

基于浮动车数据的交通状态估计精度仿真评价

唐克双, 梅雨, 李克平

(同济大学 道路与交通工程教育部重点实验室, 上海 201804)

摘要: 针对大规模复杂路网条件下基于浮动车数据的交通状态估计精度评价, 开发了基于交通仿真软件 VISSIM 的实时仿真分析方法. 以上海市陆家嘴地区的微观仿真路网为例, 通过 30 组仿真实验分析了浮动车比例和数据采样频率对路网覆盖率和平均行程车速估计精度的影响. 结果表明: 随浮动车比例和采样频率的增加, 平均行程车速估计精度与路网覆盖率逐渐提高, 当浮动车比例为 8% 和采样频率为 $1/10 \text{ s}^{-1}$ 时达到最优.

关键词: 复杂路网; 浮动车; 交通状态估计; 交通仿真

中图分类号: U491.1

文献标志码: A

A Simulation-based Evaluation of Traffic State Estimation Accuracy by Using Floating Car Data in Complex Road Networks

TANG Keshuang, MEI Yu, LI Keping

(Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China)

Abstract: This study developed a simulation-based approach to comprehensively evaluate the joint effects of penetration rate and uploading frequency of floating car data on traffic state estimation accuracy for the large-scale complex road networks. A case study based on the simulation model of Lujiazui Region in Shanghai was conducted to demonstrate the applicability of the developed approach and explore the optimal penetration rate and uploading frequency in terms of travel speed estimation accuracy and network coverage. Totally, 30 combinations of the two parameters were tested in the calibrated simulation model. Results show that the accuracy of travel speed estimation and network coverage rise as the penetration rate and uploading frequency increase. The optimal penetration rate and uploading frequency are found to be 8% and $1/10 \text{ s}^{-1}$ respectively.

Key words: large-scale complex road networks; floating car data; traffic state estimation; traffic simulation

基于浮动车(floating car data)全球定位系统(global positioning system, GPS)数据的城市道路状态估计以其成本小、精度高、覆盖率广等特点, 已经广泛应用于我国许多城市, 如北京、上海、深圳、杭州等. 从浮动车数据采集的角度看, 浮动车占全部车辆的比例和浮动车数据的采样频率对交通状态估计精度有显著影响, 而且更高的精度伴随着更高的成本. 因此, 有必要对浮动车比例和采样频率两个参数进行精度评价, 以期在较低成本的情况下达到较高的估计精度.

利用浮动车比例和采样频率估计道路行程时间在国内均有丰富的研究, 主要集中在两个方面: 基于经典数理分析的方法和基于仿真分析的方法. 经典数理分析方面, Quiroga 等^[1]、Srinivasan 等^[2]和 Ygnace 等^[3]分别采用基于路段平均速度估计精度、路段行程时间估计可靠性和浮动车蜂窝定位技术, 分析了最佳浮动车比例. 辛飞飞等^[4]和涂智等^[5]通过实证数据分析了浮动车样本容量与路网覆盖能力以及交通信息更新周期之间的关系. 但是, 针对浮动车采样频率的研究相对较少. Gössel 等^[6]采用频谱分析方法研究了保证一定估计精度下的浮动车采样频率范围. 张存保等^[7]运用信号处理中的采样原理确定了浮动车在城市道路中的最优采样频率. 朱鲤等^[8]研究了浮动车采样频率较低时的行程车速估计算法. 李清泉等^[9]研究了浮动车采样频率低时的地图匹配算法.

仿真分析方面, Li 等^[10]使用 VISSIM 软件构建了简化的干线仿真网络, 分析了浮动车数量对路段行程时间的影响. Cheu 等^[11]采用 Integration 软件进行仿真实验, 选取新加坡某市的路网作为研究对象, 分析了浮动车比例从 3% 变化到 18% 时相应路段平均车速的估计精度. 林思等^[12]提出了基于仿真实验的最佳样本容量确定方法, 结果表明当浮动车

比例为 3%~5% 时,准确度可达到 95% 以上。

现有研究中,针对数理分析的方法缺少交通工程学的验证,且分析实际交通数据时受数据源限制,很难做到多层次全面分析浮动车比例和采样频率。文献中基于仿真的分析方法通常只是针对几个交叉口或是小范围的区域路网建立仿真模型,单一分析浮动车比例或者采样频率,得到的仿真精度较高。但是,现有研究缺少针对大规模复杂路网条件下,同时考虑浮动车比例和采样频率两个变量对交通状态估计精度影响的分析。因此,当路网规模增大、交通状况变复杂时,最佳的浮动车比例和采样频率还需要进一步研究。

本文针对复杂路网条件下基于浮动车 GPS 数据的交通状态估计精度的评价问题,开发了基于交通仿真软件 VISSIM 的仿真分析平台。通过构建和标定上海市陆家嘴地区的大规模微观仿真路网,系统分析了浮动车比例和数据采样频率两个参数对于浮动车路网覆盖率和平均行程车速相对误差的影响,并给出了两个参数的最优取值建议。

1 基于浮动车数据的交通状态估计方法

基于浮动车数据的交通状态判别流程主要分为数据采集、数据预处理、地图匹配、路段平均车速计算和交通状态的判别五个部分^[13-14]。

本文使用仿真分析方法,首先模拟实际浮动车数据采集的方式,根据浮动车比例和浮动车采样频率两个参数采集仿真路网中所有浮动车的数据。通常实际浮动车数据采集会有一定的设备和传输误差,这些数据误差会对地图匹配算法造成较大影响。因此,基于浮动车数据的交通状态估计一般需要先通过数据预处理来过滤掉异常数据和填补缺失数据。但是由于本文采用的是基于仿真的研究方法,在仿真中车辆的位置和速度信息可以精确、实时地获取,不存在数据传输和通信误差。另外,影响实际数据误差的因素比较复杂,不同浮动车系统的差异较大,因此本文省略了数据预处理的步骤。仿真中可以准确地知道浮动车所在路段,不需要任何地图匹配算法即可完成地图匹配。

为与实际中常用的方法统一,计算路段平均车速采用空间平均法,即在固定的计算时间间隔输出一组各路段的平均车速^[14]。交通状态判别可根据输出的路段平均车速完成。这里固定的计算时间间隔是指实时交通信息系统中,浮动车行程时间计算系

统以固定的时间间隔向上级系统输出。考虑到实时交通信息系统的计算速度需求,时间间隔不能太短;同时太长的时间间隔也不能正确反映实时交通状态。因此,结合现行正在应用的浮动车信息系统与文献中推荐的取值^[5],本文以 5 min 作为计算时间间隔。

目前,在国内外常用的基于浮动车的道路交通状态估计方法中,空间平均法因其计算方法简单且精度较高的特点而被广泛应用^[1-2,4,10,13,15-16]。空间平均法的本质是通过计算每辆浮动车在路段的行程时间来推算整个路段在一定计算时间间隔内的平均行程时间,目前我国大部分城市的浮动车应用系统中均采用这种方法。图 1 展示了空间平均法的计算过程。五角星表示一辆浮动车 k 的行驶轨迹,点 A, B, C 和 D 为浮动车在路段上最靠近节点的数据点,经过 A, B, C 和 D 的时刻分别为 t_1, t_2, t_3 和 t_4 。假设车辆在相邻两个数据点之间匀速行驶,那么浮动车 k 经过节点 e 和 f 的时刻分别为

$$t_e = t_1 + \frac{l_{Ae}}{l_{AB}}(t_2 - t_1), t_f = t_3 + \frac{l_{Cf}}{l_{CD}}(t_4 - t_3) \quad (1)$$

式中: l_{Ae} 表示点 A 到节点 e 的距离, l_{AB} 表示点 A 到点 B 的距离, l_{Cf} 表示点 C 到节点 f 的距离, l_{CD} 表示点 C 到点 D 的距离。

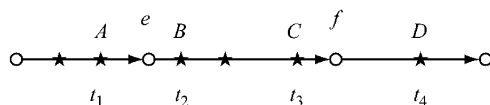


图 1 路段行程时间计算示意图

Fig. 1 Illustration of travel time calculation

对浮动车 k ,其在以节点 e 和 f 为端点的路段 j 的行程时间

$$t_{jk} = (t_f - t_e) \quad (2)$$

通常,路段平均车速以固定的计算时间间隔输出,对第 i 个计算时间间隔内所有经过路段 j 的浮动车行程时间求算术平均,即可得路段的平均行程时间 t_{ij} 及路段的平均车速 v_{ij} 。

$$t_{ij} = \frac{\sum_k t_{ijk}}{m} \quad (3)$$

$$v_{ij} = \frac{l_j}{t_{ij}} = \frac{ml_j}{\sum_k t_{ijk}} \quad (4)$$

式中: m 为经过路段 j 的浮动车总数, l_j 为路段 j 的长度。

2 精度评价方法

2.1 精度评价指标

为了全面地评价浮动车比例和采样频率对交通状态估计精度的影响,首先需要选取评价精度的相应指标。

第一个指标用以分析交通状态估计的路网全面性。考虑到路网中浮动车相对于所有车辆仅占较小比例,在较短的计算时间间隔内并不能覆盖整个路网,故定义浮动车覆盖率为在固定计算时间间隔内所有浮动车所经过的路段占路网中所有路段的比例。这里路段定义为一个方向上两个交叉口之间的道路,覆盖率计算公式为

$$H = \frac{P}{S} \times 100\% \quad (5)$$

式中: H 为路网覆盖率, P 为浮动车所经过的路段数, S 为路网中路段总数。

第二个指标用以分析浮动车技术在每个路段的估计精度。在仿真过程中,通过路网浮动车数据,根据式(4)可计算出各路段的平均车速 v_{ij} 。同时,仿真软件可以输出仿真中相应的准确车速 u_{ij} 。对每个路段,定义路段平均车速的相对误差

$$\epsilon_{ij} = \frac{|v_{ij} - u_{ij}|}{u_{ij}} \quad (6)$$

2.2 仿真路网建模

本文研究复杂路网情况下交通状态估计的精度,仿真路网的建模要同时满足复杂性和真实性要求。构建复杂路网模型并进行参数标定是十分困难的,本文同时使用宏观交通仿真软件 VISUM 和微观仿真软件 VISSIM 对上海市陆家嘴地区进行建模和模型标定^[17]。

首先,根据陆家嘴地区路网地图在 VISUM 软件中构建路网模型,使用路网中主要交叉口各个进口道的悉尼自适应交通控制系统(SCATS)线圈数据进行路网的起迄点(OD)反推,并通过实际调查的干道交通流量进行流量校准,以保证仿真路网的交通状态尽可能地与实际状态吻合。其中,SCATS 线圈数据采用 2011 年 11 月 10 日晚高峰 16:30—18:30 的数据。图 2 显示了建模的仿真路网,图中圆圈为采集 SCATS 线圈数据的交叉口。

其次,把 VISUM 软件建模的路网导入到 VISSIM 软件中,进行下一步的标定。通过实际数据调查,在 VISSIM 软件中完成了交通组成、路段车速分布和交叉口信号灯配时方案等参数的标定。其中,



图2 上海市陆家嘴地区仿真路网

Fig.2 Simulation network of Lujiazui Region in Shanghai

交通组成与路段车速分布为分析路网的交通输入路段的交通流参数,对各个输入交通流分别设定对应的交通组成与路段车速分布。

最后,运行 VISSIM 的仿真路网,通过对比核查交叉口各进口道的流量与 SCATS 线圈数据流量,确保了仿真路网流量与实际路网流量的一致性。

2.3 仿真实验设计

建模的路网是一个包含了主干路、次干路、支路、隧道和非十字交叉口的复杂路网。其中,仿真实验中标记的路段为 160 个,浮动车比例分为五组,分别为 2%,4%,6%,8%和 10%,浮动车采样频率分为六组,分别为 1/60,1/45,1/30,1/20,1/10,1/5 s^{-1} ,总计为 30 个实验组。仿真时间为 4 200 s,其中包括 600 s 的预热时间和 3 600 s 的仿真计算时间。道路状态的计算时间间隔为 5 min。

为获得每一组实验方案的浮动车路网覆盖率和各计算时间间隔内各路段的平均车速,需要对路网进行特殊处理。如图 3 所示,VISUM 路网导入到 VISSIM 软件中时已经对各个路段进行了编号,以唯一识别每个路段。用 VISSIM 的 com 接口编程,以实验要求的 30 组参数共进行 30 次仿真实验。每组实验中,按照各组的浮动车比例值随机从仿真车流中选取车辆标记为浮动车,再根据采样频率记录下每辆浮动车的位置信息,这一过程可以模拟现实中浮动车 GPS 数据的采集流程。数据采集以每个计算时间间隔为节点,计算时间间隔段内所有样本车辆的平均行程时间,从而根据式(3)和(4)计算出各路段的平均车速,这一过程也完全模拟了现实中浮动车

GPS 数据的计算过程. 实验组 n 的仿真结果为车速矩阵 V_n 和 U_n .

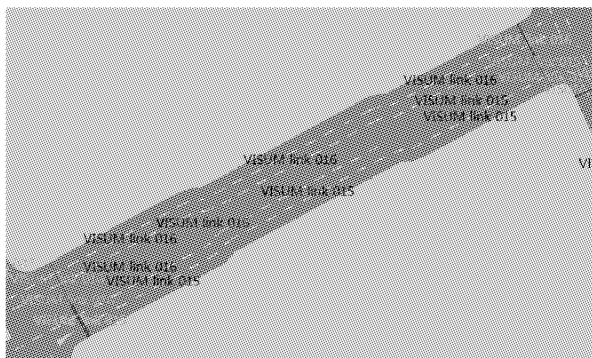


图 3 仿真路网中的路段

Fig.3 Road sections in simulation network

为获得仿真中各路段的准确平均车速,可用相同的编号规则以及计算时间间隔,通过 VISSIM 软件中的行程时间检测器记录下仿真中各路段的实际行程时间,再把路段平均行程时间转换为路段的平均车速

$$u_{ijn} = \frac{l_j}{T_{ijn}} \quad (7)$$

式中: u_{ijn} 为第 n 组实验中第 i 计算时间间隔内路段 j 的实际平均车速, T_{ijn} 为第 n 组实验中第 i 计算时间间隔内路段 j 的准确平均行程时间. 由此可得同一实验组的真实车速矩阵 U_n (u_{ijn} 为 U_n 的元素).

com 编程实现了 30 组实验的数据采集、计算和存储. 路网覆盖率 H_n 、路段平均车速 V_n 和 U_n ($n=1, 2, \dots, 30$) 的仿真结果分类保存在数据库中, 仿真结束后使用 Matlab 读取数据库进行精度评价.

图 4 为整个仿真分析系统的框架图.

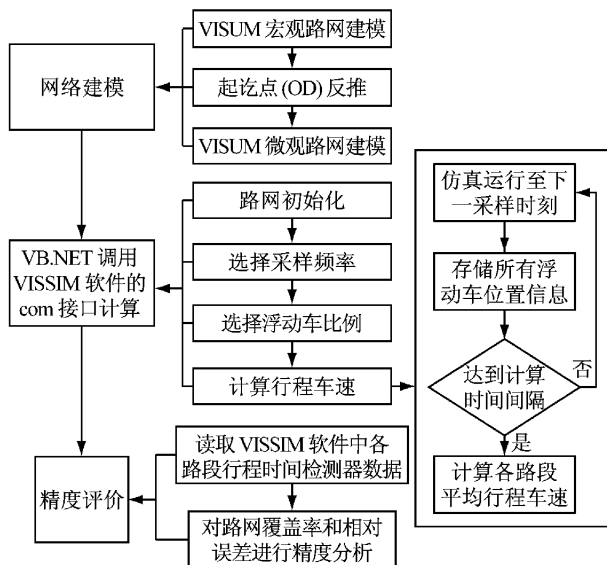


图 4 仿真分析系统框架图

Fig.4 Analysis framework of traffic simulation

3 结果分析

3.1 路网覆盖率分析

不同的浮动车比例和采样频率对应的路网覆盖率不同,越高的浮动车比例或越快的采样频率对应的路网覆盖率越大. 对不同浮动车比例和采样频率的路网覆盖率统计见表 1.

表 1 路网覆盖率统计表

Tab.1 Statistics of coverage ratios					%
采样频率/ s ⁻¹	浮动车比例/%				
	10	8	6	4	2
1/60	54	49	45	38	26
1/45	57	57	48	43	33
1/30	60	63	58	53	38
1/20	64	67	62	56	42
1/10	67	71	65	60	45
1/5	68	71	66	61	46

由表 1 可分析,采样频率对路网覆盖率有一定的影响. 当采样频率在 $1/60 s^{-1}$ 到 $1/10 s^{-1}$ 增加时,覆盖率缓慢增加;而当采样频率在 $1/10 s^{-1}$ 到 $1/5 s^{-1}$ 增加时,路网覆盖率没有明显变化. 可推断,采样频率在 $1/10 s^{-1}$ 时,路网覆盖率达到峰值,采样频率高于 $1/10 s^{-1}$ 时,路网覆盖率基本不再增加. 这个结果表明,浮动车采样频率在 $1/10 s^{-1}$ 时最优,更高的采样频率不能提高交通状态估计的精度,这与张存保等^[7]的研究结果相符. 路网覆盖率随浮动车比例增大而增大,但增速逐渐减慢. 浮动车比例增加到 8% 后,路网覆盖率变化趋势放缓. 这个结果表明,浮动车比例越大,交通状态估计的精度越高,在考虑经济成本的情况下,采用低于 8% 的浮动车比例较为合理.

3.2 误差分析

通过仿真采集的车速矩阵 V_n 和 U_n ($n=1, 2, \dots, 30$),可用式(6)计算出其相对误差矩阵 E_n . 绘制相对误差矩阵的箱图,如图 5 所示.

从图 5 可以看出,相对误差主要集中在 0~0.6 之间,有少数大偏差的异常值. 较大偏差的原因是部分路段较为拥挤,从而导致路段平均车速较低,使得相对误差的分母较小,导致即便较低的误差值也会造成较大的相对误差. 因此,取相对误差的平均值作为评价指标会造成较大偏差,故本文选取中位相对误差作为平均行程车速的精度评价指标,如表 2 所示.

由表 2 可分析,采样频率对中位相对误差影响较大. 当采样频率在 $1/60 s^{-1}$ 到 $1/10 s^{-1}$ 增加时,中位相对误差随之降低;而当采样频率在 $1/10 s^{-1}$ 到

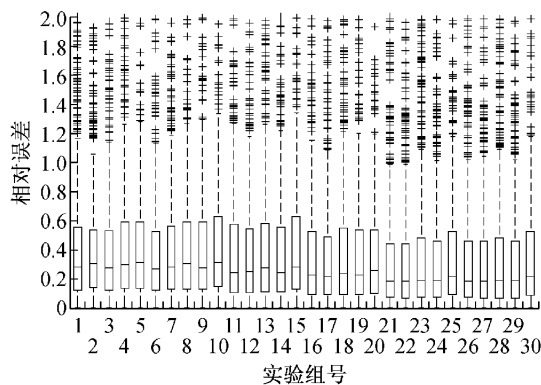


图5 相对误差的箱图

Fig.5 Boxplot of relative errors

表2 中位相对误差统计表

Tab.2 Statistics of median relative errors

采样频率/ s^{-1}	浮动车比例/%				
	10	8	6	4	2
1/60	0.28	0.31	0.28	0.30	0.31
1/45	0.27	0.28	0.31	0.28	0.31
1/30	0.25	0.25	0.28	0.25	0.28
1/20	0.23	0.21	0.24	0.23	0.26
1/10	0.19	0.19	0.19	0.19	0.22
1/5	0.18	0.18	0.19	0.19	0.22

1/5 s^{-1} 增加时,中位相对误差没有明显变化.可推断,采样频率高于 1/10 s^{-1} 时,中位相对误差基本不再降低.随着浮动车比例增加,中位相对误差逐渐降低,不过下降幅度不大.总体来看,采样频率从 1/60 s^{-1} 到 1/5 s^{-1} 变化时,整个路网平均行程车速的精度在 0.7 到 0.8 左右.

4 结论

(1) 浮动车比例越大,交通状态估计的精度越高.在考虑经济成本的情况下,采用低于 8% 的浮动车比例较为合理.

(2) 浮动车采样频率在 1/10 s^{-1} 时最优,更高的采样频率不能提高交通状态估计的精度.

(3) 基于浮动车的交通状态估计方法在对大规模复杂路网进行路段状态估计时,浮动车比例从 2% 到 10% 变化,采样频率从 1/60 s^{-1} 到 1/5 s^{-1} 变化时,整个路网平均行程车速的精度在 0.7 到 0.8 左右.

同时,本文也存在一些不足.仿真分析没有考虑实际中地图匹配的误差,也没有充分考虑出租车驾驶行为的特殊性,如行驶路径上的偏好等.上述问题将是未来研究的重点.

参考文献:

[1] Quiroga C A, Bullock D. Determination of sample sizes for

travel time studies [J]. ITE Journal on the Web, 1998, 68 (8): 92.

[2] Srinivasan K K, Jovanis P P. Determination of number of probe vehicles required for reliable travel time measurement in urban network [J]. Transportation Research Record, 1996, 1537(1): 15.

[3] Ygnace J L, Drane C. Cellular telecommunication and transportation convergence: a case study of a research conducted in California and in France on cellular positioning techniques and transportation issues [C]//Intelligent Transportation Systems, 2001. Oakland: IEEE, 2001: 16-22.

[4] 辛飞飞,陈小鸿,林航飞.基于样本容量的浮动车数据路网覆盖能力研究[J].公路交通科技,2009,26(8):140.

XIN Feifei, CHEN Xiaohong, LIN Hangfei. Study on detecting capability of floating car data in road network based on sample size [J]. Journal of Traffic and Transportation Engineering, 2009, 26(8): 140.

[5] 涂智,李昊,姚琛,等.基于最小浮动车样本数量的道路覆盖率与交通信息更新周期研究[J].中国铁道科学,2006,27(5):127.

TU Zhi, LI Hao, YAO Chen, *et al.* Study on the route coverage and the update cycle of transportation information based on the minimum samples of floating car [J]. China Railway Science, 2006, 27(5): 127.

[6] Gössel F, Michler E, Wrase B. Spectral analysis of floating car data[J]. Advances in Radio Science, 2003, 1(8): 139.

[7] 张存保,杨晓光,严新平.浮动车采样周期优化方法研究[J].交通运输系统工程与信息,2007,7(3):100.

ZHANG Cunbao, YANG Xiaoguang, YAN Xinping. Method for floating cars sampling cycle optimization [J]. Journal of Transportation Systems Engineering and Information Technology, 2007, 7(3): 100.

[8] 朱鲤,杨东援.基于低采样频率浮动车的行程车速信息实时采集技术[J].交通运输系统工程与信息,2008,8(4):42.

ZHU Li, YANG Dongyuan. Dynamic travel speed collection technology based on low frequency FCD [J]. Journal of Transportation Systems Engineering and Information Technology, 2008, 8(4): 42.

[9] 李清泉,尹建忠,贺奋琴.面向道路网的GPS浮动车覆盖率模型研究[J].武汉大学学报:信息科学版,2009,34(6):715.

LI Qingquan, YIN Jianzhong, HE Fenqin. A coverage rate model of GPS floating car for road networks [J]. Geomatics and Information Science of Wuhan University, 2009, 34(6): 715.

[10] Li Q, Miwa T, Morikawa T. Preliminary analysis on link travel time for probe-based estimation method by microscopic simulation [J]. International Journal of ITS Research, 2008, 6 (1): 21.

[11] Cheu L R, Xie C, Lee D H. Probe vehicle population and sample size for arterial speed estimation [J]. Computer-Aided Civil and Infrastructure Engineering, 2002, 17(1): 53.

[12] 林思,徐建闽,熊文华.基于仿真试验的GPS浮动车最佳样本容量确定方法研究[J].交通与计算机,2008,26(4):72.

LIN Si, XU Jianmin, XIONG Wenhua. Determination method of sample capacity of GPS floating vehicle based on simulation [J]. Computer and Communications, 2008, 26(4): 72.