

基于低频定点检测数据的交叉口交通状态估计

唐克双, 徐天祥, 董可然, 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: 针对我国中小城市数据现状, 提出了一种基于路中定点线圈低频(1/60Hz)检测数据的交叉口交通状态估计方法. 该方法基于仿真数据, 分析了不同环境变量组合条件下占有率、流量和交通状态的关系, 并提出了基于线性拟合的交通状态分界线建立方法; 又利用多元线性回归拟合出分界曲线各系数与环境变量的函数关系, 用其估计一般条件下的交通状态. 经过验证, 本方法仿真环境和实证环境下的平均估计准确率分别达到 80% 和 75% 以上, 且严重错误率均低于 2.1%.

关键词: 交通状态估计; 信号控制交叉口; 定点检测器; 低频检测数据

中图分类号: U491.1

文献标志码: A

Traffic State Estimation Based on Low Frequency Detection Data at Intersections

TANG Keshuang, XU Tianxiang, DONG Keran, LI Keping

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

Abstract: A traffic state identification method for intersections was proposed based on detection data with a low frequency of 1/60 Hz from detection on the mid of the urban roads which was applied for urban interrupted flow in medium and small cities of our country. At first, the relationship among occupation, volume and traffic state was analyzed under different parameters of circumstances based on simulation data and a method of curve fitting was proposed to build the boundaries of different traffic state. Then the functional relationship of coefficients of boundaries functions with environment variables was fit out which was later applied to general ones. The methods above were verified by simulation data with the identification rate of over 80%, and empirical data with the identification rate of over 75%, with severe mistake rates less than 2.1%.

Key words: traffic state estimation; signalized intersection;

point detector; low frequency detection data

城市道路交通状态的准确估计是交通控制与管理的基础. 随着我国交通信息化的发展, 各类交通检测技术与设备已被广泛应用于城市道路交通状态的估计与预测. 常用的城市道路交通检测技术可分为定点检测技术和移动检测技术. 其中, 定点检测器主要包括感应线圈、地磁、微波雷达、红外线、超声波等, 可以以一定的时间间隔(通常为 20~60s)采集断面流量、速度和占有率等数据; 移动检测器主要包括出租车浮动车、自动车辆识别(AVI)设备、车载或者手机导航等, 可通过车辆行驶轨迹和路径信息, 直接采集个体车辆的路径、行程、时间等. 国内外学者针对基于定点和移动检测数据的城市道路交通状态估计已开展了大量的研究, 采用的数据源几乎涉及上述所有种类的检测数据^[1].

在我国大部分的中小城市, 出租车浮动车的覆盖率小且上传频率较低(通常为 10~60 s), 难以用于城市道路交通状态的估计; AVI 的布设密度有限且受到车辆捕获率和匹配成功率的影响, 交通状态估计的可靠性较差; 车载导航数据虽然覆盖率和精度都可以满足要求, 但是一般需要政府向高德、百度等运营商购买服务, 经济成本较高. 在上述背景下, 布设密度较高的定点检测器则具有一定的差位优势.

定点检测器在我国城市道路实际工程应用中主要有 3 种典型的布设位置: ①下游检测器布设在路段下游接近停车线位置, 主要用于信号控制方案的计算和优化. ②路中检测器布设在路段中间位置, 主要用于路段交通状态参数的检测. ③上游检测器布设在路段上游靠近出口道的位置, 主要用于交叉口的溢出检测与信号控制. 下游检测器距离交叉口过

近,受到交叉口信号控制影响大,且由于其自身接入信号控制机的特点,存在通信传输和时间匹配等问题,在实际工程中往往只用于信号控制而不用于交通状态估计;上游检测器一般布设在短间距交叉口用于溢出检测,覆盖率低,难以用于交通状态估计;相比之下,路中定点检测器则具有覆盖率高、受交叉口信号控制影响小、稳定可靠等优点,在实际工程中经常被直接用于城市道路交通状态的估计。

然而,传统的基于路中检测器的交通状态估计方法绝大部分通过检测器获取的断面行程速度按照一定的阈值范围直接对应交通状态,而未充分考虑下游信号控制交叉口周期、绿信比、检测器布设位置等因素的影响.因此,本文利用我国中小城市普遍布设的路中检测器以及低频(1/60 Hz)的占有率和流量数据,构建占有率和流量二维坐标系下交通状态的划分方法,并利用仿真数据和实证数据进行验证.本文使用 5 min 集计流量与占有率,因为频率过低会失去应用价值而频率过高会造成随机性过大,该情况能够较好服务于大部分城市 5~15 min 的交通状态发布频率.同时,本文通过历史数据的不断积累而更新数据库并优化相关参数,提高估计准确率。

1 文献综述

道路交通状态的估计方法按照估计主体可以分为人工估计方法与自动估计方法(automatic congestion identification, ACI),而人工估计方法随着时代的技术进步因其诸多不足而基本被淘汰,ACI 方法从 20 世纪 60 年代末期开始逐步发展与应用. ACI 方法按照交通状态评价标准可分为基于阈值和基于模糊逻辑的交通状态估计方法,前者的基本原理是根据行程车速设定阈值以判断不同交通状态;后者针对交通状态具有的模糊性,建立交通顺畅程度的模糊推理系统,将连续交通参数模糊处理从而实现交通状态的估计^[1],文献^[2]对模糊逻辑应用于交通工程领域的前景进行了综述,文献^[3]分析了道路交通状态本身具有模糊性并将模糊逻辑引入交通状态估计。

传统的基于阈值的 ACI 经典算法包括 California 算法^[4]、指数平滑法^[5]、标准偏差法^[5]以及 McMaster 算法^[6]等,数据来源主要为线圈检测器,数理统计分析、决策树等是常用的数据处理方法^[7]. California 算法^[4]由美国加州运输部开发,属于双截面法,基于拥堵发生时上游检测界面占有率

上升而下游检测截面占有率下降的原理,用上下游占有率的差值相关量与阈值对比以判断是否发生拥堵;指数平滑法^[5]对原始交通数据进行平滑处理以去除短期交通干扰,再将其与预先设定的阈值进行比较从而判断是否发生拥堵;标准偏差法^[5]计算某时刻之前多个采样周期交通参数(流量或占有率)的算术平均值与该时刻实际交通参数数值的标准正态偏差,并与设定的阈值对比以判断是否发生偶发性拥堵. McMaster 算法^[6]基于突变理论^[8],即基于在连续流中交通状态从拥挤状态向非拥挤状态变化时流量和占有率平稳变化而速度突然变化的现象,用于判断当前交通状态是否拥堵及其原因.如图 1 所示,该算法通过大量历史数据在流量-占有率的二维坐标系下将交通状态划分为 4 个区域并以此根据不同的检测参数直接判断交通状态,其中区域 1~4 分别表示正常交通状态、下游出现偶发性拥堵、缓慢交通流、下游出现常发性拥堵.但上述传统 ACI 经典算法主要针对高速公路的交通事件检测,并不适用于城市交叉口的间断流情况。

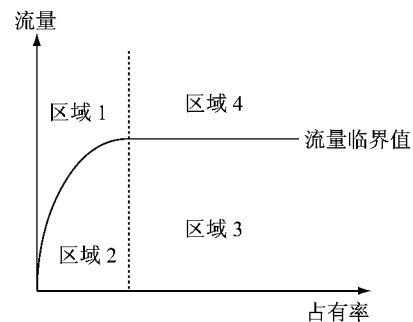


图 1 McMaster 算法交通状态估计模板
Fig.1 Traffic state estimation template of McMaster algorithm

近年来,城市交通状态估计研究得到了进一步发展,主要存在两大类:第一类是基于新型数据源或增加数据源进行估计算法开发;第二类是在定点线圈数据的基础上,开发更先进算法对数据进行深层次挖掘从而进行交通状态估计。

第一类方法中,文献^[9]开发了一种基于浮动车提供的高质量行程时间数据的交通状态估计方法,文献^[10-11]开发了基于浮动车检测与感应线圈融合数据的道路交通状态估计模型,利用神经网络算法处理以上两模块的结果从而提高道路交通状态估计的精度,但单一模块精度并不高.文献^[12]分析了固定检测器和移动检测器在数据检测、信息采集方面的差异性,提出了信息融合的技术框架,并采用神经网络对交通状态参数区间平均车速进行了信息融

合,通过软件的仿真,对数据融合的效果进行了验证.文献[13]开发了一种基于移动检测器、固定检测器以及人工报告等多源数据决策级融合的自动拥挤估计算法,该方法将多种方式获得的交通状态估计结果进行融合,进而得出交通状态的最终结果,并用实例进行了验证.

第二类方法中,文献[14]利用模糊聚类方法分析定点检测器的交通流三参数数据,将交通状态分为四类.文献[15]利用 SCOOT 系统感应线圈的交通数据为基础,开发了一种基于模糊 C 均值聚类的城市交通状态实时估计算法,并分析了采样时间间隔对状态估计的影响,最后通过仿真数据进行验证,运用其提出的交叉判断方法准确率达到 90% 以上.文献[16]基于秒级上游检测器数据和信控配时数据,开发了一种适用于拥挤状态下的实时路段行程时间估计算法并用实证数据验证了方法的有效性.文献[17]提出了一个模块化神经网络动脉事件检测模型的方法,基于上下游检测器以 20 s 为时间间隔提取速度、流量和占有率参数,验证了方法的有效性.文献[18]基于路网的下游和上游线圈检测器组的 5 min 集计数据,由路段和交叉口状态估计模型建立了区域交通状态的综合估计模型,并分析路网整体状态指标 P 与路网车辆平均速度的关系以确定 P 与交通状态的关系,并用仿真数据进行了验证.

综上所述,在城市交通状态估计的研究中,第一类主要是基于新型数据源或者多数据源,实际应用中存在如前文所述的经济成本、数据精度与频率等方面的局限;而第 2 类主要基于定点检测器,但存在 2 个问题:一是主要利用模糊逻辑、神经网络等人工智能算法对交通状态划分等级或建立评价体系以进行交通状态估计,难以适用于直接通过行程速度阈值来确定交通状态的应用场景(我国大部分中小城市目前的应用现状);二是大多以多个位置的定点检测器作为数据源而非单一的路中检测器,对检测器的布设密度要求较高.上述原因导致已有研究难以完全适应我国中小城市交通信息采集与服务的技术现状,亟需一种简单实用的基于路中检测器数据和行程速度阈值来估计城市道路交通状态的方法.

2 基于定点检测器的交叉口交通状态估计方法

2.1 基本思想

针对的城市道路交通检测数据条件如图 2 所示:道路路中各车道布设有定点线圈检测器,以 N

min 的频率(选取为 5 min,可由 5 个 1 min 数据集计得出)上传集计流量与占有率数据,下游与信控交叉口相连.利用曲线拟合的数理统计方法,基于路中检测器在不同环境变量下的大量历史数据(即为上传的长时段的 5 min 集计数据集),开发一种道路状态估计方法.

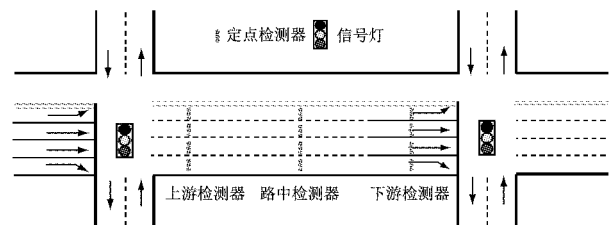


图 2 我国城市道路定点检测器典型布设方式
Fig.2 Typical layout of detectors on urban roads in China

2.2 方法流程

方法流程如图 3 所示,主要包含以下的 5 个步骤.

(1)选取交叉口及特定车道(左转、直行、右转或混合车道),提取路中检测器历史集计数据以及对应的环境变量数据和交通状态数据:①提取目标路段特定车道的路中检测器大量历史数据(5 min 集计数据,包括占有率与流量).②提取上述检测器 5 min 集计数据所对应的环境变量数据,包括下游交叉口信控周期 C (s)、绿信比 λ 、路中检测器位置 S (检测器距离下游停车线距离, m)和交通状态数据(基于对应 5 min 下的路段平均行程速度进行状态区分,分别设定状态 1 为通畅,状态 2 为拥挤,状态 3 为拥堵).

(2)确定环境变量下的二维图生成与分析:①生成二维图.根据大量历史数据,得到每种环境组合下的占有率-车辆数-交通状态二维图,示意图如图 4a 所示.其中,每个点的位置代表路中检测器记录的 5 min 集计数据,图 4b, 4c 中点的颜色代表对应时间段的交通状态与图 4a 相同.②分析二维图.

(3)交通状态分界曲线拟合与参数提取:①根据交通状态划分较为明显的 4 条边界曲线,如图 4b 所示,边界曲线 I, II, III 和 IV 分别代表状态 1, 2 和 3 的外围或中间边界.②针对每种交通状态,在纵轴方向(流量)从最低流量开始以一定的区间长度选取每区间的曲线边界拟合点(当接近曲线顶部时以更小的区间长度选取边界点),值得注意的是用于拟合曲线 II 的点包括状态 1 的右边界点和状态 2 的左边界点,同理拟合曲线 III.③在拟合过程中,根据实际情

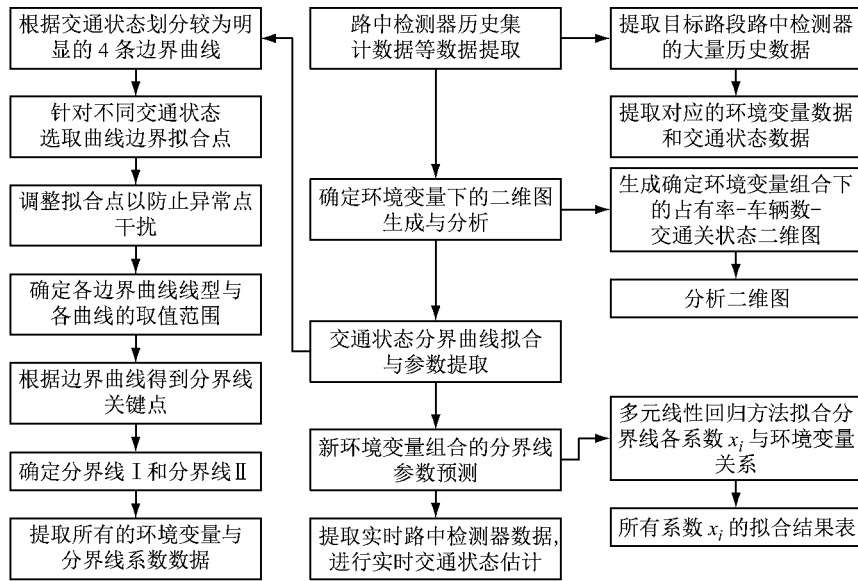


图 3 方法流程

Fig.3 Algorithm flow

况排除异常点干扰。④综合评判各线型的优劣以选取合适线型；同时，确定各曲线的取值范围。根据综合分析，确定曲线 I，II，III 的最优线型为二次函数线型，曲线 IV 的最优线型为一次函数线型，各边界曲线的函数表达式如式(1)和式(2)所示：

$$V_{ol_i} = a_i O_{cc}^2 + b_i O_{cc} + c_i \quad (1)$$

$$V_{ol_{IV}} = b_4 O_{cc} + c_4 \quad (2)$$

式中： $i=1\sim3$ ； V_{ol} 表示每 5 min 的流量； O_{cc} 表示实际占有率，%； a_i, b_i, c_i 分别代表二次项系数、一次项系数和常数项。⑤得到分界关键点。图 4b 中 A 点为曲线 I 和曲线 II 点的右交点；C 点为曲线 II 和曲线 III 的右交点；B 点为曲线 II 定义域左端点，取交通状态

2 数据点的最小占有率值；D 点为曲线 III 定义域左端点，取交通状态 3 数据点的最小占有率值。⑥在实际应用过程中，已有的历史大量数据不可能包含所有的数据点，有可能有数据点在已有曲线 I，II，III 和 IV 的范围之外，解决的方法是根据不同边界曲线的端点所确定的占有率作为曲线外区域不同交通状态的分界线，如图 4c 所示。分界线 I 和分界线 II 分别是交通状态 1 和 2 以及 2 和 3 的分界线。⑦提取所有的环境变量(C,λ,S)组合数据及其对应的分界曲线 I 和分界线 II 的各系数 x_i 。 x_i 代表第 i 条边界曲线的系数 x 。

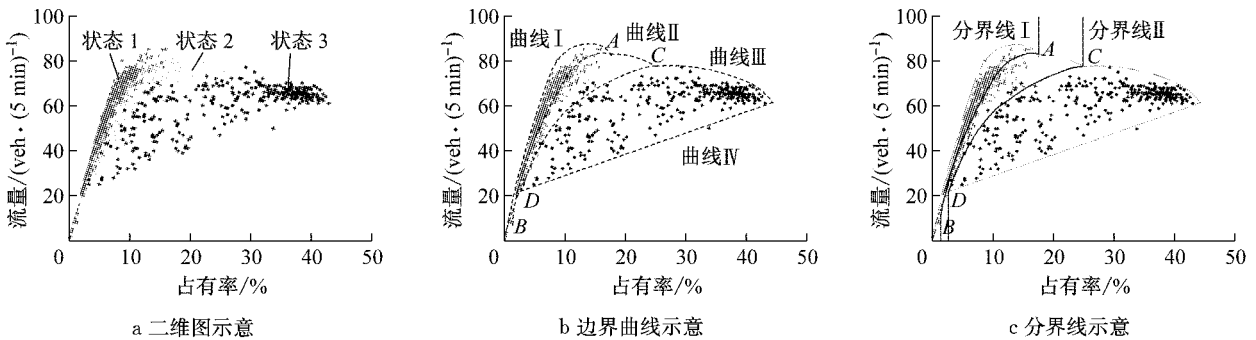


图 4 分界方法示意

Fig.4 Diagram of demarcation method

(4)新环境变量组合的分界线参数预测。①对所有的环境变量组合进行分界曲线拟合，则能得到各分界曲线各系数 x_i 与环境变量组合(C,λ,S)的数据组，从而可以利用多元线性回归拟合出各系数 x_i 与周期 $C/100$ (s)、占有率 λ 、检测器位置 $S/100$ (s) 这 3

个环境变量的函数关系，如式(3)所示：

$$x_i = F\left(\frac{C}{100}, \lambda, \frac{S}{100}\right) = f_{i0} + f_{i1} \frac{C}{100} + f_{i2} \lambda + f_{i3} \frac{S}{100} \quad (3)$$

式中： F 代表多元线性回归函数； f_{ij} 分别为常数项与

3 个环境变量的系数. ②得到所有系数 x_i 的拟合结果,可用于新环境变量组合下的边界曲线系数预测.

(5)提取实时路中检测器数据,根据上述方法,进行实时交通状态估计.在实际应用中,随着历史数据的不断增加,可以动态更新历史数据库,不断修正分界曲线的系数以适应新的情况,提高交通状态估计准确率.

3 模型建立与数据提取

本文基于仿真环境进行方法研究,并用实证数据进行方法验证.这是因为在现实环境中,交叉口相关的环境变量组合较为单一且上游路段的流量变化有限,短期内难以达到研究所需的样本量;而在仿真环境中,能够设置不同的环境变量组合和上游路段流量输入值,有效地解决了实证数据的局限,极大地

提高了研究效率.

3.1 仿真场景搭建

选取安徽省合肥市黄山路—科学大道西进口道及其上游路段作为研究对象,如图 5a 所示.该路段为城市主干道,符合本研究的典型场景,布设有路中检测器.利用 VISSIM 微观仿真软件建立如图 5b 所示的交叉口仿真模型,以路段中间的直行车道作为研究区域(选取左转、右转或其他交叉口的混合车道亦可).在交叉口上游路段 C 处设置路中检测器作为数据来源;同时以该进口道上游 A 处与下游进口道 B 处之间作为行程时间检测区域,布设行程时间检测器以检测该区域的平均行程时间.设置不同的环境变量(C, λ, S)取值如表 1 所示,分别为 8 种、5 种、3 种,每个环境变量的取值范围均能覆盖绝大部分实际交通状况.对不同的环境变量(C, λ, S)组合分别进行仿真试验,共有 $8 \times 5 \times 3 = 120$ 组试验.

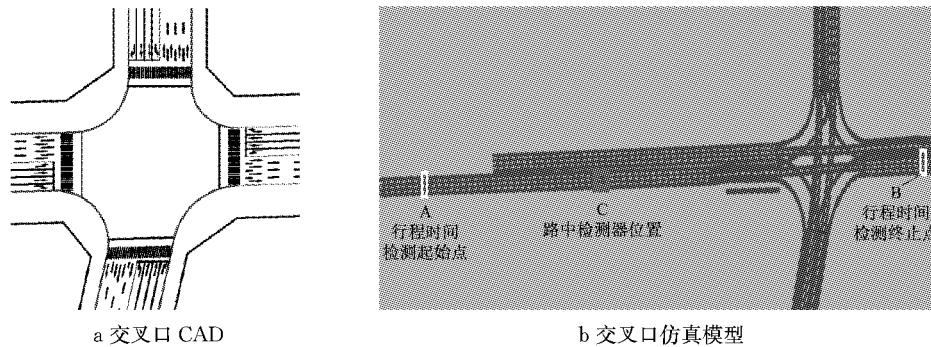


图 5 仿真场景

Fig.5 Simulation scene

表 1 环境变量(C, λ, S)取值

Tab.1 Values of environment variables of C, λ and S

参数	数值
λ	0.20, 0.30, 0.33, 0.36, 0.39, 0.42, 0.50, 0.60
C	60, 90, 120, 150, 180 s
S	200, 250, 300 m

3.2 仿真数据采集

对于上述 120 组仿真试验中的每一种环境变量(C, λ, S)组合,提取路中检测器与相对应的行程时间检测器的 5 min 集计数据.具体方法为将 VISSIM 软件连接数据库,输出 *.sdf 文件,包括路中检测器的 5 min 流量($\text{veh} \cdot (5 \text{ min})^{-1}$),5 min 占有时间($\text{s} \cdot (5 \text{ min})^{-1}$)以及 5 min 的研究路段平均行程时间(s).

将上述数据库数据进一步处理为每 5 min 的流量 V_{ol} ($\text{veh} \cdot (5 \text{ min})^{-1}$)、实际占有率 O_{oc} (%)、对应的区间(图 5 中 A, B 检测器之间)行程速度 v ($\text{km} \cdot \text{h}^{-1}$)及其对应的交通状态,其中交通状态由区间平

均行程速度决定,设定不同交通状态阈值:状态 1 ($>22 \text{ km} \cdot \text{h}^{-1}$),状态 2($16 \sim 22 \text{ km} \cdot \text{h}^{-1}$),状态 3 ($<16 \text{ km} \cdot \text{h}^{-1}$).仿真时长共计 8 000 min,即每组有 1 600 条数据.

最终得到研究所需数据,数据记录形式如表 2,其中位置代表路中检测器距离停车线的距离.

表 2 仿真数据记录

Tab.2 Simulation data records

编号	S/m	$V_{ol}/(\text{veh} \cdot (5 \text{ min})^{-1})$	$O_{oc}/\%$	$v/(\text{km} \cdot \text{h}^{-1})$	交通状态
1	300	90	10.06	7.01	3(拥堵)
2	250	11	1.11	31.18	1(畅通)
3	200	12	1.15	29.19	1(畅通)
⋮	⋮	⋮	⋮	⋮	⋮

4 方法评价

4.1 占有率、流量与交通状态基本关系分析

4.1.1 二维图分析

在不同环境变量(C, λ, S)的组合下进行了 120

组仿真试验,选取其中 6 组环境变量(C, λ, S)组合进行示例分析。

由仿真试验得出的 1 600 条数据得到如图 6a~6f 所示的占有率-流量图,其中 C 均为 90 s。图中,每个点的位置代表路中检测器记录的集计数据。从图

中可以看出,不同交通状态对应的数据点相对集中,相互之间有较为明显的位置区分,如畅通状态下的数据点集中在较低占有率的位置,而拥堵状态下的数据点所对应的占有率较高,符合实际交通情况。

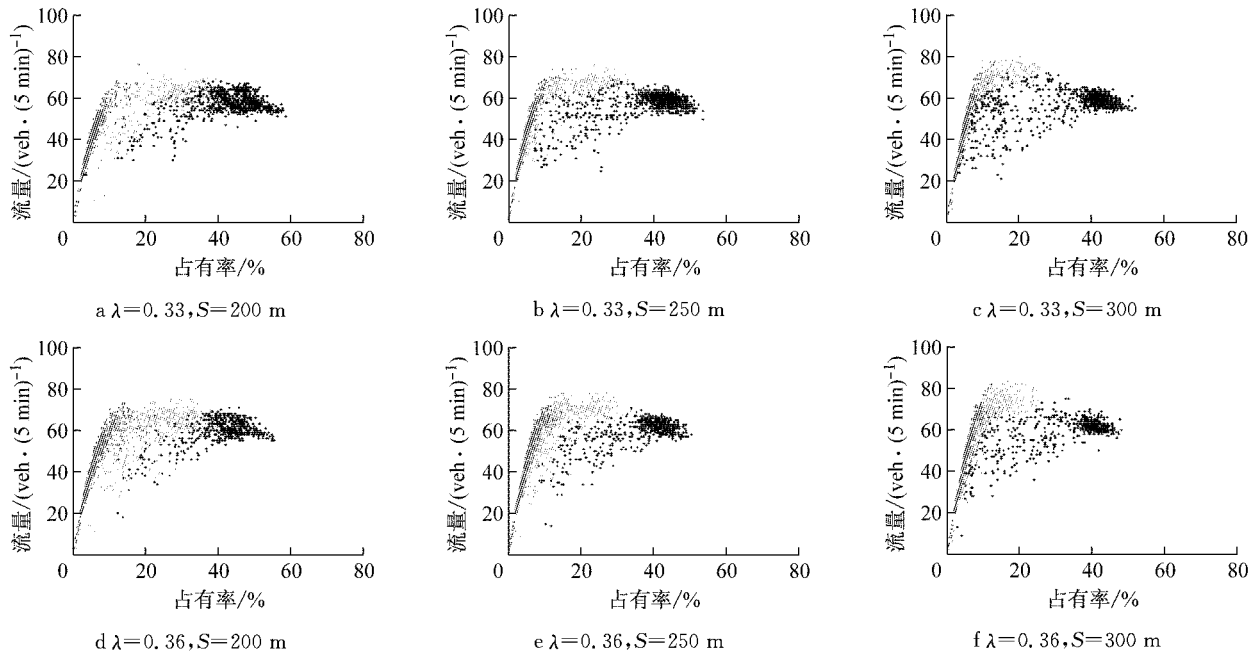


图 6 $C=90$ s 时不同情况的占有率-流量二维图

Fig. 6 Two-dimensional maps of flow-occupation on different situations with cycle of 90 s

4.1.2 三维图分析

通过固定环境变量(C, λ, S)中的 2 个参数可以分析不同环境变量对交通状态估计的影响。

(1)周期. 通过固定其他 2 个参数可以分析该参数对交通状态估计的影响。对周期进行分析,固定绿信比为 0.36 和检测器位置为 3,得到如图 7a 所示的三维图。图中可以看出,随着周期的变大,不同交通状态均发生了向占有率较小的位置偏移的现象,即在某时间段路中检测器检测到的占有率、流量数据相同的情况下,周期较大的情境下所对应的交通状态可能更拥堵。这是因为以 5 min 为集计时间段,更大的周期使得 5 min 内研究车道的车辆停止时间所占的比重更高,导致区间行程速度下降,即体现为更为拥堵。

(2)绿信比. 对绿信比进行分析,固定周期为 120 s 和检测器位置为 3,得到如图 7b 所示的三维图。从图中可以看出,随着绿信比增大,不同交通状态均发生了向占有率较大的位置偏移的现象,即在某时间段路中检测器检测到的占有率、流量数据相同的情况下,绿信比较大的情境下所对应的交通状态可能更畅通。这是因为对研究车道而言,绿信比越大,相

同周期下的通行时间越长,则整体交通越畅通。

(3)检测器位置. 对检测器位置进行分析,固定周期为 90 s 和绿信比为 0.39,得到如图 7c 所示的三维图。图中可以看出,随着检测器位置向下游移动,不同交通状态均发生了向占有率较大的位置偏移的现象,即在某时间段路中检测器检测到的占有率、流量数据相同的情况下,绿信比较大的情境下所对应的交通状态可能更畅通。这是因为检测器越接近下游,信号灯尤其是红灯对检测器处的交通运行影响越大,当检测到相同的占有率、流量数据时,则检测器更接近下游的情况下的交通可能更畅通。

4.2 基于交通状态的分界曲线确定

4.2.1 分界结果分析

采用 MATLAB 软件进行编程,根据方法流程中所述方法可以得到所有环境变量(C, λ, S)组合下的分界结果。

以其中一组环境变量(C, λ, S)组合(周期为 60 s、绿信比为 0.36、检测器位置为 1)为例,分界结果如表 3 所示,拟合效果如图 8 所示,可以看出各边界曲线对各交通状态边界点的拟合效果均较好,拟合优度 R^2 均能达到 80% 以上,能够有效区分出不同的

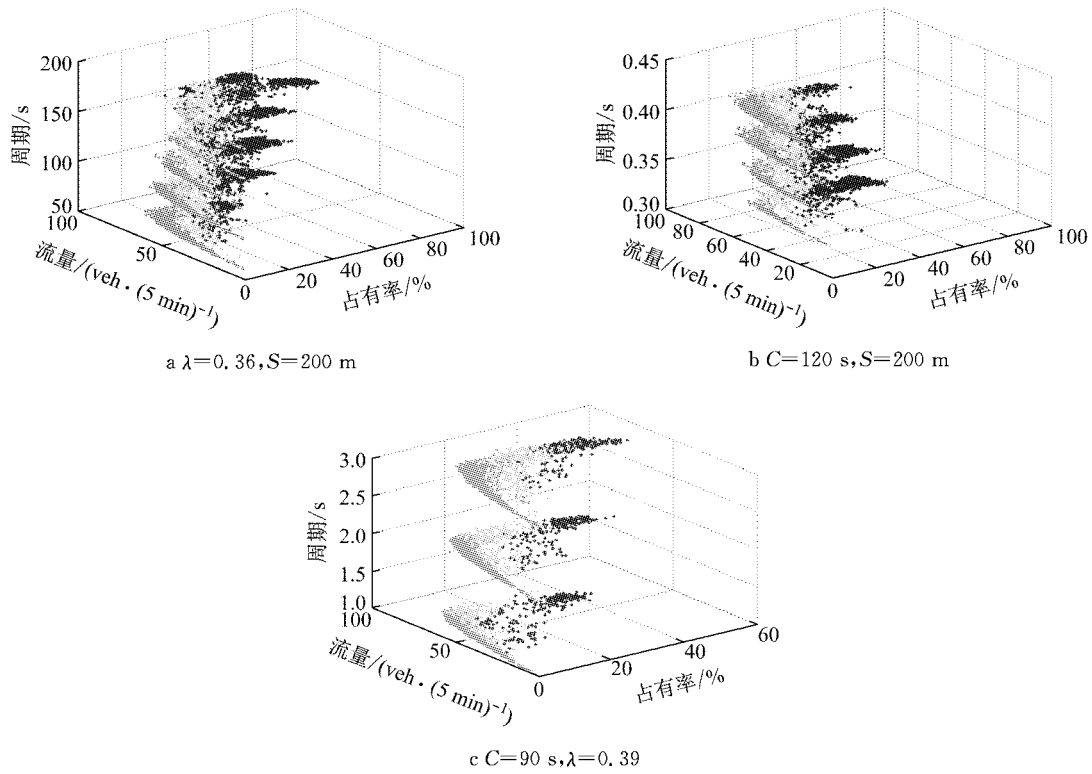


图 7 不同情况下的占有率-流量三维图

Fig.7 Three-dimensional maps of flow-occupation on different situations

表 3 各边界曲线拟合结果

Tab.3 Results of the boundary curves fitting

曲线 I			曲线 II			曲线 III			曲线 IV	
a	b	c	a	b	c	a	b	c	b	c
-0.46	12.86	-2.59	-0.23	8.47	0.53	-0.43	12.34	-8.11	5.59	5.97

注:a 表示二次项系数,b 表示一次项系数,c 表示常数.

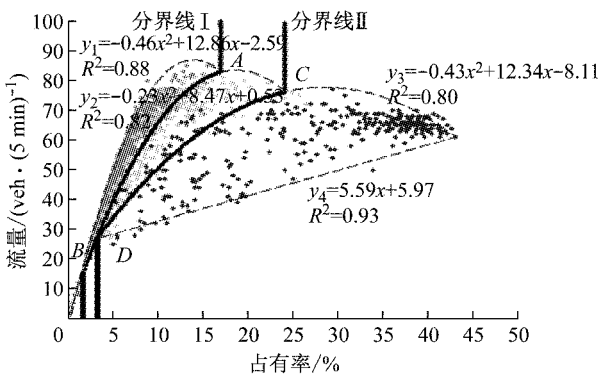


图 8 分界曲线结果

Fig.8 Results of demarcation lines

交通状态.

4.2.2 验证分析

基于仿真数据对分界方法进行验证,分别进行特例验证和综合验证.

首先进行特例验证,以 $C=60\text{ s}, \lambda=0.36, S=300\text{ m}$ 情景为例.具体方法为随机选取该情景下所有数据(1 600 条)中的 75% 数据(1 200 条)进行曲线拟合,剩下的 25% 数据(400 条)用于验证分析.

随机选取的 75% 数据的曲线拟合结果如表 4 和图 9 所示.其中表 4 将所有数据(1 600 条)的拟合结果与部分数据(1 200 条)的拟合结果进行了对比,结果显示各参数均较为接近;图 9 显示部分数据得出的拟合分界曲线能够较好地地区分出不同交通状态.

表 4 边界曲线拟合对比结果

Tab.4 Fitting results of boundary curves

数据类型	曲线 I			曲线 II			曲线 III			曲线 IV	
	a	b	c	a	b	c	a	b	c	b	c
完整数据	-0.46	12.83	-2.54	-0.24	8.82	1.16	-0.08	4.59	12.75	0.86	24.07
部分数据	-0.48	12.94	-2.74	-0.25	9.02	0.63	-0.08	4.59	12.83	1	21.15

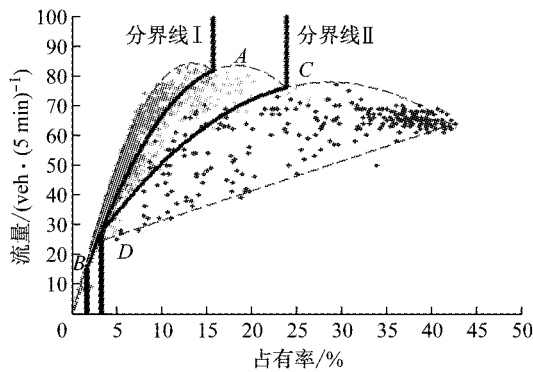


图 9 边界曲线验证图

Fig.9 Verification of boundary curves

用剩下的 400 条数据进行部分数据得出的分界线的验证分析,即利用本方法,输入检测器的占有率和流量数据得出交通状态,将其与实际交通状态进行对比.结果如表 5,由特例验证部分可见,准确率达到 86.75%,且严重错误的比例非常低,只有 0.25%.

按照特例验证所示的方法对本方法进行综合验证.具体方法为选取检测器位置 $S=300\text{ m}$ 时的所有周期 C 与绿信比 λ 组合进行综合正确率验证,即对每种组合均随机选取 75% 的数据进行拟合并验证剩下的 25% 数据.具体结果如表 5 中综合验证部分所

表 5 交通状态估计验证结果

Tab.5 Verification results of traffic state estimation

试验对象	实际交通状态	样本量/个	估计状态	估计量 ¹⁾ /个	分状态正确率/%	正确率 ²⁾ /%	严重错误率 ³⁾ /%				
特例验证	1(畅通)	277	1	243	87.73	86.75	0.25				
			2	33							
			3	1							
	2(拥挤)	73	1	7	84.93	86.75	0.25				
			2	62							
			3	4							
3(拥堵)	50	1	0	84.00	86.75	0.25					
		2	8								
		3	42								
综合验证	1(畅通)	9 873	1	8 078	81.82	81.30	1.98				
			2	1 559							
			3	236							
	2(拥挤)	2 792	1	341	80.16			81.30	1.98		
			2	2 238							
			3	213							
	3(拥堵)	2 135	1	57	80.37					81.30	1.98
			2	362							
			3	1 716							

注:1)估计量表示实际交通状态为 i 而根据本算法估计出的交通状态为 j 的样本个数.

2)“正确”表示根据本算法估计出的交通状态与实际交通状态相同.

3)“严重错误”表示本算法估计交通状态与实际交通状态相差 2 个交通状态的情况,如:本算法估计交通状态为 1 而实际交通状态为 3.

示,得出综合正确率为 81.30%,综合严重错误率为 1.98%;同时,不同实际交通状态下利用本算法估计出的交通状态正确率没有明显差异,均接近 81%.

4.3 新环境变量组合的分界线预测

4.3.1 系数预测

对所有的环境变量 (C, λ, S) 组合进行分界曲线拟合得到 120 组各分界曲线各系数 x_i 与环境变量组合 (C, λ, S) 的数据组,可以进一步拟合出分界曲线各系数 x_i 与 $C/100, \lambda, S/100$ 这 3 个环境变量的函数关系.

最终得到所有系数 x_i 拟合结果如表 6 所示,可用于新环境变量组合下的边界曲线系数预测,并进一步进行交通状态估计.

表 6 多元线性回归结果

Tab.6 Multiple linear regression results

边界曲线	系数	f_{i0}	f_{i1}	f_{i2}	f_{i3}
1	a	-0.52	-0.05	1.52	-0.16
	b	5.67	-0.47	-6.96	3.33
	c	11.12	0.68	8.21	-5.66
2	a	-0.21	0	1.14	-0.16
	b	3.26	-0.55	-13.14	3.66
	c	11.38	5.06	-2.06	-4.82
3	a	0	0.02	-0.06	-0.02
	b	0.56	-1.08	5.77	0.79
	c	3.25	8.21	-70.90	8.73
4	b	-0.17	-0.41	3.67	0.05
	c	1.65	-0.99	39.90	-0.06

4.3.2 实证分析

采用实证数据对预测方法进行验证.

选取青岛市福州南路 3 个路段为研究对象,选取 2015 年 11 月 3 日早高峰为研究时间段,采集相关实证数据,包括定点微波检测器数据、视频数据和信控配时、检测器布设数据.

视频数据用于获得该路段某 5 min 集计时间段内的路段平均车辆速度 v 以确定交通状态,具体方法为通过视频采集该路段 5 min 中所有车辆的路段行程时间,并计算出平均速度;由定点微波检测器数据可获得对应路段与集计时间段的流量 V_d 和实际占有率 O_{oc} ;由信控配时、检测器布设数据可获得对应路段与集计时间段的环境变量 C, λ 和 S . 实证数据具体参数如表 7 所示.

用上述实证数据对本文方法进行验证,对于每组验证样本,根据表 6 和环境变量预测出对应路段和集计时间段的分界曲线并得到估计交通状态,并与实际交通状态对比,验证样本量为 48,验证结果如表 8 所示,可以得出总体正确率约为 75.00%,且严重错误率约为 2.1%.同时,在不同实际交通状态下

利用本文方法估计出的交通状态正确率有一定差异,其中在拥挤状态下的估计正确率最高,达到 80% 以上;而拥堵状态下的估计正确率最低,约为 70% 左右。

表 7 实证分析的相关条件和参数

Tab.7 Relevant conditions and parameters of empirical analysis

路段名称	路段长度/m	S/m	C/s	λ	采集车辆数/veh	5 min 集计数据样本量/个
漳浦路—漳州二路	160	75	175	0.80	140	20
漳州二路—香港中路	218	125	175	0.31	98	14
漳州二路—漳浦路	160	85	175	0.83	105	14

表 8 实证数据验证

Tab.8 Verification of empirical data

实际交通状态	样本量	估计状态	估计量	分状态正确率/%	总正确率/%	严重错误率/%
1(畅通)	20	1	15	75.00	75.00	2.08
		2	4			
		3	1			
2(拥挤)	15	1	1	80.00	75.00	2.08
		2	12			
		3	2			
3(拥堵)	13	1	0	69.23	75.00	2.08
		2	4			
		3	9			

通过实证数据的验证,本分界线预测方法的总体正确率不很高且不同交通状态下的估计正确率有一定差异的主要原因有 3 个:

(1)实证数据存在一定的检测误差.微波检测器在车型较为单一、车速一致性高、车流量不高的路段上检测精度较高,而在交通拥堵的情况下因车辆间遮挡、开普勒效应等因素,检测精度会下降。

(2)系数拟合过程存在误差.因为在确定环境变量下对分界曲线系数进行拟合的过程中必然存在着一定的拟合误差,再基于这些系数数据对新环境变量下的分界线进行预测时,必然存在一定的误差。

(3)单个路中检测器的信息量存在限制.因为本文未加入其他数据源,所使用的数据仅局限于其单个路中检测器的检测数据,存在一定的信息量限制。

5 结论与展望

针对我国中小城市的道路交通检测数据条件,以城市道路交叉口间断流为研究对象,基于路中检测器的大量历史数据(低频数据,为 5 min 集计流量与占有率)开发出了一种交通状态估计方法.在实际

应用中,该算法能够根据实时的路中检测器低频集计数据准确地实现城市交叉口交通状态的估计,同时可以随着历史数据的不断积累更新数据库并优化相关参数以提高估计准确率.同时,随着多源数据应用愈加广泛,本算法可作为基于多源数据的城市交通状态估计方法的构成算法之一。

首先在确定环境变量组合(周期、绿信比、检测器位置)下由大量历史数据形成的占有率-流量-交通状态二维图进行分界曲线的拟合与建立,然后利用多组环境变量组合下的分界曲线系数拟合出其与环境变量的函数关系.经过验证,在确定环境变量组合下,交通状态估计准确率达到 80% 以上,严重错误率低于 2%;而在一般条件下,通过分界线预测方法,实证数据的交通状态估计准确率达到 75% 以上,严重错误率低于 2.1%.基本可以满足交通检测条件较差的中小城市交通状态估计的基本需求.本研究仍存在一定不足,主要在于预测得出的新环境变量组合分界线对交通状态估计准确率不够高。

参考文献:

[1] 唐克双,孙剑. 基于多源数据融合的城市道路交通控制与管理[M]. 上海:同济大学出版社,2015.
TANG Keshuang, SUN Jian. Urban traffic control and management based on fusion of multi-source data [M]. Shanghai: Tongji University Press, 2015.

[2] HOOGENDOORN S, HOOGENDOORN-LANSER S, SCHUURMAN H. Fuzzy perspectives in traffic engineering [R]. [S. l.]: TRAIL Research School, Dutch Ministry of Transport, 1998.

[3] HAWAS Y E. A fuzzy-based system for incident detection in urban street networks [J]. Transportation Research Part C: Emerging Technologies, 2007, 15(2):69.

[4] LEVIN M, KRAUSE G M. Incident detection: A bayesian approach [M]. Washington D C: Transportation Research Record, 1978.

[5] 姜桂艳. 道路交通状态判别技术与应用[M]. 北京:人民交通出版社,2004.
JIANG Guiyan. Road traffic state identification technology and application [M]. Beijing: China Communications Press, 2004.

[6] PERSAUD B N, HALL F L, HALL L M. Congestion identification aspects of the McMaster incident detection algorithm[J]. Transportation Research Record, 1990, 1287.

[7] 韩悦臻. 城市道路交通状态指标体系及判定方法研究[D]. 吉林:吉林大学, 2003.
HAN Yuezhen. Evaluation indexes and estimation for urban traffic conditions [D]. Jinlin: Jinlin University, 2013.

[8] THOM R. Structural stability and morphogenesis [M]. Boulder: Westview Press, 1989.