

简述高速公路收费站出入口交通流量分析及预测

王玉标¹ 赵宇鸿¹ 崔毅¹ 魏永福¹ 王崇宏²

(1 山西省交通规划勘察设计院有限公司 山西省太原市 030000; 2 山西省智慧交通研究院有限公司 山西省太原市 030000)

摘要: 高速公路作为我国的基础设施,对推动我国经济社会发展具有重要作用。为充分掌握高速公路收费站交通流的特征,助力智慧高速公路建设,本文重点对高速公路收费站出入口交通流量进行分析和预测。在具体的研究中,以某收费站为例,采用科学计算方法对该收费站的出入口交通流量进行分析,分析结果表明,该收费站出入口交通流量之间存在极强的正相关性,且出入口交通流量变化呈现出周期性特征。结合这一特征,通过构建长短期记忆网络模型的方式,基于收费数据和交通流量分析结果基础上对该收费站出入口交通流量进行预测,为交通管理者的管理决策提供参考。

关键词: 高速公路;收费站出入口;交通流量

引言

围绕交通强国战略,高速公路交通运营管理现代化建设,是我国公路事业发展的重要方向。与此同时,为满足国民经济发展对公路基础设施的要求,随着我国高速公路路网的不断完善,依托大数据为传统交通基础设施赋能,推进高速公路“一张网”立体布局与交通运输领域科技创新,实现高速公路交通运营管理“智慧感知、智慧管理、智慧服务”,是构建现代化综合交通体系的重要前提与保证。

而在推进智慧高速公路交通运营管理系统建设过程中,交通流量的分析及预测是重点内容。通过分析和预测交通流量,可为高速公路运营、收费等业务的开展提供支持。尤其是在高速公路全国联网的背景下,高速公路收费站作为交通路网的敏感点,时常出现拥堵的现象,在遇到节假日时,更是承受着巨大的车流压力。此外,收费站出入口会产生大量的交易数据,这些数据包含详细的车辆信息,如何高效合理地利用这些数据已经成为目前研究的热点。因此,有必要利用先进的技术手段,基于高速公路收费站出入口的收费数据,对高速公路交通流量的分析、预测和可视化展示,挖掘现有数据中可靠的信息,帮助出行者规划出行时间和路线,分散收费站出入口的交通压力,从而助力高速公路运营管理效率和水平的提升。

1、研究对象

本文以某收费站 2021 年收费数据为研究对象。以 1 个工作日为单位对该收费站出入口 2021 年全年的车流量进行统计,共统计有 500 条记录,在每条记录中含有时间、入口车流量、出口车流量、车辆详细信息等信息。

2、收费站出入口交通流量分析

2.1 分析方法

本文主要采用科学算法对该收费站出入口交通流量进行分析。科学计算方法主要是用于揭示各变量之间的相关性,通过这一方法可获得皮尔森相关系数 r 和显著性水平 p 这两个重要指标, r 可准确反映变量之间的相关程度,取值范围在 $[-1,1]$; p 具体指的是不相关的概率,是用来衡量变量之间的关系,若相关系数为 1,则 p 为 0。在统计学中,当 p 大约 0.05 时,说明数据之间的差异无显著意义,不具有统计学意义。

2.2 分析结果

(1) 相关性分析

为进一步探讨该收费站出入口交通流量的相关程度,采用科学算法计算得出该收费站出入口交通流量皮尔森相关系数。通过计算可知,该收费站出入口交通流量之间的皮尔森相关系数 r 为 0.84。在此基础上,通过 T 检验得知 $P < 0.05$,这说明该收费站出入口交通流量之间确实存在关系。而对于该收费出入口交通流量之间的相关程度,依据 r 的取值范围进行判断,具体见表 1。通过计算得知 r 为 0.84,在 $[0.8,1.0]$ 取值范围,

这说明该收费站出入口交通流量之间有着极强的正相关关系。

表 1 收费站出入口交通流量之间的相关程度

r 的取值范围	相关程度
[0,0.2]	极弱相关或不相关
[0.2,0.4]	弱相关
[0.4,0.6]	中等程度相关
[0.6,0.8]	强相关
[0.8,1.0]	极强相关

(2) 变化规律分析

基于该收费站 2021 年全年 365 天的出入口交通流量数据基础上,对其出入口交通流量变化规律进行分析,具体见图 1(均做放大处理)。通过观察图 1 可得知,该收费站出入口交通流量的变化具有周期性的规律,周期为 3~7 天。

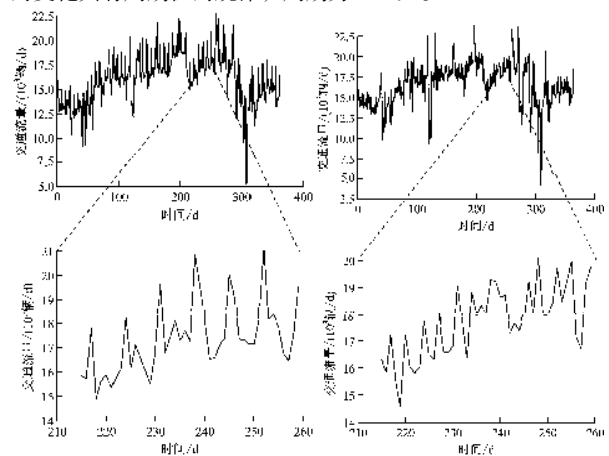


图 1 该收费站出入口交通流量变化图

为保证获得的周期性结果准确、可靠,先是将获得的时序数据转换为三角函数的线性组合,由此获得各展开项的系数,即傅里叶系数。在傅里叶系数越大的情况下,说明所对应的正弦波周期是这份数据的周期的可能性越大。将变换所得的傅里叶系数中最大的前 10 个值和对应周期,分别对应该收费站出入口交通流量数据,具体可见下表 2 所示。

表 2 出入口交通流量对应的傅里叶系数和可能的周期

序号	出口		入口	
	傅里叶系数	可能的周期/天	傅里叶系数	可能的周期/天
1	207813.673	91	192671.856	91
2	119512.017	7	10677.111	7
3	111523.063	73	102354.520	35
4	84800.780	3	98360.396	26

5	71986.968	16	96060.455	16
6	71593.976	14	93664.562	3
7	71535.134	60	90017.886	52
8	71364.663	181	85448.754	33
9	71207.828	45	85116.178	11
10	65215.390	36	75713.721	120

对表2中获得的周期进行自相关系数计算, 计算公式

$$r_a = \frac{\sum_{i=1}^{n-h} (x_i - u)(x_{i+h} - u)}{\sum_{i=1}^{n-h} (x_i - u)^2}$$

。式中, r_a 表示自相关系数; h 表示周期; u 表示序列的均值。自相关系数的计算结果见表3。

表3 收费站出入口交通流量的自相关系数和可能的周期

序号	出口		入口	
	自相关系数	可能的周期/天	自相关系数	可能的周期/天
1	0.599	7	0.666	7
2	0.578	3	0.598	3
3	0.565	14	0.538	11
4	0.431	16	0.393	15
5	0.231	36	-0.161	121
6	-0.111	182	0.144	26
7	0.097	45	0.083	30
8	0.059	91	-0.060	52
9	0.027	60	0.045	33
10	-0.004	73	-0.003	91

从表3中可得知, 该收费站出口交通流量的周期为7天、3天和14天时, 自相关系数较大; 入口交通流量的周期为7天、3天和11天时, 自相关系数较大。自相关系数的绝对值越大, 相关程度越大, 则是交通流量周期特征的可能性越大。据此便可得出, 该收费站出入口交通流量变化呈周期性规律, 且周期均为7天(1周), 以周为单位呈现重复性, 主要受到1周内工作日和休息日的影响, 尤其是在周五和周日时容易出现流量高峰。

3、收费站出入口交通流量预测

3.1 预测方法

基于该收费站出入口交通流量分析结果基础上, 对该收费站未来出入口7天内的交通流量进行预测。在交通流量预测中, 首先是通过随机森林的方式对原始数据进行预处理, 获得完整的交通流量数据。之后采用长短期记忆网络模型对收费站出入口交通流量进行预测。

考虑工作时间较统一, 同一收费站的出入口交通流量呈现出周期性的通畅、拥挤等现象, 如本周五交通模式通常与上周五的交通模式相似, 但与周日的交通模式有较大的差异。对此, 就需要在模型中添加这一现象, 以保证模型预测的精度。此外, 收费站出入口交通流量数据也并非呈现出严格的周期性特征, 可能会受到交通突发事件、交通拥堵等多种因素的影响, 也就是说收费站出入口交通流量并非严格以7天为周期变化, 所以还需要通过注意力机制来解决这一问题。

3.2 预测结果分析

通过长短期记忆网络模型对收费站出入口交通流量进行预测, 并将获得的预测值与实际值进行对比。之后采用平均绝对误差 E_{MA} 、均方根误差 E_{RMS} 、对称平均绝对百分比 E_{SMAP} 、决定系数 R^2 对预测结果进行评估, 计算实际交通流量 Y_1 和预测交通流量 Y_2 之间的存在的差异。各指标表达式如下:

$$(1) \text{平均绝对误差 } E_{MA} = \frac{1}{n} \sum_{i=1}^n |Y_1 - Y_2|$$

$$(2) \text{均方根误差 } E_{RMA} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_1 - Y_2)^2}$$

(3) 对称平均绝对百分比

$$E_{SMAP} = \frac{100\%}{n} \sum_{i=1}^n \frac{|Y_1 - Y_2|}{(|Y_1| + |Y_2|)/2}$$

$$(4) \text{决定系数 } R^2 = 1 - \frac{\sum_{i=1}^n (Y_1 - Y_2)^2}{\sum_{i=1}^n (Y_1 - Y_2)^2}$$

其中, E_{MA} 的取值范围是 $[0, +\infty]$, 该值越大, 说明模型的预测误差越大; E_{RMS} 的取值范围是 $[0, +\infty]$, 该值越大, 说明模型预测的精度越高; E_{SMAP} 是 $[0, +\infty]$, 若该值为 0, 说明构建的模型是完美模型, 若该值大于 100%, 说明构建的模型为劣质模型; 一般而言, 当 R^2 大于 0.4 时, 说明模型的拟合效果好。

该收费出入口交通流量的评估结果见表4所示。从表4中可得知, 在训练集和测试集中 R^2 均大于 0.4, 说明构建的模型拟合效果好; 平均绝对误差 E_{MA} 、均方根误差 E_{RMS} 和对称平均绝对百分比 E_{SMAP} 等指标的值均处于较理想的状态, 说明模型对收费站出入口交通流量预测效果良好, 误差在可接受的范围内。

表4 收费站出入口交通流量评估结果

评价指标	出口		入口	
	训练集	测试集	实测值	测试集
E_{MA}	258.59	399.59	316.93	484.16
E_{RMS}	211.04	251.27	439.77	496.06
E_{SMAP}	0.29	0.37	0.28	0.40
R^2	0.74	0.67	0.69	0.62

4、结论与展望

4.1 结论

本文基于某收费站收费基础上对该收费站的出入口交通流量进行分析和预测, 获得了以下结论:

(1) 该收费站出入口交通流量之间呈极强的正相关性, 出入口交通流量之间相互影响。

(2) 该收费站出入口交通流量变化呈周期性特征, 变化周期为7天。

(3) 结合该收费站交通流量周期性特征建立模型对未来7天内收费站出入口交通流量进行预测, 结果表明建立的模型可获得良好的预测效果, 可准确预测未来7天交通流量。

4.2 展望

“大数据+交通”是未来交通管理发展的大方向, 其核心在于大数据分析与管理管控的有机结合, 从多个层面为交通运营管理提供依据, 进而提高运营管理的整体效率。交通流量分析和预测作为高速公路运营管理的重点, 也应当积极引入先进的技术开展高速公路交通流量的分析和预测, 如在高速公路收费站交通流量分析中, 可借助大数据分析技术、人工智能的深度分析及专家决策等技术, 提升对高速公路收费站出入口交通状况分析的准确性。

参考文献:

[1] 吴剑云, 于安双. 基于 ConvLSTM 的高速公路交通流预测仿真研究[J]. 实验室研究与探索, 2022, 41(12): 132-137.

[2] 林培群, 周楠楠. 基于多特征 GBDT 模型的收费站短时交通流量预测[J]. 广西大学学报(自然科学版), 2018, 43(03): 1192-1199.

[3] 郭瑞军, 于景, 孙晓亮, 牛树云. 基于电子收费数据的高速公路交通流特性分析[J]. 大连交通大学学报, 2018, 39(01): 17-22.

依托项目: 山西交通科学研究院集团有限公司创新发展计划项目(21-JKCF-06)。