

基于数据挖掘的电子警察执法异常分析方法及应用研究

罗强¹, 刘宗广¹, 李友欣¹, 周沛²

(1.广州市交通管理科学技术研究所, 广东 广州 510000; 2.广州市交通规划研究院, 广东 广州 510030)

摘要: 随着城市经济的快速发展, 机动车普及进程加快, 对交通基础设施、公安交通管理部门带来不小压力。电子警察的出现符合科技强警的战略要求, 缓解了警力不足的困难, 提高了对违法行为的管控力度, 加大了对违章车辆的威慑力。文中对电子警察系统在道路管理中的发展、应用及所取得的社会效益进行分析, 建立了电子警察管理评价体系, 通过挖掘电子警察执法数据, 提出了电子警察执法异常分析方法, 并对广州市电子警察执法异常进行了分析, 以提高设备的可靠性、执法的合理性与相关设施设置的科学性。

关键词: 公路交通; 交通违法; 电子警察; 评价体系; 异常分析

中图分类号: U491.5

文献标志码: A

文章编号: 1671-2668(2017)02-0049-03

道路交通违法问题与市民工作、生活息息相关, 是交通管理的重中之重。自 20 世纪 90 年代引入电子警察以来, 对道路交通违法的取证逐步从以人工为主转为以设备为主, 电子警察系统在很大程度上解放了警力, 大大缓解了交通管理人力资源短缺的问题。但由于执法录入、设施缺陷等原因造成部分被执法者对电子警察的执法质量存在争议, 如何深入挖掘电子警察执法数据, 提出执法数据异常分析方法, 切实提高设备的可靠性、执法的合理性与相关设施设置的科学性, 成为当前电子警察发展中急需解决的一大问题。但一直以来, 关于电子警察布设合理性的研究几乎是空白, 解决争议问题的更多方式是通过行政复议等行政手段, 这种被动式管理给交通管理者带来了一些负面社会效应。因此, 有必要针对电子警察进行科学评价, 主动发现存在的问题并及时加以改进, 减少不必要的争议。该文根据电子警察系统的执法数据, 采用数据分析方法准确及时地掌握交通违法行为系统中各种动态、静态信息, 提取交通违法行为在时间、空间、违法类型上的分布特征, 从多维度对交通违法行为数据进行分析, 挖掘电子警察系统中执法不合理的点位, 主动优化电子警察的布设和使用, 提高非现场执法的合理性和可靠性, 为经济和社会发展创造和谐的道路交通环境。

1 电子警察管理评价体系

主要从设备层面、使用层面和运维层面进行分

析, 构建电子警察管理评价体系 (见图 1)。其中: 设备层面指标包括设备故障率和取证合格率两方面, 主要用于保证电子警察系统正常运作, 保障系统故障率低, 发生故障时系统恢复时间短, 采集到的违法照片清晰合规, 最大限度地提高电子警察系统的可用性。使用层面涵盖执法量异常和执法模式异常两方面, 主要用于保障电子警察系统执法数据正常, 保证不会由于执法录入和设施缺陷等造成对电子警察执法质量的争议。运维层面包括设备巡查和设备修复两方面, 主要用于对运维管理的考核和运维水平的分析。

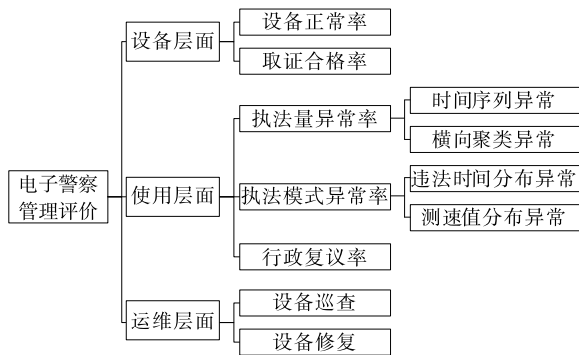


图 1 电子警察管理评价体系

由于目前电子警察管理仍处于设备功能管理阶段, 大多只关注设备是否正常拍摄、视频取证图像是否合格等, 深层次的执法管理尚未进行系统研究, 下面主要从执法管理层面研究执法数据异常判断方法, 弥补现有研究的不足。

2 执法异常分析方法

2.1 执法量异常

2.1.1 横向聚类异常判别

主要用于寻找各违法行为中执法量异常高的点位,判别方法为制作频数分布直方图。步骤如下:

(1) 确定样本顺序统计量 X_1, X_2, \dots, X_N 。

(2) 确定数据的极差 $R, R = X_N - X_1$ 。

(3) 根据数据特征确定直方图的组数 K , 一般为 10~15 较适宜。

(4) 计算组距的宽度 $h, h = R/K$ 。

(5) 计算各组的界限位。

(6) 编制频数分布表, 把多个组上下界限值分别填入频数分布表中, 并把数据表中的各个数据列入相应的组, 统计各组频数 f 。

(7) 以组距为底长、频数为高制作直方图, 通过观测直方图中不连续的样本区间, 找出明显偏离其他样本区间的点位。

2.1.2 时间序列异常判别

主要用于寻找执法量突变的点位, 通过构建各执法点位执法时间序列上的执法量曲线, 计算各时间点相较于上一时间点的变化幅度, 找出变化幅度较大的时间段。方法如下: 1) 确定样本执法时间序列统计量 x_1, x_2, \dots, x_n ; 2) 根据时间序列的数量确定执法时间序列统计量的组数 K , 一般以 10 或 15 d 为一组, 整体不少于 10 组; 3) 计算各组的日执法量平均数 \bar{x}_K ; 4) 与横向聚类异常判别方法一样, 制作频数分布直方图, 若直方图中存在不连续的样本区间, 则说明执法时间序列存在突变区间。

2.2 执法模式异常

2.2.1 违法时间分布异常判别

通过计算各点位执法量小时分布特征(数列 x) 与该违法行为总体小时分布特征(数列 y) 之间的相关系数, 考察个体点与总体特征之间的相似度, 找出相似度低(相似度低于 0.3)的违法时间分布异常点位并进行成因分析。相关系数 r 的计算公式如下:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$$

式中: x_i 为数列 x 第 i 项值; \bar{x} 为数列 x 的平均值; y_i 为数列 y 第 i 项值; \bar{y} 为数列 y 的平均值。

相关系数 r 的取值范围为 $[-1, 1]$, $r > 0$ 为正相关, $r < 0$ 为负相关。其中: $|r| \geq 0.8$ 为高度相关; $0.5 \leq |r| < 0.8$ 为中度相关; $0.3 \leq |r| < 0.5$ 为低度相关; $|r| < 0.3$ 表示相关程度极弱。

2.2.2 测速值分布异常判别

主要针对测速设备, 考察不同限速值、测速起点点位之间的超速程度分布情况, 找出与总体分布规律差异较大的点位。方法如下: 1) 查询各路段的限速值 Q 及各点位超速实测值 S ; 2) 计算超速程度 $P, P = (S - Q)/Q \times 100\%$; 3) 将超速程度分为 $< 10\%$ 、 $[10\%, 20\%]$ 、 $[20\%, 50\%]$ 、 $> 50\%$ 4 个区间, 分别计算各区间的比例; 4) 根据各点位区间比例, 找出区间比例明显偏离其他样本区间的点位。

3 广州电子警察执法异常分析

3.1 执法量异常

3.1.1 横向聚类异常

运用聚类分析法构建闯红灯违法行为电子警察各执法点位执法量直方图, 结果见图 2。从中可看出: 执法量最高的点位单独落在区间 $[2\ 000, 2\ 200]$ 内, 且区间 $[1\ 400, 1\ 600]$ 、 $[1\ 600, 1\ 800]$ 、 $[1\ 800, 2\ 000]$ 的数据量为零, 样本区间不连续。据此得单独区间 $[2\ 000, 2\ 200]$ 点位明显偏离其他样本区间, 为执法量异常点位。

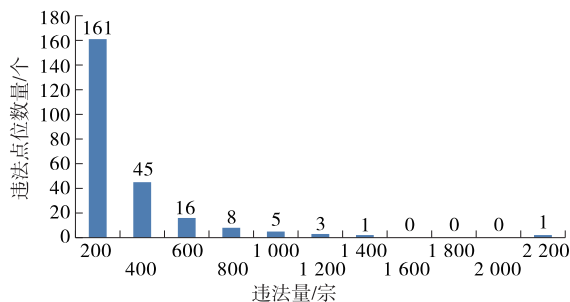


图2 闯红灯违法行为执法量直方图

3.1.2 时间序列异常

运用前文所述方法构建示例点位样本执法时间序列统计量(见图3), 根据时间序列的数量确定执法时间序列统计量的组数 $K = 11$; 计算各组的日执法量(合格数)平均数 \bar{x}_K , 得 $\bar{x}_1 = 7$ 、 $\bar{x}_2 = 65$ 、 $\bar{x}_3 = 144$ 、 $\bar{x}_4 = 68$ 、 $\bar{x}_5 = 3$ 、 $\bar{x}_6 = 5$ 、 $\bar{x}_7 = 4$ 、 $\bar{x}_8 = 3$ 、 $\bar{x}_9 = 3$ 、 $\bar{x}_{10} = 4$ 、 $\bar{x}_{11} = 5$; 制作频数分布直方图, 观测到直方图中存在不连续的样本区间, 说明该示例的执法时

间序列存在突变区间,该点位属于时间序列异常。

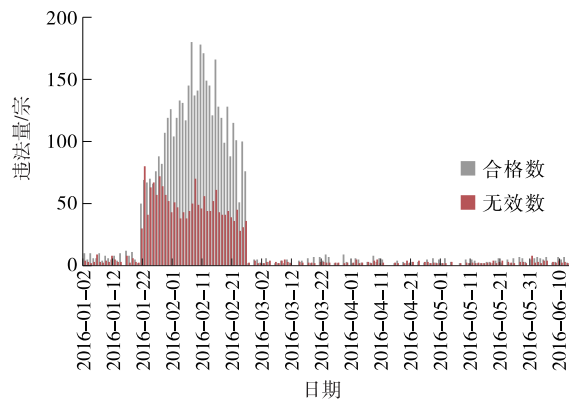


图 3 示例点位样本执法时间序列

3.2 执法模式异常

3.2.1 违法时间分布异常

示例点位小时分布特征(数列 x)与该违法行为总体小时分布特征(数列 y)见表 1,计算得示例点位相对于总体规律之间的相关系数为 -0.62 ,为负相关,示例点位为违法时间分布异常点位。

表 1 示例点位与总体规律违法量小时分布

数列 y		数列 x	数列 y		数列 x
小时	(小时违法 量/宗)	(小时违法 量/宗)	小时	(小时违法 量/宗)	(小时违法 量/宗)
0	9	36	12	33	7
1	9	23	13	36	6
2	8	38	14	38	7
3	8	37	15	38	7
4	7	21	16	36	8
5	5	12	17	29	5
6	9	12	18	18	10
7	17	13	19	11	3
8	23	6	20	13	13
9	29	12	21	13	17
10	34	8	22	12	12
11	34	10	23	13	25

3.2.2 测速值分布异常

分析不同限速值与超速起拍值点位之间超速程度的比例,结果见图 4。由图 4 可知:1) 对于限速值 80 km/h 的高速公路,超速 $20\%\sim 50\%$ 的违法行为占 95% ;对于限速值 100 km/h 、起拍值 105 km/h 的高速公路,超速 10% 以内的占 60% ,超速 20% 以内的占 90% ;对于限速值 100 km/h 、起拍值 110 km/h 的高速公路,超速 20% 以内的违法行为约占

80% ;对于限速值 120 km/h 、起拍值 125 km/h 的高速公路,超速 10% 以内的占 70% ,超速 20% 以内的占 95% ;对于限速值 120 km/h 、起拍值 130 km/h 的高速公路,超速 20% 以内的约占 90% 。均无明显异常点位。2) 仅 N 点位是限速值 110 km/h 、起拍值 120 km/h 的高速公路,其超速 $20\%\sim 50\%$ 的违法行为占 44.4% ,相比限速值 100 km/h 、起拍值 110 km/h 及限速值 120 km/h 、起拍值 130 km/h 的路段,超速 $20\%\sim 50\%$ 的违法行为所占比例偏高,相对比较异常。

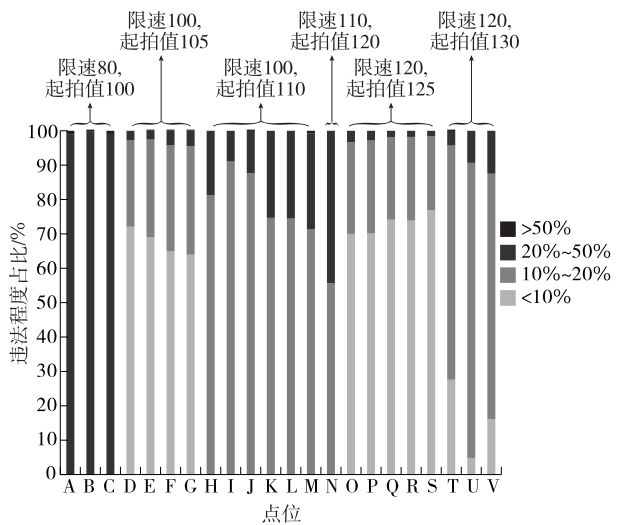


图 4 高速公路超速违法行为超速程度分布

进一步研究可发现,N 点位拍摄的车辆速度不存在 $124\sim 131\text{ km/h}$ 区间段的数据,明显不合理,可能存在设备故障(见图 5)。

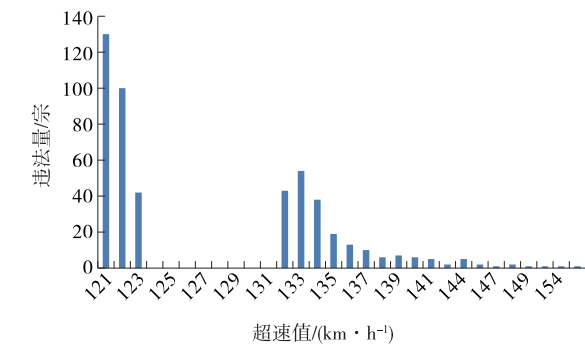


图 5 超速违法行为异常点位 N 的速度分布

4 结语

广州市交通电子警察系统经过 16 年的建设,已形成较大规模,非现场执法成为交通执法的主力军。

(下转第 55 页)

续表 2

交通大区	交通量—停车需求模型预测值			静态交通发生率模型预测值			最终预测值		
	居住停车需求	非居住停车需求	总停车需求	居住停车需求	非居住停车需求	总停车需求	居住停车需求	非居住停车需求	总停车需求
13	19 900	15 000	34 900	22 500	17 000	39 500	21 200	16 000	37 200
14	12 900	17 700	30 600	15 200	20 800	36 000	14 050	19 250	33 300
15	18 200	24 400	42 600	20 900	28 100	49 000	19 550	26 250	45 800
合计	585 300	308 500	893 800	662 900	350 500	1 013 400	624 100	329 500	953 600

根据停车需求预测结果,规划年(2020年)乌鲁木齐市主城区机动车停车泊位需求总量约为95万个,居住类需求、非居住类需求分别约为62、33万个,占总量的65%、35%。根据国外经验,综合城市发展规划、汽车保有量水平、车辆出行特征、停放特征等,合理的车位与车辆之比应为1.2~1.4个/辆。从乌鲁木齐市用地条件、经济水平、城市环境及现状停车设施供应状况来看,到2020年车位与车辆之比约为0.95个/辆,未达到要求的下限(1.2个/辆)。

为缓解主城区停车压力,抑制过大的停车需求,必须采取停车措施,包括:提高停车收费,控制需求;结合老城区改造,严格执行配建标准;建设占地少的立体停车库和地下停车场;错时利用现有配建停车场;适度增加夜间路内停车供应规模;建立停车信息化系统,提高车位利用率。

3 结语

通过对乌鲁木齐市主城区停车现状的调查与分析,掌握了其停车特征及问题所在;在分析国内外停车需求预测方法的基础上,组合交通量—停车需求预测法和静态交通发生率法预测了乌鲁木齐市主城区未来年停车需求量,为进一步进行停车规划和管

理奠定理论基础。

参考文献:

- [1] 程志华.乌鲁木齐市中心城区停车问题分析及对策研究[J].公路交通科技:应用技术版,2014(7).
- [2] 徐雷,刘冰,张涵双.停车需求预测和泊位供给策略分析[J].交通科技与经济,2014,16(2).
- [3] 过秀成.城市停车场规划与设计[M].北京:中国铁道出版社,2008.
- [4] 关宏志,王鑫,王雪.停车需求预测方法研究[J].北京工业大学学报,2006,32(7).
- [5] 杨中良,裴玉龙.哈尔滨中心城区停车现状分析及需求预测[J].哈尔滨建筑大学学报,2002,35(3).
- [6] 文雅,晏克非,崔晓天.大型车库开发的交通影响分析研究[J].现代城市研究,2004(4).
- [7] 温旭丽,孙焱飞,罗婷.基于交通特征的停车需求预测模型构建研究[J].公路交通科技:应用技术版,2014(4).
- [8] 孙吉瑞.城市混合用地停车需求预测模型研究与实践[J].交通与运输:学术版,2016(1).
- [9] 关宏志,刘小明.停车场规划设计与管理[M].北京:人民交通出版社,2003.

收稿日期:2016-10-05

(上接第51页)

通过深度挖掘电子警察运行产生的各种数据,开展交通违法数据统计、关联分析和评估,可指导电子警察在布局定点、智能化方向、执法模式等方面的发展策略制定。该文所提出的电子警察执法异常分析方法,能进一步强化设备管理,提高设备的可靠性、执法的合理性及相关设施设置的科学性,使电子警察发挥更大效用。

参考文献:

- [1] 温志刚.论电子警察在我国交通管理中的应用[J].交通

标准化,2010(13).

- [2] 李翰.交通电子警察取证问题的法律分析[D].广州:华南理工大学,2012.
- [3] 白红杰.浅论“电子警察”的应用与发展[J].科技风,2012(11)3.
- [4] 徐茜茜,朱进.规范和加强交通电子警察执法[J].江苏警官学院学报,2008,23(6).
- [5] 邵志骅,许卉莹,崔林山.公安交通管理非现场执法监管模式研究与实现[J].中国公共安全:学术版,2015,41(4).

收稿日期:2016-10-20