

# 利用车载 GPS 轨迹数据实现 公交车驾驶安全性分析

任慧君<sup>1</sup> 许 涛<sup>1</sup> 李 响<sup>1</sup>

1 华东师范大学地理信息科学教育部重点实验室,上海,200241

**摘 要:**交通拥挤在高度城市化地区正在成为一个日益严重的问题,而造成拥堵的原因之一就是由于一些不安全驾驶行为所带来的交通事故。因此,准确评价驾驶员的驾驶行为是否安全成为研究的热点。提出了一种利用车载 GPS 模块采集的车辆轨迹数据实现车辆驾驶行为安全性分析的方法,试图自轨迹数据中提取超速、急加速、急减速、急转弯这些潜在的不安全驾驶行为,进而评估驾驶行为的安全性,为交通管理部门评价和管理驾驶员的驾驶行为提供科学的依据。由于实验数据来自于公交车辆,因此,以公交车驾驶行为为潜在的研究对象。

**关键词:**驾驶行为分析;GPS;安全驾驶;轨迹数据

**中图法分类号:**P208

**文献标志码:**A

2012 年《道路交通事故统计》表明,违章驾驶造成的交通事故占有较大比例,其中驾驶员酒后驾车、疲劳驾车、超速行驶、路口违反交通信号等违章操作造成的交通事故占总数的 60% 以上<sup>[1]</sup>。据统计,去年全国驾龄不满 1 a 的驾驶员交通肇事导致死亡人数占机动车驾驶员肇事总数的 15.4%,且超速行驶、酒后驾驶等导致的事故中驾龄不满 1 a 的驾驶员肇事明显居高。因此,规范驾驶员的驾驶行为已经成为促进道路通畅和车辆行驶安全的重要保障。而其前提条件就是能够准确评估驾驶行为的安全性。随着车载 GPS 模块的普及,车辆行驶轨迹的记录和保存成为可能,由此,可以考虑从轨迹数据中探测驾驶行为的各种特征,进而评估其安全性。基于此,本文提出了一系列基于 GPS 数据的驾驶行为特征提取方法,并实验性地应用于某条公交线路上的驾驶行为分析,取得了初步的成果。

## 1 驾驶行为分析研究现状

国内外很多学者对驾驶员的驾驶行为以及影响驾驶员驾驶行为的因素进行了研究,特别是超速驾驶行为的研究。如储颖等提出综合考虑驾驶员

的驾驶行为和意图,从而提前对是否存在危险进行相应预警<sup>[2]</sup>;汪益纯等人分析了驾驶员的性别、年龄、职业、文化程度等社会经济特征以及日常驾驶情况和生活态度等因素对驾驶行为的影响<sup>[3]</sup>;庄明科等人探讨了风险驾驶行为与驾驶经验、驾驶技术、驾驶态度以及驾驶员的人格之间的关系<sup>[4]</sup>;白云等认为危险驾驶行为可以是失误性驾驶行为、错误性驾驶行为、非故意违规的驾驶行为和过激性驾驶行为<sup>[5]</sup>;孙伟等进行了疲劳驾驶行为的研究<sup>[6]</sup>;Tseng 通过对中国台湾地区出租车的超速驾驶数据的研究发现,驾驶员的行驶风格、工作时长和日行驶距离与超速驾驶之间有着密切的联系<sup>[7]</sup>。

最初,国外学者对超速驾驶行为的研究都是采用问卷调查的方式进行,主要是研究驾驶数据获取的准确性及影响驾驶行为和驾驶数据获取的因素。Gabany 等将超速驾驶行为的影响因素具体化为自我满足、冒险、时间压力以及疏忽大意<sup>[8]</sup>。Pels-macker 与 Janssens 通过建立超速驾驶行为模型,得出超速驾驶主要受习惯形成的影响<sup>[9]</sup>。Greaves 与 Ellison 认为,除了人口统计学数据(年龄、性别、受教育程度),驾驶员的个性以及风险规避的态度也与驾驶行为特别是超速驾驶行为之间有着较强的关系<sup>[10]</sup>。由于采用问卷的方式获取的数据的准确

收稿日期:2014-02-23

项目来源:教育部博士点博导基金资助项目(20130076110014);国家 863 计划资助项目(2013AA122302);国家自然科学基金资助项目(41271441);上海市自然科学基金资助项目(11ZR1410100)。

第一作者:任慧君,硕士。主要研究方向为交通地理信息系统及空间分析。E-mail:1061875201@qq.com

通讯作者:李响,博士,教授。E-mail:xli@geo.ecnu.edu.cn

性不能得到保证,他们采用了 GPS 技术来获取驾驶数据。发现驾驶员自己报告的超速驾驶数据与实际测得的超速驾驶数据之间呈正相关,并且根据 GPS 统计数据显示,大部分驾驶员约有 19% 的行程是超速行驶。Brown 等发现驾驶员的风险认知对其报告的超速驾驶数据有一定的影响<sup>[11]</sup>。

上述研究揭示,有效、系统、长期而客观的数据采集手段是进行准确的驾驶行为分析的重要支撑。此外,由于受传统数据收集方法的限制,驾驶行为的分析内容往往局限在超速行为的分析上,而忽略了驾驶员的一些潜在的危險驾驶习惯,如频繁的急加减速等。为此,考虑引入 GPS 轨迹数据,以期获取更全面的数据。

2 GPS 轨迹数据及相关数据模型

2.1 GPS 数据

车载 GPS 接收器可以实时获得车辆的当前地理位置及时间等信息。为便于介绍,将涉及的变量定义如下:  $T$  为时间,  $V$  为速度,  $W$  为纬度,  $J$  为经度,  $H$  为方向角,  $D$  为驾驶员 ID 号, 则 GPS 每个采样点的信息为  $I = \{T, V, W, J, H\}$ ,  $k$  时刻每个采样点的信息记为  $I_k = \{T_k, V_k, W_k, J_k, H_k\}$ 。用 TRIP 表示单次汽车行驶的整个行程, 其中包含了采样点信息和驾驶员信息, 第  $m$  次汽车的行程用  $TRIP_m$  表示, 则  $TRIP_m = \{D_m, I_{m1}, I_{m2}, I_{m3}, \dots, I_{mn}\}$ , 其中  $n$  为采样点数。汽车行程的集合用 TRIPS 表示, 即  $TRIPS = \{TRIP_1, TRIP_2, \dots, TRIP_p\}$ , 其中  $p$  为行程数量。

2.2 数据处理与计算

获取车辆的实时轨迹数据后, 根据任意连续采样点的信息, 既可以观察车辆的速度变化, 又可以根据给定的阈值判断车辆在两点间是否超速, 同时也可以计算车辆在  $k$  时刻的加速度值  $A_k$ :

$$A_k = (V_k - V_{k-1}) / (T_k - T_{k-1}) \tag{1}$$

然后根据给定的急加速的加速度阈值  $A^*$  和急减速的加速度阈值  $A'$  可以判断车辆是否发生急加速或者急减速。根据给定的急转弯的速度限值  $V'$  和行驶方向角改变阈值  $AN$ , 可以判断车辆在较短的时间内是否发生了急转弯。

通过以上判断获得车辆的超速、急变速以及急转弯的时间  $T$  和位置  $(W, J)$ , 并进行相关统计量的计算, 包括超速、急加速、急减速和急转弯的距离、时长和次数。各种驾驶行为信息计算中的相关变量声明如表 1 所示。

表 1 变量声明表  
Tab. 1 Variable Definition

行为	距离/m	时长/s	次数
超速	VS	VT	VN
急加速	AAS	AAT	AAN
急减速	ASS	AST	ASN
急转弯	TUS	TUT	TUN

当车辆发生超速时, 对超速行驶的距离、时长和次数进行累加, 就可以得到车辆累计超速行驶的距离和时长以及次数, 如式(2)~式(4)。同理, 根据式(5)~式(11)计算累计急变速和急转弯的距离、时长和次数。

$$VS = VS + \Delta S, V_k > V^* \tag{2}$$

$$VT = VT + T_k - T_{k-1}, V_k > V^* \tag{3}$$

$$VN = VN + 1, V_k > V^* \ \& \ V_{k-1} < V^* \tag{4}$$

$$AAS = AAS + \Delta S, A_k > A^* \tag{5}$$

$$AAT = AAT + T_k - T_{k-1}, A_k > A^* \tag{6}$$

$$AAN = AAN + 1, A_{k-1} < A^* \ \& \ A_k > A^* \tag{7}$$

$$ASS = ASS + \Delta S, A_k < A' \tag{8}$$

$$AST = AST + T_k - T_{k-1}, A_k < A' \tag{9}$$

$$ASN = ASN + 1, A_k < A' \ \& \ A_{k-1} > A' \tag{10}$$

$$TUS = TUS + \Delta S, V_k > V' \ \& \ V_{k-1} < V' \ \& \text{Abs}(H_k - H_{k+1}) > AN \tag{11}$$

$$TUT = TUT + T_k - T_{k-1}, V_k > V' \ \& \ V_{k-1} < V' \ \& \text{Abs}(H_k - H_{k+1}) > AN \tag{12}$$

$$TUN = TUN + 1, \text{Abs}(H_k - H_{k+1}) > AN \ \& \ \text{Abs}(H_k - H_{k-1}) < AN \tag{13}$$

其中,  $\Delta S$  为  $k$  时刻的位置  $(W_k, J_k)$  和  $k-1$  时刻的位置  $(W_{k-1}, J_{k-1})$  的距离, 计算公式如式(14)所示;  $\text{Abs}(H_k - H_{k+1})$  表示时间间隔为 1 s 的两个采样点方向角差的绝对值。

$$\Delta S = R \cos^{-1}(\sin(W_{k-1} \times \frac{\pi}{180}) \times \sin(W_k \times \frac{\pi}{180}) + \cos(W_{k-1} \times \frac{\pi}{180}) \times \cos(W_k \times \frac{\pi}{180})) \times \cos((J_{k-1} - J_k) \times \frac{\pi}{180}) \tag{14}$$

式中,  $R$  为地球平均半径, 取值为 6 378 137 m。

2.3 数据分析

通过以上计算得到车辆危险驾驶的情况, 就可以对驾驶员的行为进行评价。当  $VN$ 、 $AAN$ 、 $ASN$ 、 $TUN$  较大时, 表明驾驶员出现了多次超速、急变速和急转弯; 当  $VS$ 、 $VT$ 、 $AAS$ 、 $AAT$ 、 $AST$ 、 $ASS$ 、 $TUS$ 、 $TUT$  较大时, 表明超速、急变速和急转弯持续的距离和时间比较长, 即可以判断驾驶员驾驶行为的危险性较高。反之, 则表明

驾驶员驾驶过程中的危险较低或者较稳健。

### 3 实例分析

#### 3.1 数据采集

本文使用 GPS 数据记录器以时间间隔为 1 s 的频率采集实验数据,该频率已经是一般纯 GPS 位置采样设备的最高频率。从急变速的数据中可以看出,公交车在行驶过程中出现持续急变速的时间一般不超过 2 s,如果增大采样间隔,将可能降低数据分析结果的准确性。

在 VS2010 平台上使用 C# 语言编程实现对数据的提取和计算。该公交路线的行程约 30~50 min,行驶的往返路线并不完全重合,而且路线上有 9 个拐弯处,整个行程路线如图 1 所示。



图 1 路线图  
Fig. 1 Driving Route

本文选取了 4 条采集于晚高峰时间段的不同车辆的驾驶轨迹数据,分别用 TRIP<sub>1</sub>、TRIP<sub>2</sub>、TRIP<sub>3</sub>、TRIP<sub>4</sub>表示,详细信息如表 2 所示。

表 2 不同车辆的行程信息  
Tab. 2 Driving Behavior of Cars

TRIPS	驾驶员	启动时间	停止时间	行程时间/s
TRIP <sub>1</sub>	1	18:00:57	18:35:32	2 075
TRIP <sub>2</sub>	2	17:52:29	18:29:15	2 206
TRIP <sub>3</sub>	3	18:11:15	18:50:24	2 349
TRIP <sub>4</sub>	4	17:59:37	18:38:11	2 314

#### 3.2 超速行为分析

高速行驶往往会造成交通事故。在不同的道路环境中,最高安全行车速度是不同的。一般公交车的限速为 60 km/h,本文中该公交车行经多个学校和商场,且转弯路口较多,设定当速度达到 55 km/h 时,认为驾驶员有超速倾向。

图 2 中, SPEED<sub>1</sub>、SPEED<sub>2</sub>、SPEED<sub>3</sub>、SPEED<sub>4</sub> 分别对应于行程 TRIP<sub>1</sub>、TRIP<sub>2</sub>、TRIP<sub>3</sub>、TRIP<sub>4</sub> 的行驶速度。从图 2 可以看出,不同的驾驶员在晚高峰时间段的驾驶速度有一定的差异,但变化趋势大体相同,都是在开始的 1/4 行程时速度较低,在 1/2~3/4 行程上行驶的平均速度较高,到最后的 1/4 的行程时速度又较低。此外,同一位驾驶员的驾驶速度变化还表现出明显的对称性,即汽车在远离车站的路段,平均速度相对较大,在接近车站的路段,平均速度较小。根据超速倾向阈值 55 km/h,1 号驾驶员有 4 次超速驾驶倾向,2 号驾驶员和 4 号驾驶员各有一次超速驾驶倾向,3 号驾驶员在驾驶过程中没有出现超速驾驶倾向。每次汽车倾向超速行驶的详细信息如表 3 所示。1 号驾驶员超速驾驶的危险性最高,累计超速行驶 84 s,累计超速行驶距离 1 478.38 m,达累计行驶距离的 11.2%。

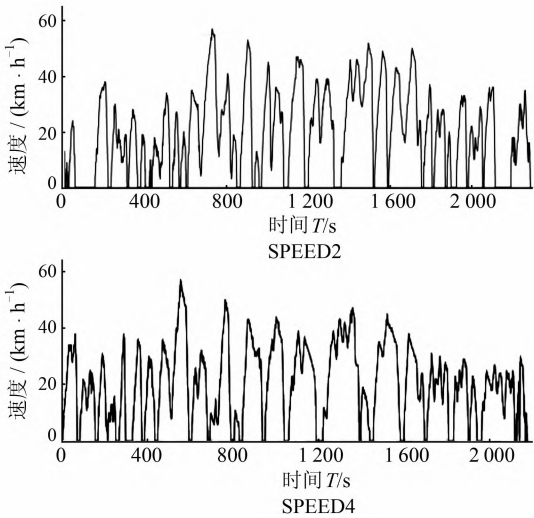
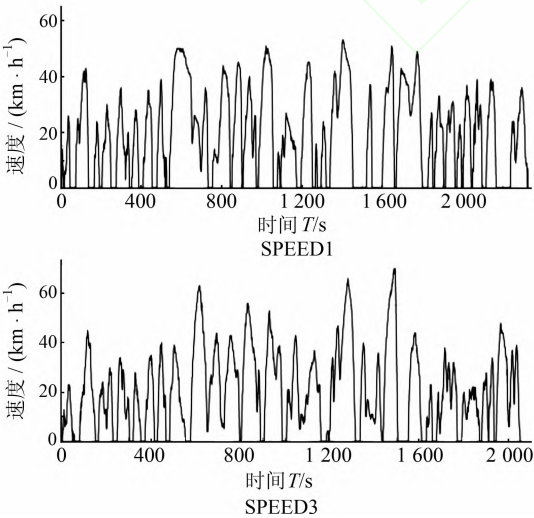


图 2 实时速度变化图  
Fig. 2 Real-time Speed Changing

图 3 是速度描述统计信息图,纵轴代表以该速度行驶时间占整个行程时间的比例。很明显,4

条行驶轨迹的低速行驶时间都占总行驶时间的 70% 左右,中速行驶时间占总时间的比值有一定



表 3 晚高峰时段超速行驶详细信息  
Tab. 3 Details of Speeding at Peak Period

TRIPS	VT/s	VS/m	VN
TRIP <sub>1</sub>	84	1 478.38	4
TRIP <sub>2</sub>	5	80.468	1
TRIP <sub>3</sub>	0	0	0
TRIP <sub>4</sub>	8	128.879	1

的差异,在 14%~21%,高速行驶的时间占总行驶时间的比例较小,其中 SPEED<sub>3</sub> 的高速行驶时间占比最长,为 14.49%,最后超速行驶时间占总时间的比例都比较少,其中 SPEED<sub>3</sub> 中没有超速, SPEED<sub>2</sub> 和 SPEED<sub>4</sub> 中超速的时间占总行驶时间的比例接近零。但值得注意的是, SPEED<sub>1</sub> 的超速行驶时间占总行驶时间的 4.09%。

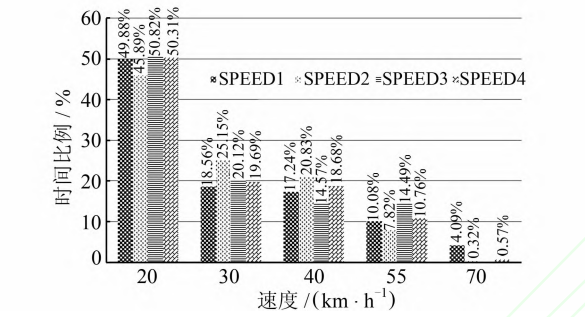


图 3 速度描述统计信息图  
Fig. 3 Statistical Information of Speed for Cars

3.3 急变速驾驶行为分析

车辆加速度急剧变化时,会给人以不舒适的感觉,特别是公共汽车,当车辆急刹车或急加速时,会造成站着的乘客失去平衡,因而急加速或急减速都存在一定的安全隐患。在本实验中,首先记录不同的乘坐感受,然后通过数据处理查看对应时刻的加速值,进而确定加速度阈值。本文设定当加速度值超过 3 m/s<sup>2</sup> 时为急加速,低于 -3 m/s<sup>2</sup> 时为急减速,大于 3 m/s<sup>2</sup> 或小于 -3 m/s<sup>2</sup> 时,就可以认为车辆发生急变速。车辆在行驶过程中以不同加速度行驶的时间占总行驶时间的比例结果如表 4 所示, TA<sub>1</sub>、TA<sub>2</sub>、TA<sub>3</sub>、TA<sub>4</sub> 分别对应于 TRIP<sub>1</sub>、TRIP<sub>2</sub>、TRIP<sub>3</sub>、TRIP<sub>4</sub> 中的加速度。

表 4 加速度值分布  
Tab. 4 Description of Acceleration

加速度值/(m · s <sup>-2</sup> )	TA <sub>1</sub>	TA <sub>2</sub>	TA <sub>3</sub>	TA <sub>4</sub>
(-5,-3]	0.00	0.05	0.00	0.00
(-3,-2]	0.15	0.14	0.21	0.26
(-2,0)(0,2)	65.71	58.99	54.99	58.53
[2,3)	0.10	0.09	0.09	0.09
[3,5)	0.00	0.09	0.04	0.00

汽车行驶过程中,加速度主要集中分布在 2

~-2 m/s<sup>2</sup> 之间。整个行驶过程中,TRIP<sub>1</sub> 的低变速行驶时间比例最大,为 65.7%,其次为 TRIP<sub>2</sub> 的 58.99%,TRIP<sub>4</sub> 的 58.53% 和 TRIP<sub>3</sub> 的 54.98%。当加速度限值设为 ±3 m/s<sup>2</sup> 时,TRIP<sub>2</sub> 中的急变速时间占比为 0.14%,TRIP<sub>3</sub> 中的急变速时间占比为 0.04%。通过分析 4 辆车在行驶过程中的加速度变化可以看出,1 号驾驶员行驶过程中最稳定,4 号驾驶员行驶较稳定,而 2 号驾驶员和 3 号驾驶员出现了不同程度的急变速驾驶,且 TRIP<sub>2</sub> 中出现了持续急变速驾驶,TRIP<sub>3</sub> 中出现了一次急加速,且急加速行驶的距离较短。急变速的详细信息如表 5 所示。

表 5 急变速的详细信息  
Tab. 5 Details of Quick Acceleration

TRIPS	AAS/m	AAT/s	ASS/m	AST/s
TRIP <sub>1</sub>	0	0	0	0
TRIP <sub>2</sub>	16.105	2	0.269	1
TRIP <sub>3</sub>	2.736	1	0	0
TRIP <sub>4</sub>	0	0	0	0

3.4 急转弯驾驶行为分析

在车辆驾驶中,急转弯不仅会给车辆带来很大的磨损,而且很有可能会造成翻车等交通事故。根据公交驾驶准则,机动车辆在通过胡同(里巷)、铁路道口、急转弯、窄路、窄桥、隧道以及掉头时,最高时速不准超过 20 km/h。因此,本文将急转弯的速度阈值设为 20 km/h,行车方向改变的角度阈值根据常规的交叉交叉路口设为 90°。本文通过式(11)~式(13)计算得出,车辆在 5 s 的时间里,平均速度为 21 km/h,行车方向改变接近 120°,即 TRIP<sub>4</sub> 中发生急转弯。车辆在某路口拐弯的 GPS 轨迹路线如图 4 所示。

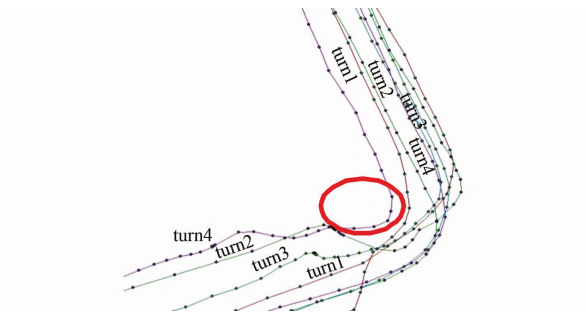


图 4 车辆的拐弯路线  
Fig. 4 Turning Routes of Cars

3.5 讨论

本文的主旨是从 GIS 的角度提出一种对驾驶行为安全性进行评定的可能方法,重点在于如何从 GPS 轨迹中获得与这些阈值进行比较的变量,因此,阈值的变化并不影响本文所提出的方法论的实践。

上述数据表明,晚高峰时段的  $TRIP_1$  中有较严重的超速驾驶行为,没有急变速; $TRIP_2$  有轻微的超速和较长的累计急变速驾驶距离; $TRIP_3$  没有超速却有轻微的急变速; $TRIP_4$  有轻微的超速和一定程度的急转弯。

经过分析可以看出,同一位驾驶员的速度变化表现出不完全对称性,这是因为该公交走的是不完全重合的往返路线。而不同的驾驶员的速度有差异,但变化趋势基本相同,说明驾驶员的驾驶行为不仅与驾驶员的驾驶习惯有关,还与当时的路况有关,如夜间行车,驾驶员视野受限且容易疲劳,所以驾驶速度会降低,但同时由于道路相对白天较通畅,所以驾驶员容易放松警惕,从而出现较严重的急变速驾驶行为。根据不同驾驶员的转弯路线可以发现,驾驶员的驾驶行为也与驾驶员的驾驶技术相关。选取的4辆汽车中,有3辆汽车出现了超速驾驶的倾向,2号驾驶员有较长距离的急变速驾驶,4号驾驶员有一定程度的急转弯驾驶行为。汽车在远离车站的路段,平均速度相对较大,在接近车站路段,平均速度较小,而且急变速也是发生在靠近车站的路上,可能是与道路周边的情况有关,因为靠近车站的地方有一个较大的商场,并且有信号交叉口,机动车和行人都比较多,因此汽车虽然平均速度较低,但有时可能会为了避让行人或红绿灯而发生急变速,而在远离车站的地方,机动车和行人较少,因此平均速度较高,但行驶过程比较平稳。

## 4 结 语

交通拥挤在高度城市化地区正在成为一个日益严重的问题,而造成拥堵的原因之一就是由于一些不安全驾驶行为所带来的交通事故。本文利用车载GPS接收器获取车辆的轨迹信息,进而计算车辆的行驶速度、加速度和转弯信息等,通过设定适当的阈值,可以为评价驾驶员的驾驶习惯、驾驶技术和安全意识提供较为客观的依据。在今后的研究中,将进一步搜集更多的车辆轨迹数据,并结合车辆实际出险情况,以期更为科学合理地制定出一系列评估驾驶行为安全性的指标体系。

## 参 考 文 献

[1] Zhong Yuanyue. Exploration on Urban Road Traffic Safety Planning[J]. *China New Technologies and Products*, 2010(19): 69-70(钟远岳. 城市道路交通安全规划探讨[J]. 中国新技术新产品, 2010(19):

69-70)

- [2] Chu Ying, Xiao Xianqiang, Zhu Jiacheng. Study on Automotive Active Safety Technology Based on Driving Behavior and Intention[J]. *Machinery Design & Manufacture*, 2011(1): 266-268(储颖, 肖献强, 朱家诚. 基于驾驶行为及意图的汽车主动安全技术研究[J]. 机械设计与制造, 2011(1): 266-268)
- [3] Wang Yichun, Chen Chuan. Research of Novice Drivers' Behavior Based on Factor Analysis[J]. *Journal of Wuhan University of Technology (Transportation Science & Engineering)*, 2012, 36(5): 1 064-1 068(汪益纯, 陈川. 基于因子分析法的初驾者驾驶行为研究[J]. 武汉理工大学学报, 2012, 36(5): 1 064-1 068)
- [4] Zhuang Mingke, Bai Haifeng, Xie Xiaofei, et al. A Study on Risky Driving Behavior and Related Factors[J]. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2008, 28(44): 475-481(庄明科, 白海峰, 谢晓非, 等. 驾驶人员风险驾驶行为分析及相关因素研究[J]. 北京大学学报(自然科学版), 2008, 28(44): 475-481)
- [5] Bai Yun, Shi Jing. Driving Behavior and Influential Factors Study in Beijing[J]. *Transportation Information and Safety*, 2010, 28(2): 114-119(白云, 石京. 北京市驾驶行为与影响因素分析[J]. 交通信息与安全, 2010, 28(2): 114-119)
- [6] Sun Wei, Zhang Weigong, Zhang Xiaorui, et al. Development of Fatigue Driving Detection Method Research[J]. *Automobile Technology*, 2009(2): 1-5(孙伟, 张为公, 张小瑞, 等. 疲劳驾驶检测方法的研究进展[J]. 汽车技术, 2009(2): 1-5)
- [7] Tseng C M. Operating Styles, Working Time and Daily Driving Distance in Relation to a Taxi Driver's Speeding Offenses in Taiwan[J]. *Accident Analysis and Prevention*, 2013, 52: 1-8
- [8] Gabany S G, Plummer P, Grigg P. Why Drivers Speed: The Speeding Perception Inventory [J]. *Journal of Safety Research*, 1997, 28(1): 29-36
- [9] De Pelsmacker P, Janssens W. The Effect of Norms, Attitudes and Habits on Speeding Behavior: Scale Development and Model Building and Estimation[J]. *Accident Analysis and Prevention*, 2007, 39: 6-15
- [10] Greaves S P, Ellison A B. Personality, Risk Aversion and Speeding: An Empirical Investigation[J]. *Accident Analysis and Prevention*, 2011, 43: 1 828-1 836
- [11] Brown S L, Gould R V. A Prospective Study of Relationships Between Propositions About Risk and Driver Speeding[J]. *Accident Analysis and Prevention*, 2012, 46: 1-7

Driving Behavior Analysis Based on Trajectory Data Collected with Vehicle-mounted GPS Receivers

REN Huijun<sup>1</sup> XU Tao<sup>1</sup> LI Xiang<sup>1</sup>

1 Key Laboratory of Geographical Information Science, Ministry of Education, East China Normal University, Shanghai 200241, China

**Abstract:** Traffic congestion is becoming an increasingly serious problem in highly urbanized areas. One reason for congestion are frequent happened traffic accidents caused by risky driving behaviors. The accurate evaluation of a driver behavior as well as safety analysis has therefore become a research hotspot. In this paper, a new method of using the vehicle-mounted GPS model to collect the vehicle track data and analyzing the safety of driver's behavior is presented. In the approach, information about driving over the speed limit, quick acceleration, sharp slowdown and sharp turns is extracted from the trajectory data. Then, the security of a drivers' behavior assessed to provide scientific evidence for the transportation department managers when evaluating and managing a drivers' driving skill.

**Key words:** driving behavior analysis;GPS;safe driving; trajectory data

**First author:** REN Huijun, master, specializes in the GIS for transportation and spatial analysis. E-mail: fjbsm\_jcl@163.com

**Corresponding author:** LI Xiang, PhD, professor. E-mail: xli@geo.ecnu.edu.cn

**Foundation support:** The Doctoral Fund of Ministry of Education of China, No. 20130076110014; the National 863 Program of China, No. 2013AA122302; the National Natural Science Foundation of China, No. 41271441; the National Natural Science Foundation of Shanghai in China, No. 11ZR1410100.

(上接第 738 页)

Personal Profile Mining Based on Mobile Phone Location Data

CHEN Jia<sup>1</sup> HU Bo<sup>2</sup> ZUO Xiaoqing<sup>3</sup> YUE Yang<sup>4</sup>

1 Fujian Provincial Geomatics Center, Fuzhou 350003, China

2 Chongqing Survey Institute, Chongqing 420020, China

3 Faculty of Land Resource Engineering, Kunming University of Science and Technology, Kunming 650093, China

4 Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, China

**Abstract:** Understanding personal profiles like preferences, income levels, and geographical areas is the basis of providing a person with personalized and accurate services. In order to acquire personal profiles we propose a reasonable technical route that first extracts the geographic regions from personal mobile phone location data based on a density-based clustering algorithm. Then, the geographic regions are tagged with semantic meaning and we analyze house descriptions by NLP(Natural Language Processing). A division method for people's daily time is given, based on the assumed the activity patterns of people. At last, an individual's taste for something or his income levels is analyzed using the statistics for POIs and house prices in the extracted places. An experiment with real data shows that this method is an effective solution to mining personal profiles.

**Key words:** data mining; mobile phone location data; user profile; location based services

**First author:** CHEN Jia, postgraduate, specializes in digital city and mobile trajectories mining and analysis. E-mail: chenjia305@126.com

**Corresponding author:** YUE Yang, PhD, associate professor. E-mail: yueyang@szu.edu.cn

**Foundation support:** The National Natural Science Foundation of China, Nos. 41231171, 41171348, 41061043; Shenzhen Scientific Research and Development Funding Program, Nos. JCYJ20121019111128765, JCYJ20130329144141856; CCF-Tencent ARG20130115.