计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.191200165

汽车大数据应用模式与挑战分析

葛雨明[1] 韩庆文[2] 王妙琼[3] 曾令秋[2] 李 璐[1]
1 中国信息通信研究院技术与标准研究所 北京 100191
2 重庆大学微电子与通信工程学院 重庆 400030
3 中国信息通信研究院云计算与大数据研究所 北京 100191
(geyuming@caict.ac.cn)

摘 要 随着车联网技术的不断演进,人、车、路、云全方位连接,催生出了大量应用服务,覆盖汽车生产制造、汽车产品网联化、汽车后服务、智能出行服务等多个环节。这些应用的核心是海量的汽车联网数据。对汽车大数据的有效利用可能成为未来汽车产业转型升级的重要突破口。为了推进车联网场景下汽车大数据的应用,文中对相关工作进行了综述。从汽车大数据的内涵与架构出发,对汽车大数据的数据源及相关应用进行了详细分析,介绍了汽车大数据支撑的各类应用,包括汽车产业类应用、网联类应用和后市场服务类应用,从数据采集、数据处理与分析、计算资源、隐私保护 4 个方面分析了汽车大数据的关键技术,并从政策法规和技术两个层面加以探讨,分析其发展现状,并展望未来的应用趋势。

关键词:车联网;大数据;架构;应用
中图法分类号 U467.5+25

Application Mode and Challenges of Vehicular Big Data

GE Yu-ming[1], HAN Qing-wen[2], WANG Miao-qiong[3], ZENG Ling-qiu[2] and LI Lu[1]
1 Technology and Standards Research Institute, China Academy of Information and Communications Technology, Beijing 100191, China
2 School of Microelectronics and Communication Engineering, Chongqing University, Chongqing, 400030, China
3 Cloud Computing and Big Data Research Institute, China Academy of Information and Communications Technology, Beijing 100191, China

Abstract With the technical evolution of connected vehicles, people-vehicle-road-cloud are all connected, and a large number of application services emerge which cover many parts such as manufacturing, connected vehicle products, vehicle service market and intelligenttravel service. The core of these applications is big data of vehicles. The effective utilization of vehicular big data may be an important breakthrough in the transformation and upgrading of automotive industry in the future. To promote the application of vehicular big data in connected vehicles, related works are reviewed in this paper. According to the application demands, this paper starts from the connotation and architecture of vehicular big data, and analyzes the characteristic of data sources and corresponding applications, such as manufacturing, connected vehicle products and vehicle service market, etc. Then the key technologies of vehicular big data are discussed from four aspects, which are data collection, data processing and analysis, computing resource and privacy protection. Based on a comprehensive analysis fits development status of policy and technology, this paper anticipates the future application trend of vehicular big data.

Keywords Connected vehicles, Big data, Architecture, Application

1 引言

近年来,汽车的智能化与网联化已成为汽车行业发展的大趋势。2015 年提出的《中国制造 2025》发展战略将智能网联汽车列为重点发展名录。智能网联汽车包括两个核心问题,即智能化和网联化,其内涵在于利用 V2X(Vehicle to Everything)和多接入边缘计算(Multi-access Edge Computing, MEC)等技术提供强大的连接和计算能力,以实现人-车-路-云间多元数据的共享、交互和融合,从而提升汽车的智能化水平[1]。随着消费互联网的升级以及车联网技术的逐步成熟,汽车已经不再是单纯的交通工具,而是可以提供出行、生活、娱乐及办公等服务的智能载体[2]。车辆共享和分时租赁等汽车共享经济成为新常态,车联网保险等后服务市场的新业务不断涌现。这些与日俱增的应用需求都需要高质量的数据作为支撑。2019 年 12 月 9 日,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)》,以推动大数据

到稿日期:2019-11-26 返修日期:2020-01-27 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。
基金项目:国家重点研发计划(2018YFB0105200)
This work was supported by the National Key R&D Program of China (2018YFB0105200).
通信作者:王妙琼(wangmiaoqiong@caict.ac.cn)

与综合交通的深度融合，汽车作为综合交通运输的重要参与者，智能交通系统需在汽车本地数据与交通数据深度融合的前提下演进发展，汽车大数据应用将建立在跨行业、多协同的基础上，其数据的规范性、兼容性、开放性和安全性[3]等对于提升综合交通运输服务水平将起到至关重要的作用。

2018年，中国汽车市场处于发展的十字路口，汽车市场销量遭遇28年来的首次负增长，迫使整车企业开始积极探索新的业务增长点，而大数据的应用将成为重要的突破口[4]，但汽车大数据应用仍处于起步阶段，且存在诸多问题：

1)各种数据源互通性低下，呈现出数据孤岛特性；

2)数据源质量参差不齐，数据可信度不高；

3)缺乏必要的数据安全保障和管理机制。

总体而言，目前汽车大数据发展势头迅猛但成熟度不足，有必要对其技术研究及应用现状进行分析总结。

本文第2节重点介绍汽车大数据的内涵及架构，并详细分析数据源的特征和分类；第3节重点阐述汽车大数据的应用需求，并在第4节分析其面临的挑战；第5节讨论汽车大数据在智能网联汽车中的研究现状及发展趋势，为汽车大数据应用的落地推广提供了可行的方法论及实施参考。

2 汽车大数据的内涵及架构

2.1 汽车大数据的内涵

汽车大数据，即汽车及相关领域数据加上相关技术和应用的总和，其内核是生产、业务、人-车-路互联和外部跨界数据，同时又外延至相关技术及应用。就汽车大数据本身而言，其横向跨越产品设计、制造、销售和使用等环节，纵向涉及企业、产业链和生态链[5]、行业管理3个层面的内容；而汽车大数据技术及应用则需面向特定场景，通过数据采集、传输、存储、挖掘及可视化等一系列技术方法，将汽车大数据中包含的价值信息释放出来，以提升生产运作和经营管理水平，促进产品及商业模式的创新[6]。

汽车大数据具有鲜明的特点，主要体现为[7]：

1)关联性，不同数据源数据表现出了强关联性，如车辆数据与交通流数据之间的强关联性，道路历史数据与交通实时数据之间的强关联性；

2)动态性，实时处理的数据量呈现动态变更特性，主要体现为车辆数据的动态性和交通流数据的动态性；

3)多维度，汽车大数据应用需在综合多维度描述的基础上获得支持，如综合车辆数据、交通流、交通控制数据的车辆轨迹策划应用。

2.2 汽车大数据架构

经典大数据架构包括4个层次，即数据采集层、数据存储层、数据处理分析层和数据应用层。理论上此架构也适用于汽车大数据，但汽车大数据应用独特性明显，主要体现在以下两个方面。

1)经典大数据架构支撑多元异构数据，面临各种数据源采集汇聚的技术困难和政策性障碍；但汽车大数据面临更大的困难，不仅涉及企业敏感数据和用户敏感数据，还包含各种政府数据，其数据互联互通困难不仅体现在技术层面，还体现在政策层面，政府不完善目前已经成为汽车大数据应用推进中的严重障碍，必须予以单独考量。

2)经典大数据架构中的数据处理一般在本地集群或云端完成，但汽车大数据应用，特别是安全类应用，对处理实时性的要求较高，集中化的处理难以满足其应用需求，需以边缘计算方式予以支撑。

因此，我们认为汽车大数据架构需特别强调应用瓶颈，即数据互联互通问题和数据处理方式问题，前者旨在促进政策层面的突破，后者则旨在为服务提供方(如车企、IT业、运营商、交通管理部门等)定义服务介入方式。根据汽车大数据的特征，本文提出的汽车大数据架构如图1所示。

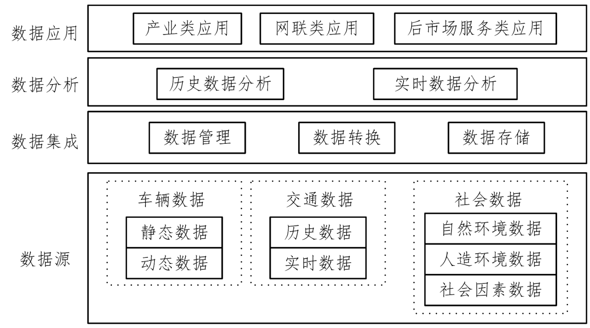

图1 汽车大数据架构

Fig. 1 Architecture of vehicular big data

从图1可以看出，汽车大数据体系架构包括4个层次的内容：1)数据源，包括车辆数据、交通数据和社会数据；2)数据集成，汽车大数据的数据源结构复杂，主要体现在数据格式和数据接口的复杂性表现，互联互通必须在有效的数据转换机制和数据抽象化机制的基础上合理建立；3)数据分析，汽车大数据分析融合需在多个层面进行，包括支持实时应用的边缘端计算分析和支持非实时应用的云端计算分析；4)数据应用服务，汽车大数据的应用服务呈现多维度、广覆盖的特性，重点表现为支持跨行业应用。

3 汽车大数据的来源与应用

汽车大数据应用可深入汽车研发、生产、销售和使用等各个环节，可分为产业类应用、网联类应用、后市场服务类应用等。不同的数据应用服务对数据类型、数据处理模式、数据处理表现的需求也各不相同，大数据应用与数据源之间存在需求对应关系，本节将从数据源入手，分析需求来源，解析数据源与应用之间的关联对应关系。

3.1 汽车大数据的数据源

如图1所示，汽车大数据的数据源表现出了多源异构特性，不仅包含车辆在设计、制造、销售和使用过程中产生的相关联网数据，还包括与车辆相关的人、事、物产生的数据。

与普通大数据应用不同，汽车大数据应用一方面涉及工业、社会数据；另一方面，应用类型与产业期望密切关联，即整个产业链都期待能通过汽车大数据发掘经济增长点，因此数据源类型与产业链之间存在强关联关系。

表1列出了汽车大数据的数据源类型、特征及其提供方。可以看出，汽车大数据数据源的复杂性不仅体现在数据源类型，更体现在涉及部门的复杂性。

表 1　汽车大数据数据源类型、特征及提供方

Table 1　Types, features and providers of vehicular big data

数据类型	数据特征	数据内涵	数据提供方
车辆数据	静态数据	车辆技术参数、研发生产数据、车辆基本信息	车企
	缓变动态数据	车辆的折旧损耗特性参数、二手车交易数据、维修保养数据等[8]	维修店，二手车交易市场
	快变动态数据	驾驶行为数据、车辆实时行驶数据	车辆个体
交通数据	道路数据	道路属性、路侧设施、通信网络、公共资源等[9]	路政部门
	交通数据	实时交通流数据、交通灯数据	交通管理部门
社会数据	自然环境数据	天气、温度、污染、噪声	环境部门
	人造环境数据	路面环境数据、住宅、停车场、加油站等 poi 数据	市政管理部门
	社会数据	交通违法数据、居民受教育程度数据、其他渠道获得的用户画像数据[10]	公安部门、民政部门

2020 年伊始，工信部等十一部委联合下发“关于印发《智能汽车创新发展战略》的通知”，从另一个侧面证明了汽车大数据应用需在政策支持的前提下，由多部委协同联合推进。

3.2　汽车大数据的应用

数据应用服务对数据类型、数据处理模式、数据处理表现的需求也各不相同，主要体现在应用响应实时性和数据处理有效性两个方面。根据图 1，本文将应用类型分为汽车产业类应用、网联化应用和后市场服务应用[11]，各应用类型与相关数据源及处理实时性和处理可靠性之间的关系如表 2 所列。

表 2　应用类型与相关数据源要求

Table 2　Types of applications and requirements of data sources

应用类型		数据准确度需求						处理能力需求	
		车辆静态数据	车辆动态数据	交通历史数据	交通实时数据	社会数据		实时性	可靠性
						基础数据	特种数据		
汽车产业类		高	高	低	低	中	中	低	高
网联应用类	安全应用	高	高	高	高	高	高	高	高
	信息服务	低	低	高	高	高	高	中	高
后市场服务类		高	高	低	低	高	高	低	高

3.2.1　汽车产业类应用

汽车产业类应用主要为车企提供服务，包括产品质量管控类应用和市场调查类应用。

质量管控类包括故障反向诊断，以及以之为基础的正向设计优化，主要涉及的数据源包括车辆固有数据及行驶特征数据，旨在实现车辆的售后跟踪，以确保车辆安全运行，并为优化设计和提升产品质量提供必要的反馈数据[12]。

市场调查类旨在收集市场信息，使得汽车制造企业能通过主动爬取互联网信息或与互联网公司进行数据合作，在汽车的设计阶段即结合当前用户的诉求和市场的供需情况作出准确的判断[13]，以提升汽车市场的竞争力。

提升汽车制造企业的市场匹配能力是汽车大数据的重要任务，打通各环节数据的汽车制造将更加自动化和智能化，能为汽车产业带来制造升级和用户体验提升，如汽车定制化服务，以最大限度满足消费者的需求[14]。

汽车产业类应用对数据准确度的需求较高，数据源质量问题将成为影响汽车产业类应用的重要瓶颈。

3.2.2　网联类应用

网联类应用可分为两类，即安全类应用和信息服务类应用，其中安全类应用又可分为实时安全类应用和非实时安全类应用。实时安全类应用与车辆控制及实时行驶策略密切相关，对数据处理实时性和准确性均有较高的要求；非实时安全类应用和信息服务类应用不涉及车辆实时行驶策略问题，对准确性的要求较高，而对实时性的要求较低。

由表 1 可知，实时安全类网联应用涉及多种数据源，需在云端数据处理、边缘端数据处理和车辆本地数据处理的严密协同下执行，且需在驾驶人员、汽车、道路、交通和公安等数据有效整合、共享、交互的基础上确保应用的响应质量。若要达成此目的，在数据抽象化、数据接口、处理方式和计算资源调动方面均需理论研究和计算硬件条件的支持，这是目前影响安全类网联应用推进进程的重要因素。

非实时安全类应用对处理实时性的要求不高，但对准确性有要求，非实时安全类应用主要面向汽车质量安全类服务，旨在提前发现因汽车质量问题而导致的安全隐患，如电池故障隐患消除等。该类应用需要有效的平台数据的支持，即对数据质量的要求较高。

信息服务类应用目前的使用率较高，主要是一些提高驾乘体验的车载信息类应用，如导航服务[15]、堵点提醒[16]、充电引导[17-19]、停车引导[20]等，这类应用因涉及与人的互动，对数据的实时性要求较高，又因为引导式服务的需求，其对数据的准确性要求也很高。此类应用的执行度较好，未来可致力于应用服务质量的提升，以确保应用服务的有效性。

3.2.3　后市场服务类应用

汽车大数据的后市场应用服务包括汽车保险、车辆服务和出行服务等。

汽车后市场保险业需要汽车行驶数据、违章数据、维修数据等的支撑，以期实现保费合理评估，并以此增进投保数量及保险质量[21]。目前已提供的相关服务有中国汽车技术研究中心和中国保险信息技术有限责任公司提供的“事故车定损云平台”服务、Metromile 的自动理赔助手 Ava 等[22]。Lamberti 提出将区块链引入汽车保险服务[23]，对汽车大数据质量和数据处理能力提出了更高的要求。

车辆服务包括维修服务、清洗服务、加油服务、配件购置和救援服务等，需通过收集服务提供方的信息来进行合理评级，

并为用户推荐可靠的服务方。此类应用对评级方的公正性要求较高,需要在确保评级可靠的前提下保证服务质量[24]。

目前,典型的出行服务方式包括“车辆共享”[25]与“分时租赁”,基于数据智能化的分时租赁模式在很大程度上节约了人工成本,租车、解锁、驾驶、支付和归还等操作都能直接通过手机应用完成,用户体验良好。

4 汽车大数据关键技术

4.1 数据采集

数据采集是各种大数据应用的基础,也是汽车大数据的首要关键技术。汽车大数据数据源类型复杂,且数据质量参差不齐。2018 年 8 月 2 日,新能源汽车国家大数据联盟发布《新能源汽车国家监管平台车辆数据质量分析报告》,报告显示,2018 年 1 月至 6 月全国接入车辆 969107 辆,其中 627461 辆车出现过错误,占比 64.75%,数据错误种类达到 72 类。另外,尽管目前实时数据通过人工审核的车辆占比约 73.18%[26],但在实际运行中仍会出现多项错误。报告显示:超过 83.5%的车企半数以上的车都会出现数据错误。数据质量问题目前已成为制约汽车产业类应用的重要原因,因此需要在数据采集环节严控数据质量。目前汽车大数据采集的主要途径如下:

1)车企采集数据。通过 CAPP,MES,ERP,PDM 等企业信息化系统[27]采集获取设计研发数据、生产过程中各类零部件数据、生产管理数据及整车基本数据等,此类数据可采用 WebServGice、数据 ETL、可扩展标记语言 XML 等数字化集成工具进行数据集成化处理。

2)社会数据采集。其包括经销商数据和政府管理数据等,通过数据管理系统采集车辆销售数据、维修数据、二手车交易数据、天气数据、违章建筑数据、人口数据等,一般通过专业 MIS(Management Information System)系统实现[28]。

3)车辆实时行驶数据。基于车载传感器系统进行采集,并通过数据接口和车载 OBU 进行数据处理,生成车辆行驶数据,该数据通过 C-V2X 网络,或收集汇聚至数据中心,或进行车间信息交互,以提供安全服务。目前,研究者们重点关注车辆实时数据的处理和存储挖掘,文献[29]提出采用分布式深度拓扑结构来提高数据的健壮性和有效性;文献[30]则设计了一种智能车内数据采集终端,提出采用实时车辆数据管理系统进行车辆信息共享,并提供安全服务。

4)交通数据采集。交通数据大多基于路侧传感器和摄像头进行采集,目前交通数据采集及其应用自成系统,即智能交通系统,其数据采集与上传汇聚方式与物联网基本相同。文献[31]以交通数据为对象,开发了交通关键数据的感知与处理平台;文献[32]则设计了车辆数据采集平台,并将其引入智能交通系统的设计。

综上所述,汽车大数据数据种类多、来源广、体量大,特别是网联化后,数据标准不一致,传输技术参差不齐,如何保证数据的质量将成为应用推进的一大难点[24]。

4.2 数据处理与分析

汽车大数据处理的表现将直接决定其相关应用的表现,目前来看,汽车大数据分析面临的两大主要问题,即数据融合问题和数据处理效率问题。

1)数据融合问题

汽车大数据的数据源类型繁多,有源于车企的汽车基础信息数据,源于运行中车辆的各类传感器的数据,源于用户的偏好数据,源于互联网的舆情数据,源于交管及其他相关部门的城市基建数据,交通数据及路况数据以及源于气象部门的自然环境数据等,各数据的提供方均根据各自行业标准和企业标准定义数据格式,数据体现为高度多元异构特性,互操作性差,从而导致计算效率低下,延时过长,影响辅助/自动驾驶应用的表现。

若要实现多源异构数据的互联互通,首先需要对不同数据源提供的数据进行抽象化处理,以应用需求为基础进行数据格式的转换,以确保数据的互联互通性,并基于应用需求支持部分互操作性。

目前,针对大数据的数据控制、融合及预处理方案已较为成熟,有针对多源数据集成的平台级异构设备数据摄入模型[33],也有卡尔曼滤波、小波变换、神经网络等数据融合方法。

对于汽车大数据而言,其数据融合技术可沿用较为成熟的大数据处理方法,而主要应用障碍体现在政策层面。如前所述,定义统一的数据接口是实现多元异构数据互联互通的关键,数据源归属部门间缺乏统一协调,导致数据交互困难,特别体现在社会数据库、交通数据库与车辆数据库之间的交互过程中,此问题的关键在于需要定义合理的数据接口。目前各大城市的民用平台数据库正在进行一体化改造,但数据交互过程复杂,可能难以满足辅助/自动驾驶服务的实时性要求,但对于非实时性服务的支撑力度明显增加。另外,由于各部门都有各自的保密机制,其对数据的脱敏处理方式也各不相同,在脱敏处理过程中难免会造成有效信息的丢失,从而导致数据可信度降低,这对于实时性和非实时性服务均会造成不利影响。

2)数据处理效率问题

数据处理难度主要体现在两个方面,即处理效率问题和数据可信度问题。

汽车大数据形态多样,有各类传感器产生的结构化数据,日志和资料等半结构化数据,来源于摄像头的图像和视频数据等[14]。如何对这些海量异构数据加以管理,使数据能够被高效利用,值得深入研究。

目前,经典大数据处理模型 MapReduce[34]可以很好地支撑汽车大数据的批量分析,该模型在处理大体量历史数据时表现出了较好的计算能力和处理效率,但对于实时数据的处理响应速度不足,需要用到 Apache Spark 或 Apache Flink 这类更实时的计算框架。另外,特定的分析场景需要结合特定的算法,如自然邻居、K-means 等,可实现交通流预测等应用。对于实时车辆数据处理能力的提升,除了计算框架的革新之外,还需要通过计算资源的合理分配,来提升处理的实时性和有效性。

4.3 计算资源

汽车大数据支持实时应用和非实时应用,只有当数据处理能力能够满足应用实时性需求时,才能确保应用响应有效。传统的数据处理方式包括客户端处理和集中云处理。随着汽车类应用需求水平的提升,数据处理对计算资源的需求也越来越大,对其处理表现的要求也越来越高,MEC 的引入能够有效支持实时类应用升级。该技术能够将原本集中在云端的计算压力分摊到边缘节点,以提供低时延、高带宽的网络环境以及高性能的服务能力,但 MEC 的应用也对计算资源调度的合理性提出了新的挑战,特别是如何合理调度资源以支持不同等级的应用服务需求,以及对不同业务搭载方的支持策略问题,目前已经成为业界的研究重点。2019 年初,IMT2020(5G)C-V2X 工作组推出了《MEC 与 C-V2X 融合白皮书》,系统阐述了 MEC 与 C-V2X 融合的内涵、特性以及典型应用场景,分析了 MEC 在汽车大数据应用中面临的技术挑战[35]。

近年来,学术界特别关注 MEC 的相关研究,包括面向车联网的 MEC 框架研究[36]、MEC 计算资源调度算法研究[37-38],并将软件定义网络[39]、特种车辆节点[40]引入 MEC 设计。产业各界也在持续推动 MEC 的发展,攻关诸如低抖动低时延的可靠传输、跨运营商业务连续性、数据与业务安全性等一系列技术难题;规范车载终端与 MEC 平台之间、MEC 平台与部署业务之间的接口协议,明确 MEC 与路侧基础设施的部署、运营、管理权限与责任,丰富 MEC 平台上车联网业务的应用,构建全面完整的车联网 MEC 生态。总体来看,MEC 框架下的各种理论研究进展顺利,但应用层面依旧存在问题:1)MEC 部署方案及其与应用的映射关系缺乏必要的研究基础;2)MEC 的部署及经营归属权依旧不明朗,导致实时类应用水平提升困难,影响安全类大数据应用的表现。以上两个问题需在学术研究与政策导向综合推进的前提下解决。

4.4 数据隐私保护

大数据涉及的隐私保护问题一直是业界和学术界的关注重点[41],汽车大数据也不例外。Nawrath 详细分析了网联汽车数据涉及的隐私敏感数据[42],认为需要对涉及隐私问题的数据进行过滤脱敏,并以之为基础提出了加密算法。Endo 则分析了驾驶员对隐私数据共享的接受度问题,为相关数据脱敏提供了分析依据[43]。文献[44]则搭建了一种用于研究隐私安全过滤的测试平台,将隐私数据脱敏之后用于提升鉴定能力。在车联网场景中,信息收集一般包括以下几种模式:1)在用户知情的情况下收集与身份无关的信息,如车辆的位置信息,用于统计某条路的车流量,但并不知道某个对应车主的位置;2)在用户知情的情况下收集与身份相关的信息用于身份认证,如汽车应用的账号、密码以及实名认证信息,这些信息一般会在平台做脱敏加密处理;3)在用户知情的情况下收集用户身份相关的信息用于车联网服务,如与车主信息关联的车型、位置和驾驶行为等,仅用于内部提供定制化的个人服务;4)在用户知情的情况下收集用户身份相关的信息,但将用户信息故意或无意的泄露给第三方;5)在用户不知情的情况下偷偷收集用户信息,进行一些恶意的数据变现行为。

很明显,在上述 5 种情况下,一般用户可以接受前 3 种信息采集,并认为后两种侵犯到了个人隐私。在服务性的数据应用中,必然会涉及与人相关的数据,通过技术或是管理的手段来保障用户的个人隐私不被侵犯是无法避免的问题[45]。

现阶段,我国个人信息保护法仍在制定中,但很多相关法律和管理办法中关于个人信息保护的内容对车联网中的数据隐私保护均有借鉴意义。

5 研究现状及未来的研究方向

汽车大数据应用问题引起了业界和学术界的普遍关注,针对其应用问题的研究可归纳为两个方面,即政策法规研究和关键技术研究。

5.1 政策法规研究

汽车大数据应用需要政策法规的支持,主要体现在两个层面,即数据平台化需求和隐私保护需求。

汽车大数据数据源类型繁多,需要建立统一的数据接口以保证数据提取和计算过程的顺利展开,但目前各大数据源归属部门不同,需要在政策法规层面进行统一协调。2017 年,我国将新能源车纳入统一平台管理,虽然数据质量依旧存在问题,但不失为一次成功的尝试。2019 年提出的交通大脑使得汽车大数据应用平台化路径更加清晰,实现数据平台化需要政府和业界密切合作,打通数据管道,破除数据孤岛。2019 年底发布的《推进综合交通运输大数据发展行动纲要(2020-2025 年)》中提到,要深入推进国家综合交通运输信息平台建设,促进政务信息资源的共享;推动建立行业大数据平台,以鼓励行业企业开展大数据融合应用。2020 年 2 月发布的《智能汽车创新发展战略》中提到,要建设国家智能汽车大数据运控基础平台,逐步实现车辆、基础设施、交通环境等领域的基础数据融合应用。

隐私保护方面,2016 年我国发布的《中华人民共和国网络安全法》,对网络运营者及其他主体保护用户个人信息的责任作出了更具体严格的规定。2019 年 5 月 28 日,国家互联网信息办公室发布《数据安全管理办法(征求意见稿)》,以保护公民、法人和其他组织在网络空间的合法权益,保障个人信息和重要数据安全。同一天在贵阳数博会上,众盟数据联合中国信息通信研究院、安徽省大数据产业联盟、四川省大数据产业联合会等多家机构,共同发布了《线下大数据行业自律公约》,以推动我国线下大数据行业的规范发展,促进线下大数据行业自律机制的建设,保障线下大数据用户的权益。《智能汽车创新发展战略》中也提到要加强数据安全监督管理,对重要数据实行分级分类管理,确保用户信息、车辆信息、测绘地理信息等数据安全可控。

5.2 关键技术研究

从技术层面来看,目前国内外研究者们重点关注的热点问题包括数据处理算法、加密算法和计算资源调度算法等。对于汽车大数据而言,算法固然重要,但由于汽车大数据处理引入了分层处理架构,需要对本地处理资源、MEC 计算资源、

云计算资源进行合理调度,以确保处理效率和处理精度;此外,因汽车大数据涉及跨行业跨领域的数据合作,对于数据的共享流通相关技术也有很强的需求,面向汽车大数据应用的计算需求、通信需求和数据管理需求都需要进一步的明确,有待后续研究成果的支持。

结束语 汽车大数据种类多、来源广、数量大,数据应用需求和场景丰富,能为车企带来新的业务增长点。但在推进数据应用过程中也会遇到数据质量无法保障、数据隐私缺乏监管等问题,这些需要企业、行业和政府部门共同协作探讨,从技术、管理和监管多个层面共同推动问题的解决,为车联网场景下的汽车大数据应用提供良好的发展环境。

参考文献

[1] China Academy of Information and Communications Technology. White Paper of Connected Vehicles[R]. 2019.

[2] GRULICH P M, ZUKUNFTO. Bringing Big Data into the Car: Does it Scale? [C]// 2017 International Conference on Big Data Innovations and Applications. 2017: 9-16.

[3] BRAHIM M B, MENOUAR H. Optimizing V2X Data Collection and Storage for a Better Cost and Quality Trade-off[C]// 2017 6th IEEE International Conference on Advanced Logistics and Transport. 2017: 7-12.

[4] LIU H, QIAO C L, ZHANG Y P, et al. Influence of connected vehicles on Automotive Industry[J]. Shanghai Auto, 2016(1): 31-37.

[5] ZHU X L, ZHAO S, ZHANG L. Research on Data Integration of Automotive Industry Chain Driven by Big Data[J]. Auto Industry Research, 2016(1): 4-11.

[6] CAI J B. Innovative Application Path of Vehicular Big Data under the Environment of Connected Vehicles[J]. Automotive Electronics, 2018(17): 239-239.

[7] JOHANSON M, BELENKI S, JALMINGER J, et al. Big Automotive Data: Leveraging large volumes of data for knowledge-driven product development[C]// 2014 IEEE International Conference on Big Data. 2014: 736-741.

[8] ILLING B, WARWEG O. Achievable revenues for Electric vehicles according to current and future energy market conditions [C]// 2016 13th International Conference on the European Energy Market. 2016: 1-5.

[9] ZHANG Q W, LI J L, SUN X H. Approaches of Big Data Innovative Application Based on the Internet of Vehicles[J]. Packaging Engineering, 2017(20): 79-85.

[10] ZENG L, ZHANG K, HAN Q, et al. Research of Path Planning Model Based on Hotspots Evaluation[C]// 2019 IEEE Intelligent Vehicles Symposium. 2019: 2429-2434.

[11] LIU F X, WANG Y Z, ZHU Q Q. Research on Business Model of Automotive Industry in Big Data Era[J]. Telecom World, 2015(18): 244-245.

[12] BRAKATSOULAS S, PFOSER D, TRYFONA N. Practical data management techniques for vehicle tracking data[C]// 21st International Conference on Data Engineering. 2005: 324-325.

[13] ZHANG H, ZHANG L, CHENG X, et al. A novel precision marketing model based on telecom big data analysis for luxury cars[C]// 2016 16th International Symposium on Communications and Information Technologies. 2016: 307-311.

[14] TALEB I, SERHANI M A. Big Data Pre-Processing: Closing the Data Quality Enforcement Loop[C]// 2017 IEEE International Congress on Big Data. 2017: 498-501.

[15] ZHAO Y, YANG Z, SONG C, et al. Vehicle dynamic model-based integrated navigation system for land vehicles[C]// 2018 25th Saint Petersburg International Conference on Integrated Navigation Systems. 2018: 1-4.

[16] ZENG L, HU X C, HAN Q W. Abnormal hotspots detection method based on region real-time congestion factor[C]// 2016 IEEE 19th International Conference on Intelligent Transportation Systems. 2016: 749-753.

[17] ZHANG J, YUAN R, YAN D, et al. A Non-Cooperative Game Based Charging Power Dispatch in Electric Vehicle Charging Station and Charging Effect Analysis[C]// 2018 2nd IEEE Conference on Energy Internet and Energy System Integration. 2018: 1-6.

[18] QAREBAGH A J, SABAHI F, NAZARPOUR D. Optimized Scheduling for Solving Position Allocation Problem in Electric Vehicle Charging Stations[C]// 2019 27th Iranian Conference on Electrical Engineering (ICEE). 2019: 593-597.

[19] NOUR M, SAID S M, ALI A, et al. Smart Charging of Electric Vehicles According to Electricity Price[C]// 2019 International Conference on Innovative Trends in Computer Engineering (ITCE). 2019: 432-437.

[20] KRASNER G, KATZ E. Automatic parking identification and vehicle guidance with road awarenes[C]// 2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE). 2016: 1-5.

[21] BADER L, BÜRGER J C, MATZUTT R, et al. Smart Contract-Based Car Insurance Policies[C]// 2018 IEEE Globecom Workshops (GC Wkshps). 2018: 1-7.

[22] YUN L B, YOON S Y, JOO B H. Development of insurance server system based on vehicle driving information[C]// 2012 7th International Conference on Computing and Convergence Technology (ICCCT). 2012: 156-159.

[23] LAMBERTI F, GATTESCHI V, DEMARTINI C, et al. Blockchains Can Work for Car Insurance: Using Smart Contracts and Sensors to Provide On-Demand Coverage[J]. IEEE Consumer Electronics Magazine, 2018, 7(4): 72-81.

[24] ZHANG M, WO T, XIE T. A Platform Solution of Data-Quality Improvement for Internet-of-Vehicle Services[C]// 2018 IEEE International Conference on Pervasive Computing and Communications (PerCom). 2018: 1-7.

[25] LUAN X, CHENG L, ZHOU Y, et al. Strategies of Car-Sharing Promotion in Real Market[C]// 2018 3rd IEEE International Conference on Intelligent Transportation Engineering (ICITE). 2018: 159-163.

[26] LIANG J, LAN G H, ZHANG W W, et al. Study on Five Di-

mensional Quantitative Evaluation Method of Driving Active Service System by Vehicle Networking Environment[J]. Journal of Chongqing University of Technology(Natural Science),2018,32(1):58-67.

[27] ZHAO Y,HOU J J,YU C L. Study and Application of Industrial Big Data in Production Management and Control [J]. Computer Science,2019,46(S1):45-51.

[28] ABU A M,AL S M J,NASER S S A. The Role of Knowledge-Based Comput erized Management Information Systems in the AdministrativeDecision-Making Process[J]. International Journal of Information Technology and Electrical Engineering,2017,6(2):1-9.

[29] THORNBURG D,SCHMOTZER J,THROOP M. SAE Technical Paper Series [SAE International WCX 17:SAE World Congress Experience-(APR. 04,2017)] SAE Technical Paper Series-Vehicle Deep Data:A Case Study in Robust Scalable Data Collection[OL]. https://xueshu. baidu. com/usercenter/paper/show?paperid=cd2483190f0093f8d8d03711a60644e5&site=xueshu_se.

[30] WANG P,YU H,WANG L,et al. Development of a Novel Vehicle-Mounted Data Collection Terminal with Multisource Information Fusion for Connected Vehicles[C]// CICTP 2017. 2018:544-553.

[31] JAN B,FARMAN H,KHAN M,et al. Designing a Smart Transportation System:An Internet of Things and Big Data Approach[J]. IEEE Wireless Communications,2019,26(4):73-79.

[32] CHAQFEH M,MOHAMED N,JAWHAR I. Vehicular Cloud data collection for Intelligent Transportation Systems [C] // Smart Cloud Networks & Systems. IEEE,2017.

[33] JI C,SHAO Q,SUN J,et al. Device Data Ingestion for Industrial Big Data Platforms with a Case Study[J]. Sensors,2016,16(3):279.

[34] ABIRAMI U,SRIDEVI S. Traffic flow prophecy with mapreduce job for big data driven[C]// Eighth International Conference on Advanced Computing. IEEE,2017.

[35] IMT-2020(5G) C-V2X. MEC C-V2X White Paper [R]. 2019.

[36] WANG H,LI X,JI H,et al. Dynamic Offloading Scheduling cheme for MEC-enabled Vehicular Networks[C]// 2018 IEEE/CIC International Conference on Communications in China (ICCC Workshops). IEEE,2018.

[37] ZHOU S,NETALKAR P P,CHANG Y,et al. The MEC-Based Architecture Design for Low-Latency and Fast Hand-Off Vehicular Networking[C]// IEEE Vehicular Technology Conference (VTC). IEEE,2018.

[38] XIE J D,JIA Y J,CHEN Z C,et al. Efficient Task Completion for Parallel Offloading in Vehicular Fog Computing[J]. China Communications,2019,16(11):42-55.

[39] HUANG C M,CHIANG M S,DAO D T,et al. V2V Data Offloading for Cellular Network based on the Software Defined Network (SDN) inside Mobile Edge Computing (MEC) Architecture[J]. IEEE Access,2018,6:17741-17755.

[40] ZENG L,ZHANG J,HAN Q,et al. A Bus-Oriented Mobile FCNs Infrastructure and Intra-Cluster BSM Transmission Mechanism[J]. IEEE Access,2019,7:24308-24320.

[41] SHRIVASTVA K M P,RIZVI M A,SINGH S. Big Data Privacy Based on Differential Privacy a Hope for Big Data[C]// 2014 International Conference on Computational Intelligence and Communication Networks. 2014:776-781.

[42] NAWRATH T,FISCHER D,MARKSCHEFFEL B. Privacy-sensitive data in connected cars[C]// Internet Technology & Secured Transactions. IEEE,2017.

[43] ENDO T,NAWA K,KATO N,et al. Study on privacy setting acceptance of drivers for data utilization on connected cars[C]// 2016 14th Annual Conference on Privacy,Security and Trust (PST). IEEE,2016.

[44] KOCH A,ALTSCHAFFEL R,KILTZ S,et al. Exploring the Processing of Personal Data in Modern Vehicles—A Proposal of a Testbed for Explorative Research to Achieve Transparency for Privacy and Security[C]// 2018 11th International Conference on IT Security Incident Management & IT Forensics (IMF). 2018:15-26.

[45] XU H T. Research on Personal Data Protection of Automobile Consumers in Big Data Era[J]. Journal of Hubei University of Automotive Technology,2018,32(1):72-76.

GE Yu-ming,born in 1985,Ph.D,senior engineer. His main research interests include connectivity,MEC,cybersecurity in the areas of connected and automated vehicles,and industrial internet.

WANG Miao-qiong,born in 1991,postgraduate,Engineer. Her main research interests include big data infrastructure,DBMS and industrial applications of big data technology.