

DOI: 10.16661/j.cnki.1672-3791.2205-5042-1487

# 交通大数据的应用现状与未来发展方向

孟鸿程

(中移(上海)信息通信科技有限公司 上海 201206)

**摘要:** 基于交通大数据的全生命周期流程,从数据采集、数据处理与数据应用角度分析交通大数据的应用模式,说明数据从采集到应用整个过程的数据工作以及所需要构建的数据体系。该文以交通态势分析系统为例,说明大数据在交通领域的具体应用方式,具体包括所采用的大数据技术组件在系统中发挥的作用以及系统利用交通大数据所实现的功能应用。之后阐明推动交通大数据领域发展的参与方格局,分析交通大数据的未来发展方向与重点工作,应当完善交通大数据标准体系,加强交通大数据全生命周期管理。指出应强化数据采集,完善数据管理体系并深化数据应用,并结合实例说明应完善和推广现有应用、发展和强化关键应用、促进和试点创新应用。

**关键词:** 智慧交通 交通大数据 交通态势 数据管理 生命周期

**中图分类号:** TP311.13;U495

**文献标识码:** A

**文章编号:** 1672-3791(2023)04-0010-05

## The Application Status and Future Development Direction of Traffic Big Data

MENG Hongcheng

(China Mobile(Shanghai) ICT Co., Ltd., Shanghai, 201206 China)

**Abstract:** Based on the whole life cycle of traffic big data, the application mode of traffic big data is analyzed from the perspective of data collection, data processing and data application, and the data work in the whole process from data collection to application and the data system to be constructed are explained. Taking the traffic situation analysis system as an example, this paper explains the specific application mode of big data in the traffic field, including the role of the adopted big data technology components in the system and the functional application realized by the system using traffic big data. After that, the pattern of the participants promoting the development of traffic big data is clarified, and the future direction and key tasks of traffic big data are analyzed. The standard system of traffic big data should be improved, the whole life cycle management of traffic big data should be strengthened, data collection should be strengthened, data management systems should be improved and data applications should be developed. Combined with examples, it is explained that existing applications should be improved and promoted, key applications should be developed and strengthened, and innovative applications should be promoted and piloted.

**Key Words:** Intelligent transportation; Traffic big data; Traffic situation; Data management; Life cycle

大数据时代的到来为交通管理带来了巨大的机遇与挑战,大数据一方面描绘了交通管理的美好愿景,打通多源异构交通数据,为政府管理、公民出行带来便利,提升管理效能与出行体验;另一方面交通大数据也面临着发展的问题,如何汇聚多源异构交通数据、如何

进行数据优化处理以实现交通应用。

国家高度重视交通大数据的发展,2015年8月,国务院印发《促进大数据发展行动纲要》,2019年9月,中共中央、国务院印发的《交通强国建设纲要》提出要“推进数据资源赋能交通发展”“构建综合交通大数据中心

作者简介: 孟鸿程(1996—),男,硕士,研究方向为智慧交通、交通大数据。

体系”,2019年12月,交通运输部印发《推进综合交通运输大数据发展行动纲要(2020—2025年)》。

近年来,交通运输大数据取得了长足发展,但对照国家要求和行业发展需求,还存在着数据资源基础不够扎实、数据共享开放需深入推进、大数据应用不够广泛、数据安全存在挑战、管理体系尚需健全等问题,需要进一步深入推进解决。

该文通过分析交通大数据从采集、处理到应用的全生命周期过程,研究大数据在交通领域的应用,说明交通大数据的发展与应用现状,同时分析交通大数据的发展方向。

## 1 交通大数据应用模式

大数据的广泛定义来自其5V属性,即,Volume数据量大、Variety数据种类多、Value数据价值密度低、Velocity数据增长速度快、Veracity数据准确性低。在交通领域,数据来源复杂,并且随着时间不断累积,最终形成海量多源的交通数据,具体包括公路、铁路、航空、水运等多行业的设施、设备、人员、安全保障等基础数据,还包括综合交通网络中人、车、路、环境等相互作用的关系数据<sup>[1]</sup>。

目前,大数据在交通领域的应用模式如下图1所示。数据从采集到处理直至最后的应用,首先通过多种形式的设备采集多源交通数据,之后对数据进行处理,建立相应的指标模型,最后实现交通大数据的应

用。与此同时,还需要做好数据标准、数据安全与数据治理体系建设工作。

### 1.1 数据采集

数据采集是交通大数据应用的基础,通过数据采集获取最原始的交通数据,为数据分析和数据应用服务。高质量的原始数据能够简化后续的数据清洗工作,提高数据发挥价值与被有效利用的可能性<sup>[2]</sup>。交通数据类型多样,数据采集设施也在不断升级进化,为交通大数据应用提供了更多可能性。

#### 1.1.1 数据源

交通大数据的数据源十分丰富,以城市交通领域为例,基于中国智能交通协会发布的城市交通时空大数据格式标准<sup>[3]</sup>,数据分为城市交通基础地理数据、城市交通设施数据、城市交通系统管理数据、城市交通系统运行状态数据、城市交通系统智能感知数据、城市交通系统外关联系统数据这几大类型。这些数据的来源、数据类型、数据格式等都存在差异,如何综合利用多源数据是交通大数据需要解决的问题。

#### 1.1.2 数据采集设备

传统数据采集设备多为单点式采集,数据的网络连通性较差。随着大数据与云计算技术的发展,交通领域的数据采集方式、数据类型越来越丰富,数据采集和传输效率、数据可获取性愈发增强。

以道路交通流数据采集为例,目前被广泛应用的



图1 交通大数据应用架构图

有感应线圈、地磁、微波、超声波、红外线、视频、激光雷达与毫米波雷达等<sup>[4]</sup>,它们都各有其适用场景和局限性。传统检测设备具备一定的优势,但因为需求与技术的变化,其适用场景越来越少。以感应线圈检测方法为例,这种方法技术成熟、效果较好且成本较低,但是需要埋入车道,且只能获取交通流量信息。视频识别技术尽管易受天气影响,但是在图像识别技术飞速发展的背景下,视频能够提供道路上更丰富的信息,因而愈发为交通管理者所青睐。目前,国内较为先进的摄像头设备本身已具备AI识别能力,并能够输出结构化数据。通过结合深度学习与计算机视觉技术,视频能够检测车辆轮廓和形态、识别车牌与车身信息、检测交通事件等<sup>[5]</sup>。

现如今,摄像头被越来越多地与激光雷达以及毫米波雷达结合,雷视一体机被快速应用于道路交通检测领域,尤其在车路协同等领域发挥了巨大的作用。数据采集设备根据不同的数据要求,在采集阶段已经具备了数据处理能力,例如:摄像头的图像识别能力与雷视一体机的融合感知能力,它们不仅能够输出原始的视频数据与雷达点云数据,还能够基于其算法能力输出结构化数据,更便捷地为数据应用服务,同时也能够减少对带宽的要求。

## 1.2 数据处理与应用

尽管数据采集设备愈发先进,所采集的数据质量愈发受到重视,但原始数据仍然无法直接发挥作用,需要对数据进行处理和分析,才能最终实现应用。由于交通数据种类丰富,需要进行良好的数据管理,才能实现数据综合运用,具体工作包括元数据管理、主数据管理、数据集成与互操作、数据模型管理和数据归档管理等。数据预处理的4个主要任务是数据清洗、数据集成、数据变换和数据规约<sup>[6]</sup>。

交通大数据研究的最终目的是实现数据应用,在该阶段,需要建立与实际情况相符合的交通指标与模型。统一标准的指标能够直观地展示交通运行水平,提供交通管理的优化方向。交通模型能够有效提升对交通运行状况的感知能力,从而提高交通管理能力,为提升交通安全与效率服务。基于交通大数据构建交通模型能够结合动态变化数据,为解释交通现象、城市活动的特性及其演化机理提供了重要的观察手段与研究保障<sup>[7]</sup>。

## 2 大数据在交通领域的典型应用——交通态势分析系统

交通大数据在交通领域中的一大典型应用是交通

态势分析系统。道路交通系统产生的海量数据蕴含着能够准确描述复杂交通系统运行态势的信息<sup>[8]</sup>,交通态势分析系统能够利用实时道路交通运行数据实现道路交通运行状态监测,系统对数据的实时性要求以及数据量的庞大决定了其交通大数据属性,对大数据技术的需求较为显著。

### 2.1 交通态势系统的数据处理

交通态势分析系统需要获取道路交通实时运行数据,为交通管理服务,需求方多为交通管理部门,其动态数据源一般是电警卡口数据,静态数据则包括设备信息、设备与车道关联信息和地图数据等。实现数据之间的关联需要进行数据表结构的设计与接口设计,之后利用大数据技术组件进行数据处理。

系统利用Flume日志采集组件获取海量多源交通数据,通过Kafka消息队列输出至Flink进行实时计算,实现对交通的实时监测,并将交通实时数据存储至Redis、HBase、ES和MySQL等数据库中,实现不同的数据应用。而利用Spark的弹性分布式数据集RDD与机器学习算法,可以实现交通预测等能力。在大数据领域,Flume是一种被广泛使用的日志收集组件,它是一种分布式、可靠和可用的服务,能够高效收集、聚合、移动不同数据源的海量日志数据,并存储到一个中心化数据存储系统中<sup>[9]</sup>。Kafka是一个分布式的发布/订阅系统,它可以被作为消息中间件用于流式数据分析,从而确保交通大数据的实时应用能力。Flink是一个分布式计算框架,能够实现数据流的实时处理,保证了流计算的低延迟。而基于Spark进行机器学习与深度学习能够实现更好的分布式训练和数据传输。

### 2.2 交通态势系统的数据应用

为了实现良好的数据应用,首先需要建立合理的交通数据指标和模型。在交通态势分析领域,所关注的交通指标基本以GB/T 33171-2016<sup>[10]</sup>等国家和行业标准为参考,从路口、路段、道路和路网不同等级设定评价指标,例如路口平均延误、路段交通运行状况、道路和路网平均行程速度以及交通拥堵指数等。除了这些行业统一的指标外,基于数据建立交通模型也是一个重要方向,基于交通数据可以建立交通拥堵评价模型、交通仿真、拥堵预测等模型,为顶层的交通应用服务。

基于多源交通数据,应用大数据技术,交通态势系统能够实现交通运行监测、拥堵预警、交通信息发布、指挥调度等高实时性应用,也能够实现车流分析、交通预测、信号优化和重点车辆监管等非实时性应用,此



外,也可以提供检索统计、交通报告等个性化应用。现阶段,机器学习、深度学习等人工智能技术也更多地被应用于交通大数据领域<sup>[11]</sup>,为交通拥堵识别、交通流预测等应用提供了新的应用方向。

### 3 交通大数据的发展方向

随着技术的发展,交通大数据的数据来源愈发丰富,数据处理流程要求更高,数据应用领域更为广泛,而这些都对交通大数据领域的从业者提出了更高的挑战。

#### 3.1 交通大数据发展格局

交通大数据领域业已形成了政府统筹规划、企业推动应用、高校技术支撑的整体格局。政府注重顶层设计,从宏观视角调查分析交通大数据发展状况,提出推动交通大数据发展的指导方法和意见。以交通运输部印发的《推进综合交通运输大数据发展行动纲要(2020—2025年)》<sup>[12]</sup>为例,纲要以党中央关于国家大数据战略部署和推进交通强国建设为指导,提出推动交通大数据发展的五大行动和21项具体任务,各项任务分别指定交通运输部属下各司及各省级交通运输主管部门负责,层层推进,逐步丰富和细化推动交通大数据发展的工作。

企业是市场的主体,也是切实推进交通大数据发展的中坚力量,在落实国家政策,推动交通大数据应用方面发挥着至关重要的作用。交通大数据从采集到处理分析直到应用的整个过程中,都需要依赖企业发挥作用,交通大数据的产业链包括智能交通设施、数据处理、智慧交通软件研发等相关企业。

除了政府和企业需要发挥作用,高校等科研机构也在交通大数据领域发挥着重要作用。高校基于研究体系与经验不断深化交通大数据的研究,在现有技术的基础上深入研究,拓宽技术边界。同时也在标准和专利方面发挥着作用,通过与合作企业发挥科研优势,校企融合促进应用。

#### 3.2 交通大数据发展方向

尽管交通大数据已经获得了一定的发展,但仍然存在着一些亟待解决的问题,例如交通数据共享开放有待加强、数据管理和治理体系尚未完善、数据应用效果仍需提升等。这些问题需要依赖政府、企业、科研机构之间的共同合作,在交通大数据采集到应用的全生命周期开展建设与研究工作。

##### 3.2.1 完善交通大数据标准体系

由于交通大数据覆盖领域广泛,从数据采集到数

据应用涉及的参与方众多,企业之间若无统一的标准,将会造成重复建设与资源浪费。如果要切实有效推进交通大数据发挥作用,需要构建和完善交通大数据标准体系。

构建交通大数据标准,应当推动硬件规格、产品开发、数据采集、数据存储、数据传输、数据交换、数据处理、数据质量、数据安全、数据模型和系统架构等多个环节的标准建设工作。具体需要依靠相关社会组织和产业技术联盟制定团体标准,并将其中具有广泛适用性的标准升级为行业标准、地方标准直至国家标准,通过多级标准体系为交通大数据领域的发展建设提供指导。同时注重国际化发展,促进国内标准与国外标准的融合与借鉴,构建高效且广泛适用的交通大数据标准。

完善交通大数据标准体系,对现有标准进行周期性梳理,开展实施情况评估与动态调整,对其中落后于现行发展状况的标准进行修订或废止,同时结合发展需求确定新标准规划方向。对于国家标准,需要强化强制性标准“保基本、兜底线”的作用和技术法规地位<sup>[13]</sup>,优化提升推荐性标准。此外,还需要促进交通大数据标准与交通行业及大数据领域相关标准的协调衔接,促进国家标准、地方标准、行业标准之间的协调衔接。

##### 3.2.2 加强交通大数据全生命周期管理

(1)强化数据采集。数据采集是交通大数据发展的基础,为推动交通大数据发展,需要强化数据采集工作。应当加强基础设施建设,包括数据采集设备与通信传输设备。将基础设施建设与交通工程建设结合到一起,在新建和优化的交通基础工程中完成数据采集相关设施建设工作。对于已建设设施,需要监测设备运行情况,保证已有设备正常采集与传输数据,并根据数据需求实现升级换代。数据采集基础设施将逐步提升智能化水平,设备端将会集成更成熟的数据处理方法及交通算法,并且可以实现自动化采集与传输。

(2)完善数据管理体系。数据管理涉及领域众多,包括数据治理、数据架构、数据建模和设计、数据存储和操作、数据安全、数据集成和互操作、文件和内容管理、参考数据和主数据、数据仓库和商务智能、元数据和数据质量等。在交通大数据领域,还未系统性地开展这些数据管理工作,因此需要完善交通大数据管理体系,以保证数据在采集、转换、存储和应用全流程的完整和准确,从而为数据应用提供支撑。

数据管理中,数据安全是其中至关重要的一环,能够确保对数据资产的适当访问,防止未经许可的非法访问,确保公民隐私得到保护。为此,在交通大数据的发展过程中,应当牢牢把握数据安全,构建交通大数据安全保障体系,具体工作包括识别数据安全需求、制定并完善数据安全制度、定义并分配数据保密等级和数据监管类别等。

现阶段,交通大数据管理的一个突出问题是数据资源共享和开放水平不足,各地区之间、政府部门之间暂未形成高效统一的数据共享机制,信息孤岛现象明显。随着国家推进交通大数据共享开放,未来交通领域的数据融合与开放创新程度将显著提升。政府部门之间的数据融合将有效推动交通大数据的创新应用,交通数据集的逐步开放也将为交通大数据领域的学术与课题研究带来新的发展动能。

(3)深化数据综合应用。交通大数据的发展最终是为了服务于所有交通参与者,因此需要不断深化交通大数据应用,具体包括完善和推广现有应用、发展和强化关键应用、促进和试点创新应用。

交通大数据现有应用中,相对成熟的有交通运行监测调度中心(Transportation Operations Coordination Center, TOCC)和交通态势分析评价系统,这些系统已经应用了较长时间,但在覆盖的地域范围上相对有限,尤其是TOCC所覆盖的城市屈指可数。未来,应当不断完善现有交通大数据应用,提升应用效果,并推动已有应用在各市、县甚至区级行政区域的应用。

此外,还需要发展和强化交通大数据领域的关键应用,例如:发挥交通大数据在应急管理和指挥调度中的能力,构建应急管理大数据,把控交通运行态势和资源分布,为应急管理服务。也有必要将大数据能力应用于交通运输生产监测中,针对重点车辆、重点交通设施和重点领域的安全风险进行监控,提升安全水平。

对于交通大数据领域的创新应用,更是需要通过试点等方式促进其发展,例如:出行即服务(MaaS, Mobility as a Service)模式,需要打通共享单车、公交、出租、地铁和航空等多种交通运输方式,需要利用大数据优化交通出行供给方的服务形式、了解交通参与方的支付意愿等。

#### 4 结语

该文结合交通大数据全生命周期流程分析了交通大数据的应用模式,并结合交通态势分析系统说明了大数据在交通领域的具体应用,最后分析了交通大数

据的发展格局以及未来的发展方向,应当完善标准体系、强化数据采集、完善数据管理体系、深化数据应用。现阶段,智能交通和大数据技术仍在发展过程中,需要结合更先进的技术,依靠政府、企业、科研院所的合作,共同推动交通大数据的发展应用,为交通参与者带来更多便利。

#### 参考文献

- [1] 刘晓波,蒋阳升,唐优华,等.综合交通大数据应用技术的发展展望[J].大数据,2019,5(3):55-68.
- [2] DAMA国际.DAMA数据管理知识体系指南[M].北京:机械工业出版社,2020.
- [3] 中国智能交通协会.城市交通时空大数据格式标准:T/CITSA 10-2021[S/OL].(2021-07-26)[2023-01-04].<http://www.its-china.org.cn/TTBZ/20210727/0002.pdf>.
- [4] 叶腾茂.面向新基建的路网交通运行状态感知体系构建研究:以厦门为例[J].无线互联科技,2021,18(10):109-111.
- [5] 赵新勇,李珊珊,夏晓敬.大数据时代新技术在智能交通中的应用[J].交通运输研究,2017,3(5):1-7.
- [6] 陈娇娜.大数据驱动下的高速公路交通运行状态评价与分析[D].西安:长安大学,2016.
- [7] 汪光焘.大数据时代城市交通学发展的机遇[J].城市交通,2016,14(1):1-7.
- [8] 包杰.基于多源数据的城市路网交通事故风险研究[D].南京:东南大学,2019.
- [9] THE APACHE SOFTWARE FOUNDATION. Flume 1.9.0 User Guide-Apache Flume[EB/OL].(2022-02-22).<https://flume.apache.org/releases/content/1.9.0/FlumeUserGuide.html#flume-1-9-0-user-guide>.
- [10] 全国城市客运标准化技术委员会.城市交通运行状况评价规范:GB/T 33171-2016[S].北京:国家质量监督检验检疫总局,国家标准化管理委员会,2016.
- [11] 朱杰.基于出行大数据的城市重点区域交通状态识别方法与应用研究[D].北京:北京交通大学,2021.
- [12] 交通运输部.交通运输部关于印发《推进综合交通运输大数据发展行动纲要(2020—2025年)》的通知[EB/OL].(2019-12-16)[2022-02-22].[http://www.gov.cn/xinwen/2019-12/16/content\\_5461503.htm](http://www.gov.cn/xinwen/2019-12/16/content_5461503.htm).
- [13] 交通运输部.交通运输标准化“十四五”发展规划[EB/OL].(2021-01-11)[2022-02-22].<https://xxgk.mot.gov.cn/2020/jigou/kjs/202111/P020211112599582853785.pdf>.