

基于卡口车牌数据的交通拥堵改善方案研究

马柱, 吴寻, 陈福临
(浙江省城乡规划设计研究院, 浙江 杭州 310030)

摘要: 卡口车牌数据含有极其丰富的交通信息, 可以获取传统调查方法无法获取的指标, 是交通分析的理想数据。提出卡口车牌数据用于交通拥堵分析的思路和方法, 即通过卡口车牌数据获取路段流量、转向流量、路段平均行程时间、车辆OD等指标, 了解交通特征, 进而有针对性地提出改善方案。在原始数据预处理方面, 提出基于路网拓扑结构和大数据思想的数据补全方法; 在交通分析方面, 根据不同分析目的, 将研究问题分为明确路径的轨迹分析、明确起(终)点的轨迹分析以及不明确起终点的轨迹分析三类, 并分别探讨了其适用性。最后, 以千岛湖镇新安东路改造为例阐述具体分析过程。

关键词: 交通工程; 交通设计; 大数据; 卡口车牌数据; 路网拓扑; 交通拥堵

Traffic Congestion Improvement Based on Vehicle License Plates Data at Checkpoints

Ma Zhu, Wu Xun, Chen Fulin

(Zhejiang Urban & Rural Planning Design Institute, Hangzhou Zhejiang 310030, China)

Abstract: Vehicle license plates data collected at checkpoints is ideal for traffic analysis because it contains lots of valuable information that is unobtainable with the traditional data collection methods. This paper proposes the traffic congestion analysis methods using license plate data to estimate traffic flow, turning volume, average travel time and vehicles' OD, which reveals traffic characteristics and thereby helps to develop the corresponding improvement measures. To handle the initial data processing, the paper proposes the complementary methods of big data based on the topological structure of roadway networks. According to different analysis objectives, the paper divides the research questions into track analysis for clear paths, track analysis with clear starting or ending points, and track analysis without starting and ending points, and then discusses their applicability. Finally, taking Xin'andong Road improvement in Qiandaohu Town as an example, the paper demonstrates the proposed analysis process.

Keywords: traffic engineering; transportation design; big data; vehicle license plate data collected at checkpoints; road network topology; traffic congestion

收稿日期: 2020-04-27

作者简介: 马柱(1989—), 男, 安徽芜湖人, 硕士, 工程师, 主要研究方向: 交通规划与道路设计。E-mail: mz_pillar@163.com

0 引言

随着智慧城市的快速建设, 高清摄像卡口相机已经遍布城市各个交叉口和路段, 全天候记录经过的车辆信息, 产生了海量的出行数据。卡口车牌数据质和量的提升为其在交通分析方面的应用创造了条件。相对于传统的调查方法, 卡口车牌数据获取方便且含有极其丰富的出行信息, 不仅可以获取一系

列传统的道路交通指标, 还可推导传统交通调查手段无法获取的交通信息, 例如车流轨迹特征、车辆出行OD等。卡口车牌数据由于数据量大、路径可追踪等特征, 能实时动态反映道路交通状态和数据间的关联性, 是进行交通分析的理想数据。

目前, 卡口车牌数据的分析和应用主要集中在交通管理方面, 例如违法车辆抓拍、套牌车检测和可疑车辆识别等^[1]。近些年,

一些学者也开始尝试利用卡口数据进行交通分析方面的研究。文献[2]基于上海市车牌识别数据,采用聚类算法识别通勤车辆,得到潜在的有通勤特征的车辆数和地图分布。文献[3]基于车牌识别数据分析研究车辆居住地的识别方法、外地车的出行特征,并利用聚类算法对外地车进行分类。文献[4]研究多源数据支撑的OD计算模型,对原始卡口数据中的车辆行程信息进行修补及迭代扩样,得到快速路网接近全量的车辆出行OD以及分类行程路径。文献[5]基于卡口车牌数据分别从毗邻区域交通量、车流轨迹和道路运行状况层面进行分析,提出一系列传统交通调查手段无法获取的交通指标。关于卡口数据在交通方面的研究主要围绕个体出行特征和路网交通状态进行,而在具体交通拥堵治理方面的应用较少。

卡口车牌数据含有极其丰富的实时交通信息,可以得到现状路网的交通运行特征,揭示交通拥堵发生的原因,且能为交通仿真提供较为真实的原始数据,在现状道路拥堵治理方面具有较大优势。而实际的卡口车牌数据中存在较多的粗差数据,极大地限制了卡口数据的应用,因此,如何对原始数据进行校正是交通分析的基础。卡口车牌数据与传统调查方法最大的不同在于不同卡口之间的数据可以通过车牌号进行关联,追踪行车轨迹,所以行车轨迹如何应用于交通拥堵分析也是需要探讨的重点。

本文提出了基于卡口车牌数据进行交通拥堵治理的基本思路和方法,并对分析过程中的两个重点问题——原始数据预处理和车流轨迹分析进行探讨。最后以浙江省淳安县千岛湖镇新安东路改造工程为例,阐述具体的分析过程。

1 技术方案

卡口数据库中一条数据一般包括卡口位置、车牌号、拍摄时间、车牌颜色、车型、车道、车辆速度等。由于设备的差异,车道、车辆速度等数据与真实差距较大,不予采用,本文主要使用字段包括卡口位置、进口道、车牌号、经过时间和车型。

由于原始数据存在错误、重复、丢失等问题,在数据分析前需要对原始数据进行预处理,错误数据和重复数据均容易判断,丢

失数据可以基于车辆行驶路径受路网时空约束的特点,利用路网的拓扑结构进行完善补充。

数据预处理后,可以对数据进行分析。按照车牌将每辆车经过的卡口以及经过的时刻按先后顺序排列后,可得到所有车辆的卡口序列和对应的时刻序列,即车辆出行链。对出行链进行分析可获得需要的交通指标,包括传统交通调查手段无法获取的交通指标。首先,对卡口车牌数据按字段进行统计分析,可得到卡口所在道路截面交通量、车型比例等常规交通调查指标。第二,基于出行链,获取经过分析路段的车辆集合,通过车辆经过该路段时间差,获得路段行程时间。第三,对每个进口道车辆分别进行分析,通过获取该车辆经过的下一个交叉口位置可以判断该车辆转向情况,进而获取交叉口转向流量。第四,通过追踪进入分析路段车辆的行车轨迹,获取车辆的OD交通量。

交通特征指标提取后,可以通过路段流量和平均行程时间判断拥堵路段;通过拥堵路段内交叉口流量、流向可以提出针对交叉口的改造方案;通过车辆OD分析可以在宏观上提出车辆诱导方案,在微观上提出路段的改造方案。最后,可以利用获取的OD数据对改善方案进行交通仿真。交通分析主要流程如图1所示。

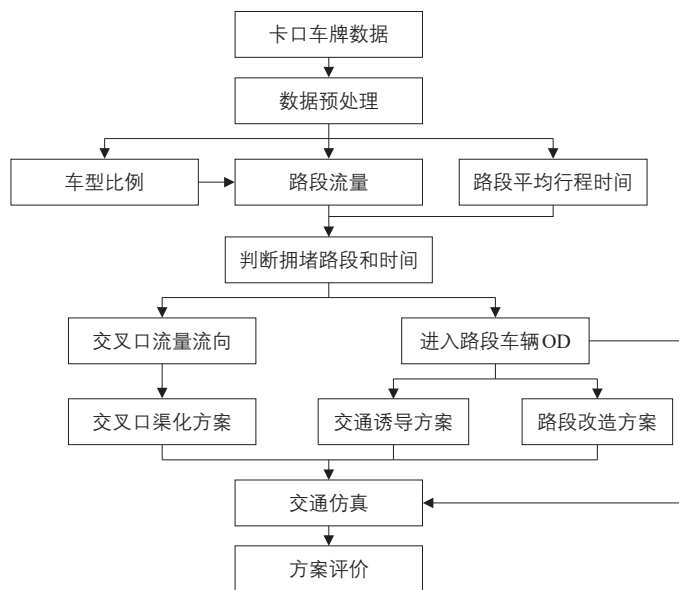


图1 基于车牌卡口数据的交通改善方案流程

Fig.1 Flow chart of traffic improvements based on vehicle license plate data collected at checkpoints

2 基于路网拓扑结构的数据预处理

本文用到的卡口车牌数据包括城市路网中各个信号交叉口卡口和路段卡口车牌数据。在车牌数据采集过程中,受到设备识别率、信号传输干扰、车辆遮挡、套牌车等的影响,数据往往存在错误数据、重复数据及丢失数据,需要进行数据的预处理。

2.1 错误数据处理

错误数据主要包括数据乱码、检测错误以及逻辑错误(包括套牌车等)^[6]。

由于逻辑错误在错误数据中的比例极小,本文暂不考虑,处理方法可参考其他文献^[7]。数据乱码、检测错误产生的错误数据因缺乏其他有效的信息较难进行还原、矫正,本文采取直接删除、删除后按丢失数据进行处理策略。错误数据的识别方法为:若检测车牌的字符长度以及字母数字比例不符合车牌设置规则,则定义为异常车牌。

2.2 重复数据处理

卡口连续抓拍、检测数据重复上传等因

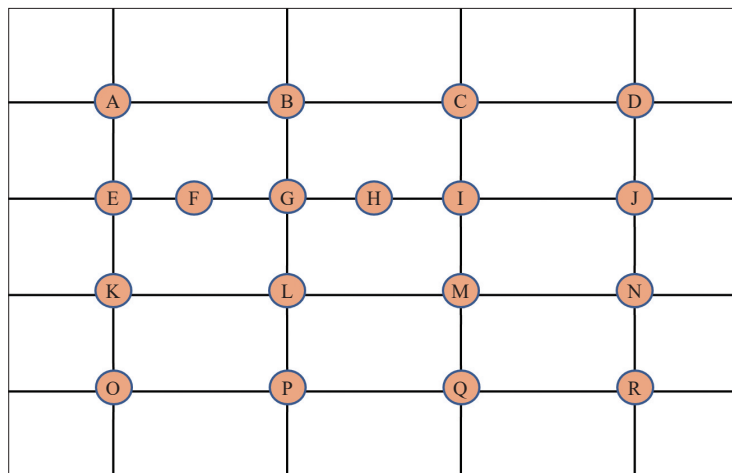


图2 卡口点位示意

Fig.2 Location of vehicle checkpoints

表1 对应卡口邻接表

Tab.1 Adjacency matrix of vehicle checkpoints

卡口	东进口	西进口	南进口	北进口	卡口类型
G	H	F	L	B	交叉口
H	I	G			路段
I	J	H	M	C	交叉口
...

素会导致重复数据的产生,需要对其进行去重处理,具体方法为:按车牌、抓拍时间进行排序,车牌号一致、抓拍时间间隔小于设定阈值(取该点位交叉口的最小红灯时间)的数据视为重复数据,进行去重处理^[3]。

2.3 丢失数据处理

在实际的卡口车牌数据中,数据丢失现象较为严重。数据丢失的主要原因有两类:一类为随机性原因,包括设备识别率和信号传输干扰等;另一类为系统性原因,包括不抓拍右转弯车辆数据、部分重要点位无卡口设备等。数据丢失直接降低了数据的可靠性,特别是系统性原因产生的数据丢失,可能直接导致产生错误的分析结果。

由于车辆只能在路网中行驶,受到路网的时空限制,所以可以通过建立路网拓扑结构对行车轨迹进行校正,进而对丢失的数据进行补全。具体步骤如下:

1) 建立路网拓扑结构。

由于每个进口道对应唯一的上游点位(包括路段卡口或交叉口卡口),为每个卡口配置各个进口道的连接属性,构建路网卡口之间的邻接表,卡口点位示意图2,与之相对应的邻接表见表1。对于存在5路以上交叉口或跨线桥等特殊情况下应增加邻接表字段,本文暂不考虑这一情况。

邻接表的建立可以通过大数据的方法实现,具体实现方法为:分析某一时段的车辆出行链(为减小多次出行对出行链的影响,建议采取高峰时段),获取出行链中需要分析的某进口道上一个卡口名称,遍历所有出行链取出现频率最高的卡口作为该进口道上游卡口,填入邻接表。通过对所有卡口邻接状态的建立,初步获取整个路网的拓扑结构。

由于存在部分畸形交叉口或者某些点位缺失的情况,还需要手动对现有的路网拓扑结构进行校正。然后,为每个卡口填入经纬度数据,这样就可以把卡口点位和连接情况清楚地显示在地图上,并快速发现错误的连接情况。对于缺失的交叉口,可能会导致路网较为紊乱,并在没有数据的交叉口或路段增加虚节点,修改连接情况后建立完整的路网拓扑结构。

卡口邻接表可以通过人工或大数据的方法建立。在初次建立邻接表时,建议采用大数据的方法,一方面可以大大提高效率,另

一方面可以及时发现设备属性信息填写错误以及设备故障，防止因设备信息错误对后期交通分析产生较大影响。

2) 出行链分离。

由于一辆车在一日中可能有多次出行，需要将出行链按照出行次序进行分离。为降低算法的运算复杂度，本文将出行链分离进行简化，将一辆车停留时间过长视为2次出行。考虑到未被多个连续卡点检测到的概率比较小，提取出行链中相邻卡口的检测时间差 Δt 。当 Δt 大于设定的阈值 t_0 时即分割为两条出行链(t_0 可根据两个卡口在某一个时段内的行程时间分布进行确定，本文暂取 10 min)。

3) 数据补全。

对每条出行链从末端向起始端进行拓扑检查。以出行链 A—B—C—D 为例，拓扑检查方法为：先以 D 作为检查对象，邻接表中 D 某进口道上游卡口为 C，则认为 C 与 D 间无数据丢失，停止本次搜索，再以 C 为检查对象，如此类推。如果不相同，则认为 C 与 D 之间存在数据丢失，将邻接表中上游卡口 F 加入出行链(如果 F 为路段卡口，则继续将邻接表中 F 的上一个卡口加入出行链，这里暂不展开路段卡口的描述)，出行链变为 A—B—C—F—D。再检查 F 的 4 个进口道上游交叉口是否有 C，如果有，填入 F 相应的进口道信息，如果没有，则 F 进口道信息空缺。再以 C 为检查对象，如此类推。这样，对整个出行链进行了第一次补充，数据得到了极大的补全，特别是虚节点和原先没有右转数据的节点产生了大量的出行数据。第一次数据补充流程如图 3 所示。

在进行第一次数据补充后，仍有一些出行链因为存在连续多次漏检而没有得到有效补全，对于这种情况需进行第二次数据补充。仍以上例说明，由于 F 已经缺失了进口道的信息，C 与 F 间的道路路径可以有多种，不可以用拓扑关系进行数据的补全，一般采取最短路径的方法^[6]，但最短路径方法算法较为复杂，可行性较差。本文第二次数据补充采取大数据的方法：对出行链进行重新遍历，获取所有包含 C 与 F 的片段，以频率次数最高的出行片段作为最短路径，对 C 与 F 间缺失的数据进行补充。

基于大数据的数据预处理，数据量越大，数据处理的质量越高，所以在卡口选取

的时候，尽量把能获取的所有的卡口车牌数据都添加进去。对于城市外围卡口较为稀疏的情况，数据补全的效果会受到较大影响，可以通过适当增加卡口密度来提高数据的质量。而交通拥堵往往存在于市区段，卡口分布密度大，数据恢复的效果比较好。

卡口的数据预处理不仅可以补充因卡口识别问题而丢失的数据，甚至可以为没有数据的点位添加数据，大大提高了数据的质量和范围，为下一步的交通分析提供了保证。同时，数据处理结果又可以对各个设备状况进行反馈，排查故障设备。

3 车流轨迹分析

相对于传统调查方法，卡口车牌数据的最大优点是可以追踪车辆的运行轨迹，进而了解路段车流状态，能较好地分析交通拥堵产生的原因，为交通改善方案提供依据。所以，分析车流轨迹是数据分析的重点。

根据分析目的的不同，本文将行车轨迹分析分为明确路径的轨迹分析、明确起(终)点的轨迹分析和不明确起终点的轨迹分析三类。

1) 明确路径的轨迹分析，即起点、终点和路径都已给定，分析该确定路径的交通量。以图 4 为例，通过统计某时段出行链中含 A—B—E 序列的数量得到该时段内 A—B—E 路径的交通量。确定路径的轨迹结果定量表达了某一路径上的交通量，可以为交

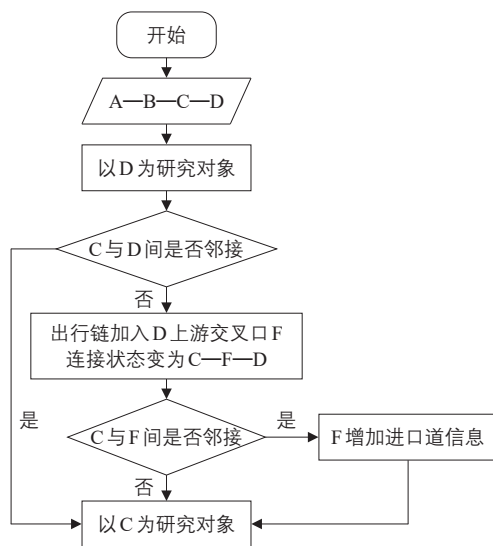


图3 第一次数据补充流程图

Fig.3 Flow chart for the first time data complementary

通仿真提供流量数据。

2) 明确起(终)点的轨迹分析, 即起(终)点已经给定, 分析经过该点的所有车辆的后序路径(或前序路径)。以图4为例, 需要得到某时段内通过A东进口道车辆后序路径。

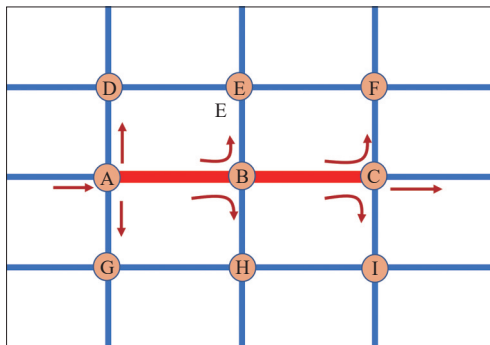


图4 路段和交叉口示意

Fig.4 Intersections and road segments

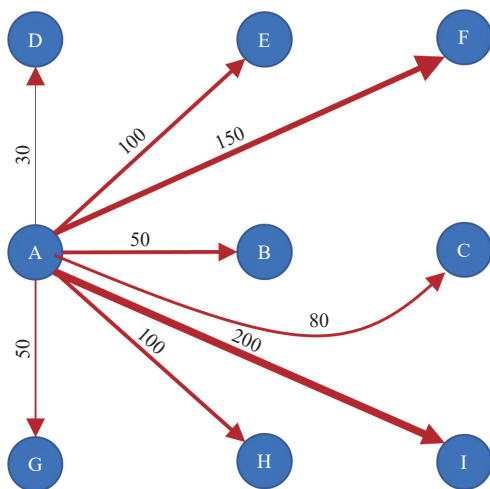


图5 确定起点的OD结果示例

Fig.5 OD results with clear starting points

通过获取某时段内经过A东进口道所有车辆的出行链, 统计出行链的后序路径和对应的频数(在实际分析中, 常常只关心OD, 分析结果为经过分析点的OD, 见图5)。获取主流向行车轨迹还可以分析主要车流是过境交通还是集散交通, 可以发现主要的交通流向, 为道路改造提供依据。有些时候, 终点(或起点)过多, 不利于发现规律, 可以框定一个范围, 对出行链进行修剪, 不分析出行链中超过范围的部分, 有助于成果的归集和展示。

3) 不明确起终点的轨迹分析, 即不给定分析车辆的起点和终点, 统计所有需要分析车辆的实际路径(常常以OD进行表示)。对于城区拥堵路段, 来源比较分散的车辆宜选取经过拥堵路段作为出行链的筛选条件。筛选高峰期所有通过该路段的出行链, 进而得到经过拥堵路段所有车辆的出行路径, 这样有助于了解拥堵路段的交通特征。在局部道路治理中, 常常并不关心车辆全部路径, 与确定起(终)点的轨迹分析相同, 需要框定一个影响范围, 修剪出行链, 对出行路径进行归集。

在交通拥堵治理方面, 明确路径的轨迹分析适用于交通仿真; 明确起(终)点的轨迹分析适用于某一流量较大交叉路口的路径分析; 不明确起终点的轨迹分析适用于某拥堵路段的路径分析。

4 千岛湖镇新安东路改造实例分析

4.1 项目概况

新安东路中段是主要的进出城通道, 改造前是千岛湖镇最大的拥堵点, 晚高峰新安东路—明珠路交叉口东进口道(进城方向)排队长度可达600 m, 路段持续拥堵时间超过50 min, 是城区主要的交通拥堵路段(见图6)。

现状新安东路中段为双幅路, 双向4车道, 中间分隔带宽6 m, 单幅路面宽10 m, 两侧为商铺, 红线无法拓宽, 人行道也不宜占用。

4.2 数据来源及格式

数据来源于淳安县交警指挥中心, 共包括18个交叉口点位和14个路段点位(见图7), 是某工作日全天所有卡口点的车牌识别数据, 数据总量约60万条。数据字段包括车牌号、车型、拍摄时间、卡口位置、车型等。



图6 千岛湖镇现状路网

Fig.6 Roadway network of Qiandaohu Town

4.3 数据预处理

整个城区现状存在较多的路段卡口，点位密集，为交通分析创造了条件。对原数据进行分析后发现，车牌未识别车辆比例达10%，且交叉口现有电子警察无法抓拍专用右转车道的车辆，需要对原始数据进行处理。

传统的出行轨迹补全算法考虑距离、行程时间、轨迹偏好等一系列因素^[6]，计算非常复杂且参数选取具有较大的不确定性。采取本文提出的网络拓扑结构和大数据的方法，算法简单且准确性较高。通过大数据建立路网，除无信号控制交叉口外，路网结构基本与现状一致，说明各卡口点位数据的有效性。针对在无信号控制交叉口出现路网错乱的情况，在3个无信号控制交叉口增加虚节点。路网拓扑结构如图8所示。

原始数据有596 638条，对数据预处理后，共删除数据66 367条，增加数据94 563条。其中，新安东路中段晚高峰进城(东向西)流量原始数据1 932条，删除数据241条，恢复352条，恢复后达2 043条。通过观看视频人工得到实际高峰小时经过车辆2 163辆，恢复后数据准确率达95%，已满足分析要求。

4.4 交通分析

因为交通拥堵主要发生在晚高峰进城方向，本次对新安东路中段晚高峰进城数据进行分析。从进城流量来看，流量最大时段出现在16:30—17:30(见图9)，而通过路段(本次选择环湖北路公路大厦卡口至新安东路—新安大街交叉口卡口区段作为分析区段，全长约1.3 km)平均行程时间可以看出，主要的拥堵时段在17:00—18:00(见图10)，说明

用平均行程时间判断拥堵状态与实际较为一致，用流量判断拥堵状态会出现偏差，因为交通拥堵已经造成路段的通行能力下降。本次选择17:00—18:00作为晚高峰分析时段。

新安东路—明珠路交叉口和新安东路—龙门路交叉口原本没有右转数据，数据恢复后交叉口流量流向如图11所示。虽然两个交

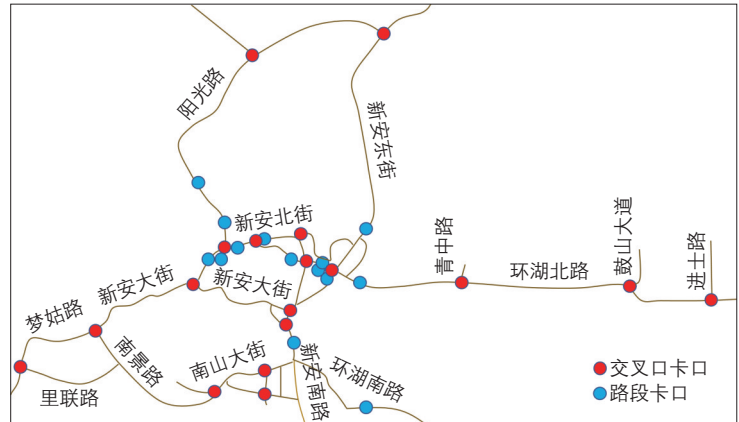


图7 各卡口位置
Fig.7 Location of vehicle checkpoints

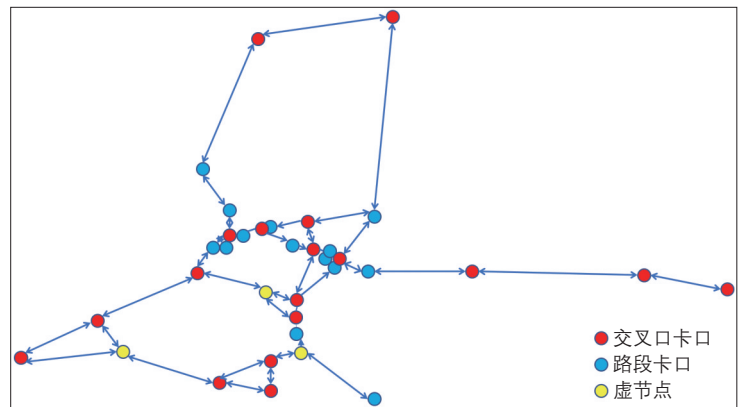


图8 路网拓扑结构
Fig.8 Topological structure of roadway network

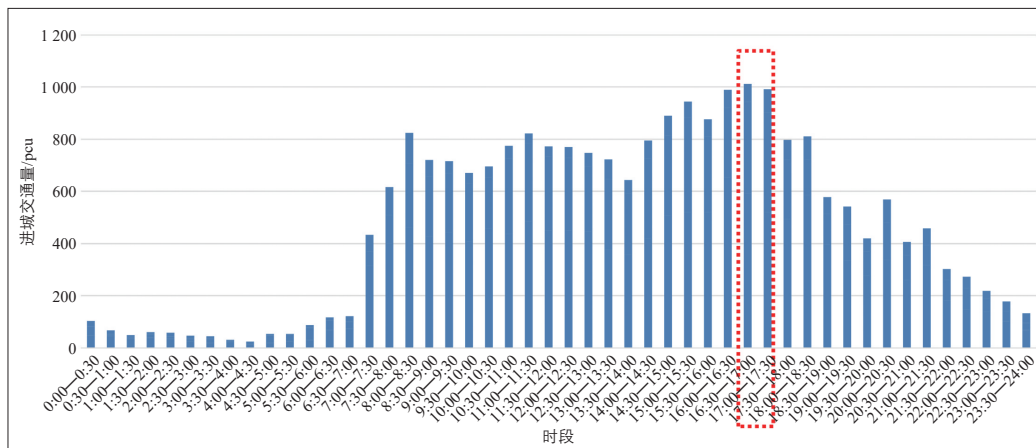


图9 新安东路中段交通量分布
Fig.9 Traffic volume on the middle section of Xin'andong Road

叉口右转均未观察到明显的拥堵,但从恢复结果看,新安东路—明珠路交叉口北进口道右转和新安东路—龙门路交叉口东进口道右转占比较大,与直观结果有较大偏差,因此右转车辆不能忽略。

首先,对排队长度最长的新安东路—明珠路交叉口东进口道(东侧进城方向)车辆进行明确起点(进城方向)的行车轨迹分析(见图12),发现有56.7%的车辆是龙门路方向,并不是普遍认为的新安南路方向,分析产生了意想不到的结论。

再对新安东路中段采取不明确起终点的行车轨迹分析,分析范围选择各向外拓展一个交叉口(见图13)。从车辆OD来看,有30.3%是由环湖北路至龙门路,有33.3%是新安东路北段至新安东路南段,这部分车辆占总车辆的比例达63.6%,是主要的交通流向,与原始认识有较大差异;而这两股车流

产生交织,进一步降低了路段通行能力。

4.5 改善方案

在未对数据分析之初,主要观点为进城通道较少,需增加进城通道,建议将新安东路中段及环湖北路排队600 m区间的中间分隔带全部拆除,拓宽车道。

通过对车辆轨迹的分析可以看出,环湖北路上游的交通量并不大,未处于饱和状态,产生排队的主要原因是新安东路中段处于交通瓶颈,北侧和东侧的交通全部汇集于此,而拆除环湖北路绿化带(进城口)没有必要。因此,从路网、路段、交叉口三个方面进行道路改造:

1) 对于路网,主要解决两股交织车流从其他道路绕行。高峰时段将新安东路北段进入新安东路南段的车辆诱导至从望江垄路绕行,而将环湖北路进入龙门路车辆诱导至

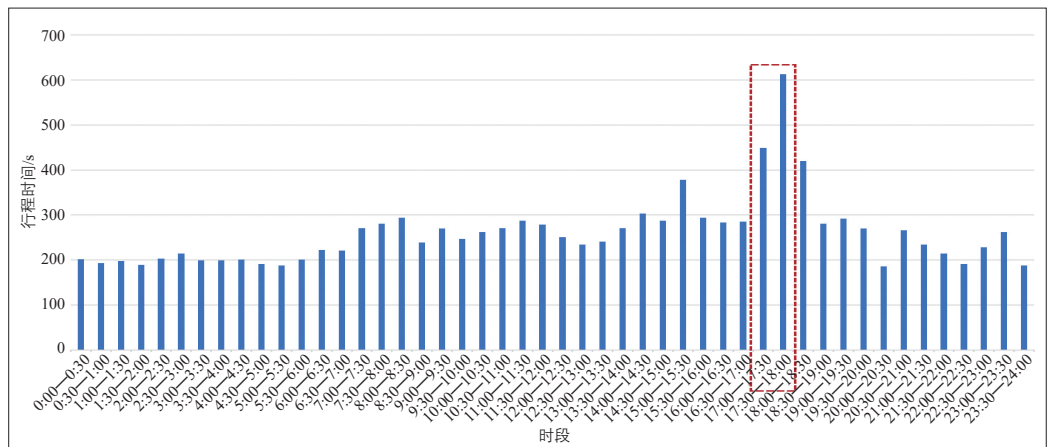


图10 分析路段平均行程时间分布

Fig.10 Distribution of average travel time on the analyzed sections

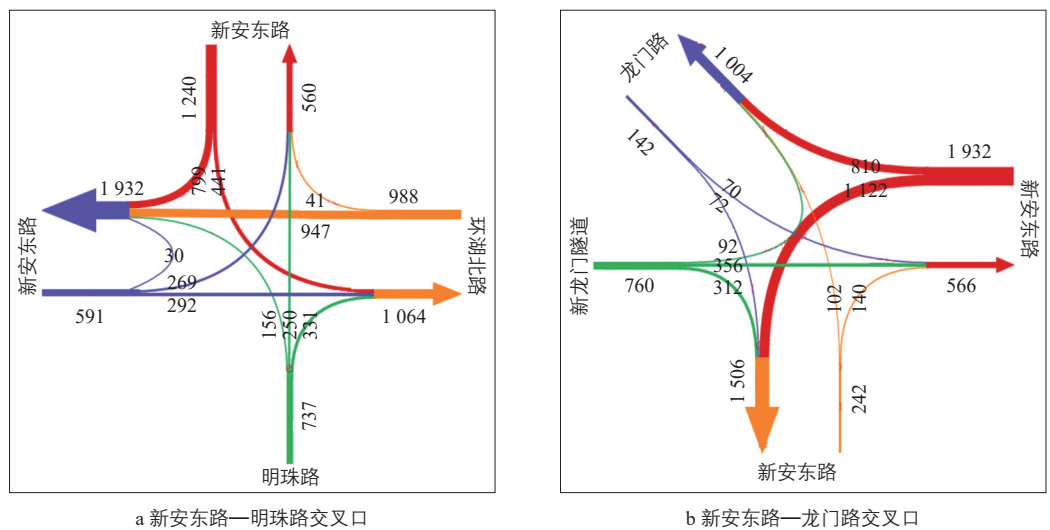


图11 晚高峰交叉口流量流向

Fig.11 Traffic flow and turning volume at intersections during evening peak

从青春路经过望江垄路绕行。

2) 对于路段，着重增加进城通道通行能力。首先，将新安东路中间分隔带拆除后新增2个进城车道(见图14)；第二，新增公共汽车站台，消除公共汽车与非机动车冲突；第三，拆除新安东路—明珠路西北角交通岛，减小缘石半径，增加交织段长度。

3) 对于交叉口，根据恢复后的交叉口流量流向进行合理渠化，这里不做展开。

4.6 交通评价

对分析后形成的改善意见进行交通仿真，仿真数据来源于实际车辆轨迹分析数据。仿真结果显示，按照改善方案对路段以及交叉口进行重新组织和渠化后，大大改善了交通拥堵(见表2)。

2019年10月，该路段按照本方案进行改造完工，高峰时段车辆排队长度仅为50m，基本解决了新安东路的交通拥堵状况。

本工程通过卡口车牌数据分析，明确新安东路拥堵的主要原因，有针对性地提出改善方案，一方面解决了交通拥堵问题，另一方面减小了投资成本，产生了较好的社会效益和经济效益。

5 结语

本文提出利用卡口车牌数据进行交通拥堵改善的思路与方法。针对原始数据丢失严重的问题，提出基于路网拓扑关系，运用大数据进行数据补全方法。根据分析目的的不同，将行车轨迹分析分为明确路径的轨迹分析、明确起(终)点的轨迹分析以及不明确起终点的轨迹分析三类，并对其分别进行探讨。

本文在数据预处理上，对于车牌识别错误的数据进行直接删除，丢失了很多可以利用的数据，而对于添加的丢失数据仍无法添加关于时间的信息，对这两个问题应进行进

一步研究，以提高原始数据的准确性和完整性。

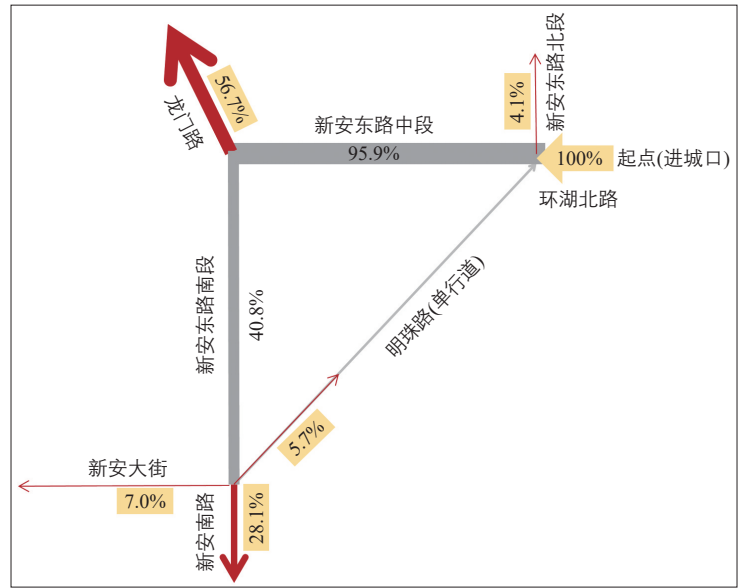


图12 新安东路—明珠路交叉口东侧进城方向行车轨迹比例
Fig.12 Ratio of vehicles to downtown areas from the eastern side at the intersection of Xin'andong Road and Ming Zhu Road

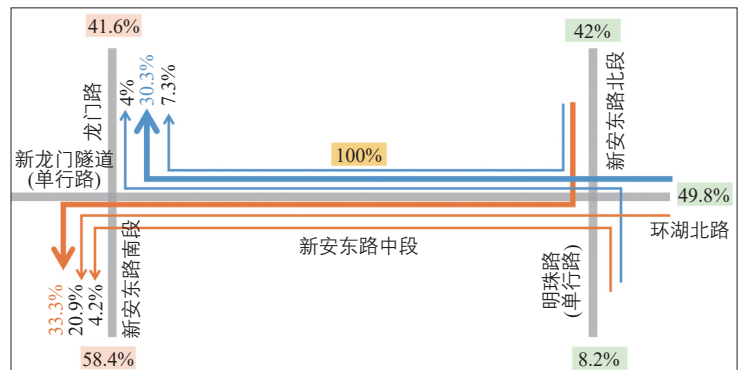


图13 新安东路中段行车轨迹比例
Fig.13 Vehicle track through the middle section of Xin'andong Road

表2 路段行程时间对比

分析路段	改造前	仿真结果	改造后
环湖北路公路大厦至新安东路—新安大街交叉口	613	396	420

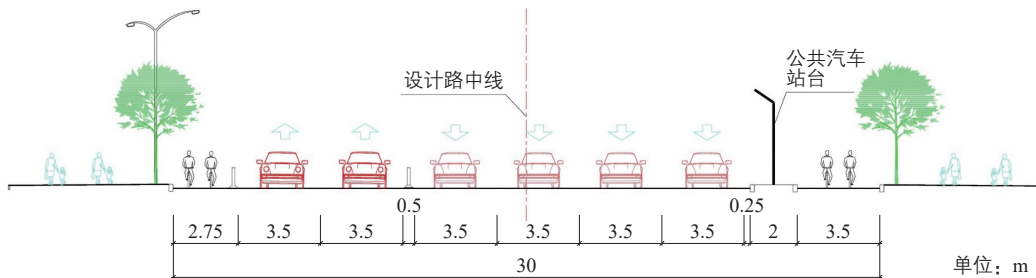


图14 改造后道路横断面
Fig.14 Roadway cross-section after improvement

(下转第78页)