

基于手机浮动车的高速公路平均路段行程速度提取技术

杨庆芳, 魏领红, 郑黎黎

(吉林大学 交通学院, 长春 130022)

摘要:提出了一种新的高速公路路段行程速度提取技术。因为在提取过程中路段划分是高速公路手机浮动车数据提取的基础和前提,首先提出了基于Fisher聚类分析法的手机浮动车高速公路路段长度划分方法,并基于此方法建立了高速公路平均行程速度参数提取模型,最后利用SPSS软件对相关数据进行了聚类分析,并采用VISSIM仿真软件对高速公路手机浮动车进行了模拟。模拟结果表明,模型在交通数据采集时有很好的精度,具有一定的应用价值。

关键词:道路工程;高速公路;手机浮动车;Fisher聚类分析;平均行程速度

中图分类号:U491.2 **文献标识码:**A **文章编号:**1671-5497(2009)Sup.1-0104-05

Extraction technology of road link mean speed on highway based on cell phone floating car

YANG Qing-fang, WEI Ling-hong, ZHENG Li-li

(College of Transportation, Jilin University, Changchun 130022, China)

Abstract: Road link travel speed can effectively reflect the state of freeway traffic flow, and its extraction technology is particularly important. During the extraction, the sections of the length of highway is the foundation and prerequisite for the floating car data from the phone. Firstly, the sections of the length of highway on phone floating car based on Fisher cluster analysis was introduced, and based on the method, the model of highway road link travel speed was established. Finally, SPSS software was used to perform the cluster analysis on related data. This algorithm was tested with the micro simulation tool VISSIM. The results indicate that model has a very good precision for data collection, so has a certain value of application.

Key words: road engineering; highway; phone floating car; Fisher cluster analysis; road link mean speed

近年来,移动通讯技术的发展和手机使用率的增加,使得手机覆盖率、定位数据的定位频率和定位精度已可以满足交通参数估计的要求。从而实现对路网范围内多个路段交通状态的低成本、

全天候、“动态”实时检测^[1]。而以往的固定型检测器只能检测到某断面的基础交通参数数据,具有很强的限制性。采集整个路网上的交通信息需要在路网上全面布设检测器,大大增加了交通信

收稿日期:2008-07-08.

基金项目:“863”国家高技术研究发展计划项目(2007AA11Z245).

作者简介:杨庆芳(1966—),女,教授.研究方向:智能交通运输系统. E-mail: yangqf@163.com

©1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

息采集的成本。随着定位技术和无线通信技术的发展, 基于浮动车的交通信息采集技术由于覆盖范围、实施成本等方面的优势, 已经成为国内外研究的热点。

对于基于手机浮动车的交通信息采集技术, 国内还处于起步阶段, 在高速公路中的应用也少有成果。作者首先针对基于手机浮动车高速公路路段行程速度提取技术中传统路段划分方法的不足, 以手机浮动车数量作为确定路段划分的依据, 以手机浮动车按路段的状态分布作为聚类阈值, 提出了基于 Fisher 聚类分析法的手机浮动车高速公路路段长度划分方法。基于此方法提出了手机浮动车高速公路路段行程速度提取方法, 最后运用统计软件和仿真软件对实例进行分析和模拟验证。通过实例验证, 基于手机浮动车的高速公路路段行程速度提取技术能够动态、经济、全面地采集到路网的路段平均行程速度及固定型检测器检测不到的交通信息, 为分析拥堵成因、改善交通组织、智能导航及位置服务提供宝贵的基础数据^[2-3]。

1 路段划分

1.1 路段划分方法的提出

路段长度划分是进行基于手机浮动车高速公路实时交通参数建模的前提条件, 直接影响交通参数提取精度。传统的路段划分有两种^[4-6]: 一是定长法, 即路段以固定的长度间隔进行划分, 通常划分间隔为 0.5 km 或 1 km, 但是由于划分形式过于简单, 对道路交通状况实际区别的反映灵活性欠佳; 二是不定长法, 主要以交通量、限速或者线性指标作为分段依据, 考虑道路特征点较多时计算比较繁琐, 容易造成分段过短或者是过长。因此, 传统方法都不能很好地适用于高速公路。

高速公路上手机浮动车是随机零散分布的, 如果不进行路段划分或简单划分, 会造成高速公路匝道之间(路段)过长或过短^[7], 在基于手机浮动车矢量地图进行大规模行程时间数据采集时, 不利于手机浮动车数据的提取, 也无法对道路的交通进行更详细地分析, 更无法得出较为精确的数据模型, 如图 1 所示。基于上述原因, 提出了本文方法来进行精确的路段划分, 以保证每一辆工作中的手机浮动车能够精确高效地实现数据采集任务。本文方法能较为准确地判断手机浮动车的分布情况, 分段时能够很好地考虑手机浮动车的

分布, 把手机浮动车分布相近的邻近路段划分到同一研究样本内, 有助于准确地量化影响因素的作用; 且划分后的手机浮动车易于服从某种概率分布, 有利于交通模型的构建。

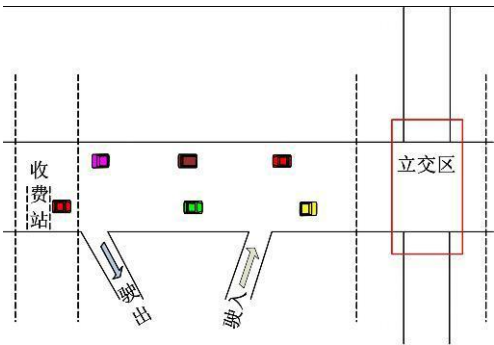


图 1 手机浮动车高速公路路段长度划分路段示意图

Fig. 1 Sections of length of highway on phone floating car

1.2 方法的原理

“聚类分析”是指统计对象(元素个体)没有顺序, 按一定指标进行聚类。有序样本即为按照一定顺序排列的样本, 其次序不可打乱。有序样本聚类方法与一般的聚类方法区别在于各样品的“地位”不相同, 最常用的是 Fisher 最优分割方法^[4-5], 该方法的本质是把有序样本分成 n 段, 使得段内离差平方和最小, 段外离差平方和最大。手机浮动车的空间分布就是有序样本, 其按里程分布的手机浮动车指标次序是不可以打乱的。

基于高速公路路网特性及手机浮动车特性, 通过路段划分确定明确的路段进出口边界, 建立路段间的联通关系及方向等属性, 为实时提取高速公路交通状态参数奠定基础。图 2 为基于手机浮动车的有序聚类分析划分高速公路路段的方法。图 2 中所示分段形式把相邻的手机浮动车数比较接近的路段划分为一段, 把手机浮动车分布集中和手机浮动车分布稀少的相邻区段分开来考虑。从理论上讲, 这种分段形式有利于考虑交通流的运行状况, 提高交通参数模型的精度, 该方法的实际应用也证明了这一点。

1.3 方法的步骤

(1) 计算各段类直径

G_{ij} 的直径用 $D(i, j)$ 表示

$$D(i, j) = \sum_{l=i}^j (x_l - \bar{x}_{ij})(x_l - \bar{x}_{ij}) \quad (1)$$

式中: 样本依次为 x_1, x_2, \dots, x_n (每个 x 都可以是 m 维的向量), 某一类 G_{ij} 是 $\{x_1, x_2, \dots, x_j\}$ ($j >$

$i)$, 它们的均值记成 \bar{x}_{ij}

$$\bar{x}_{ij} = \frac{1}{j-i+1} \sum_{l=i}^j x_l \quad (2)$$

(2) 计算误差函数

可通过下式计算这种分类的目标函数为

$$e[p(n, k)] = \sum_{j=i}^k D(i_j, i_{j+1} - 1) \quad (3)$$

将 n 个样品分成 k 类, 设某种分法是: $p(n, k): \{x_{i_1}, x_{i_1+1}, \dots, x_{i_2-1}\}, \{x_{i_2}, x_{i_2+1}, \dots, x_{i_3-1}\}, \dots, \{x_{i_k}, x_{i_k+1}, \dots, x_n\}$ 或简记成 $p(n, k): \{i_1, i_1+1, \dots, i_2-1\}, \{i_2, i_2+1, \dots, i_3-1\}, \dots, \{i_k, i_k+1, \dots, n\}$, 其中分点为 $1=i_1 < i_2 < \dots < i_k < i_{k+1} \leq n$, 当 n, k 固定时, $e[p(n, k)]$ 越小表示各类的离差平均和越小, 分类是合理的, 因此要寻找一种方法使目标函数 $e[p(n, k)]$ 达到极小, 下面的 $p(i, j)$ 表示是 $e[p(n, k)]$ 达到极小时的分类。

(3) 求精确最优解

当要分成 k 类时, 首先找 j_k , 使其满足式(4)最小

$$e[p(n, k)] = \min_{k \leq j \leq n} \{e[p(j-1, k-1)] + D(j, n)\} \quad (4)$$

于是 $G_k = \{j_k, j_{k+1}, \dots, n\}$, 然后再找 j_{k-1} , 使它满足

$$e[p(j_k-1, k-1)] = \min \{e[p(j_{k-1}-1, k-2)] + D(j_{k-1}, j_k-1)\} \quad (5)$$

由式(5)可以得到 $G_{k-1} = \{j_{k-1}, \dots, j_k-1\}$, 依此类推, 可得到 G_1, G_2, \dots, G_k , 这就是欲求的最优解^[4]。

2 行程速度提取模型

基于上述路段划分方法, 手机浮动车经过高速公路 i 路段, 即经过节点 i 和 $i+1$ 时的时刻分别是 T_i, T_{i+1} , 建立如下手机浮动车的行程速度模型

$$v_{\text{probe}(j)} = \frac{T_{i+1} - T_i}{D} \quad (6)$$

式中: $v_{\text{probe}(j)}$ 为浮动车 j 在路段 D 上的速度; T_{i+1} 为路段 i 下游时间; T_i 为路段 i 上游时间。

多手机浮动车路段距离平均速度表示如下

$$\bar{v}_{\text{probe}} = \frac{\int_0^\infty vk(v)[D + vT] dv}{\int_0^\infty k(v)[D + vT] dv} \quad (7)$$

式中: \bar{v}_{probe} 为浮动车路段平均车速; T 为时间间隔; D 为路段长度^[8]。

因此, 路段距离平均速度可通过如下公式获得

$$\bar{v}_{\text{probe}} = \frac{D\bar{v}_{\text{SMS}} + T\bar{v}_{\text{TMS}}\bar{v}_{\text{SMS}}}{D + T\bar{v}_{\text{SMS}}} \quad (8)$$

式中: \bar{v}_{SMS} 为空间平均速度; \bar{v}_{TMS} 为时间平均速度。

路段平均速度和时间平均速度的相关关系如下

$$\bar{v}_{\text{probe}} = \frac{D\bar{v}_{\text{SMS}} + T(\bar{v}_{\text{SMS}}^2 + \sigma^2)}{D + T\bar{v}_{\text{SMS}}} \quad (9)$$

式中: σ^2 为浮动车时间平均速度样本方差; 其中, 时间平均车速

$$\bar{v}_{\text{TMS}} = \frac{\sum_{i=1}^n v_i}{n} \quad (10)$$

式中: \bar{v}_{TMS} 为时间平均速度; v_i 为第 i 辆车的点速度; n 为观测的车辆数。

空间平均车速表示为

$$\bar{v}_{\text{SMS}} = \frac{1}{1/n \sum_{i=1}^n (1/v_i)} \quad (11)$$

对其中的某指定的手机浮动车进行了人工调查, 确定其在路段划分后的路段 D 上的 \bar{v}_{SMS} 和 \bar{v}_{TMS} 。

3 方法验证

利用手机浮动车的 Fisher 有序聚类分析方法, 以单公里手机浮动车数的空间分布为聚类指标, 对长春市某高速公路进行了分段分析, 在此方法的基础上由行程速度提取模型提取手机浮动车平均行程车速, 该高速公路全长 126 km, 手机浮动车数为 1000 台。根据建模样本量的需求, 由 SPSS 软件实现 Fisher 聚类将 2 个方向分成了 141 段, 部分路段如图 2 所示, 统计指标如下: 平均 1.75 km, 标准偏差 1.38 km, 最小值 0.8 km, 最大值 7 km, 求和 252 km, 计数 141 km。分段后手机浮动车的统计结果如图 2 所示。

可以看出, 图 2 的手机浮动车分布是倾向于离散的 Normal 分布的。经过 χ^2 检验, 图 2 在置信度为 0.95 的条件下服从标准正态分布。相关研究表明用 Poisson、NB、ZIP 或 ZINB 建立模型也是合适的。

经计算可得手机浮动车的路段平均时间和路段平均速度, 为了方便与模拟数据的对比, 采用 10 组时间数据, 所得数据如表 1 所示, 分别得到了 \bar{T} 和 \bar{v} 的 10 组数据。

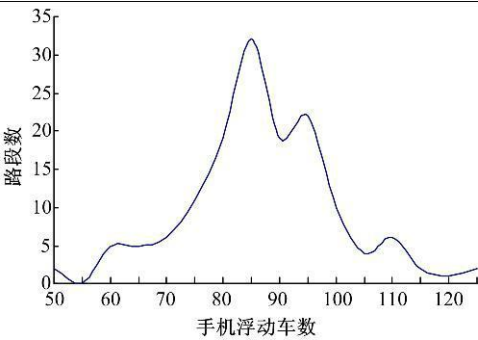


图 2 手机浮动车路段分布直方图

Fig. 2 Distribution histogram of phone floating car sections

表 1 实际值与模拟值的对比分析

Table 1 Comparative analysis of truth and simulation value

时间/s	实际值		模拟值		误差 1 /%	误差 2 /%
	平均行程时间/s	平均行程速度/(km·h ⁻¹)	平均行程时间/s	平均行程速度/(km·h ⁻¹)		
0~60	87	72.41	78.65	65.33	10.62	10.83
60~120	91.12	69.14	82.05	61.59	11.05	12.25
120~180	67.84	92.87	61.80	83.53	9.78	11.18
180~240	48.16	130.81	43.74	117.80	10.11	11.04
240~300	56.32	111.86	49.79	98.33	13.12	13.76
300~360	81.24	77.55	74.33	70.42	9.30	10.13
360~420	56.01	112.48	50.61	101.50	10.68	10.82
420~480	60.07	104.88	54.51	94.17	10.20	11.37
480~540	38.68	162.87	35.50	149.26	8.97	9.12
540~600	43.56	144.63	39.04	129.01	11.57	12.10
平均值	—	—	—	—	10.54	11.26

注: 误差 1 为平均行程时间模拟值与实际值的相对误差; 误差 2 为平均行程速度模拟值与实际值的相对误差。

在路段划分的基础上, 考虑了交通量、路段长度、是否是立交区(或收费站、服务区), 以及道路线形指标等影响因素, 建立了该高速公路路段平均速度的估计模型, 还可建立该高速公路的路段交通量、占有率、车头时距等交通参数估计模型^[9-10]。

通过模拟对比验证, 测试工具采用 VISSIM 4.2 模拟软件^[11], 模拟路网如图 3 所示, 通过定义路段交通量, 利用手机浮动车的 Fisher 有序聚类分析方法, 以单公里手机浮动车数的空间分布为聚类指标, 设置模拟参数道路长度 126 km, 双向 4 车道, 单车道手机浮动车生成量 800 veh/h, 饱和度 $\lambda=0.6$ (路况参数设置根据实际路段设置), 进行手机浮动车高速路段的仿真模拟, 仿真采集

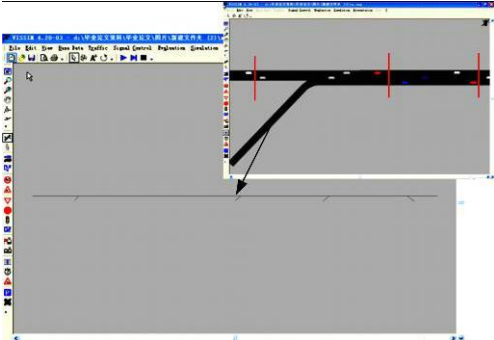


图 3 部分路段划分结果

Fig 3 VISSIM simulation of some section of highway

时间为 $t=600\text{ s}$, 分别得 $0\sim60$ 、 $60\sim120$ 、 $120\sim180$ 、 $180\sim240$ 、 $240\sim300$ 、 $300\sim360$ 、 $360\sim420$ 、 $420\sim480$ 、 $480\sim540$ 、 $540\sim600\text{ s}$ 的模拟数据, 见表 1, 模拟结果得 $t=600\text{ s}$ 的时间内的手机浮动车的几组 $\overline{T'}$ 和 $\overline{v'_{probe}}$ 的值, 模拟结果还可得手机浮动车路段平均时间和路段平均速度模型误差分别是 10.54% 、 11.26% (见表 1), 就目前手机浮动车提取参数数据模型已大大满足交通信息数据的采集。这里需要指出的是, 本模型仅适用于道路条件良好、交通流畅、手机浮动车正常运行, 与样本道路、交通条件及手机浮动车分布相似的高速公路, 对于其他线形指标变化较大的高速公路, 本方法还有待进一步的验证和比较, 也是下一步努力的方向。

4 结束语

采用基于手机浮动的高速公路路段行程速度提取技术中的路段划分方法, 把手机浮动车运用于 Fisher 聚类分析法中, 对手机浮动车高速公路路段长度划分进行了进一步的研究, 为手机浮动车高速公路路段行程速度的准确提取做了有效的准备工作。手机的高普及率, 使得交通参数提取方便廉价, 通过实例验证, 本文提取技术具有一定的应用价值。本研究目前对手机浮动车高速公路路段行程速度提取进行了试探性的分析, 因此, 该方法还有待于进一步地应用检验。除了继续用高速公路的样本验证该方法的适用性以外; 还正致力于和用其他提取方法建立的模型进行深入的比较分析, 进一步提高基于手机浮动车高速公路交通参数提取模型的精度。

参考文献:

[1] 马丽. 基于手机浮动车及聚类的高速公路实时交通

- 参数估计研究[D]. 重庆: 重庆大学, 2005: 21-25.
- Ma Li. Estimating real-time traffic variables on freeway based on mobile phone locating and clustering [D]. Chongqing: Chongqing University, 2005: 21-25.
- [2] 董升. 山东省高速公路事件自动检测系统研究[D]. 长春: 吉林大学交通学院, 2007: 1-9.
- Dong Sheng. Research on automatic incident detection of freeway in Shandong province [D]. Changchun: Jilin University, College of Transportation, 2007: 1-9.
- [3] 张可, 王笑京. 交通信息提取计算理论及其技术框架与发展策略[J]. 公路交通科技, 2007, 24(4): 134-139.
- Zhang Ke, Wang Xiao-jing. Transport information granular computing introduction technical architecture and development strategy [J]. Journal of Highway and Transportation Research and Development, 2007, 24(4): 134-139.
- [4] 钟连德, 陈永胜. 基于有序聚类分析的高速公路路段长度划分研究[J]. 武汉理工大学学报, 2008, 32(1): 43-46.
- Zhong Lian-de, Chen Yong-sheng. Research on section division of freeway with ordinal clustering method [J]. Journal of Wuhan University of Technology, 2008, 32(1): 43-46.
- [5] 王佳, 陈舟宇. 路段变长划分的动态聚类方法[J]. 山西科技, 2005(6): 105-106.
- Wang Jia, Chen Zhou-yu. Method of dynamic cluster for pavements sectionalized with units in different lengths [J]. Shanxi Science and Technology, 2005 (6): 105-106.
- [6] 蒋红妍. 高速公路路面养护管理系统研究[D]. 西安: 长安大学, 2006: 33-48.
- Jiang Hong-yan. Study on management system of expressway pavements maintenance [D]. Xi'an: Chang'an University, 2006: 33-48.
- [7] 张晓华. 高速公路路面管理系统[D]. 成都: 西南交通大学, 2002: 7-10.
- Zhang Xiao-hua. Freeway pavement management system [D]. Chengdu: Southwest Jiaotong University, 2002: 7-10.
- [8] Marcel Westerman, Remco Litjens, Jean-Paul Linnartz. Integration of probe vehicle and induction loop data-estimation of travel times and automatic incident detection [R]. Berkeley: University of California, 1996: 113-117.
- [9] 元海英, 杨孝宽, 张金喜, 等. 高速公路自由流状态下对运行车速模型的研究[J]. 交通标准化, 2005(5): 73-75.
- Yuan Hai-ying, Yang Xiao-kuan, Zhang Jin-xi, et al. Research on operating speed in free flow of freeway [J]. Communications Standardization, 2005 (5): 73-75.
- [10] 张和生. 利用GPS数据估计路段平均行程时间方法研究[J]. 吉林大学学报: 工学版, 2006, 36(5): 533-537.
- Zhang He-sheng. Research on estimation approaches of average link travel time using GPS data [J]. Journal of Jilin University (Engineering and Technology Edition), 2006, 36(5): 533-537.
- [11] PTV. VISSIM User Manual Version 4.20 [M/OL]. [2006-10-26]. <http://www.english.ptv.de/cgi-bin/traffic/traf-vissim.pl>