个人的回忆提供丰富的信息,也能够很方便地和朋友分享生活经历^[2]。此外,通过交通出行方式的识别,还能够进行交通状况的分析与预测,如特定时间里,在道路上(非停车场)用户乘坐的公交长时间没有移动,就可以判断该路段可能发生了交通堵塞。基于多个用户的交通方式识别,还能够检测城市里的热点区域和

收稿日期:2014-04-23.

基金项目:国家自然科学基金资助项目(61262089,61262087);新疆教育厅高校教师科研计划重点资助项目(XJEDU2012109);新疆大学博士毕业生科研启动基金资助项目(BS110127).

通信作者:肖艳丽.E-mail:xiaoyanli1314@163.com.

经典路线。

快速发展的无线定位技术能够使人们更加方便地获得移动数据,这些数据不仅可以用于定位、跟踪,而且能够表示空间上的信息。例如,将单独的数据点按照时间顺序排列,可得到用户的轨迹,这条轨迹反映了个人的移动状态,而它随着不同的交通方式改变^[3]。人们在生活工作中经常结合多种交通方式出行,例如去工作时,首先从家步行到公交车站,然后乘坐公交车,下车后再步行至公司。在运动过程中,如何识别多种交通出行方式、它们之间如何转换引起了许多研究者的关注。

表 1 不同传感器性能比较

Table 1 Performance comparison of the different sensors

传感器	表达方式	精确度	能耗强度	覆盖范围	是否受设备影响
GPS	地理坐标	高	强	户外	是
WiFi	接入口 ID+信号强度或坐标	中等	强	接入口 100 m 内	是
GSM	接入口 ID+信号强度	中等	中等	1.5~10 km	是
GIS	空间数据、属性数据	高	强	户外	是
加速度计	二维、三维加速度值	较高	弱	±2~200 g 不等	否

交通出行方式的识别,一般是用不同传感器收集多个用户的移动数据,然后从用户的历史数据中提取出特征,利用这些特征以不同的学习方法训练出一个分类模型,当遇到其他用户的移动数据时,该模型可自动识别不同的交通方式。

根据收集移动数据所使用的设备以及是否考虑能耗问题,将交通出行方式的识别方法分为 4 类:

- 1) 基于移动手机的方法;
- 2) 基于 GPS 的方法;
- 3) 基于多种传感器混合的方法;
- 4) 基于能源效率的方法。

1.1 基于移动手机的方法

基于手机的方法大部分是通过估计速度来识别的,速度是通过信号强度变化、通信时间、无线接入口的数量等特征值来估计的。随着智能手机的快速发展,手机上携带了许多内置传感器,如加速度计和螺旋仪等。

文献[4-6]利用手机上携带的加速度计收集数据,首先从数据中抽取不依赖于方向的特征,如加速度信号的大小和加速度计水平、垂直方向上的信号强度。这些方法在决策树和水平、垂直的加速度信号强度结合下能够达到较好的识别精确度。

文献[7]使用 AirSage 收集的手机呼叫详细记录,该记录由匿名的手机信令消息数据构成。用 AirSage 的无线信号抽取技术首先抽取位置信息,再

1 交通出行方式识别方法分类

随着无线网络技术的快速发展,移动数据能够被多种传感器进行收集,如移动手机、全球定位系统 GPS(global positioning system)、加速度计等。通过对移动数据的分析,可以获得不同的交通出行方式信息。如表 1 所示,不同传感器在表达方式、精确度、覆盖范围等方面有明显的区别,收集的数据具有不同的特征,因此通过不同传感器数据识别交通出行方式的方法各异。

按照时间的先后顺序确定用户的运动路径,根据路径确定运动的起始点、目的地、运动开始和结束时间。相同起点和终点上的用户,使用 K-means 聚类方法分组,计算每组所使用的平均时间,将平均时间与谷歌地图上对应路程所提供的时间进行对比,能够估计这一组人所使用的交通出行方式,然后通过粗略计算在相应路线上使用同种交通方式的人数,可推断每种交通方式在相应道路上所占比例。

上述几种识别方法没有考虑实时的交通状况,文献[8]对交通正常和堵塞两种情况分别进行识别,正常情况下,通过分析平均速度,采用速度分配率构造隐藏马尔可夫模型识别,精确度达到了 94%。堵塞情况下,利用历史数据对交通情况进行预测,并使用经济学里的累计前景理论(CPT)分析在不确定交通状况时用户对交通方式的选择,识别的精确度在 70%~89%之间,精确度变化的原因是,堵塞情况下,用户可能会改变平常乘车的习惯,继续使用以前的历史数据预测当前交通出行方式会出现偏差。

除了利用移动手机携带的内置传感器收集数据之外,研究者还通过全球移动通信系统(global system for mobile communication, GSM),进行数据的收集。文献[9-10]运用信号强度波动和服务小区数量识别交通方式。通常情况下,速度越快,与静止或慢速时相比,其信号强度变化一般比较快,而且与它相连的服务小区的数目也比较多。利用隐藏马尔可夫

模型识别,达到了82%的精确度。文献[11]计算两次连续数据点的欧几里得距离估计 GSM 的信号强度变化,然后采用二阶分类模式分别对静止、移动、行走、乘车4种出行方式进行识别。

1.2 基于GPS的方法

与移动手机收集的数据相比,GPS收集的定位数据更加精确、内容更加丰富。越来越多的人,开始使用基于GPS数据的交通出行方式识别。

文献[12]利用手机中的GPS收集数据,运用神经网络原理自动识别交通出行方式。研究者首先从GPS数据中抽取特征,如平均速度、停留时间等。仅将移动特征作为神经网络的输入,不能够保证较高的识别精确度。研究者发现输入数据量大,不仅增加了识别的计算能力,而且还消耗了移动设备的资源。为了提高识别的精确度和减少输入的数据量,又提出了一个关键点算法,算法能够过滤一些不必要的GPS数据,而只保留最重要的点的信息。利用关键点和移动特征作为神经网络的输入,使精确度达到了91.23%,同时保存了移动设备的资源。

文献[13]利用移动数据与真实世界一般常识相结合的方法来识别,并引入了一个3层的分类模型,其中模型包括低级过滤、局部过滤和真实地图3部分,低级过滤能够连续地纠正传感器的错误信息,粒子滤波能够运用状态转换空间识别多种模式,真实地图指导高级转换模型。由于考虑了实际的地图信息,因此所设计的方法具有一定的局限性,只能识别该地图区域内的交通方式。为了使方法能够适用于其他的GPS数据集,文献[14-16]通过不考虑实际地图信息的方法来识别,这些方法的预处理过程都是相似的,即将用户轨迹用基于变化点的方式分段。在抽取特征方面,文献[14]选择基本特征,如距离、平均速度等。文献[15-16]除了利用基本特征之外,还抽取了一些高级特征,如速度改变率、停车率等。一些基于实例的分类方法根据对应的特征识别不同交通方式,然而基于变化点的分段方法结合决策树算法进行识别能够得到较高的精确度,识别精确度达到了76%。由于在训练和后处理阶段都运用了人工标记数据,这使得该方法的工作量较大。

与以上提出的方法不同,为了减少人工标记数据的工作量,文献[17]中运用了无监督方法对无标记的GPS数据进行识别,这种方法基于两个假设:(1)相同方式的速度分布是相似的,(2)不同方式的速度分布不一样。首先将GPS轨迹分段,给定任一子段和它的速度分布,并给每一个速度值分配一个概率。然后通过电学当中的加权自举技术选择最高

的速度值,并基于重新选择的速度值,估计每一段的速度累计分布,最后通过速度分布就能够识别不同的交通出行方式。

文献[18]通过扩展关系马尔可夫网络自动的识别用户的不同行为,这里的关系马尔可夫网络指的是基于条件随机场^[19]方法构建的条件概率模型。只要给定了GPS数据的位置和时间信息,就能够识别购物、吃饭、访问、工作等活动,但不能识别交通方式。文献[20]在文献[18]的基础上,运用最大期望算法首先识别用户的模式转化位置,最后应用分层马尔可夫模型估计在重要位置上的交通方式。

除了识别行走、公交车、自行车等交通出行方式之外,文献[21]对道路上的卡车和轿车进行了分类。首先在主干道上分别用GPS收集轿车和卡车的定位数据,然后对两个数据集进行预处理,将轨迹分为相似长度的子轨迹。从数据集中,抽取与速度相关的加速度、减速度、加速度和减速度标准偏差4个特征值,然后选用内核支持向量机SVM对车辆进行分类,其误分类率仅为4.2%。文献[22]通过提取相应出行属性数据,利用模糊推理机制也实现了对机动车、自行车、步行等交通出行方式的识别。

1.3 基于传感器混合的方法

采用一种传感器收集数据,由于传感器故障等原因造成收集的数据不可靠,导致交通方式无法正确识别。因此交通出行方式的识别有时需要多种传感器混合收集数据,这样能够保证识别方法的可靠性。当一种传感器失效时,可以借助其他传感器记录数据。

1.3.1 传统的传感器混合方法

文献[23-24]使用20多种传感器,收集了大量的基础数据,并从时间和空间两个方面抽取位置信息、光强度、呼吸强度等特征。由于抽取出的特征值较多,利用二次抽取方式去除了一些不必要的特征值,最后选取6种特征值作为分类的依据,然后利用自动生成决策树进行识别。这种方法能够达到较高的识别精确度,并以较高的精确度识别不同行为,但由于采集数据时需要将多种传感器随身携带,在现实生活中实现比较困难。为了简化收集数据的工作,文献[25]使用5个携带在人不同部位的二轴加速度计收集数据。从数据当中抽取加速度大小、频率、速度等特征值,利用决策树自动识别,达到了较高的精确度。文献[26]从不同角度出发,使用一个多模式的传感器板收集数据。虽然传统传感器混合方法在识别方面达到了理想的精确度,但由于携带的设备太多,往往会给数据采集造成许多不便,针对

这个问题,研究者又提出了许多识别方法,结合少量设备用以减少采集数据的工作量。

1.3.2 基于GPS与加速度计的方法

文献[27]运用传感器分别收集移动数据。首先从GPS数据中抽取特征,识别精确度只有86.5%,然后从加速度数据中抽取特征,识别精度为90%。最后,将GPS与加速度数据特征相结合识别精确度达到了93%。方法表明了结合GPS和加速度计收集数据,能够更好地识别交通出行方式。

文献[28]使用连续隐藏马尔可夫模型和决策树,分别与一阶连续隐藏马尔可夫模型结合,形成组合模型识别交通方式。使用决策树与一阶离散隐藏马尔可夫模型的结合对交通方式识别有较高的精确度。不过该方法只区分了行走、自行车和机动车,而不能有效地对机动车的类别进行区分。

已有的单独设备识别方法大部分假设在GPS信号不会丢失的情况下对交通方式进行识别,而这种方法只能识别少量的交通方式。针对这种情况文献[29]提出用GPS和加速度数据共同识别,这样即使在GPS信号丢失的情况下,也能够区分交通方式,并且能够识别多种交通方式。方法首先从数据集中提取速度、加速度、减速度、角速度4个特征,然后使用随机分类器和隐藏马尔可夫模型的结合自动识别交通出行方式。最终不仅识别了多种交通出行方式,而且还区分了机动车的类别。

1.3.3 基于GSM和WiFi的结合

考虑到GPS、加速度器等传感器在大型环境下,成本和收集数据的复杂度较高,研究者开始选用适合于大型环境下传感器结合的方法。由于WiFi和GSM覆盖面广,成本较低,许多基于这两种传感器的方法开始出现。

文献[30]分别从GSM和WiFi数据集中抽取识别网络、区域、接入端口、信号强度等特征。将抽取出的特征,作为分类模型的输入数据,并使用决策树进行识别,与只用GSM数据识别的精确度相比,基于GSM和WiFi结合的方法有了很大提高。

文献[31]提出了一个移动性分类方法。在识别阶段采用C.4.5决策树作为分类模型。决策树作为分类模型使得推断过程的处理复杂度降低,运用GSM和手机上收集的WiFi数据则保证了相对较低的能耗,而且,由于GSM和WiFi的覆盖范围比较广,能够保证用户数据的持久性。方法能够区分出静止,走路和骑车3种移动状态,并且能够达到80.2%的精确率。

1.3.4 基于GPS和GIS的结合

GSM、WiFi、加速度计等收集的数据精确度不高,很多情况下,识别交通方式需要比较精确的信息。GPS与全球信息系统GIS能够提供比较精确的地理信息,结合这两种传感器能够获得较为精确的数据,从而能够提高识别方法的精确度。

文献[32]使用概率矩阵识别。通过自行车所有权,即用户是否含有自行车来判断自行车的权重,如果没有将权重设置为0,并通过轨迹的多种速度特征来识别其他的交通方式。不同的权重根据速度的范围分配给不同的交通方式,然后结合GIS信息,将机动车辆划分为特定的类别,通过GIS收集的公交车站和轨道信息及平均速度识别公交车与火车两种方式。但是该方法存在一定的不足,当GPS失效时,将会导致无法识别交通方式。除此之外,在交通堵塞的情况下,也不能区分机动车辆的类别。

为了区分机动车辆的类别,文献[33]使用GPS和真实交通网络信息对交通方式进行识别。研究者除了使用基本特征,如平均速度、最大速度等,还引入了平均车辆亲密度、平均铁路亲密度、平均候选车辆亲密度等额外特征用于识别。文章首次提出了区分机动车类别的方法,也是第一次将火车看做是一种交通出行方式,并以93.5%的精确度识别出汽车、公交等6种交通出行方式。

文献[34]提出了两个版本的分析系统,第一种用一个只依赖GPS信息的系统,也就是仅仅将GPS信息当做输入,运用模糊逻辑的算法识别交通方式,第二种则利用集成的GPS和GIS地理信息作为输入,利用模式匹配的识别子系统进行识别。方法整体识别的精确率达到了91%,GPS数据结合GIS地理信息中的公交线路信息对公交车的识别从76%提高到了80%。此外,用GIS中的公共交通路线与GPS轨迹相匹配,使有轨电车的识别精确率达到了88%。

文献[35]将GPS轨迹根据交通方式转化位置进行分段,通过基于地理信息的算法识别行走、公交等5种交通方式,其精确率达到82.6%。文献[36]设计了一个用户独立的层次马尔可夫模型,从用户历史数据中,分析用户的出行目的地和路程,推断行走和乘车。这个方法的精确度可以达到84%。

文献[37]提出了GIS地图匹配算法和基于规则的算法来识别交通出行方式,精确度达到了91.7%。周小红等^[38]也利用GIS与匹配二维地图的方式,在客户端利用HTML5作为渲染引擎,对交通方式的识

别进行可视化操作,结果较为直观。

1.4 基于能量效率的方法

已经存在的许多识别方法很少考虑能量效率问题,如内存的大小、电池的寿命等。近几年来,有很多识别方法都开始考虑能量效率的问题。提出一个普遍适用、能耗较低、同时保持较高识别精确度的方法,成为目前新的挑战。

文献[39]提出一个自适应加速度活动识别方法,用于对连续的活动、移动状态进行识别。该方法的核心是连续的处理当前或下一时刻的数据,然后动态的调整抽样频率和分类特征两个参数,并用于最优的识别活动。对于特定的活动,通过调整相适应的分类特征,以保证特征值的抽取既符合当前的活动识别,又避免抽取多余不必要的特征,从而达到节约能量这个目标。

文献[40]提出了一个高效节能移动传感系统(EEMSS)识别人们日常生活中的活动与交通出行方式。高效节能移动传感系统的核心是通过分层管理移动设备上的传感器,有选择地使用有用的传感器,关闭暂时不用的传感器,从而达到节约电量的目标。除此之外,还对加速度计和移动手机收集的数据进行了实时的分类,通过分类算法能够识别出移动物体的状态,而且还能使电量保持在 75%左右。

文献[41-42]的目标是设计一个低功耗低成本的交通方式探测器。为了达到这个目标,使用了一个低功耗低成本的感测器集线器,该设备由处理单元、限制的内存和一系列的传感器构成,本身就具有低功耗低成本的特点。该方法首先通过简化数据集和模型以缩小分类器大小,使其符合感测器集线器的内存限制,达到能量消耗最少的目标。然后通过数据的低阶多项式映射和权值向量的计算,比较决策树、自适应增强和支持向量机 3 种分类器的识别性能,可以得出支持向量机不仅在能量上消耗最少,而且还保持了较高的精确度。该方法的能耗在 5%左右,并且对 5 种交通方式识别的精确度达到了 90%。

无线网络协议 Zigbee 具有近距离、低复杂度、低功耗、低成本等特点,加速度计相比于其他传感器,能量消耗也相对较少。文献[43]结合本身低功耗低成本的 Zigbee 和加速度计,使得识别过程消耗较低的能量,并且保证了较高的识别精确度。

2 交通出行方式识别方法的比较

上述提出的 4 种交通出行方式识别方法,从一

定程度上来说,并不是绝对的,因为有些方法,除了考虑用不同传感器收集移动数据之外,还考虑了其他因素,如:抽取特征不同,将导致划分种类不同;使用不同分类方法,也会导致识别的多样性。为了更好地说明上述 4 种方法的特点,分析不同方法的精确度,从以下几个方面对不同方法进行比较,分析比较结果如表 2、表 3 和表 4 所示。

1)从实验数据方面:由于收集数据的传感器存在很大差异,见表 1,导致收集到的具体数据差别较大。在几种实验数据中,由于手机的普及,移动手机数据最容易获得。GPS 数据较为精确,在记录数据时通常以经纬度形式表现,但在室内收集会受到限制。GSM 和 WiFi 数据通常以信号强度、频率等形式存在,容易受到外界环境(如基站等)的影响。加速度计、螺旋仪、磁力计等数据使用相对较少,且精确度相对较差。在识别过程中,首先根据不同数据抽取具体特征,然后根据抽取的不同特征选择合适分类器,最后对交通方式进行具体识别。数据差异是造成识别方法多样性的主要原因。

2)从传感器方面:由于不同传感器在收集移动数据类型、表达方式、能源消耗等方面都具有明显的不同,因此对于交通方式的识别,不同传感器收集的移动数据具有不同的特性。

3)从特征提取方面:由于不同的交通出行方式是以特定形式表示的,因此抽取的特征不同,识别结果也会不一样。

4)从分类器选择方面:现在常用的分类器主要有 BP 神经网络、贝叶斯网络、马尔可夫模型、决策树、支持向量机等。贝叶斯网络和马尔可夫模型都是基于统计分类的,需要的样本量大、经验风险小、非线性拟合能力较强,这 2 种从理论上都解决了最优分类器的设计问题,但是在实施时,首先必须要解决概率密度估计问题。BP 神经网络,直接通过观测数据,从训练集中学习,是一种启发式技术。决策树、支持向量机都是基于结构分类的,需要样本少、效率较高,但是不适合划分多类。在表 3、表 4 中,用相应缩写字母表示不同的分类器:隐藏马尔可夫(HMM),混淆矩阵(CM),决策树(DT),神经网络(NNs),支持向量机(SVM),一阶马尔可夫(FM),随机森林(RF),概率矩阵(TPM)。

5)是否考虑能耗:由于考虑能耗的交通方式识别方法是近几年才提出的,还不太成熟,因此本文先从收集移动数据的传感器角度将其分类,分别在表 2、表 3 中列出。

6)精确度:表示在一次识别过程中,得到的识别 别方法越好。
结果与真实值之间的接近程度,精确度越高,说明识

表 2 方法使用不同实验数据的比较

Table 2 Comparison of the different experiment data used by the method

数据	文献方法									
	[9]	[10,12]	[13]	[16,17]	[26]	[28,29]	[30]	[32,33]	[39]	[41]
移动手机数据	×		×		×				×	
GSM 数据		×					×			
GPS 数据			×	×	×	×		×		
WiFi 数据							×			
GIS 数据								×		
加速度计数据					×	×			×	×
螺旋仪数据					×					×
磁力计数据										×

表 3 基于单一传感器识别方法比较

Table 3 Comparison of the identification methods based on single sensor

方法	传感器	提取特征	分类器	是否考虑能耗	精确度/%
文献[9]	移动手机	平均速度、用户的历史通话数据	HMM	否	94.00
文献[42]	移动手机	加速度大小、信号强度大小等	CM	是	81.75
文献[10,11]	GSM	信号强度大小、服务小区数量	HMM	否	82.00
文献[12]	GSM	数据点欧几里得距离、信号塔数目等	DT	否	85.00
文献[13]	移动手机	最大速度、车站间距离、停留时间等	NNs	是	91.23
文献[16,17]	GPS	方向改变率、停车率、速度改变率等	DT	否	76.00

表 4 基于多种传感器混合识别方法比较

Table 4 Comparison of the identification methods based on various sensors

方法	传感器	提取特征	分类器	是否考虑能耗	精确度/%
文献[40,41]	多种传感器	加速度标准差、距离等	SVM	是	90.000
文献[26]	多种传感器	能量、速度、谱熵等	HMM	否	90.000
文献[39]	多种传感器	瞬时速度、接入口数目等	EEMSS	是	90.365
文献[28]	GPS、加速度计	速度变化量、移动距离等	DT、FM	否	93.600
文献[29]	GPS、加速度计	加减速速度、角速度等	RF、HMM	否	91.000
文献[30]	GSM、WiFi	端口数目、信号强度等	DT	否	88.000
文献[32]	GPS、GIS	轨迹距离、站点位置信息	TPM	否	95.000
文献[33]	GPS、GIS	平均车辆、铁路亲密度等	RF	否	93.500

3 结束语

影响交通出行方式识别的因素很多,目前已有的方法大部分是根据具体应用设计的,具有一定的局限性。通过多种随身携带仪器收集的移动数据,选择适合特定种类的移动特征进行识别限制了交通出行方式识别的多样性。除此之外,由人携带多种

仪器收集数据,在实际情况中不太可取。运用移动手机、GPS、加速度计等少量传感器收集数据的方法,这种方法减少了收集数据的难度,并对机动车辆进行了识别,使得应用的范围更加广泛。以上方法,大多没有考虑能耗问题,近几年来,许多研究者在交通出行方式的识别中开始考虑能耗问题,虽然已经提出了几种解决方法[39-43],但还并不成熟。因

此,设计一个能耗较低、精确度较高的识别方法是未来研究的重点。

参考文献:

- [1] MIAO Lin, HSU Wenjing. Mining GPS data for mobility patterns: a survey [EB/OL]. [2014-06-20] <http://www.sciencedirect.com/science/article/pii/S1574119213000825>.
- [2] 谢幸,郑宇.基于地理信息的用户行为理解[J].计算机学会通讯, 2008, 4(10): 45-51.
XIE Xing, ZHEN Yu. To understand user behavior based on geographical information [J]. Communications of the CCF, 2008, 4(10): 45-51.
- [3] 张治华.基于 GPS 轨迹的出行信息提取研究[D].上海: 华东师范大学, 2010: 15-51.
ZHANG Zhihua. Deriving trip information from gps trajectories [D]. Shanghai: East China Normal University, 2010: 15-51.
- [4] YANG Jun. Toward physical activity diary: motion recognition using simple acceleration features with mobile phones [C]//Proceedings of the 1st International workshop on Interactive multimedia for consumer electronics. Beijing: ACM Press, 2009: 1-10.
- [5] RAVI N, DANDEKAR N, MYSORE P, et al. Activity recognition from accelerometer data [C]//Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence. Pittsburgh: AAAI Press, 2005: 1541-1546.
- [6] WANG S, CHEN C, MA J. Accelerometer based transportation mode recognition on mobile phones [C]//2010 Asia-Pacific Conference on Wearable Computing Systems (APWCS). Shenzhen, China, 2010: 44-46.
- [7] WANG H, CALABRESE F, DI LORENZO G, et al. Transportation mode inference from anonymized and aggregated mobile phone call detail records [C]//2010 13th International IEEE Conference on Intelligent Transportation Systems (ITSC). Piscataway: IEEE, 2010: 318-323.
- [8] XU Dafeng, SONG Guojie, GAO Peng, et al. Transportation modes identification from mobile phone data using probabilistic models [C]//Proceedings of the 7th International Conference on Advanced Data Mining and Applications. Beijing: China, 2011: 359-371.
- [9] MULLER IAH. Practical activity recognition using GSM data [C]//Proceedings of the 5th International Semantic Web Conference (ISWC). Athens, 2006: 1-8.
- [10] ANDERSON I, MULLER H. Exploring GSM data in pervasive environments [J]. International Journal of Pervasive Computing and Communications, 2008, 4(1): 8-25.
- [11] SOHN T, VARSHAVSKY A, LAMARCA A, et al. Mobility detection using everyday gsm traces [C]// UbiComp 2006: Ubiquitous Computing. Orange County: Springer Berlin Heidelberg, 2006: 212-224.
- [12] GONZALEZ P, WEINSTEIN J, BARBEAU S, et al. Automating mode detection using neural networks and assisted GPS data collected using gps-enabled mobile phones [J]. Intelligent Transport Systems, 2010, 4(1): 37-49.
- [13] PATTERSON DJ, LIAO Like, WANG Longhao, et al. Inferring high-level behavior from low-level sensors [C]// UbiComp 2003: Ubiquitous Computing. Seattle: Springer Berlin Heidelberg, 2003: 73-89.
- [14] ZHENG Yu, LIU Like, WANG Longhao, et al. Learning transportation mode from raw GPS data for geographic applications on the web [C]//Proceedings of the 17th International Conference on World Wide Web. Beijing: ACM, 2008: 247-256.
- [15] ZHENG Yu, CHEN Yukun, LI Quannan, et al. Understanding transportation modes based on GPS data for web applications [J]. ACM Transactions on the Web (TWEB), 2010, 4(1): 1-36.
- [16] ZHENG Yu, LI Quannan, CHEN Yingying, et al. Understanding mobility based on GPS data [C]//Proceedings of the 10th International Conference on Ubiquitous Computing. South Korea ACM, 2008: 312-321.
- [17] LIN M, HSU W J, LEE Z Q. Detecting modes of transport from unlabelled positioning sensor data [J]. Journal of Location Based Services, 2013: 1-19.
- [18] LIAO L. Location-based activity recognition [D]. Washington: University of Washington, 2006: 1-35.
- [19] LAFFERTY J, MCCALLUM A, PEREIRA FCN. Conditional random fields: Probabilistic models for segmenting and labeling sequence data [C]//Proceedings of International Conference on Machine Learning. Williams Town: ACM, 2001: 282-289.
- [20] LIAO L, PATTERSON DJ, FOX D, et al. Learning and inferring transportation routines [J]. Artificial Intelligence, 2007, 171(5): 311-331.
- [21] SUN Zhanbo, BAN Xungang. Vehicle classification using GPS data [J]. Transportation Research Part C: Emerging Technologies, 2013, 37: 102-117.
- [22] 张德欣, 刘小明, 孙立光, 等. 基于大规模定位数据的出行方式模糊判别研究 [J]. 交通信息与安全, 2011, 29(2): 1-4.
ZHANG Dexin, LIU Xiaoming, SUN Liguang, et al. Fuzzy classification of trip mode based on large scale location data [J]. Journal of Transport Information and Safety, 2011, 29(2): 1-4.
- [23] PARKKA J, ERMES M, KORPIA P, et al. Activity classification using realistic data from wearable sensors [J]. IEEE Transactions on, 2006, 10(1): 119-128.
- [24] ERMES M, PARKKA J, MANTYJARVI J, et al. Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions [J]. IEEE Transactions on Information Technology in Biomedicine, 2008, 12(1):

- 20-26.
- [25] BAO L, INTILLE SS. Activity recognition from user-annotated acceleration data[C]// Proceedings of the 2nd International Conference on Pervasive Computing. Vienna: Springer Berlin Heidelberg, 2004: 1-17.
- [26] LESTER J, CHOUDHURY T, BORRIELLO G. A practical approach to recognizing physical activities[M]// Pervasive Computing. Berlin: Springer Berlin Heidelberg, 2006: 1-16.
- [27] TROPEL P J, OLIVEIRA M S, MATTHEWS C E, et al. Prediction of activity mode with global positioning system and accelerometer data[J]. Medicine and Science in Sports and Exercise, 2008, 40(5): 972-979.
- [28] REDDY S, MUN M, BURKE J, et al. Using mobile phones to determine transportation modes[J]. ACM Transactions on Sensor Networks(TOSN), 2010, 6(2): 13-40.
- [29] MILUZZO E, LANE ND, FODOR K, et al. Sensing meets mobile social networks: the design, implementation and evaluation of the CenceMe application[C]// Proceedings of the 6th ACM Conference on Embedded Network sensor Systems. Raleigh: ACM Press, 2008: 337-350.
- [30] MUN M, ESTRIN D, BURKE J, et al. Parsimonious mobility classification using GSM and WiFi traces[C]// Proceedings of the 5th Workshop on Embedded Networked Sensors (HotEmNets). Charlottesville: ACM, 2008: 1-5.
- [31] MUN MY, SEO YW. Everyday mobility context classification using radio beacons[C]// Consumer Communications and Networking Conference. Las Vegas: IEEE, 2013: 112-117.
- [32] STOPHER P, FITZGERALD C, ZHANG J. Search for a global positioning system device to measure person travel[J]. Transportation Research Part C: Emerging Technologies, 2008, 16(3): 350-369.
- [33] STENNETH L, WOLFSON O, YU P S, et al. Transportation mode detection using mobile phones and GIS information[C]// Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. Chicago: ACM, 2011: 54-63.
- [34] TSUI S Y A, SHALABY A S. Enhanced system for link and mode identification for personal travel surveys based on global positioning systems[J]. Transportation Research Record: Journal of the Transportation Research Board, 2006, 1972(1): 38-45.
- [35] GONG H, CHEN C, BIALOSTOZKY E, et al. A GPS/GIS method for travel mode detection in New York City[J]. Computers, Environment and Urban Systems, 2012, 36(2): 131-139.
- [36] LIAO L, PATTERSON D J, FOX D, et al. Learning and inferring transportation routines[J]. Artificial Intelligence, 2007, 171(5): 311-331.
- [37] CHUNG E H, SHALABY A. A trip reconstruction tool for GPS-based personal travel surveys[J]. Transportation Planning and Technology, 2005, 28(5): 381-401.
- [38] 周小红, 刘松. 基于 GIS 的城市交通出行方式自动识别模型[J]. 现代测绘, 2013, 36(3): 46-48.
- [39] YAN Z, SUBBARAJU V, CHAKRABORTY D, et al. Energy efficient continuous activity recognition on mobile phones: an activity-adaptive approach[C]// 2012 16th International Symposium on Wearable Computers (ISWC). Newcastle: IEEE, 2012: 17-24.
- [40] WANG Y I, LIN JIALIU, ANNAVARAM M, et al. A framework of energy efficient mobile sensing for automatic user state recognition[C]// Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services. ACM, Kraków, 2009: 179-192.
- [41] YU MC, YU T, LIN C J, et al. Low power and low cost sensor hub for transportation-mode detection[EB/OL]. [2013-02-03]. Studio Engineering, HTC, Tech Rep, 2013. <http://www.csie.ntu.edu.tw/~cjlin/papers/transportation-mode/transportation.pdf>.
- [42] YU T, WANG S C, YU M C, et al. Careful use of machine learning methods is needed for mobile applications: a case study on transportation-mode detection[R]. Studio Engineering, HTC, 2013.
- [43] MUNISANKAR M, NAGENDRA R. Transportation Mode Recognition with Zigbee Using Accelerometer[J]. International Journal of Emerging Technology and Advanced Engineering, 2012, 2(5): 155-163.

作者简介:



肖艳丽,女,1989年生,主要研究方向为数据挖掘、模式识别技术等。



张振宇,男 1964 年生,副教授,主要研究方向为模式识别技术、数据挖掘、移动对等网络、机会网络等。



杨文忠,男,1971年生,博士,副教授。主要研究方向为,物联网、数据挖掘、信息安全等。