

A Standard Cloud Platform Technology of Traffic Performance Index Based on Multi-Source Data Fusion

Jiandong Qiu¹, Lijian Zhuang¹, Yong Zhou², Zhongyuan Duan²

¹Shenzhen Urban Transport Planning Center Co., Ltd., Shenzhen Guangdong

²Traffic Information Engineering & Technology Research Center of Guangdong Province, Shenzhen Guangdong
Email: qiujd@sutpc.com

Received: Sep. 3rd, 2018; accepted: Sep. 14th, 2018; published: Sep. 21st, 2018

Abstract

Traffic index is the most basic and intuitive evaluation method in urban traffic management. Multi-source data fusion technology can effectively improve the accuracy of traffic operation evaluation. Real-time data from the Internet is used to improve spatial dimension deviation of the index, especially in remote areas, and fixed point detection technology to check the time dimension deviation of the index. Based on years of big data working practice, the concept of standardized cloud platform was proposed for the first time. This paper describes the complete solutions and practices such as “multi-source data access and processing-software hardware architecture-online dynamic publishing”. The transfer of cloud platform technology makes each city’s traffic researchers free from the sea of cumbersome big data processing and IT technology, so that they can easily have their own big data analysis system, and focus more on the technology and algorithm innovation of traffic itself.

Keywords

Big Data, Traffic Operation Index, Standardization, Cloud Platform Technology

多源数据融合的交通指数标准化云平台技术

丘建栋¹, 庄立坚¹, 周 勇², 段仲渊²

¹深圳市城市交通规划设计研究中心有限公司, 广东 深圳

²广东省交通信息工程技术研究中心, 广东 深圳

Email: qiujd@sutpc.com

收稿日期: 2018年9月3日; 录用日期: 2018年9月14日; 发布日期: 2018年9月21日

摘要

交通指数是城市交通治理中最基础、最直观的评估方法。多源数据融合技术,能有效提高交通运行评估的精度。引入互联网实时数据提升了指数的空间维度偏差,尤其是偏远区域;引入定点检测技术则有效校核了指数在时间维度的偏差。本文基于多年的大数据工作实践,首次提出标准化云平台的概念,并阐述了“多源数据接入与处理-软件硬件架构-在线动态发布”等完整的解决方案与实践,通过云平台技术转移,使各城市交通研究者从繁琐的大数据处理和IT技术无边的海洋中解放出来,轻松拥有自身的大数据分析系统,把更多的精力投注于交通本身的技术和算法创新。

关键词

大数据, 交通运行指数, 标准化, 云平台技术

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

交通指数的概念是由德克萨斯交通研究院(Texas Transportation Institute, TTI)在建立道路交通机动性和交通拥堵相关的评价指标体系,每年发布《城市畅通性报告》中提出的,其报告选择了交通拥堵指数等指标,定期评估并向管理部门和公众发布[1]。日本对指数体系的构建是最为完整的,从交通拥堵在时间和空间的分布状态出发,提出了道路交通拥堵评价指标,包括行程车速、排队长度和拥堵持续时间等[2]。目前,以美国、日本为代表的国外先进国家,已形成了对综合交通的全方位监测,并在评价标准、数据融合、交通预测、交通仿真等进行了大量研究[3],但其理论在中国实际交通情况中的应用仍存在适用性不足等问题。

在国外,主要从道路V/C比、拥堵时间比、拥堵时间经济价值等不同角度出发,开发了适用于当地的交通运行指数模型[4]。在区域宏观层面比较受用的指数模型有:被美国加州35个城市采用的拥堵持续指标(lane kilometer duration index, LKDIF) [5] [6] [7]以及在日本的道路交通形势调查中使用的拥堵度指标(degree of congestion, DC) [8]。在微观层面比较受用的指数模型有:Tom Tom公司作为商用的基于时间比的交通拥堵评价模型[5] [9]等。同时,欧洲的部分研究开始注重从经济价值角度研究交通拥堵评价,如OECD道路交通研究组织将交通与经济学相结合,提出了包括出行者平均耗费、出行者交通系统满意度、出行预测费用与实际费用比等15项评价指标[4]。

在国内,指数模型的计算主要可以归纳为基于严重拥堵里程比、出行时间比和混合评价等几类[4]。如北京[10]、广州[11]、南京等城市从宏观交通拥堵程度角度出发,其算法的核心指标为拥堵里程比例;而上海、深圳则是从出行者角度出发,关注延误,均与车速直接相关,且其核心指标呈指数关系[12]。

目前,运用浮动车大数据做交通运行评估的应用在国内刚刚兴起,中国各大城市的交通研究机构、高德百度等地图公司均已开展实际应用。2006年,深圳市城市交通规划设计研究中心建立“深圳市城市交通仿真系统”,成为中国较早运用动态数据,实时评估交通运行状态的城市[13] [14] [15]。2010年,北京交通发展研究中心推出“北京市道路交通指数”,把复杂的道路评估用简单的指数形式呈现给政府

和市民，交通数据由政府走向民间[16]。2014年，高德交通季度分析报告基于海量历史路况数据分析出不同区域在不同时间段内的拥堵延时指数，首次推出主要城市拥堵延时指数排名[17]。各地城市相继建成交通运行指数系统，以实时掌握全市道路交通运行状态，实现全市交通拥堵评估、交通运输应急指挥和交通投资合理配置，支撑城市政府的科学评估与决策。随着交通信息发布精准化与交通管理精细化的需求日益增长，基于单一数据源的交通运行指数系统在空间覆盖和精度上存在一定不足，且交通指数的建设与维护成本较大，制约了交通指数在城市政府决策中发挥更大的作用，多源数据的有效融合与应用是目前提高城市交通运行评估的基础。

2. 多源数据融合

2.1. 数据融合总体框架

在单一数据的指数系统的基础上，构建多源数据融合模型，系统按照“数据接入→数据预处理→特征提取→数据匹配→数据融合→数据存储→路况计算→融合结果”的总体思路进行构建。系统逻辑框架见图1。

2.2. 数据融合内容

以深圳为例，在深圳市的新一轮的交通运行指数研发中，构建多类型车辆GPS数据、固定检测器数据、手机数据与视频数据的多源融合校验模型，增加指数系统的监测覆盖范围，提高指数系统的评估精度，融合内容包括：GPS数据、固定检测器数据、手机数据及视频数据。各类数据必要信息见表1。

2.3. 数据融合方法

按照数据源的分布特征和数据特点，将深圳市城市道路交通运行指数系统划分为城市道路指数和高速公路指数，两部分叠加形成深圳市整个市域和高速公路的道路交通运行指数系统，见图2。

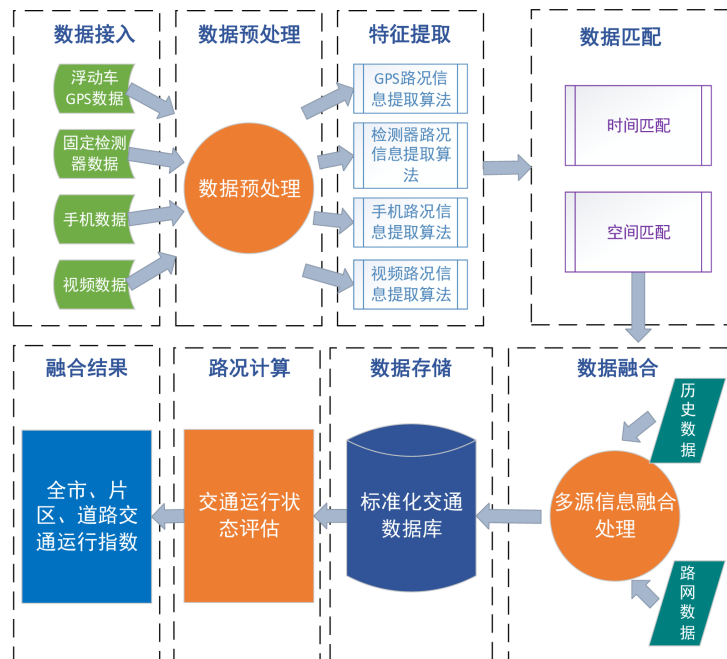


Figure 1. Design diagram of multi-source data fusion framework
图1. 多源数据融合框架设计图

Table 1. Necessary information for the data source
表 1. 数据源必要信息

序号	信息类别	所需数据内容
1	各类型车辆 GPS	日期、时间、车牌号码、经度、纬度、行驶角度、行驶速度、定位描述、空重车描述(仅出租车)、状态、状态描述等
2	固定检测点	编号、日期、时间、道路、道路等级、流量、速度、占有率、大车数量等
3	手机信令	设备唯一标识编号、基站块编号、基站编号、时间
4	视频(特征提取后)	编号、日期、时间、道路、道路等级、车牌、车辆类型、流量、速度等
5	系统基础地图	全市域 GIS 基础地图, 路段属性、片区属性等

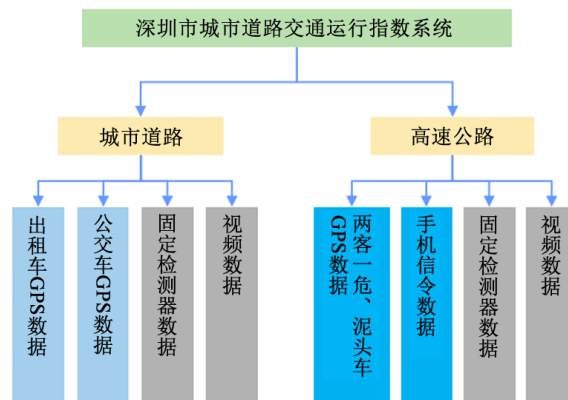


Figure 2. Regional division of road index system
图 2. 道路指数系统区域划分

2.4. 数据融合层次

数据融合是将从同一目标获得的多组传感器数据进行多级别、多方面、多层次的处理和组合, 产生新的有意义的信息。数据融合可分为检测级融合、属性目标识别级融合及决策级融合等几种融合级别, 见表 2。

检测级融合: 检测级融合是在采集到的原始信息层次上的融合, 即在多传感器的信息未进行预处理之前进行分析和综合, 先确定信号的有效性, 再进行数据联合。其优点是保持了尽可能多的环境信息, 缺点是信息处理量大、处理时间长、实时性差。

属性级融合: 也称为目标识别级的融合。它是指对来自多个传感器的目标识别数据进行组合, 以得到对目标身份的联合估计。属性层融合利用多个传感器采集观测目标的数据, 再进行特征提取和数据联合。将同一目标进行分组、再利用融合算法将同一目标的分组数据进行合成, 最后得到目标的联合属性判决即得到目标的类型和类别。该层融合的优点是实现了信息的压缩利于实时处理并且融合结果直接与决策分析相关。

决策级融合: 决策级融合是指在融合之前, 各传感器数据源都经过变换并获得独立的身份估计。信息根据一定准则和决策的可信度对各自传感器的属性决策结果进行融合, 最终得到整体一致的决策。

2.5. 数据融合结果

1) 在浮动车 GPS 融合方面, 5 分钟内不同类型 GPS 数据对深圳路网的覆盖情况, 见图 3、图 4, 展示的主要结论包括:

- 出租车存在明显的空间分布不均匀性, 在中心城区密集。

- 公交车、百度导航在非中心城区的覆盖率大大优于出租车。
- 公交车由于其稳定性运营的因素，在非中心城区的覆盖率是最高的。
- 不同类型的数据存在波动，需要数据层融合在不同时间、不同空间维度采用不同的权重。

Table 2. Data fusion classification at different information levels

表 2. 不同的信息层次上的数据融合分类

类型	检测级融合	属性级融合	决策级融合
所属层次	最低层次	中间层次	高层次
主要优点	原始信息丰富，并能提供另外 2 个融合层次所不能提供的详细信息，精度最高。	实现了对原始数据的压缩，减少了大量干扰数据，易实现实时处理，并具有较高的精确度。	所需要的通信量小，传输带宽低，容错能力比较强，可以应用于异质传感器。
主要缺点	所要处理的传感器数据量巨大，处理代价高，耗时长，实时性差 原始数据易受噪声污染，需融合系统具有较好的容错能力。	在融合前必须先对特征进行相关处理，把特征向量分类成有意义的组合。	判决精度低，误判决率升高，同时数据处理的代价比较高。
主要方法	HIS 变换，PCA 变换，小波变换及加权平均等。	聚类分析法，贝叶斯估计法，信息熵法，加权平均法，D-S 证据推理法，表决法及神经网络法等。	贝叶斯估计法、专家系统、神经网络法、模糊集理论、可靠性理论以及逻辑模板法等。
主要应用	多源图像复合、图像分析和理解。	多传感器目标跟踪领域，融合系统主要实现参数相关和状态向量估计。	其结果可为指挥控制与决策提供依据。

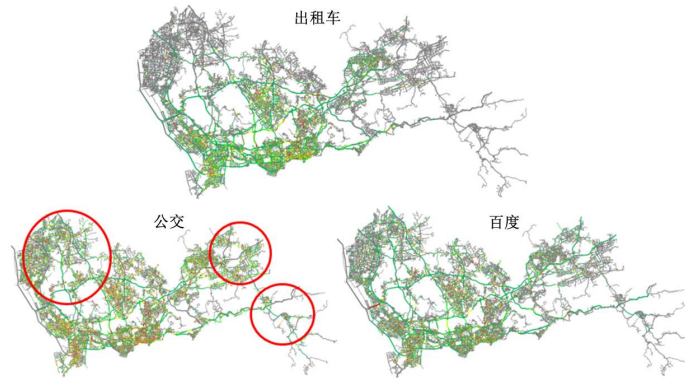


Figure 3. Difference of data coverage of different GPS data

图 3. 不同 GPS 数据覆盖差异

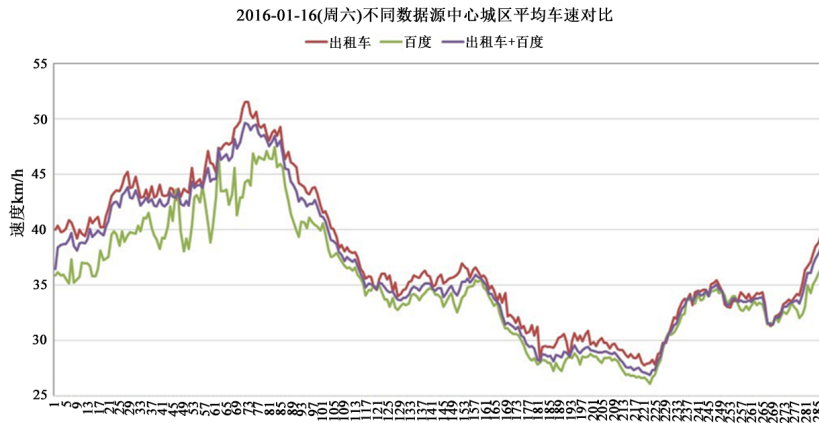


Figure 4. Difference in speed values of different GPS data

图 4. 不同 GPS 数据速度值差异

2) 公交车与出租车速度存在周期性的关系, 进行数据融合前需进行对比训练, 获取特征映射关系, 见图 5。

3) 定点地磁线圈检测速度, 有其特殊性, 主要表现为地磁流量和速度扰动较少, 但为点速度, 速度偏高, 无法直接用于校核。实际应用时可将道路状态分为 5 级, 对应指数 0~10。其中指数 0~6 对应畅通 - 缓行状态, 流量从 0 变化到饱和流量; 指数 6~10 对应较拥堵 - 拥堵状态, 流量位于较高位, 速度从最高位变化到最低位。利用 1 个月的历史数据, 统计各路段的地磁流量、地磁速度的最高位和最低位, 并建立指数与路段旅行时间(出租车)的关系, 见图 6。

- 4) 数据融合有效提升了样本的覆盖率, 见图 7, 主要体现在:
- 采用出租车、公交车、百度数据融合后, 覆盖率显著提升。
 - 罗湖、福田、南山提升 16%, 覆盖率达 91%。
 - 外围区域提升 50%, 覆盖率达 70%以上。



Figure 5. Difference of data coverage of different GPS data
图 5. 不同 GPS 数据覆盖差异

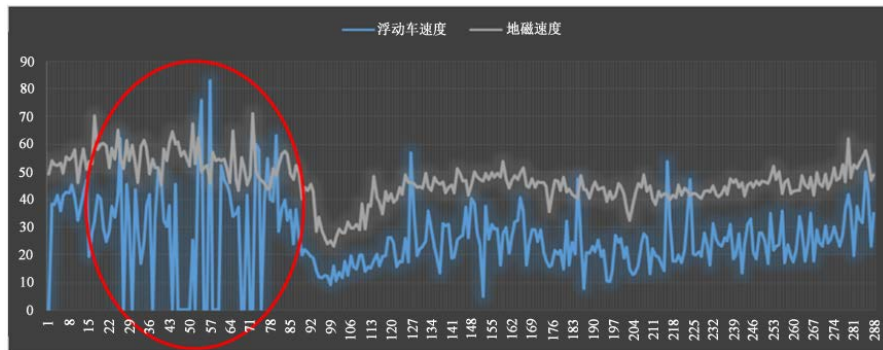


Figure 6. Difference of data coverage of different GPS data
图 6. 不同 GPS 数据覆盖差异

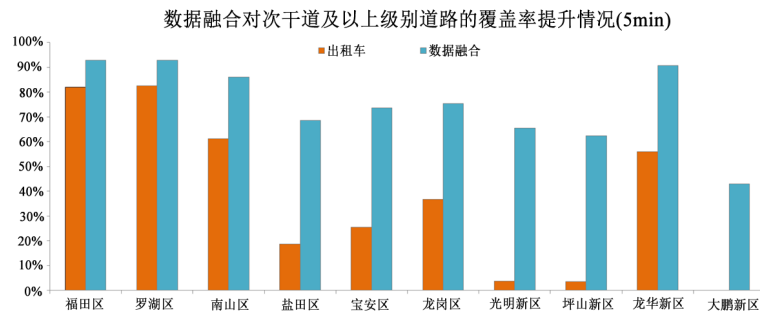


Figure 7. Improvement of GPS data coverage in different regions
图 7. 不同区域 GPS 数据覆盖提升情况

5) 数据融合有效提升了指数的准确率, 见图 8, 主要体现在:

- 认为路段经过的样本车数量越多, 精度越高。
- 在出租车已有覆盖的前提下, 准确率有提升的路段占比 83%。
- 原先低样本(仅有 1 辆车经过)情况下有显著提升的比率超过 70%。

3. 标准化云平台技术

3.1. 云平台技术

1) 云平台部署架构及优点

云平台部署结构见图 9。云平台部署优点如下:

- 屏蔽了硬件层自身的差异和复杂度;
- 共享基础架构资源池;
- 灵活扩展和收缩、弹性的虚拟化资源池;
- 应用系统流程化、自动化部署和管理;
- 优点: 高利用率高可用性低成本、低能耗、快速部署易于扩展智能管理。

2) 云平台网络拓扑

私有云搭建的拓扑图见图 10。其中若干台专用服务器、存储服务器和光纤网络组成了服务器群集, 所有服务器均可通过 VCenter Server 等服务器管理软件实现统一管理。服务器群集使用是把资源形成一个巨大的资源池并虚拟化, 支持云平台的快速复制与部署。标准化云平台包括了硬件系统、软件系统和指数运算的配置模块, 在虚拟池中可以根据需要灵活配置, 实现秒级新建和上线。

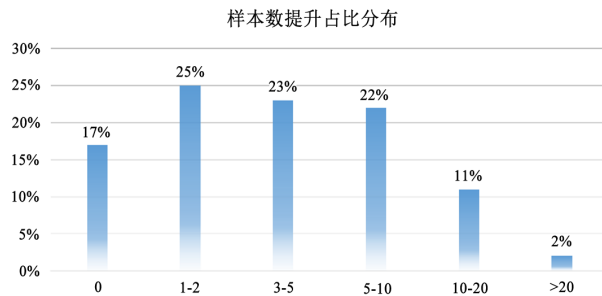


Figure 8. Percentage increase of sample data of different sections
图 8. 不同路段样本数据提升百分比

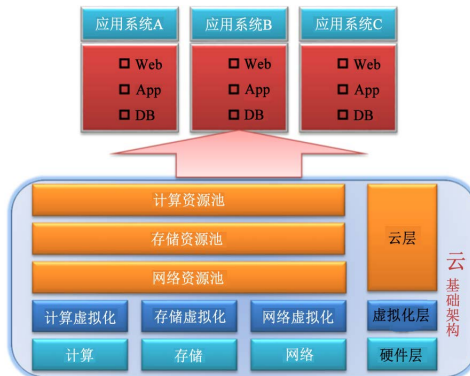


Figure 9. Cloud platform deployment architecture
图 9. 云平台部署架构

3.2. 标准化功能

1) 交通运行指数标准化云平台功能框架体系

以深圳市为例，标准化云平台大幅降低建设成本，云技术实现了在线托管技术，本地化技术参数，共享指数软件平台同步升级，所有过程无需软硬件的购买与维护，深圳市云平台使用框架体系见图 11。

2) 指数系统功能包括：

- 实时交通运行监测模块，实现宏观、中观、微观的全方位、多层次的实时监控，见图 12。
- 历史数据查询模块，实现指标项、时间、空间等 3 大类 128 个维度的交叉查询，见图 13。
- 专业报告自动生成模块，实现专业交通报告(周报、月报等)的极速生成，见图 14。
- 专业交通分析评估模块，实现交通拥堵治理与影响评价的决策分析，见图 15。

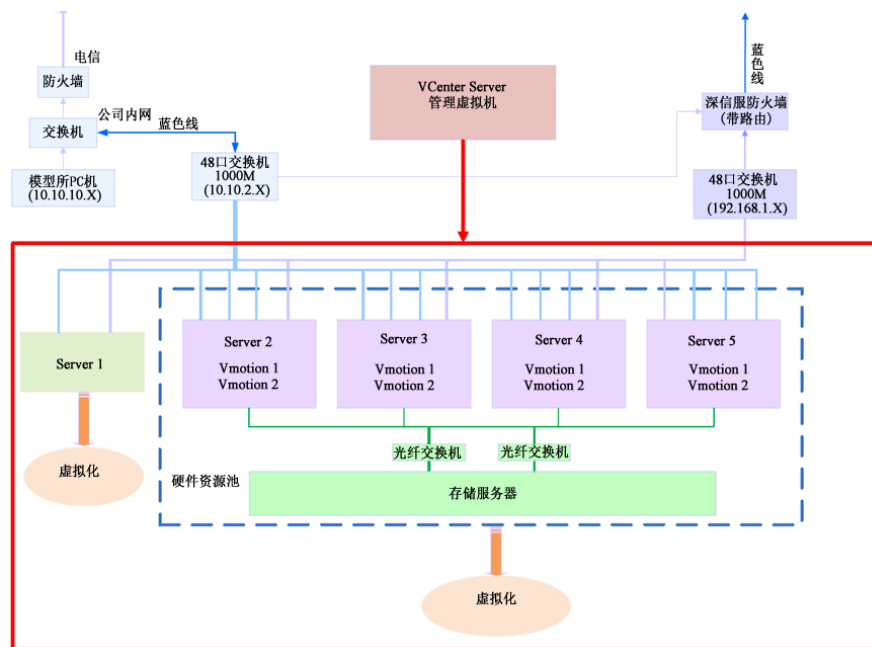


Figure 10. The network topology of the private cloud

图 10. 私有云网络拓扑

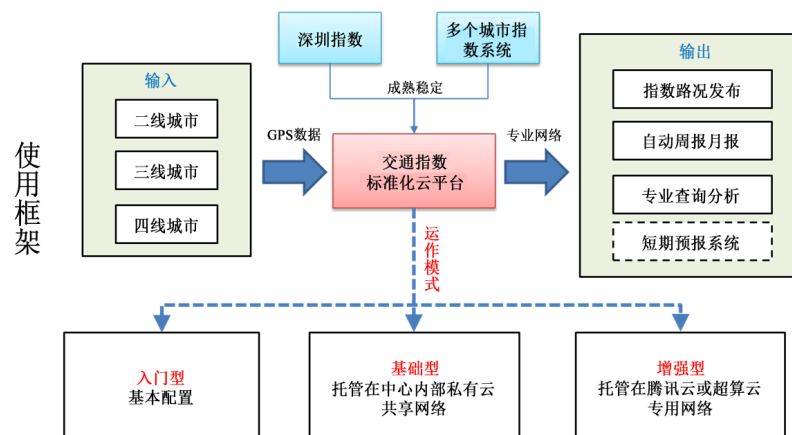


Figure 11. Framework of cloud platform use in Shenzhen

图 11. 深圳市云平台使用框架体系



Figure 12. Real-time traffic operation monitoring module
图 12. 实时交通运行监测模块

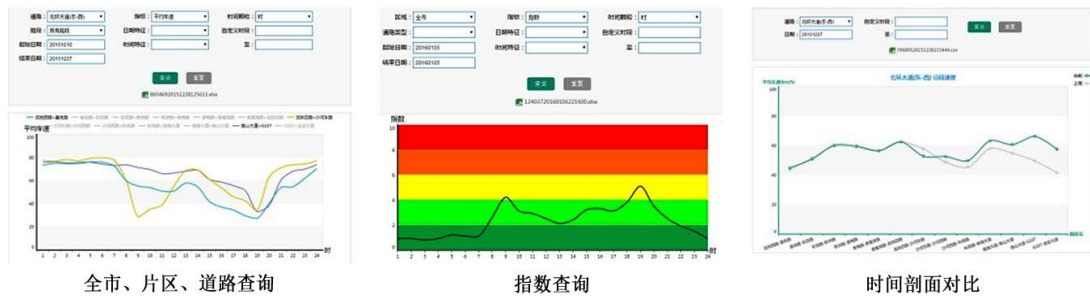


Figure 13. Historical data query module
图 13. 历史数据查询模块

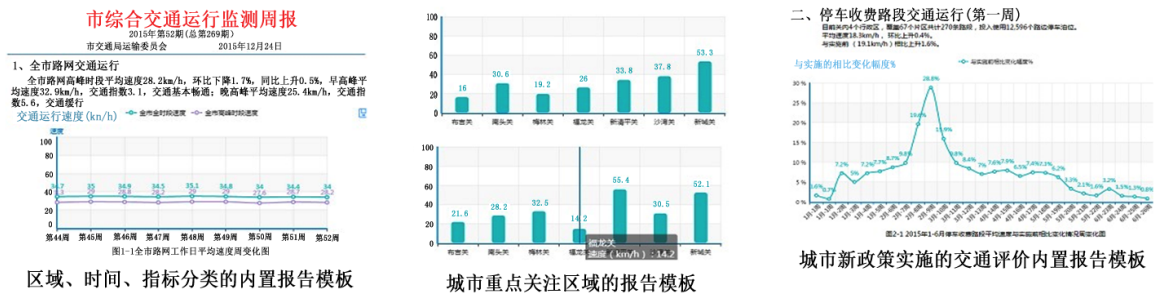


Figure 14. Automatic generation module of professional report
图 14. 专业报告自动生成模块

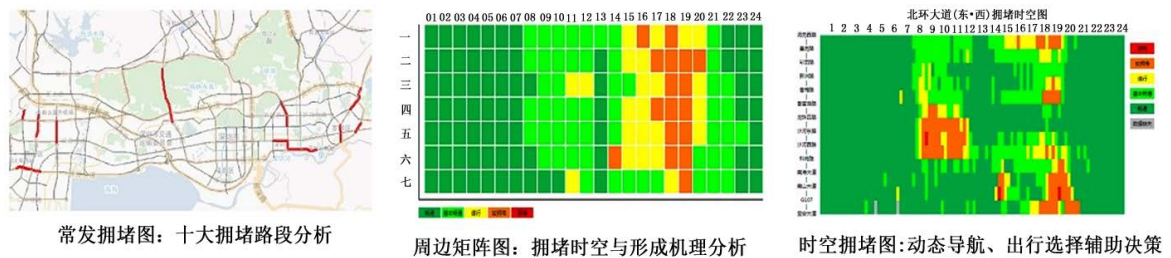


Figure 15. Professional traffic analysis and evaluation module
图 15. 专业交通分析评估模块

3.3. 云平台技术特点

1) 云平台极速响应

动态接入城市浮动车的实时 GPS 数据, 通过云平台实现快速部署应用; 专业化查询与自动化的周报、

月报、年报秒级响应，快速上线。

2) 先进的数据清洗与数据融合

高效、稳定、安全的多通道、多方式的数据接收软件，灵活的异常数据清洗算法，基于卡尔曼滤波，融合历史与实时数据的全方位数据拟合及修补技术，系统匹配清洗日志及数据接收异常自动报警系统。

3) 精准的地图匹配与路径规划

自主研发基于浮动车地图投影与匹配、空间索引的最优路径搜索等核心算法，采用多线程与分布式并行技术，10万级 link 路段 35 秒内完成地图匹配。

4) 专业的车速建模与指数运算

首创双层次路网车速、交叉口节点车速、路网平均车速(行程时间)算法及模型，具备多源异构交通数据深度融合的功能，实现海量数据的交通指数实时处理与精准计算。

5) 多元的指标查询与评估发布

基于“面-线-点”多层次、多维度的道路交通畅通性和可达性统计分析 with 评估方式，涵盖指标项、时间、空间 3 大类共 128 个维度的交叉查询分析，系统功能模块灵活扩展，具备专业网站、移动终端、电视、车载等载体的多方式发布模式。

3.4. 云平台技术应用实践

重庆交通指数依托于云平台，在指数标准化云平台的基础上实现重庆指数的快速部署和上线。依托于指数云平台的弹性扩展，实现了对桥隧、重要节点和流量的实时监控等个性化定制功能，见图 16。

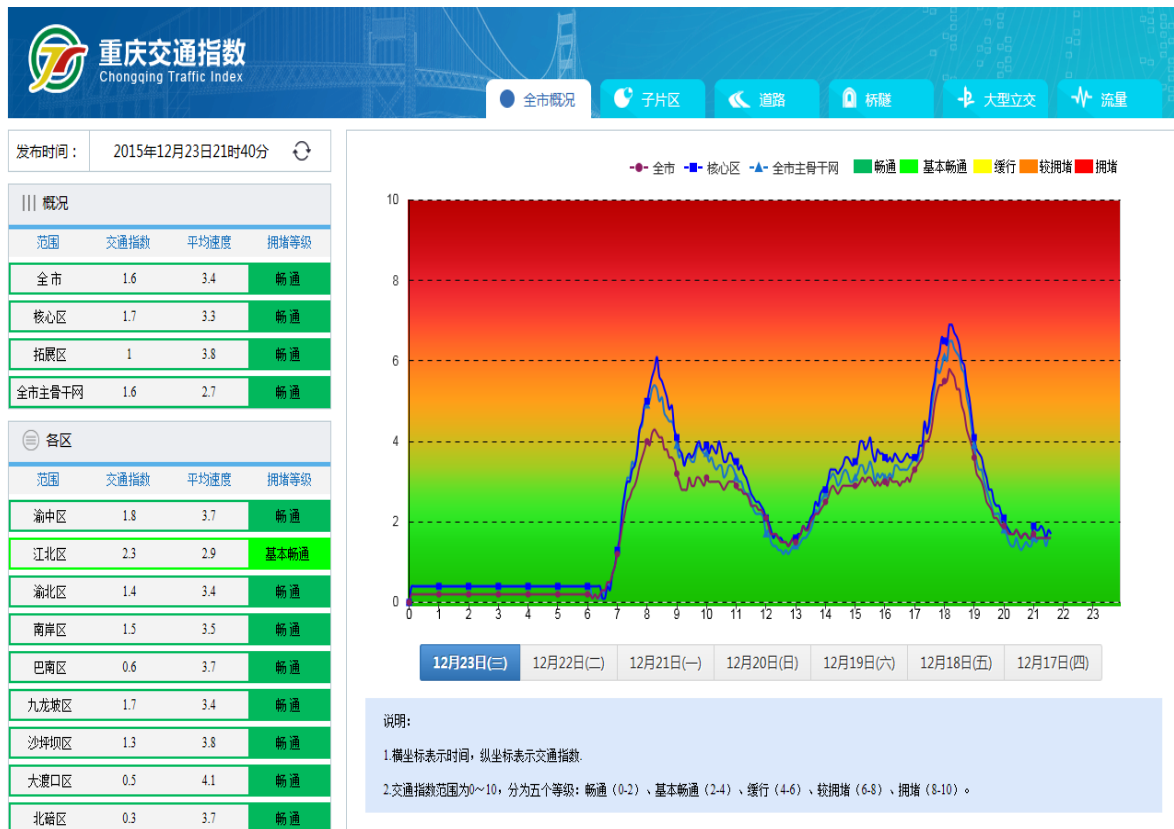


Figure 16. Chongqing traffic index based on cloud platform technology

图 16. 基于云平台技术的重庆交通指数

4. 结语

按照国内外城市经验,以先进成熟的理论研究为基础建立交通运行评估指标体系,以智能化自动化的交通信息采集和处理技术为基础开展道路交通评估,能够全面准确地评估道路拥堵状况、动态监测变化趋势,为研究拥堵产生机理、分析交通系统存在问题、制定改善和治理方案等工作提供定量化的分析手段和依据,是交通管理部门制定缓解交通拥堵各项政策措施、合理安排基础设施建设时序、重大事件应急处理等工作的技术基础,有助于提升交通运行管理的科技化和信息化水平。本文提出的多源数据融合方法与标准化云平台技术,将进一步提升交通指数的覆盖面与精度,同时有效地降低交通指数的建设与维护成本,支撑系统在城市政府决策中发挥更大的作用。

致 谢

感谢深圳市科技计划项目(项目编号 GGF2016033017241891,项目名称“深圳市交通大数据公共技术服务平台”)和深圳市战略性新兴产业发展专项资金 2017 年第一批扶持计划(项目名称:深圳市交通碳排放工程实验室,批复文号:深发改(2017) 550 号)的资助。

参考文献

- [1] 崔燕鸣. 基于关联性分析的城市交通评价指标体系研究[J]. 森林工程, 2009, 25(4): 55-57.
- [2] Jiang, Y.H., Xiao-Zhou, H.E. and Guo, X.C. (2008) Discussion on the Evaluation Index System of Urban Green Traffic Planning. *Journal of Hefei University of Technology*, 31, 1399-1402.
- [3] Jia, H., Tan, Y. and Fu, X. (2010) Research and Application of Urban Traffic Congestion Evaluation System. 2010 *International Conference on Optoelectronics and Image Processing*, Haikou, 11-12 November 2010, 177-180. <https://doi.org/10.1109/ICOIP.2010.126>
- [4] 王璐媛, 于雷, 孙建平, 宋国华. 交通运行指数的研究与应用综述[J]. 交通信息与安全, 2016, 34(3): 1-9+26.
- [5] 郑淑鉴, 杨敬锋. 国内外交通拥堵评价指标计算方法研究[J]. 公路与汽运, 2014(1): 57-61.
- [6] Vazir, I.M. (2002) Development of Highway Congestion Index with Fuzzy Set Models. *Transportation Research Record*, Paper No. 02-2229, 16-22.
- [7] Ishimaru, J., Hallenbeck, M. and Nee, J. (2000) Weekend Freeway Performance and the Use of HOV Lanes on Weekends. Washington State Transportation Center, Washington DC, 1-39.
- [8] 郑淑鉴, 杨敬锋. 国内外交通拥堵评价指标计算方法研究[J]. 公路与汽运, 2014(1): 57-61.
- [9] Tom Tom (2015) Tom Tom Traffic Index Measuring Congestion Worldwide. https://www.tomtom.com/en_gb/trafficindex/list?citySize=LARGE&continent=ALL&country=ALL
- [10] 北京交通发展研究中心. 城市道路交通运行评价指标体系: DB11/T785-2011 [S]. 北京: 北京市质量技术监督局, 2011.
- [11] 广州市交通运输研究所. 城市道路交通运行评价指标体系: DBJ440100/T164-2013 [S]. 广州: 广州市质量技术监督局, 2013.
- [12] 但媛. 重庆市主城区交通运行指数系统构建及应用[C]//中国城市规划学会城市交通规划学术委员会. 2017 年中国城市交通规划年会论文集. 中国城市规划学会城市交通规划学术委员会, 2017: 14.
- [13] 林群, 关志超, 杨东援, 等. 深圳市城市交通仿真系统[R]. 深圳: 深圳市城市交通规划设计研究中心, 2006.
- [14] 林群, 李锋, 等. 深圳城市交通仿真系统建设实践[J]. 城市交通, 2008, 5(5): 22-27.
- [15] 陈蔚, 段仲渊, 周子益, 等. 基于出行时间的道路交通运行指数算法与应用研究[C]//中国城市交通规划 2012 年年会暨第 26 次学术研讨会, 2012.
- [16] 郭继孚, 温慧英, 全永桑, 等. 北京市道路交通指数研究[R]. 北京: 北京交通发展研究中心, 2011.
- [17] 高德交通. 高德交通第一季度分析报告[R/OL]. <http://trp.autonavi.com/traffic/>, 2014-07-22.

知网检索的两种方式：

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择：[ISSN]，输入期刊 ISSN：2326-3431，即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：ojtt@hanspub.org