

文章编号:1671-1637(2014)06-0100-10

## 路段行程时间估计的浮动车数据挖掘方法

李慧兵<sup>1</sup>, 杨晓光<sup>2</sup>, 罗莉华<sup>1</sup>

(1. 上海海事大学 交通运输学院, 上海 201306; 2. 同济大学 交通运输工程学院, 上海 201804)

**摘要:**基于浮动车数据,提出一种信号配时信息缺失下的路段行程时间估计方法,由交叉口范围动态划分、路段影响范围划分、浮动车数据提取与路段行程时间估计4个模块组成,每个模块的实现均需借助于前一模块的输出。根据交叉口信号控制下的车辆行驶状态,在交叉口范围动态划分与路段影响范围划分2个模块中,利用密度法将单元路段划分为不同区域。根据路段行程时间估计原理,利用浮动车数据提取模块过滤掉受信号控制影响较大的浮动车数据,提取路段行程时间估计的目标数据。利用路段行程时间估计模块挖掘历史浮动车数据,根据浮动车目标数据点存在区域的不同,将浮动车数据分为3类,并对不同类型数据采取相应的断面通过时刻估计方法,建立基于不同数据条件下的行程时间估计模型。利用VISSIM软件对路段行程时间估计方法进行仿真验证,并与直接法和间接法进行对比分析。分析结果表明:对于粗粒度浮动车数据,路段行程时间估计方法的平均绝对误差和平均相对误差分别为12 s和8.67%,优于传统的直接法与间接法。

**关键词:**智能交通系统;路段行程时间估计;浮动车数据;信号配时;粗粒度数据

**中图分类号:**U491.1

**文献标志码:**A

## Mining method of floating car data based on link travel time estimation

LI Hui-bing<sup>1</sup>, YANG Xiao-guang<sup>2</sup>, LUO Li-hua<sup>1</sup>

(1. School of Transport and Communications, Shanghai Maritime University, Shanghai 201306, China;

2. School of Transportation Engineering, Tongji University, Shanghai 201804, China)

**Abstract:** Based on floating car data, a link travel time estimation method without signal timing data was proposed. The method consisted of four modules, which were intersection boundary dynamic partition module, link influence range partition module, floating car data extraction module, and link travel time estimation module, and the implementation of each module relied greatly on the output of previous one. According to vehicle travel state under the influence of signal control, link unit was divided into different segments by using density method in intersection boundary dynamic partition module and link influence range partition module. According to link travel time estimation mechanism, floating car data that were seriously affected by signal control were filtered off in floating car data extraction module, so the target floating car data could be obtained. Historical floating car data were excavated in link travel time estimation module, and floating car data were divided into 3 types according to different existed regions of target data. Corresponding section travel time estimation methods were used for different types of data, and corresponding section travel time estimation models were established. Link travel time estimation method was simulated and verified by using software VISSIM, and its result was compared with the results of direct and indirect methods. Analysis result indicates that for coarse-

收稿日期:2014-06-28

基金项目:国家自然科学基金项目(61304203);上海市科研计划项目(12ZR1444800)

作者简介:李慧兵(1983-),男,山西吕梁人,上海海事大学讲师,工学博士,从事智能交通系统研究。

grained floating car data, the average absolute error and average relative error of link travel time estimation method are 12 s and 8.67% respectively, so it performs better than traditional direct and indirect methods. 3 tabs, 22 figs, 20 refs.

**Key words:** intelligent transportation system; link travel time estimation; floating car data; signal timing; coarse-grained data

**Author resume:** LI Hui-bing(1983-), male, lecturer, PhD, +86-21-38282347, hbli@shmtu.edu.cn.

## 0 引言

间断流路段行程时间预测是构建动态车辆诱导系统、动态交通信息服务系统与信号协调控制系统的重要基础<sup>[1]</sup>,预测结果的准确性与间断流行程时间估计值的准确度有直接关系,因此,采用有效的方法准确预测间断流行程时间已经成为国内外学者研究的热点。在国外,基于浮动车数据(Floating Car Data, FCD)的间断流路段行程时间预测是一个重要的研究热点<sup>[2-3]</sup>。目前,中国各大城市均已建立了ITS平台,并配置了大量基于出租车的浮动车设备,获取的浮动车数据量十分庞大,且该数据在间断流路网上覆盖范围大,不涉及隐私,因此,利用丰富的出租车GPS数据估计路段行程时间是可行的。

面向间断流的传统浮动车路段行程时间估计方法主要分为2类:理论解析法和实际工程模型。

理论解析法相对复杂,使用的模型包括神经网络模型<sup>[4-6]</sup>、路段聚合模型<sup>[7]</sup>、卡尔曼滤波模型<sup>[8-10]</sup>、虚拟浮动车模型<sup>[11]</sup>、极大似然模型<sup>[12]</sup>和数据融合模型等<sup>[13-14]</sup>。Zheng等借助于复杂的神经网络模型估计路段行程时间<sup>[4-6]</sup>,其中Bi等还需要把交叉口的信号信息作为一个输入参数才能对路段行程时间进行有效估计;Han等提出一个路段聚合模型,以此减小车辆在相邻路段行驶时产生的较大行程时间误差,模型的输入参数为交叉口信号配时信息、车辆到达时刻等<sup>[7]</sup>;Jiang等利用卡尔曼滤波模型计算路段行程时间<sup>[8-10]</sup>,模型复杂,且输入参数多;Liu等提出一种基于交通事件的虚拟浮动车模型<sup>[11]</sup>,模型提出大量假设,需要大量在实际工程中无法获取的额外数据,其中交叉口信号配时信息也是其输入参数之一;Jenelius等提出把粗粒度数据下的路径行程时间分解为路段行程时间,并利用极大似然法估计路段行程时间,模型的输入参数为道路特征参数(限制速度、道路等级等)、道路情况(周期内日变化交通量、当前季节和天气等)、交叉口信号配时信息等<sup>[12]</sup>;李嘉等采用小波神经网络建立数据融合模型来预测路段行程时间,采用遗传算法优化小波神经

网络的初始网络参数,模型输入参数为浮动车行程时间、微波行程时间、浮动车样本量和交通流量<sup>[13]</sup>;方路平等融合线圈数据、信号配时信息与浮动车数据来估计路段行程时间,利用线圈流量数据计算周期内的动态排队长度,根据最大排队长度确定受交叉口影响的区域,剔除该区域内的浮动车数据,计算浮动车在区域外的行程时间与区域内的延误,最终得到路段平均行程时间<sup>[14]</sup>。理论解析法假设条件多,建模复杂,且大部分并不适用于实际工程。此外,理论解析法大多需要借助于交叉口信号配时数据,由于目前中国城市交通管理部门之间交通数据共享不畅,很多重要数据(特别是交叉口信号配时数据)经常无法获取,在很大程度上制约了方法的实施。

实地工程模型相对简单,主要包括直接法与间接法。直接法利用路段边界两侧GPS定位点的位置坐标,采用内插的方式估算车辆经过路段边界的时刻,进而计算单车路段行程时间。间接法利用GPS定位点的瞬时速度序列估算车辆经过路段的平均速度,进而计算单车路段行程时间。姜桂艳等比较了直接法与间接法2种路段行程时间估计方法,分析表明随着GPS数据采样间隔的不断增大,利用直接法与间接法获得的路段行程时间估计误差均呈总体增加趋势,当采样间隔小于7s时,直接法和间接法的估计误差非常接近,当采样间隔不小于7s时,直接法的估计效果明显优于间接法<sup>[15]</sup>。

直接法假设相邻GPS定位点之间的车辆匀速行驶,其实施条件为2个相邻的GPS定位点分别位于路段边界的两侧<sup>[15-17]</sup>。间接法也假设相邻GPS定位点之间的车辆匀速行驶,其实施条件为路段边界附近存在GPS定位点,且GPS数据采样间隔较小,数据序列能够覆盖绝大部分路段<sup>[15]</sup>。然而在很多情况下,车辆在两定位点(特别是当它们位于信号交叉口两侧时)之间的行驶速度是不断变化的。

实地工程模型简单易行,假设条件少,不需要借助于信号配时数据,适用于工程实践,然而其相邻GPS定位点之间车辆是匀速行驶的假设使此类方

法只适用于浮动车数据采样间隔较小的情况<sup>[15]</sup>。在中国的大、中城市中,实际浮动车数据的平均采样间隔一般为 30 s(粗粒度数据),此时,行程时间估计值的准确度无法得到保证。

基于实际工程模型,本文提出一种信号配时信息缺失下的浮动车路段行程时间估计方法,方法由交叉口范围动态划分、路段影响范围动态划分、浮动车数据提取、路段行程时间估计 4 个模块组成,每个模块的实现均需借助于前一模块的输出。

交叉口范围动态划分与路段影响范围动态划分 2 个模块根据交叉口信号控制下的车辆行驶状态将单元路段划分为不同的区域;浮动车数据提取模块过滤掉受信号控制影响较大的浮动车数据,提取路段行程时间估计的目标数据;路段行程时间估计模块对落入不同区域的浮动车目标数据点采取不同的断面通过时刻估计方法,挖掘历史浮动车数据,根据不同的数据条件建立相应的行程时间估计模型,以此估计路段行程时间。

本文根据交叉口信号控制下的车辆行驶状态将路段划分为不同的区域,认为路段边界两侧(不同区域内)相邻 GPS 定位点之间的车辆不是匀速行驶的,适用于粗粒度数据下的路段行程时间估计。本文方法过滤掉受信号控制影响较大的浮动车数据,因此,路段行程时间估计不需要借助于精细的信号配时信息。同时方法的实施条件为路段边界某一侧的区域内存在浮动车目标数据点,因此,模型的数据需求显著减少,适用范围显著变广。

本文方法在很大程度上改进了传统实际工程模型,可有效解决在交叉口信号配时信息缺失下,基于粗粒度浮动车数据的路段行程时间估计问题。

## 1 交叉口范围动态划分

交叉口延误是间断流路段行程时间的一个重要组成部分,然而,在利用浮动车数据进行间断流路段行程时间估计时,往往受限于对车辆进入交叉口速度发生突变位置的判断,浮动车延误估计常出现较大误差,因此,路段与交叉口范围的划分将直接影响浮动车行程时间计算的稳定性和可靠性。

本文首先分析平面交叉口的车流受阻特性,然后挖掘江苏省南京市某路段的历史实测浮动车数据,实现基于密度法<sup>[18-19]</sup>的动态路段划分,以 15 min 为划分间隔。

### 1.1 平面交叉口的车流受阻特性

车流到达交叉口具有随机性且存在冲突与让行

延误,以信息控制交叉口为例,在 1 个信号周期内,红灯总会对部分车流有阻滞作用,停车线前的车辆排队也因此产生,随着绿灯的开启,排队车辆开始消散。按进入交叉口车流的运行轨迹,车辆在停车线前的状态可分为完全停车、不完全停车与不停车<sup>[19]</sup>。对于通过交叉口的车辆来说,其速度变化一般表现为匀速、减速-停车-加速、减速-不停车-加速。

假如通过路段动态划分能确定车流在交叉口的开始减速位置和加速完成位置,将可以判断车流在 2 个交叉口间路段不同区域的行驶状态,对计算车辆在交叉口的控制延误与车辆的行程时间具有重要意义。基于 GPS 车辆定位技术的浮动车数据,能为此提供及时的交通流信息。

### 1.2 交叉口范围动态划分

选取南京市北京西路(主干道)上从宁海路至上海路的直行路段为研究对象,时间范围为该路段上 2010 年 7 月 11 日至 2010 年 7 月 14 日每天 7:00~9:00(本文研究时段均为上午),对浮动车数据进行预处理,主要包括 2 个步骤<sup>[20]</sup>。

Step 1:数据清洗。填充缺失数据,光滑噪声并识别离群点,纠正数据中的不一致。

Step 2:数据转换。将数据转换或统一为适合于挖掘的形式。

以预处理后的浮动车数据为研究数据,基于密度法研究交叉口范围动态划分<sup>[18-19]</sup>。假设在 15 min 内,间断流上的交通状态保持稳定,车辆在下游交叉口的排队和延误特征相似,将研究路段上 7:00~9:00 的交通状态分解为 8 个子状态,不同子状态的属性见表 1。挖掘研究路段上 2010 年 7 月 11 日到 2010 年 7 月 14 日的浮动车数据,分析每个子状态下的路段交通特性。

表 1 子状态属性

Tab. 1 Sub-state properties

子状态编号	浮动车数据
1	7:00~7:15 内的浮动车数据
2	7:15~7:30 内的浮动车数据
3	7:30~7:45 内的浮动车数据
4	7:45~8:00 内的浮动车数据
5	8:00~8:15 内的浮动车数据
6	8:15~8:30 内的浮动车数据
7	8:30~8:45 内的浮动车数据
8	8:45~9:00 内的浮动车数据

### 1.3 试验结果分析

利用以上方法分别对 8 个子状态下的路段交通

状态进行模拟,结果分别见图 1~8,浮动车数据点(负数无意义)在路段空间范围内服从类似的分布,其分布的空间范围大致为以下 3 个区域。

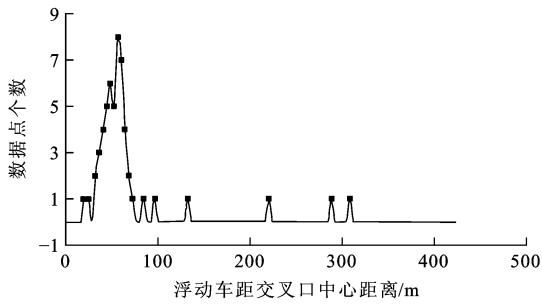


图 1 子状态 1 的路段划分

Fig. 1 Link division of sub-state 1

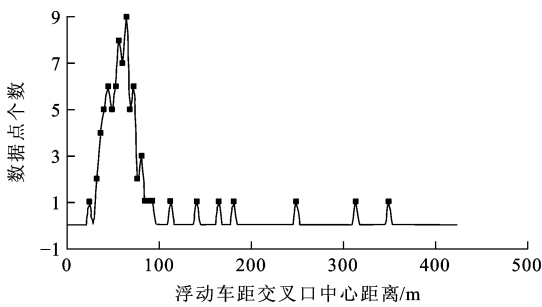


图 2 子状态 2 的路段划分

Fig. 2 Link division of sub-state 1

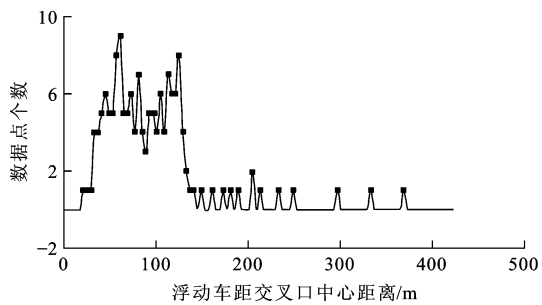


图 3 子状态 3 的路段划分

Fig. 3 Link division of sub-state 3

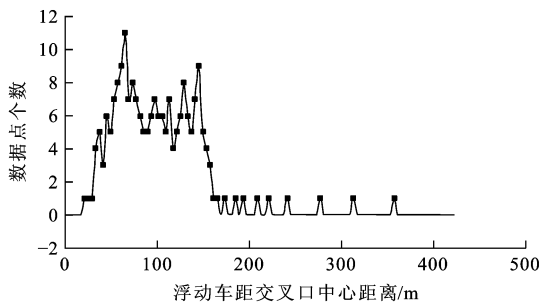


图 4 子状态 4 的路段划分

Fig. 4 Link division of sub-state 4

(1)区域 1。区域 1 为由交叉口中心点开始到交叉口停车线处结束的区域,浮动车数据点几乎为 0。在区域 1 内,车辆以较快的速度运行,由于数据

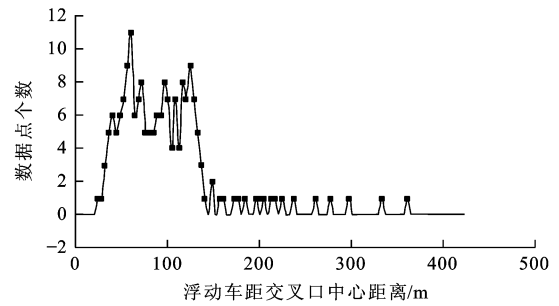


图 5 子状态 5 的路段划分

Fig. 5 Link division of sub-state 5

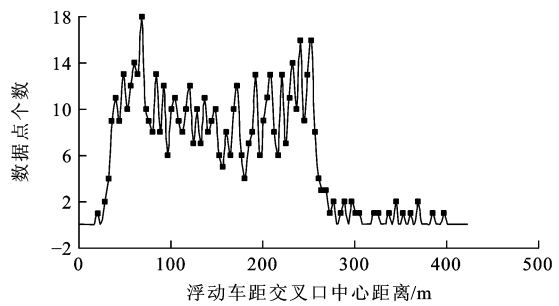


图 6 子状态 6 的路段划分

Fig. 6 Link division of sub-state 6

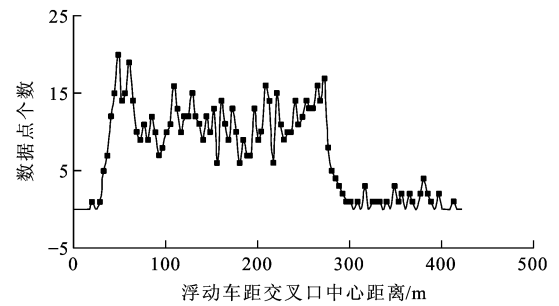


图 7 子状态 7 的路段划分

Fig. 7 Link division of sub-state 7

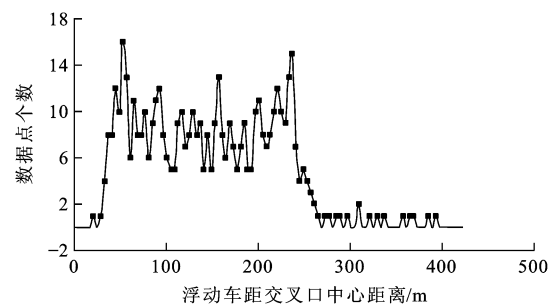


图 8 子状态 8 的路段划分

Fig. 8 Link division of sub-state 8

的采样间隔较大(约为 30 s),因此,区域 1 内几乎不存在浮动车数据点,数据点的密度很小。

(2)区域 2。区域 2 为由交叉口停车线处开始到某位置处结束的区域,浮动车数据点从交叉口停车线处开始快速增多,其数量在该区域内维持在一一定的较大范围内。在区域 2 内,车辆受到交叉口信

号控制延误的影响较大,因此,浮动车数据点较多,数据点的密度较大。

(3)区域 3。区域 3 为由交叉口处开始到上游交叉口出口处结束的区域,浮动车数据点从该位置处快速减少,其数量在该区域内维持在一定的较小范围内。在区域 3 内,车辆受到交叉口信号控制延误的影响较小,运行速度较快,由于浮动车数据的采样间隔较大,因此,浮动车数据点较少,浮动车数据点的密度较小。

将浮动车数据点快速减少的位置定义为交叉口平均最大排队位置(车流开始减速的平均位置),该位置处于区域 2 的末尾或区域 3 的开始。图 9 为 8 个子状态下的平均最大排队长度,可以看出,在 7:00~9:00 间,交叉口的平均最大排队长度有逐渐增大的趋势,并且从 8:15 开始交叉口平均最大排队长度急剧增加,子状态 6~8 的平均最大排队长度一直维持在一个较高的水平上。这个结论基本符合实际的间断流路段交通状态特征。

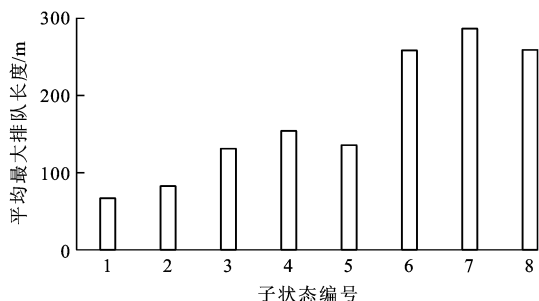


图 9 平均最大排队长度

Fig. 9 Average maximum queuing lengths

## 2 路段影响范围动态划分

在交叉口范围动态划分的基础上,对路段影响范围进行动态划分。研究的单元路段起始于上游交叉口出口处距离出口断面 15 m 的位置,终止于下游交叉口出口处距离出口断面 15 m 的位置,见图 10,其中,下游交叉口的延误被包括在该单元路段的行程时间内。

### 2.1 浮动车路段行程时间估计原理

浮动车路段行程时间为

$$T = t_1 - t_0 \quad (1)$$

式中: $T$  为某浮动车通过单元路段的行程时间; $t_0$  为浮动车通过起始断面的时刻; $t_1$  为浮动车通过终止断面的时刻。

由式(1)可以看出,浮动车的行程时间取决于该车分别通过单元路段终止断面和起始断面的时

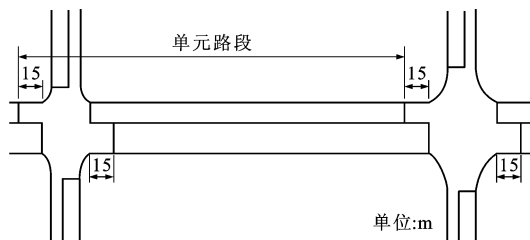


图 10 单元路段

Fig. 10 Link unit

刻,因此,只有获取浮动车通过单元路段前后 2 个断面的时刻,才能对该车的路段行程时间进行有效估计。

### 2.2 路段影响范围划分

根据交叉口范围动态划分方法,把两交叉口间的区段动态划分为排队区域和畅通区域,见图 11。排队区域由下游交叉口停车线处开始,在下游交叉口平均最大排队位置处结束。畅通区域由下游交叉口平均最大排队位置处开始,在上游交叉口停车线处结束。排队区域为该路段交通拥堵的区域,而畅通区域属于该路段交通相对畅通的区域。

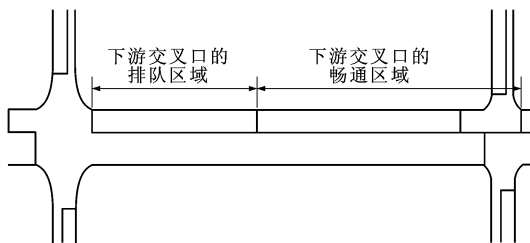


图 11 交叉口间的区段划分

Fig. 11 Segment division between two intersections

图 12 为路段影响范围划分。为估计浮动车经过起始断面的时刻,需要借助于畅通区域 1 与排队区域 1 内的浮动车数据点。为估计浮动车经过终止断面的时刻,需要借助于畅通区域 2 与排队区域 2 内的浮动车数据点。

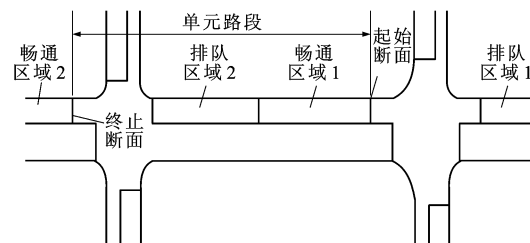


图 12 路段影响范围划分

Fig. 12 Division of link influence scope

## 3 浮动车数据提取

预处理浮动车原始数据后,数据量依然很大。因为并不是所有浮动车数据均可对路段行程时间进

行有效估计,因此,本文将进一步过滤浮动车数据,提取其中有效的浮动车数据作为目标数据。以单辆浮动车数据作为数据统计单元,对浮动车数据进行提取,分为以下4个步骤。

Step 1:将每个浮动车数据点撒布到对应单元路段的畅通区域1、排队区域1、畅通区域2、排队区域2等4个区域内。在每个区域内设置1个关键断面,排队区域的关键断面为交叉口停车线处,畅通区域的关键断面为距离交叉口出口15 m处(该路段的起始断面或终止断面处)。

Step 2:选取各区域内距离其关键断面最近的1个数据点作为该区域的研究数据。

Step 3:将畅通区域2与排队区域2合并为1个大区域,将畅通区域1与排队区域1合并为1个大区域。如果某浮动车数据点仅落在其中1个大区域内,则该浮动车数据点不能用于路段行程时间估计,因为仅借助于1个大区域内的浮动车数据点只能估算到该浮动车通过其中1个路段断面的时刻,而无法估计到该浮动车通过另1个路段断面的时刻。

Step 4:过滤掉不适合进行行程时间估计的浮动车数据,将剩余数据作为路段行程时间估计的目标数据。

## 4 路段行程时间估计

挖掘浮动车目标数据,进一步对浮动车路段行程时间估计方法进行研究。

### 4.1 不同区域内浮动车数据特征

分别分析排队区域和畅通区域内浮动车的行驶特征,进而得出这2种区域内的浮动车数据特征。

#### 4.1.1 排队区域的浮动车数据特征

由于排队区域的浮动车受交叉口信号控制延误的影响较大,因此,首先分析排队区域内目标数据特征与交叉口信号配时方案之间的关系。

研究路段为南京市北京西路上从上海路至宁海路的直行路段(东西直行),研究路段区域为北京西路与宁海路交叉口的排队区域,研究目标交叉口为北京西路与宁海路交叉口。由于南京市采用的是SCATS(Sydney Coordinated Adaptive Traffic System)信号控制系统,因此,目标交叉口的信号配时方案属于方案选择式。表2为目标交叉口在研究时段(6:30~12:00)的信号配时方案。

表2中目标交叉口的信号周期为160 s,其东西直行相位的绿灯时长为20 s,红灯时长为137 s。由

表2 目标交叉口的信息配时方案

Tab. 2 Signal timing plan of target intersection

相位序列	控制方向	绿灯时长/s	黄灯时长/s	红灯时长/s
0	南北直	88	3	69
1	南北左	20	3	137
2	东西直	20	3	137
3	东西左	20	3	137

于南京市浮动车数据的采样间隔一般为30 s,可以得出:某浮动车在排队区域内发送了最后一条浮动车数据(即排队区域内的目标数据)之后,不超过30 s即可顺利通过目标交叉口停车线。统计、分析浮动车目标数据后得出,排队区域内的目标数据点仅会落在以下2个时间范围内。

(1)在绿灯时间间隔内,浮动车可通过交叉口停车线,其数据点速度一般不为0,该车基本没有受到交叉口信号控制的影响,以较高的速度通过交叉口停车线,全程无延误。

(2)在红灯末尾或黄灯间隔内,浮动车数据点速度一般为0,在较短的时间内(不大于30 s)可通过交叉口停车线,红灯信号很快结束,车辆受到交叉口信号控制的影响。行驶行为表现为先停车,当绿灯间隔开始后,以较快的速度通过交叉口停车线。

由于南京市中心区主要交叉口的信号周期一般均大于100 s,并且对任一相位来说,其红灯时间也均大于30 s,因此,上述结论对南京市中心区的浮动车数据均适用。

#### 4.1.2 畅通区域的浮动车数据特征

畅通区域的浮动车基本不受信号控制延误的影响,以较快的速度通过畅通区域,区域内的浮动车目标数据点速度一般不为0。对粗粒度的浮动车数据来说,点速度为0的数据代表该浮动车处于停车状态,在下个采样时刻之前,很难准确获取其行驶状态,因此,依靠点速度为0的浮动车数据无法准确估计该浮动车的断面通过时刻。

### 4.2 浮动车目标数据分类

对浮动车目标数据进行分类,根据不同的数据类型建立相应的路段断面通过时刻估计方法。删除所有点速度为0的浮动车目标数据。由于大区域为路段断面通过时刻估计的一个重要区域范围,以大区域为研究区段,统计、分析剩余数据,并将其划分为以下3种类型。

#### 4.2.1 数据类型1

数据类型1为仅存在于相邻排队区域内的数据,见图13,仅依靠该排队区域内的浮动车目标数

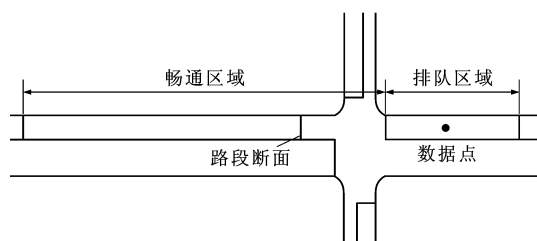


图 13 数据类型 1

Fig. 13 Data type 1

据估计路段断面通过时刻。

#### 4.2.2 数据类型 2

数据类型 2 为仅存在于相邻畅通区域内的数据,见图 14,仅依靠该畅通区域内的浮动车目标数据估计路段断面通过时刻。

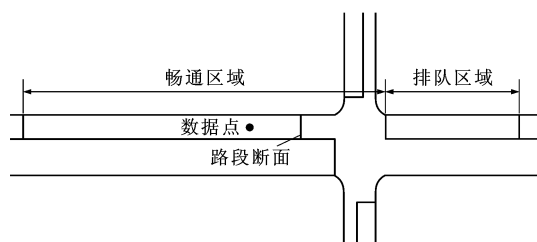


图 14 数据类型 2

Fig. 14 Data type 2

#### 4.2.3 数据类型 3

数据类型 3 为相邻排队区域与畅通区域内的数据同时存在,见图 15,融合这 2 个区域内的浮动车目标数据估计路段断面通过时刻。

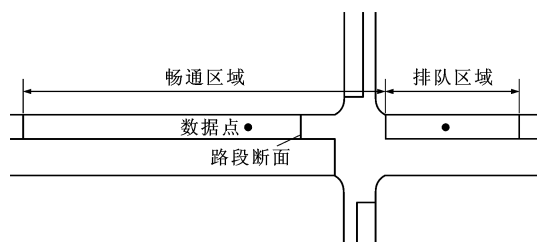


图 15 数据类型 3

Fig. 15 Data type 3

### 4.3 路段断面通过时刻估计方法

对 3 种数据类型下的路段断面通过时刻进行估计。

#### 4.3.1 数据类型 1

数据类型 1 的路段断面通过时刻估计方法由参数标定和估计模型 2 部分组成。

##### (1) 参数标定

估计路段断面通过时刻之前,需要对其中的模型参数进行标定。模型参数包括排队区域的区段范围参数与区段平均速度参数。

在排队区域区段划分时,以排队区域起始断面

为基准位置,每隔 4 m(接近于小汽车平均长度)设立一个断面,见图 16,则在该排队区域内可设立  $n_0$  个断面,  $n_0$  为

$$n_0 = \text{fix}\left(\frac{L_0}{4}\right) \quad (2)$$

式中:  $\text{fix}(\cdot)$  为 MATLAB 中的向零取整函数;  $L_0$  为排队区域的长度。

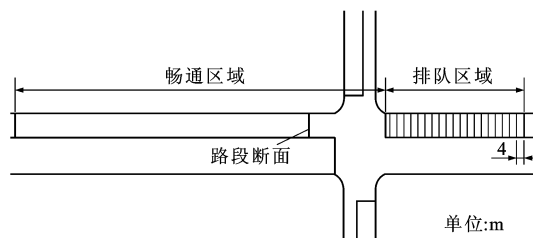


图 16 排队区域的区段划分

Fig. 16 Segment division of queuing region

将排队区域内每个断面与下游路段断面之间的区段视为 1 个小区段,可生成  $n_0$  个小区段,其中,断面 1 与下游路段断面之间的区段为区段 1,断面 2 与下游路段断面之间的区段为区段 2,以此类推,断面  $n_0$  与下游路段断面之间的区段为区段  $n_0$ 。

在区段平均速度标定时,利用历史浮动车数据(预处理后的数据,非目标数据),计算每个小区段内的浮动车平均点速度。区段  $k$  ( $k=1, 2, \dots, n_0$ ) 的平均速度  $\bar{V}_k$  为

$$\bar{V}_k = \frac{1}{l} \sum_{j=1}^l V_j \quad (3)$$

式中:  $V_j$  为区段  $k$  内的历史浮动车点  $j$  的速度。

##### (2) 估计模型

假如排队区域内的某浮动车目标数据点位于排队区域的断面  $k, k+1$  之间,且该数据点距离断面  $k$  较近,则认为该浮动车从当前位置以  $\bar{V}_k$  的速度匀速通过下游路段断面,此时该浮动车的路段断面通过时刻  $T_0$  为

$$T_0 = t_2 + \frac{L}{\bar{V}_k} \quad (4)$$

式中:  $t_2$  为浮动车数据点的采集时刻;  $L$  为浮动车数据点距下游路段断面的距离。

同理,如果该数据点距离断面  $k+1$  较近,则认为该浮动车从当前位置以  $\bar{V}_{k+1}$  的速度匀速通过下游路段断面,此时该浮动车的路段断面通过时刻为

$$T_0 = t_2 + \frac{L}{\bar{V}_{k+1}} \quad (5)$$

#### 4.3.2 数据类型 2

根据断面通过时刻估计机理的不同,将畅通区

域划分为3个大区段,分别研究基于不同大区段内数据的断面通过时刻估计方法。

#### (1) 参数标定

参数标定由畅通区域区段划分与区段平均速度标定两部分组成。

畅通区域区段划分中,将畅通区域划分为3个大区段,分别为畅通区段1、2、3,其中:区段1起始于上游交叉口停车线处,终止于上游交叉口出口处,该区段长度为 $L_1$ ;区段2起始于区段1的终止断面,终止于距离其起始断面30 m处的位置,该区段长度 $L_2$ 为30 m(假设在这30 m的区段内,车辆以较快的速度近匀速运行,速度波动较小,全程无阻滞);区段3起始于区段2的终止断面,终止于畅通区域的终止断面,该区段的长度为 $L_3$ 。畅通区域的区段划分见图17,畅通区域的总长度 $L_4$ 为

$$L_4 = L_1 + L_3 + 30 \quad (6)$$

以某畅通区域的起始断面(上游交叉口停车线处)为基准,每隔4 m设立1个断面,在区段1内可设立 $n_1$ 个断面。以区段3的起始断面为基准,每隔4 m设立1个断面,在区段3内可设立 $n_3$ 个断面。区段1与区段3的断面数 $n_1$ 、 $n_3$ 的确定方法见式(2)。区段平均速度标定方法与4.3.1节相同。

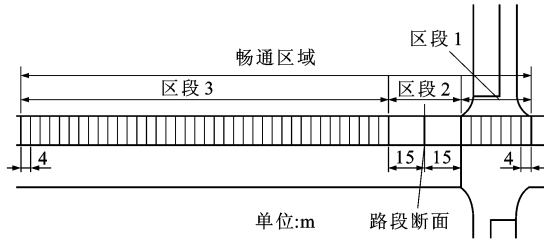


图17 畅通区域的区段划分

Fig. 17 Segment division of unblocked region

#### (2) 估计模型

基于区段1内数据的断面通过时刻估计模型与4.3.1节相同。

基于区段2内数据的断面通过时刻估计模型中,假如某浮动车目标数据点位于区段2内,且该数据点位于路段断面的下游,则该浮动车的路段断面通过时刻为

$$T_0 = t_2 - \frac{L}{V} \quad (7)$$

式中: $V$ 为该浮动车数据点的点速度。

假如该数据点位于路段断面的上游,则该浮动车的路段断面通过时刻为

$$T_0 = t_2 + \frac{L}{V} \quad (8)$$

基于区段3内数据的断面通过时刻估计模型中,假如畅通区域区段3内的某浮动车数据点位于该区段的断面 $i, i+1$ 间,且该数据点距离断面 $k$ 较近,则认为该浮动车从上游路段断面以 $\bar{V}_i$ 的速度匀速行驶至当前位置,则该浮动车的路段断面通过时刻为

$$T_0 = t_2 - \frac{L}{\bar{V}_i} \quad (9)$$

如果该数据点距离断面 $i+1$ 较近,则认为该浮动车从上游路段断面以 $\bar{V}_{i+1}$ 的速度匀速行驶至当前位置,则该浮动车的路段断面通过时刻为

$$T_0 = t_2 - \frac{L}{\bar{V}_{i+1}} \quad (10)$$

#### 4.3.3 数据类型3

根据浮动车目标数据点在相邻排队区域和畅通区域内所处位置的不同,把数据类型3分为3种情景:情景1中1个目标数据点位于相邻排队区域内,另1个目标数据点位于相邻畅通区域的区段1内;情景2中1个目标数据点位于相邻排队区域内,另1个目标数据点位于相邻畅通区域的区段2内;情景3中1个目标数据点位于相邻排队区域内,另1个目标数据点位于相邻畅通区域的区段3内。3种情景分别见图18~20。3种情景的断面通过时刻估计模型如下。

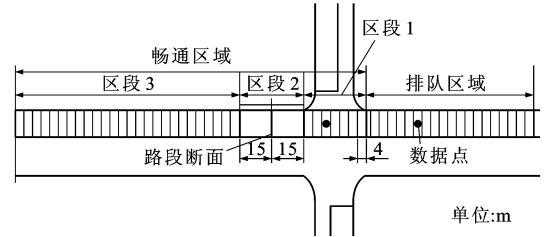


图18 情景1

Fig. 18 Scenario 1

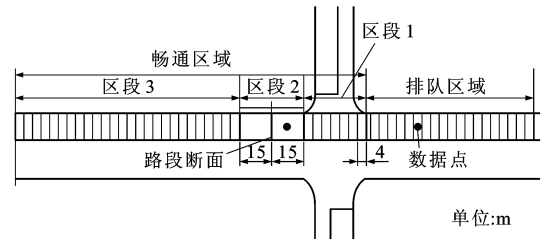


图19 情景2

Fig. 19 Scenario 2

#### (1) 情景1的断面通过时刻估计模型

在情景1中,路段断面通过时刻估计需要借助于畅通区域区段1内的浮动车目标数据点,具体模



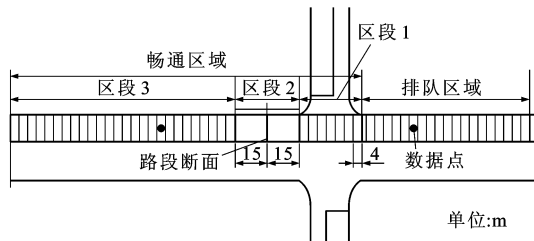


图 20 情景 3

Fig. 20 Scenario 3

型见 4.3.2。

(2)情景 2 的断面通过时刻估计模型

在情景 2 中,路段断面通过时刻估计需要借助于畅通区域区段 2 内的浮动车目标数据点,具体模型见 4.3.2。

(3)情景 3 的断面通过时刻估计模型

确定排队区域内的目标数据点所隶属的小区段,如果该小区段为排队区域的区段  $k$ ,则计算该区段的点速度平均值  $\bar{V}_k$ 。确定畅通区域区段 3 内的目标数据点所隶属的小区段,如果该小区段为区段 3 内的小区段  $i$ ,则计算该小区段的点速度平均值  $\bar{V}_i$ 。情景 3 的断面通过时刻估计模型为

$$T_1 = t_3 - t_4 = T_2 + T_3 \quad (11)$$

$$\alpha \bar{V}_i T_2 = L_5 \quad (12)$$

$$\alpha \bar{V}_k T_3 = L_6 \quad (13)$$

$$t = t_4 + T_3 = t_3 - T_2 \quad (14)$$

式中: $t$  为某浮动车的路段断面通过时刻; $t_3$  为畅通区域内数据点的时刻; $t_4$  为排队区域内数据点的时刻; $T_1$  为畅通区域内数据点与排队区域内数据点之间的时间差,即该浮动车通过两数据点之间区段的行程时间; $T_2$  为浮动车从路段断面到畅通区域内数据点位置所行驶的时间; $T_3$  为浮动车从排队区域内数据点位置到路段断面所行驶的时间; $\alpha$  为修正系数; $L_5$  为畅通区域内数据点距路段断面的距离; $L_6$  为排队区域内数据点距路段断面的距离。

## 5 实例分析

本次试验的研究对象为南京市某路网,采用同济大学车路协同试验车进行实地跟车,试验车可采集到时间间隔为 1 s 的数据,这些数据被实时存入数据库,以备后续分析处理。研究时段为早高峰时段 7:00~9:00。2010 年 7 月 12 日的试验路线方向依次为北京西路(虎踞路与太平北路之间)、太平北路(北京西路与中山东路之间)、中山东路(太平北路

和虎踞路之间)、虎踞路(中山东路与北京西路之间)、中山路(北京西路与中山东路之间)和广州路(虎踞路与太平北路之间),2010 年 7 月 13 日的试验路线方向与 2010 年 7 月 12 日的试验路段方向相反。

为模拟南京市路网浮动车数据采集的真实情况,对试验车数据进行筛选处理,处理方法为:将试验车数据按采集时间排序,每隔 30 s 从试验车数据中提取一组数据,最终提取数据即为以 30 s 为时间间隔的粗粒度数据,然后分别利用本文方法、直接法与间接法对浮动车路段行程时间进行估计。

选择 16 个路段上总路段行程时间(实测行程时间、本文方法行程时间估计值、直接法行程时间估计值和间接法行程时间估计值)作为试验比较数据,路段行程时间样本见表 3。使用绝对误差和相对误差来评价方法的有效性,分别表示为

表 3 路段行程时间

Tab. 3 Link travel times

s

数据样本 编号	实测行程 时间	本文方法 行程时间	直接法 行程时间	间接法 行程时间
1	136	145	108	158
2	111	99	128	152
3	511	529	545	462
4	152	158	167	192
5	284	265	277	261
6	93	86	123	46
7	128	159	151	208
8	101	108	114	82
9	316	298	334	285
10	179	192	162	86
11	196	199	177	233
12	54	47	145	47
13	121	110	163	144
14	98	111	89	133
15	119	105	156	80
16	116	124	104	187

$$a = |A - E| \quad (15)$$

$$r = 100 \frac{|A - E|}{A} \quad (16)$$

式中: $A$ 、 $E$  分别为路段的实测行程时间和行程时间估计值; $a$ 、 $r$  分别为该路段行程时间估计值的绝对误差和相对误差。

图 21、22 分别为不同方法的绝对误差与相对误差,可知:本文方法行程时间估计值的平均绝对误差和平均相对误差分别为 12 s 和 8.67%;直接法行程时间估计值的平均绝对误差和平均相对误差分别为 22 s 和 17.85%;间接法行程时间估计值的平均绝对误差和平均相对误差分别为 41 s 和 29.80%。可见,本文方法计算精度明显优于直接法和间接法。

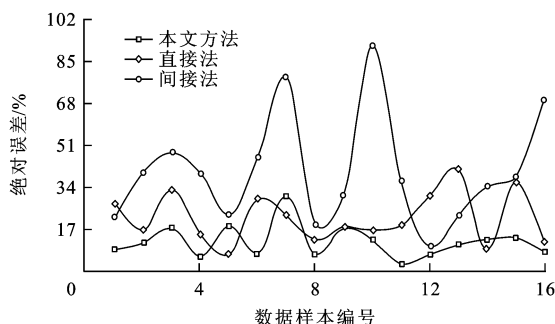


图 21 绝对误差比较

Fig. 21 Comparison of absolute errors

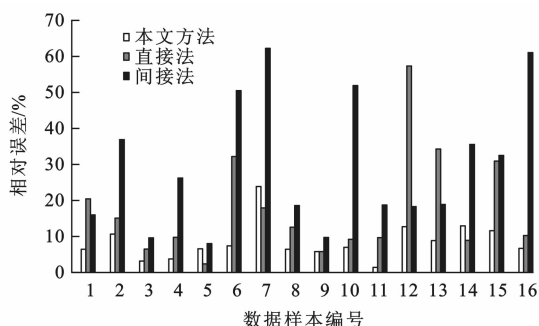


图 22 相对误差比较

Fig. 22 Comparison of relative errors

## 6 结 语

(1)本文方法可有效解决粗粒度(30 s 左右)浮动车数据的路段行程时间估计问题,且不需要借助于交叉口信号配时信息。

(2)根据交叉口信号控制下的车辆行驶状态将路段划分为不同区域,认为路段边界两侧(不同区域内)的 GPS 定位点之间车辆不是匀速行驶的,适用于粗粒度数据下的路段行程时间估计;过滤掉受信号控制影响较大的浮动车数据,因此,路段行程时间估计不需要借助于精细的信号配时信息。

(3)本文方法的实施条件为路段边界某一侧的区域内存在浮动车目标数据点,因此,数据需求显著减少,适用范围显著变广。

(4)本文方法行程时间估计值的平均绝对误差和平均相对误差分别为 12 s 和 8.67%,明显优于传

统实际工程模型。

## 参考文献:

### References :

- [1] WANG Jian-qiang, NIU Hui-min. Graded-information feed-back strategy in two-route systems under ATIS[J]. Journal of Traffic and Transportation Engineering: English Edition, 2014, 1(2): 138-145.
- [2] WU C H, HO J M, LEE D T. Travel time prediction with support vector regression[J]. IEEE Transactions on Intelligent Transportation Systems, 2004, 5(4): 276-281.
- [3] SUN Lu, YANG Jun, MAHMASSANI H. Travel time estimation based on piecewise truncated quadratic speed trajectory[J]. Transportation Research Part A: Policy and Practice, 2008, 42(1): 173-186.
- [4] ZHENG Fang-fang, ZUYLEN H. Comparison of urban link travel time estimation models based on probe vehicle data[C]// MAO Bao-hua, TIAN Zong-zhong, HUANG Hai-jun, et al. Seventh International Conference on Traffic and Transportation Studies. Kunming: ASCE, 2010: 615-626.
- [5] ZHENG Fang-fang, ZUYLEN H. Urban link travel time estimation based on sparse probe vehicle data[J]. Transportation Research Part C: Emerging Technologies, 2013, 31(1): 145-157.
- [6] BI Song, WANG Zhi-jian, HAN Cun-wu, et al. Estimation of left-turning travel time at traffic intersection[J]. The Journal of China Universities of Posts and Telecommunications, 2013, 20(S1): 10-14.
- [7] HAN Shu, LIN Hang-fei, CHEN Xiao-hong. The application of link aggregating algorithm in travel time estimate on signalized arterial network[C]//PENG Qi-yuan, WANG K C P, QIU Yan-jun, et al. Second International Conference on Transportation Engineering. Chengdu: ASCE, 2009: 2315-2321.
- [8] JIANG Zhou, ZHANG Cun-bao, XIA Yin-xia. Travel time prediction model for urban road network based on multi-source data[J]. Procedia-Social and Behavioral Sciences, 2014, 138: 811-818.
- [9] 李进燕,朱征宇,刘琳,等.基于简化路网模型的卡尔曼滤波多步行程时间预测方法[J].系统工程理论与实践,2013, 33(5):1289-1297.  
LI Jin-yan, ZHU Zheng-yu, LIU Lin, et al. Multi-step Kalman filtering travel time estimation method based on simplified road network model[J]. Systems Engineering—Theory and Practice, 2013, 33(5): 1289-1297. (in Chinese)
- [10] ZHENG Fang-fang, WAN Yu, WU Ping-heng. Link travel-time prediction using extended exponential smoothing and Kalman filter in dynamic networks[C]//LIU Rong-fang, ZHANG Jin, GUAN Chang-qian. The Eighth International Conference of Chinese Logistics and Transportation Professionals. Chengdu: ASCE, 2008: 3753-3759.

(下转第 116 页)

- traffic responsive signal systems[R]. Austin: Texas A&M University, 1998.
- [18] SHELBY S G, BULLOCK D M, GETTMAN D. Transition methods in traffic signal control[J]. Transportation Research Record, 2006(1978): 130-140.
- [19] 栗红强, 陆化普, 刘 强. 城市交通线控多时段配时方案过渡算法研究[J]. 武汉理工大学学报: 交通科学与工程版, 2008, 32(5): 779-781, 852.
- LI Hong-qiang, LU Hua-pu, LIU Qiang. Research on transition algorithms for timing plans of arterial time in-day traffic control[J]. Journal of Wuhan University of Technology: Transportation Science and Engineering, 2008, 32(5): 779-781, 852. (in Chinese)
- [20] 卢 凯, 徐建闽, 郑淑鉴, 等. 交通信号协调控制方案过渡优化算法[J]. 交通运输工程学报, 2012, 12(6): 97-103.
- LU Kai, XU Jian-min, ZHENG Shu-jian, et al. Optimization algorithm of coordinated control scheme transition of traffic signal[J]. Journal of Traffic and Transportation Engineering, 2012, 12(6): 97-103. (in Chinese)
- [21] WANG Jian-qiang, NIU Hui-min. Graded-information feedback strategy in two-route systems under ATIS[J]. Journal of Traffic and Transportation Engineering: English Edition, 2014, 1(2): 138-145.
- [22] SARLE W S. Cubic clustering criterion[R]. Cary: SAS Institute Inc., 1983.
- [23] LIAO Fei-xiong, ARENTZE T, TIMMERMAN H. Multi-state supernetworks: recent progress and prospects[J]. Journal of Traffic and Transportation Engineering: English Edition, 2014, 1(1): 13-27.
- [24] 姚 佼. 基于车辆行驶数据的交叉口交通控制机制研究[D]. 上海: 同济大学, 2011.
- YAO Jiao. Study on vehicle probe data based traffic control mechanism at isolated intersection[D]. Shanghai: Tongji University, 2011. (in Chinese)
- 
- (上接第 109 页)
- [11] LIU H X, MA Wen-teng. A virtual vehicle probe model for time-dependent travel time estimation on signalized arterials[J]. Transportation Research Part C: Emerging Technologies, 2009, 17(1): 11-26.
- [12] JENELIUS E, KOUTSOPOULOS H N. Travel time estimation for urban road networks using low frequency probe vehicle data[J]. Transportation Research Part B: Methodological, 2013, 53: 64-81.
- [13] 李 嘉, 刘春华, 胡赛阳, 等. 基于交通数据融合技术的行程时间预测模型[J]. 湖南大学学报: 自然科学版, 2014, 41(1): 33-38.
- LI Jia, LIU Chun-hua, HU Sai-yang, et al. A travel time prediction model based on traffic data fusion technology[J]. Journal of Hunan University: Natural Sciences, 2014, 41(1): 33-38. (in Chinese)
- [14] 方路平, 陈仕骅, 赵飞帆. 基于小样本浮动车系统的平均行程时间估计[J]. 计算机仿真, 2012, 29(9): 367-370.
- FANG Lu-ping, CHEN Shi-xiao, ZHAO Fei-fan. Average link travel time estimation based on floating car of small sample size[J]. Computer Simulation, 2012, 29(9): 367-370. (in Chinese)
- [15] 姜桂艳, 常安德, 张 玮. 基于 GPS 浮动车的路段行程时间估计方法比较[J]. 吉林大学学报: 工学版, 2009, 39(增 2): 182-186.
- JIANG Gui-yan, CHANG An-de, ZHANG Wei. Comparison of link travel-time estimation methods based on GPS equipped floating car[J]. Journal of Jilin University: Engineering and Technology Edition, 2009, 39(S2): 182-186. (in Chinese)
- [16] 姜桂艳, 常安德, 张 玮, 等. 基于 GPS 浮动车的自然路段行程时间估计方法[J]. 公路, 2009(11): 87-90.
- JIANG Gui-yan, CHANG An-de, ZHANG Wei. Physical link travel-time estimation method based on GPS equipped floating car[J]. Highway, 2009(11): 87-90. (in Chinese)
- [17] 姜桂艳, 常安德, 吴超腾. 基于 GPS 浮动车的交通信息采集方法[J]. 吉林大学学报: 工学版, 2010, 40(4): 971-975.
- JIANG Gui-yan, CHANG An-de, WU Chao-teng. Traffic information collection method based on GPS equipped floating car[J]. Journal of Jilin University: Engineering and Technology Edition, 2010, 40(4): 971-975. (in Chinese)
- [18] 韩 舒, 林航飞, 辛飞飞. 浮动车采集系统中城市道路分段方法研究[J]. 交通与计算机, 2007, 25(5): 105-109.
- HAN Shu, LIN Hang-fei, XIN Fei-fei. Method of merging urban road subsections based on floating car data[J]. Computer and Communications, 2007, 25(5): 105-109. (in Chinese)
- [19] 熊英格, 徐卓立, 刘好德. 基于浮动车数据的交叉口范围动态划分方法[J]. 交通信息与安全, 2009, 27(5): 38-43.
- XIONG Ying-ge, XU Zhuo-li, LIU Hao-de. Methods of intersection dynamic subsections based on floating car data[J]. Journal of Transport Information and Safety, 2009, 27(5): 38-43. (in Chinese)
- [20] 胡小文. 基于探测车数据和固定检测器数据的路段行程时间估计[D]. 上海: 同济大学, 2008.
- HU Xiao-wen. Link travel time estimation based on probe data and fixed detector data[D]. Shanghai: Tongji University, 2008. (in Chinese)