DOI:10.19789/j.1004-9398.2015.02.020

基于浮动车数据的城市学校周边交通拥堵分析

刘瑶杰 张 晶* 张富华 翟 瀚
(首都师范大学 资源环境与旅游学院,北京 100048)

摘 要

交通拥堵是当今各大城市的普遍现象. 城市学校在上下学时间段对其周边的道路产生了明显的交通压力. 本文将开学后某一周的浮动车数据与道路路网进行匹配,计算出道路的平均车速和道路拥堵级别. 选取具有代表性的5所北京市小学作为研究对象,根据不同学校,不同时间的各种情况对学校周边道路进行拥堵分析. 采用空间统计分析中的空间自相关性分析方法研究学校对道路的空间上的影响. 在时间尺度上,采用统计分析和对比分析探讨学校对其周边道路的时间上的影响. 研究结果表明: ①学校对道路的影响具有空间自相关特征,且呈现距离扩散的特征; ②学校对道路的影响时间上体现在上下学时间段; ③人数越多的学校对周边道路的影响也越大,周边道路的等级越高其缓解拥堵的能力也越高.

关键词: 浮动车数据,学校,交通拥堵,数据挖掘.
中图分类号: U121

0 引 言

交通拥堵实际上是指同一辆车经过规定距离时所要花费的时间长,其主要的考核是时间的长短. 现阶段,道路的交通拥堵分析的主要方向包括: 关键位置或整体路网的拥堵度分析、交通拥堵度判定或检测分析、交通拥堵模型分析. 国外研究交通拥堵的时间较早,Bando M. 提出了基于交通动力学模型的考虑每辆车的运动方程的交通拥堵演变模型(Bando M et al,1995). Downs A. 提出了交通高峰不同交通环境下的交通拥堵模式(Down A,1962). 国内的最新研究有: 杨祖元、鲁小丫、陈阳舟、祁宏生等对交通拥堵点判别进行了研究(杨祖元等,2008; 鲁小丫等,2012; 陈阳舟等,2011; Hong et al,2011); 高自有、刘友军、朱琳、张水潮等对城市交通拥堵的特征分析进行了研究(高自有等,2011、刘友军等,2013、朱琳等,2011、张水潮等,2009).

浮动车获得准确实时的动态交通信息的优势有成本低且效率高,实时性强,覆盖范围大. 浮动车信息(FCD)采集技术是目前国际上采集道路交通信息的先进技术手段,FCD经过汇总、处理后生成反映道路拥堵情况的交通信息,能够为交通管理部门和公众提供动态、准确的交通控制、诱导信息. 在我国交通拥堵比较严重的大城市,比如北京、上海、广州、深圳等地均开始了对浮动车技术的深入研究和应用推广(北京市交通信息中心,2008). 利用浮动车数据可以获得城市交通模式,并从中解释城市内的土地利用变化(刘瑜等,2012),可以分析区域内交通拥挤的成因以及交通拥挤的时空演变规律(孙健等,2011). 利用浮动车数据还可以分析城市居民的出行模式(童晓君等,2012; 贾宁等,2013). 利用浮动车数据还可以分析降雨、内涝对城市道路交通的影响(曾伟良等,2011; 刘军等,2013)以及对城市道路交通异常事件进行检测(曹成孝,2013). 本文从分析学校拥堵的角度出发,对使用浮动车获取路况信息的技术进行了研究,并运用道路路况信息分析不同学校在不同时间段对周边道路拥堵度的影响. 利用地学分析中的空间叠加分析和空间统计分析等方法分析出交通拥堵的时空差异性特征,为交通管理、校车规划等提供参考.

1 拥堵时空特征的分析方法

在像北京这样的特大城市中,交通拥堵分为常

收稿日期: 2014-02-01
* 通讯作者

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

态拥堵和偶然拥堵. 常态拥堵就发生在特定的时间和空间范围,并且,常态的拥堵具有十分明显的时空分布特征. 对空间和时间的拥堵特征的进行研究可以分析出拥堵从形成到消失和传播的时空规律. 对于学校而言,其相对较为固定的通勤时间和人数为拥堵的时空特征研究提供了良好条件.

浮动车数据信息中包括经纬度位置和运行时间的信息. 通过将出租车轨迹数据与城市道路网络以及其他社会经济属性进行叠加,能够挖掘出一些潜在有价值的信息,比如居民出行时间、居民出行热点地区(即居民进出量较大的区域)、人均停留时间、人均出行距离和时间以及道路的平均车速信息(即道路拥堵度信息).

学校上下学对周边道路状况的时空分布影响是本文主要研究内容. 学校上学放学时间对道路的影响反映了学校的时间影响因子,学校的校址分布反映了学校的空间影响因子,学校的排名、人数等属性对路况的影响反映了学校的属性影响因子. 因此研究学校对周边道路拥堵度的影响有利于疏导城市交通引导发展校车的规划.

1.1 拥堵的空间特征分析方法

学校产生通勤人员的区域出现在校门的位置上,因此可以将学校的空间信息存储在一个位于校门位置的点中.

由于道路本身的属性和背景路况的影响,要对路况进行归一化处理. 归一化后的路况数据能够很好的反映出道路的路况变化情况. 学校对道路路况的影响体现在两者的空间关系上. 将离学校最接近的道路与学校的距离设为 1,则与这条道路相连通的道路的距离为 2,其他的道路离学校的距离都设为 3(如图 1). 然后通过空间统计分析方法分析道路网络的空间相关性以及由道路与学校之间的距离引起的道路路况的变化特征.

空间自相关是空间统计分析的前提条件,是认识空间分布特征的一种常用方法(何宗贵等,2008). 空间自相关分析是研究空间中某位置的观察值与其相邻位置的观察值是否相关以及相关程度的一种空间数据分析方法(陈小勇等,2000).

空间权重矩阵是时空自相关性度量的前提,是对空间对象邻接关系的定量测度,表达了不同空间对象之间的邻接关系. 线状数据的空间权重矩阵是根据线与线之间邻接关系规则确定的,规则可表达为:

$$w_{ij}=\begin{cases}1, & \text{当 } i \text{ 和 } j \text{ 相邻}\\ 0, & \text{其他}\end{cases} \tag{1}$$

根据该式,图 1 中的十字路口的 a、b、c、d、e、f、g、h 八条路的空间权重矩阵为如图 2 所示.

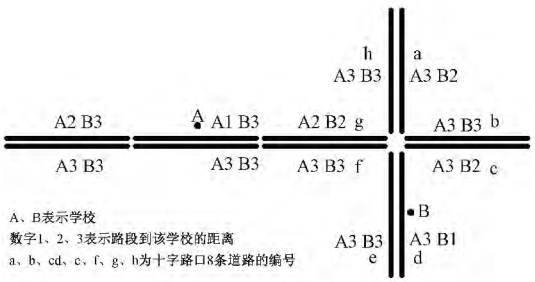


图 1 学校与道路之间的距离

	a	b	c	d	e	f	g	h
a	0	1	0	1	0	1	0	0
b	1	0	0	0	1	0	1	0
c	0	0	0	1	0	1	0	1
d	1	0	1	0	0	0	1	0
e	0	1	0	0	0	1	0	1
f	1	0	1	0	1	0	0	0
g	0	1	0	1	0	0	0	1
h	0	0	1	0	1	0	1	0

图 2 十字路口 8 条道路的空间权重矩阵

常用的全局空间自相关指数为 Moran's I 统计量,具体计算为(王远飞等,2007):

$$I=\frac{n}{\sum_{i=1}^{n}\sum^{n}\sum_{j=1}^{n}w_{ij}}*\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{2}$$

式中 n 表示观测单元数目,w_{ij}表示互为邻居单元之间交互权重值,x_i 和 x_j 表示单元观测值,$\bar{x}$ 表示所有观测值得均值. Moran's I 取值范围在(-1,1),如果空间过程不相关,则 I 的期望接近 0,当 I 取负值时,一般表示负的自相关,取正值表示正的自相关. I 的绝对值越大,属性与空间的相关性越大,I 的绝对值越小则两者的相关性越小. 计算出 I 指数后,还需要对其进行检验,一般采用 Z 检验:

$$Z_I=\frac{I-E(I)}{\sqrt{Var(I)}} \tag{3}$$

其中 E(I) 为空间自相关指数的期望值,Var(I) 为标准差. 通过对比不同时刻下的全局空间自相关指数可以发现学校对周边道路的距离影响程度.

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

1.2　拥堵的时间特征分析方法

在确定了学校对周边道路的空间分布范围后，我们可以针对学校影响最大的道路在时间轴上进行分析. 由于学校的通勤时间相对固定，对拥堵的分析可以集中在工作日的上午上学时段和下午放学时段以及周末的相同时段.

工作日时，对上下学时段进行统计分析，通过分析学校上下学时间与道路路况变化时间节点的时间间隔同学校人数、学校位置之间的关系得到学校对道路路况影响的时间特征. 根据上学和放学时间段的交通背景不同、通勤特点的不同解释学校对道路路况的影响特征. 学校的周末通勤特点与工作日相差甚多，因此通过分析两者之间的不同可以得到学校对路况影响的程度.

2　实验数据的处理与结果

2.1　实验数据

实验利用北京市 10357 辆出租车从 2008 年 4 月 14 日到 20 日共 7 天的 GPS 轨迹数据，结合北京市 5 所学校对其周边道路交通拥堵度的影响进行分析. 其中 4 月 19 日和 4 月 20 日为双休日时间，学校放假.

浮动车数据的特点是 GPS 轨迹数据只包含了轨迹点的时间和坐标，没有运动方向和车辆的运行状态. 出租车轨迹数据由一系列的轨迹点 P_i 组成，即 P_i 为其中的一条记录. 其中(x，y) 表示轨迹点的空间经纬度位置，t 表示运营时间. 在进行数据分析之前，先要对出租车轨迹数据进行预处理，剔除错误数据.

实验对象是北京市不同区域位置的 5 所小学. 各所小学之间特点不同: 1、人数不同. 人数从大到小的顺序为海淀实验小学(3 530 人) 、海淀五一小学(3 364 人) 、光明小学(1 157 人) 、板厂小学(934 人) 、六道口小学(378 人) . 按人数可以分为大中小三个档次. 2、周边道路类型不同. 海淀实验小学周围是西三环主干道路，海淀五一小学和六道口小学周围道路是较宽的道路，光明小学和板厂小学周围的道路则相对较窄. 3、地理区位不同. 光明小学和板厂小学在二环内的西南角，海淀实验小学在三环周边，海淀五一小学西四环以外，六道口小学在北四环以外.

2.2　浮动车数据的处理方法

2.2.1　浮动车数据的道路匹配方法

道路匹配技术是 FCD 处理的关键技术之一. 只有判断出车辆在哪条道路上行驶，才能将 GPS 数据转化为道路的交通状态(WANG et al，2009) . 对 FCD 轨迹数据进行道路匹配首先需要进行数据预处理. 实际中，因为车辆被建筑物遮挡等客观因素，个别 GPS 数据会产生漂移误差. 因此，在计算前对接收到的 GPS 数据进行预处理十分必要. 数据的预处理包括两个方面: 自身错误过滤和静态漂移过滤(如图 3 中第 5、6 点) .

地图匹配算法是基于轨迹点的缓冲区与路网的叠加分析，直接以轨迹在道路的垂足为道路匹配点进行匹配(图 3) ，从而减少了匹配时的运算量. 但这样同样也降低了轨迹的匹配精度，它会导致道路平均车速获取时，忽略 GPS 精度误差引起的逆行的现象(如图 3 中的 11、12 点) .

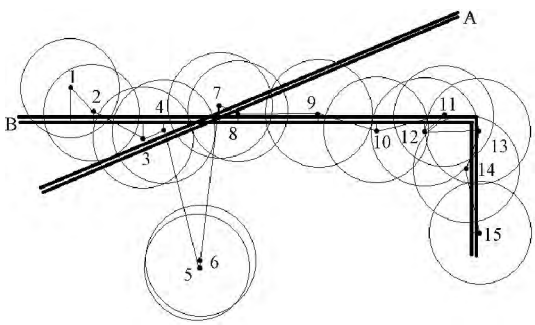


图 3　地图匹配算法的基本原理

2.2.2　道路平均车速以及道路路况的获取方法

浮动车轨迹数据通过地图匹配后，可以通过 P_i 前后点的时间和距离算出 P_i 前后点之间车辆的平均速度，并将该值作为 P_i 的速度. 由于存在多辆车驶过同一路段的情况，即该路段对应多条运行速度记录，可用式(1) 计算该路段的平均运行速度:

$$\overline{V} = \sum_{i=1}^{n} L_i / \sum_{i=1}^{n} (L_i / V_i) \qquad (4)$$

式中: $\overline{V}$ 为平均运行速度/(km/h) ; L_i 为第 i 条记录中车辆在该路段上行驶的距离/km; V_i 为第 i 条记录中车辆在该路段上的行驶速度/(km/h) (朱丽云等，2008) . 误差中的逆行误差是指图 3 中第 12 点，通过时间判断 12 点在 11 点之后，但由于本算法的原因，匹配时忽略逆行的 12 点，对其不进行匹配.

得到每个轨迹点的速度后，可以通过公式(4) 获得每条道路的平均车速. 将车速按照 < 10 km/h、10 km/h ~ 35 km/h、> 35 km/h 分为拥堵、缓行、畅通. 在道路路况表示中分别以 3、2、1 表示.

2.3　浮动车处理的实验结果

实验的结果主要是浮动车数据经过地图匹配后

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

的结果以及路网道路的路况信息. 属性表中的 Time 字段是出租车运行时间与初始时间的秒差,Speed 字段是由公式(5)计算得出.

第 i 点的速度 = 前后点距离 / 前后点时间差 (5)

前后点的距离由两点的坐标经过高斯投影后,在路网中计算两点之间的路网距离. 得到所有点的平均速度后,利用公式(4)计算出所有道路的平均车速,再按照车速区间划分道路路况. 最终得到的实验结果是 5 所学校周边道路一周的路况信息.

3 拥堵时空特征的分析

3.1 拥堵的空间特征

3.1.1 拥堵的空间自相关特征

同一时间段,学校对其周边道路路况的影响体现在学校周边道路和学校之间的距离以及连通性的差异. 道路与学校的距离不同、连通性不同,学校对道路影响也不同.

利用 Moran 指数公式计算,得到上学时段不同时刻的全局 Moran's I 指数(如图 4). 从中,我们可以发现在 7:55 和 7:57 时刻,全局 Moran's I 指数最高,达 0.3 以上. 这表明在这几分钟内的道路拥堵程度明显空间正自相关的,也就是道路拥堵程度具有距离衰减的特征,即道路离学校越远路况越好,离学校越近路况越差. 同样也可以发现在 7:45 时刻出现第三大值,其原因是 7:55 和 7:57 时刻的 2 个大值是由 4 所学校上学引起的路段拥堵,而 7:45 时刻出现的第三大值是由于 1 所学校上学引起的.

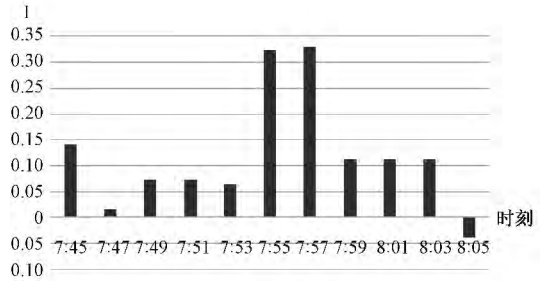
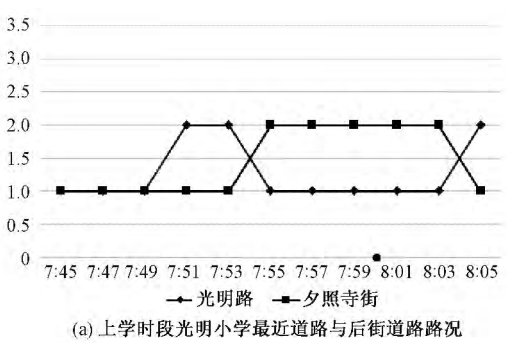


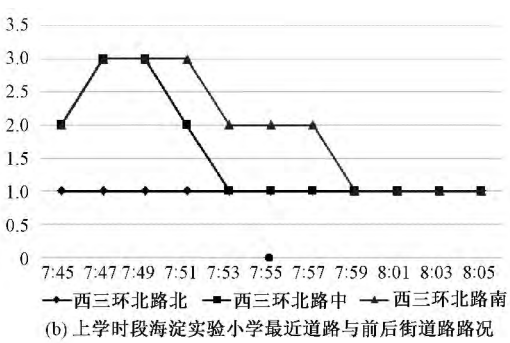
图 4 不同时刻的全局 Moran's I 指数

3.1.2 拥堵的距离扩散特征

与学校最近的道路的前后接道路的相关性也不是很高,但相互之间存在明显的时间推移特征与距离扩散特征. "夕照寺路"是"光明路近"的后接道路(如图 5-a),可以发现"夕照寺路"的道路路况在4 分钟之后呈现于"光明路近"相同的级别,但其拥堵的持续时间却比"光明路近"更长. "西三环北路北"和"西三环北路南"分别是"西三环北路中"的前接和后接道路(如图 5-b),这幅图进一步证实了图 5-a 反映的特点. 这两幅图表现出了明显的道路拥堵的距离扩散现象,在校门最近道路的拥堵会传导至后面的道路,且拥堵的消散需要 5 至 7 分钟的时间.



(a) 上学时段光明小学最近道路与后街道路路况



(b) 上学时段海淀实验小学最近道路与前后街道路路况

图 5 学校影响的距离扩散特征(横轴上的黑点为上学时间)

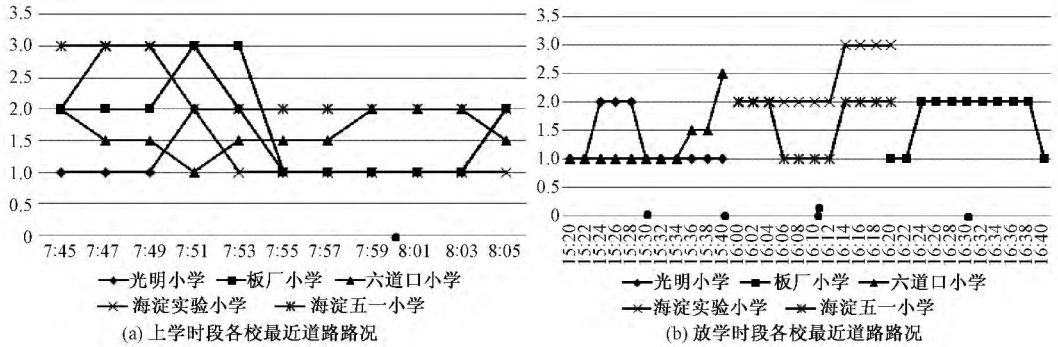
由上述分析可得,学校对其最近的道路的影响更大,因此我们分析各校对其最近道路的路况的影响可以代表学校对周边道路的影响程度. 同时,由学校产生的周边道路拥堵会扩散至下游道路,且拥堵消散时间需要一定时间.

3.2 拥堵的时间特征

3.2.1 一天内学校对周边道路的影响特征

学校对一段特定时间段内对周边道路的影响是有特点的. 上学时刻前的 5 分钟至 15 分钟,每所学校的门口道路路况均开始形成各种程度的拥堵. 当过了上学时间后,学校门口的交通拥堵程度在几分钟之内降低. (图 6-a)放学时间段内,由于各学校之间的差异,时间轴线拉的较长. 光明小学和六道口小学由于放学较早,其共同的规律较少,但可以看出学校放学还是会在道路上产生不同形态的拥堵. (图 6-b)

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net



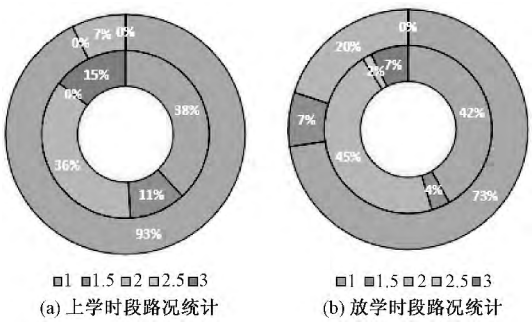
(a) 上学时段各校最近道路路况　　(b) 放学时段各校最近道路路况

图 6　上学和放学时间段的各个学校最近道路的路况

从学校规模上看,六道口小学的人数最少,因此引起的交通拥挤程度最小. 光明小学和板厂小学的人数是相对较多的,因此交通拥挤程度相对较高. 海淀实验小学和海淀五一小学的人数是最多的,反映在交通拥挤程度上也呈现出最高的拥堵状态. 同样的规模下,海淀实验小学由于位于环路的一侧,交通疏导能力比海淀五一小学强,因此交通拥堵的时间较短.

3.2.2　工作日与双休日学校对道路拥堵影响特征

除了同一天中学校对道路有一定的影响规律,由于学校的影响,工作日与周末的路况也有不同规律. 图 7-a 是周末与工作日各校周边道路在上学时段的路况对比,图 7-b 是周末与工作日各校周边道路在放学时段的路况对比,两图可以明显发现周末的路况比工作日好. 进一步对比上下学之间的路况,发现上学时段差异很大,道路畅通的情况变化达到 55%,而放学时段的道路畅通差异为 31%,由此说明,即便上学时段的道路拥堵因为整体交通由于受到早高峰的影响变得更严重,但学校对其周边道路的影响还是十分重大的.



(a) 上学时段路况统计　　(b) 放学时段路况统计

图 7　周末与工作日各校上下学时段道路路况对比
(内圈为工作日路况统计,外圈为周末路况统计)

4　结　论

本文提出了利用浮动车轨迹数据,结合路网分析了学校对道路拥堵状况的影响的方法,对学校对周边道路的影响时空特征进行分析. 实验表明: 学校对周边道路的主要影响特征有: (1) 在上下学时段学校对周边道路影响具有空间自相关的特征,自相关度在产生通勤流的时刻达到最大值. 道路离学校越近,受到学校的影响越大. 道路拥堵的特征呈现距离扩散的特征. (2) 工作日时,学校周边拥堵特征体现在上学前高峰和放学高峰. (3) 人数越多的学校,其对周边道路的影响也越大. 周边道路的等级越高其缓解拥堵的能力也越高.

进一步的研究方向是增加寒暑假的对比,并且增加学校的样本数和更多的分析因素,从而进一步分析学校对周边道路路况的影响的时空特征.

参 考 文 献

[1] Bando M, Hasebe K. Dynamical model of traffic congestion and numerical simulation [J]. Phys Rev E, 1995, 51(2): 1035 – 1042.

[2] Downs A. The law of peak-hour expressway congestion [J]. Traffic Quarterly, 1962, 16(3).

[3] 杨祖元, 黄席樾, 杜长海, 等. 基于 FFCM 聚类的城市交通拥堵判别研究 [J]. 2008.

[4] 鲁小丫, 宋志豪, 徐柱, 等. 利用实时路况数据聚类方法检测城市交通拥堵点 [J]. 地球信息科学学报, 2012, 14(6): 775 – 780.

[5] 陈阳舟, 田秋芳, 张利国. 基于神经网络的城市快速路交通拥堵判别算法 [J]. 计算机测量与控制, 2011, 19(001): 167 – 169.

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

[6] Qi H, Wang D, Song X. On the critical conditions of traffic jams [J]. Journal of Southeast University, 2011, 27(2): 180 – 184.
[7] 高自友，龙建成，李新刚. 城市交通拥堵传播规律与消散控制策略研究 [J]. 上海理工大学学报，2011，33(6)：701 – 708.
[8] 刘有军，田聪. 不同路网形态下城市交通拥堵特性分析 [J]. 中国公路学报，2013，26(1)：163 – 169，190.
[9] 朱琳，于雷，宋国华，等. 基于浮动车数据的交通拥堵时间维度特征 [J]. 北京交通大学学报，2011，35(6).
[10] 张水潮，任刚，王炜. 面向交通规划的城市道路交通拥堵度分析模型 [J]. 吉林大学学报：工学版，2009，39(9)：111 – 115.
[11] 北京市交通信息中心. 基于卫星定位的交通信息采集技术及应用研究 [EB/OL]. [2008 – 06 – 26]. http://www.gpsbao.com/datum/1132.html.
[12] Liu Y, Kang C, Gao S, et al. Understanding intra-urban trip patterns from taxi trajectory data [J]. Journal of Geographical Systems, 2012, 14(4): 463 – 483.
[13] 孙健，刘琼，彭仲仁. 城市交通拥挤成因及时空演化规律分析——以深圳市为例 [J]. 交通运输系统工程与信息，2011，11(5)：86.
[14] 童晓君，向南平，朱定局. 基于出租车 GPS 数据的城市居民出行行为分析 [J]. 电脑与电信，2012，1.
[15] 贾宁，杜盛楠. 基于浮动车数据的城市居民出行行为规律分析 [J]. 科技创新导报，2013，16：157.
[16] 曾伟良，龚峻峰，何兆成，等. 降雨对城市道路交通行驶速度的影响分析 [J]. 环境科学与技术，2011，2.
[17] 刘军，宋国华，赵琦，等. 基于 FCD 的桥下积水导致城市交通拥堵点段识别方法研究 [J]. 交通信息与安全 ISTIC，2013，31(3).
[18] 曹成孝. 基于浮动车的城市道路交通异常事件检测的研究 [J]. 科技风，2013 (13).
[19] 何宗贵，韩世民，崔道. 空间自相关分析的统计量探讨 [J]. 中国血吸虫病防治杂志，2008，20 (4)：317.
[20] 陈小勇，林鹏. 我国红树植物分布的空间自相关分析 [J]. 华东师范大学学报：自然科学版，2000 (3)：1042109.
[21] 王远飞，何洪林. 空间数据分析方法 [M]. 北京：科学出版社，2007：110 – 119.
[22] Wang Dong Zhu, Dong Jiming, Li Yameng, et al. Map Matching Method Based on Zero-speed Points in Floating Car Data [J]. Journal of Transport Information and Safety, 2009, 27(6): 38 – 42.
[23] 朱丽云，温慧敏，孙建平. 北京市浮动车交通状况信息实时计算系统 [J]. 城市交通，2008，6(1)：77 – 80.

Analysis of Traffic Congestion around School Based on Taxi Trajectory

Liu Yaojie* Zhang Jing Zhang Fuhua Zhai Han

(College of Resource Environment and Tourism, Capital Normal University, Beijing 100048)

Abstract

Traffic congestion is a common problem of big cities. Schools in the city produce a significant traffic pressure on its surrounding roads on the time of going or leaving school. In this paper, the average road speed and road congestion level can be calculated by using floating car data of a week`s time since school started. Selecting 5 representative primary school in Beijing as research objects, the traffic congestion around schools can be analyzed by depending on a variety of situations of different schools and different time. The method of spatial auto-correlation analysis in spatial statistical analysis can be used on the research of the spatial impact of the road around schools. On the time scale, using statistical analysis and comparative analysis we can explore the temporal impact of the road around schools. The results showed that: (1) the impact of school on the road is spatial auto-correlative, and the feature is diffused by distance; (2) the impact of school on the road is reflected in the time of going and leaving school; (3) the more the number of schools, the bigger the affection on the road is, and the ability of alleviate congestion is better if the level of the road is higher.

Key words: floating car data, school, traffic congestion, data mining.

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net