



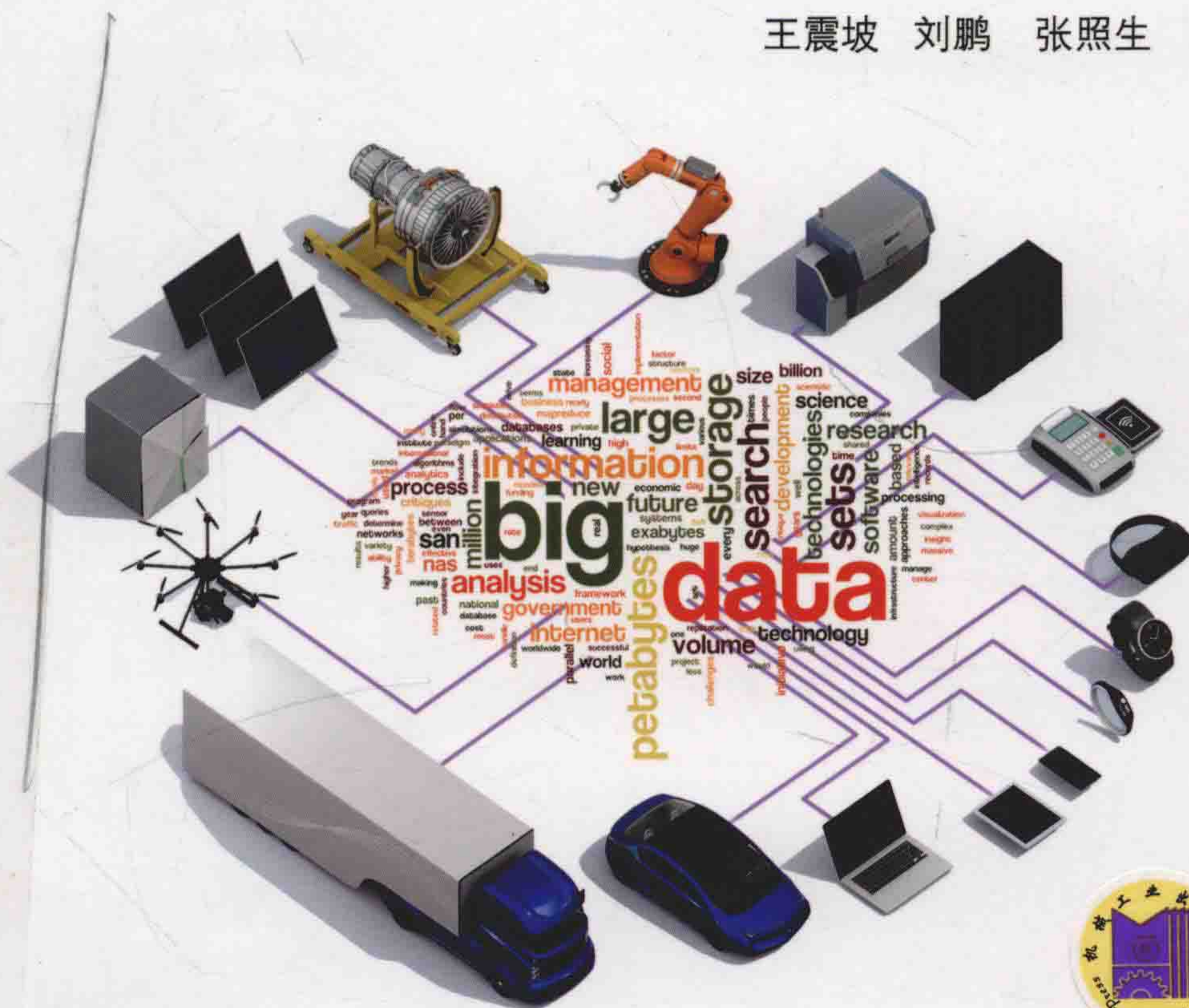
新能源汽车关键技术研发系列

新能源汽车

大数据分析与应用技术

Analysis and Application
Technology of Big Data for EVs

王震坡 刘鹏 张照生 编著

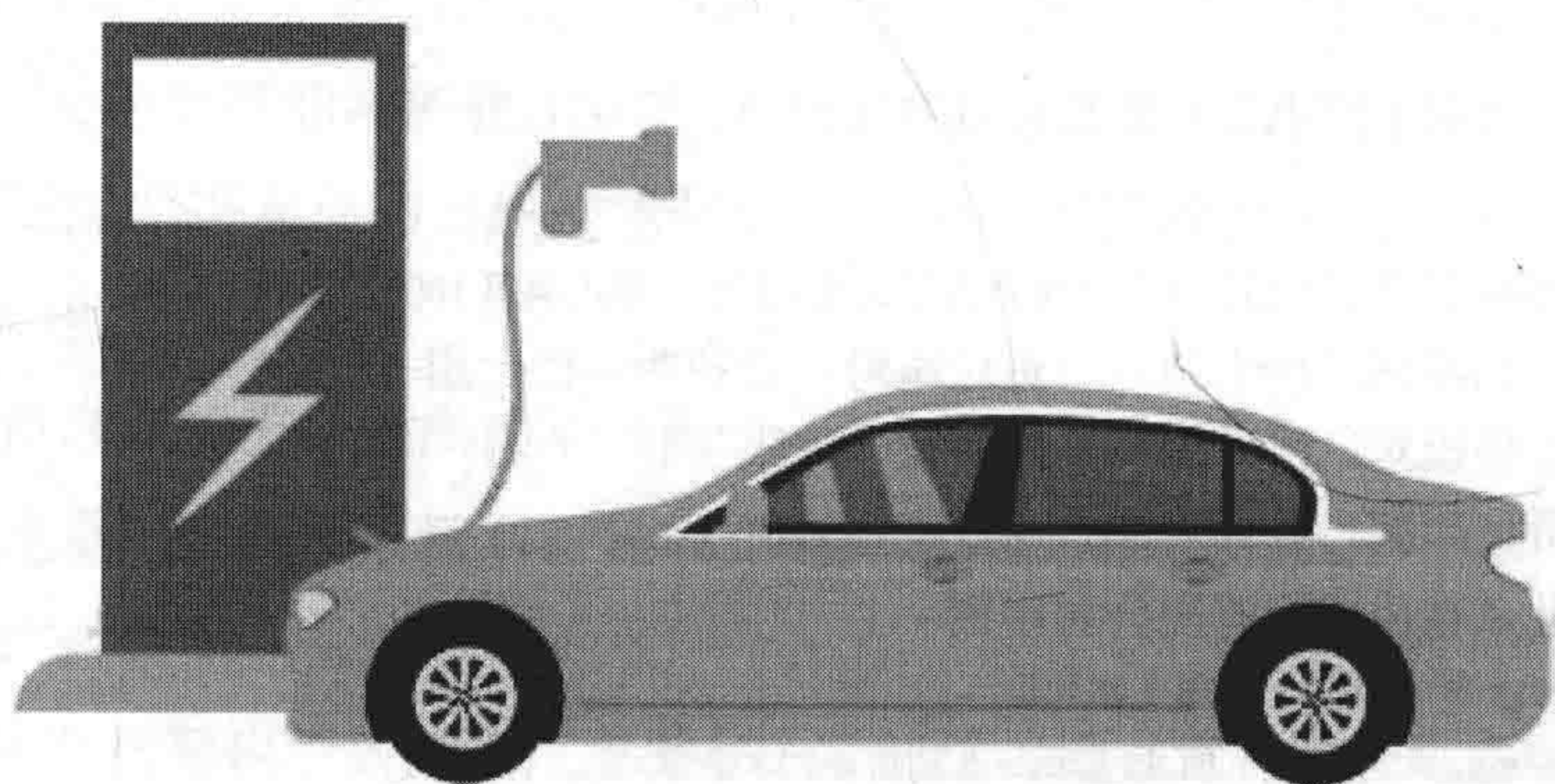


 机械工业出版社
CHINA MACHINE PRESS

新能源汽车关键技术研发系列

新能源汽车 大数据分析 与应用技术

王震坡 刘 鹏 张照生 编著



机械工业出版社

本书涵盖了新能源汽车的车联网技术、大数据应用的业务需求、大数据分析基础理论、大数据的采集与处理、车辆运行大数据统计分析与应用实例等内容,详细介绍了车联网车载数据采集、网络通信等方法的实现和应用,并以较多实例展示了新能源汽车大数据的应用,可为从事车联网和大数据分析工作的研究人员和相关工作者提供借鉴和参考,对新能源汽车大数据的平台构建、数据管理、数据应用有很实用的指导意义。

图书在版编目(CIP)数据

新能源汽车大数据分析与应用技术 / 王震坡, 刘鹏, 张照生编著. —北京: 机械工业出版社, 2018.5

(新能源汽车关键技术研发系列)

ISBN 978-7-111-59638-7

I. ①新… II. ①王…②刘…③张… III. ①互联网络-应用-新能源汽车-研究 IV. ①U469.7

中国版本图书馆 CIP 数据核字 (2018) 第 071838 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑: 何士娟 责任编辑: 张利萍 何士娟

责任校对: 潘蕊 封面设计: 张静

责任印制: 常天培

北京铭成印刷有限公司印刷

2018 年 6 月第 1 版第 1 次印刷

169mm×239mm·14.75 印张·2 插页·277 千字

0001—3000 册

标准书号: ISBN 978-7-111-59638-7

定价: 89.90 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

电话服务

网络服务

服务咨询热线: 010-88361066

机工官网: www.cmpbook.com

读者购书热线: 010-68326294

机工官博: weibo.com/cmp1952

010-88379203

金书网: www.golden-book.com

封面无防伪标均为盗版

教育服务网: www.cmpedu.com

作者简介

王震坡，工学博士，北京理工大学教授、博导，新能源汽车领域专家，现任电动车辆国家工程实验室副主任、新能源汽车国家大数据联盟秘书长，在新能源汽车大数据分析挖掘、动力电池系统集成控制与成组技术等领域具有丰富的经验。

刘鹏，工学博士，新能源汽车领域专家，现任北京电动车辆协同创新中心办公室主任、新能源汽车国家大数据联盟副秘书长，长期从事新能源汽车大数据分析挖掘工作，参与多项新能源汽车行业相关政策与标准的研究和制定。

张照生，工学博士，新能源汽车领域专家，长期从事新能源汽车及交通领域大数据分析挖掘工作，参与了多项新能源汽车行业相关政策与标准的研究和制定。



Analysis and Application
Technology of Big Data for EVs

近几年来，随着计算机与信息技术的迅猛发展和普及应用，行业应用系统的规模迅速扩大，行业应用所产生的数据呈爆炸式增长。新能源汽车运行所产生的数据规模极其庞大，车辆技术和大数据技术的结合，能挖掘出更有价值的资源和财富。在未来发展中，大数据与车辆技术的结合将会成为非常有前景的学科与研究领域，需要大量的人才与资源进行支撑。

本书系统地介绍了新能源汽车数据采集、存储、处理与挖掘流程，辅以丰富的大数据分析案例对数据挖掘流程进行详细的展示，借助大数据方法和工具从不同角度探索新能源汽车海量数据背后隐藏的价值。希望本书能够起到抛砖引玉的作用，为从事车联网与大数据分析工作的研究人员与相关工作者提供借鉴和参考。



Analysis and Application Technology of Big Data for EVs

丛书序



在新能源汽车成为战略新兴产业之一等国家战略的背景下，以纯电动汽车和燃料电池汽车、插电式混合动力汽车为代表的新能源汽车，作为能源网络中用能、储能和回馈能源的终端，成为我国乃至经济新体系中的重要组成部分。我国经过4个五年计划的科技攻关，基本掌握了新能源汽车的整车技术和关键零部件技术，实现了跨越式发展，并逐步实现了产业化。

但是，在世界这个完全开放的市场中，中国新能源汽车核心关键技术尚未彻底突破，技术竞争压力越来越大，加快新能源汽车持续创新，推进中国汽车产业转型升级，是中国科技发展的重大战略需求。尽管我们头顶着全球最大新能源汽车市场的光环，但中国的新能源汽车产业正遭遇着成长的烦恼：

1. 与国际先进水平和市场需求相比，中国的新能源汽车技术水平及产品性能需要进一步提高。
2. 示范区域的市场发展尚不平衡。
3. 充电基础设施发展滞后，建成充电桩总体使用率较低。
4. 推广政策尚存在缝隙。

本套丛书将聚焦于新能源汽车整车、零部件关键技术，以及与新能源汽车配套的科技体系和产业链，邀请行业内各领域一直从事研究和试验工作的产品第一线技术人员编写，内容系统、科学，极具实用性，希望能够为我国新能源汽车的持续发展提供技术支撑和智力支持。

前 言



随着电子信息技术的进一步发展，电气化、智能化及网联化成为当今汽车工业发展的主旋律。本书是在新能源汽车国家大数据联盟的支持下出版的一本跨领域的专著，是新能源汽车+车联网+大数据知识的全面融合。

近几年来，随着计算机和信息技术的迅猛发展和普及应用，行业应用系统的规模迅速扩大，行业应用所产生的数据呈爆炸性增长。动辄达到数百太字节（1TB=1024GB）甚至数十至数百拍字节（1PB=1024TB）规模的行业/企业大数据已远远超出了现有传统的计算技术和信息系统的处理能力。因此，寻求有效的大数据处理技术、方法和手段已成为业界的迫切需求。

车联网（Connected Vehicles）是由车辆位置、行驶速度、行驶路线等构成的信息交互网络，是一种向信息通信、环保、节能、安全方向发展的车-网联合技术。数据是车联网的基础，车联网打通了传统数据孤岛，在人—车—基础设施之间进行信息交互，连点成线，连线成网，实现人—车—环境的信息通信无缝融合。

因为结构简单、电子设备配置齐全，所以新能源汽车是车联网的最佳载体。新能源汽车运行所产生的数据规模极其庞大，车辆技术和大数据技术的结合，能挖掘出更有价值的资源和财富。在未来发展中，大数据与车辆技术的结合将会成为非常有前景的学科和研究领域，需要大量的人才和资源进行支撑。

目前全世界的专家、学者、相关行业人员，以及政府、高校、公司等各大组织机构都在致力于研究和开发车辆大数据相关领域，并推进该领域的进步和发展。

本书涵盖了新能源汽车的车联网技术、大数据应用的业务需求、大数据分析与应用基础理论、大数据的采集与处理、车辆运行大数据统计分析与应用实例等内容，详细介绍了车联网车载数据采集、网络通信等方法的实现和应用，并以较多实例展示了新能源汽车大数据的应用，可为从事车联网和大数据分析工作的研究人员

和相关工作者提供借鉴和参考，对新能源汽车大数据的平台构建、数据管理、数据应用有很强的指导意义。国内还没有像这样将新能源汽车行业 and 大数据应用结合起来的图书。本书适合与新能源汽车行业相关的从业人员，从事相关科研项目或相关专业的教师、学者，以及本科生、研究生阅读。

本书在撰写过程中得到了科技部高新技术发展及产业化司，国家自然科学基金委员会，北京市科委双新处、科室处、新能源汽车发展促进中心以及其他有关单位和领导的大力支持与帮助，在此对他们表示衷心的感谢。本书是项目“大数据驱动的服务运营系统性优化与管理——以新能源汽车为例”（项目批准号：91746210）、课题“基于大数据的新能源公交客车运行监测与管理关键技术研究”（课题编号：2017YFC0840205）和课题“电动汽车产业培育要素分析研究”（课题编号：2171100003217074）的研究成果之一。

限于作者水平和条件所限，书中难免有不妥和错漏之处，恳请读者批评指正。

编著者

2018年1月

目 录



丛书序

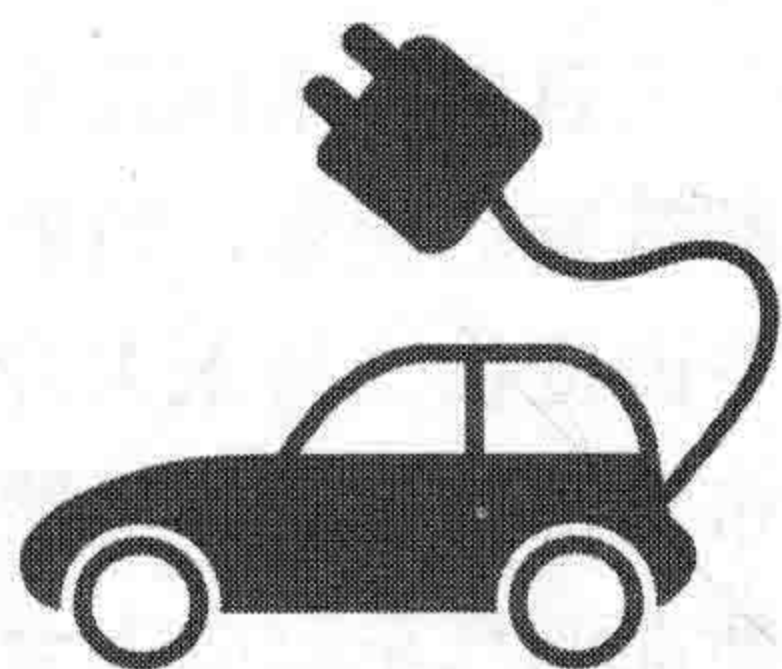
前言

第 1 章	导论	1
1.1	汽车电气化与智能化	1
1.1.1	电气化	1
1.1.2	智能化	2
1.2	信息化的车联网	3
1.2.1	车联网技术	3
1.2.2	国内外发展历史及现状	7
1.2.3	车联网发展趋势	10
1.3	大数据简介	11
1.3.1	大数据的产生与发展	12
1.3.2	大数据的概念与特征	13
1.3.3	大数据的价值与挑战	15
1.4	车辆大数据与应用	18
1.4.1	汽车行业大数据应用	19
1.4.2	新能源汽车大数据应用	20
第 2 章	新能源汽车车联网技术	23
2.1	新能源汽车与数据采集	24
2.1.1	纯电动汽车	24
2.1.2	混合动力电动汽车	26
2.1.3	燃料电池电动汽车	30
2.2	车辆数据通信技术	31
2.2.1	CAN 总线的数据交换原理	32
2.2.2	CAN 总线的特征和分层结构	34
2.2.3	CAN 总线在汽车控制系统中的应用	35
2.2.4	FlexRay 总线	35

2.3	车载设备应用	36
2.3.1	汽车厂商领域	36
2.3.2	公共交通领域	37
2.3.3	私人交通领域	39
2.4	新能源汽车车联网大数据平台	40
2.4.1	新能源汽车大数据平台的应用背景	40
2.4.2	新能源汽车大数据平台的架构	42
2.4.3	新能源汽车大数据平台的功能	43
2.5	新能源汽车车联网应用实例分析	44
2.5.1	车联网技术在智慧交通方面的应用	44
2.5.2	车联网在新能源汽车上的应用	46
第3章	新能源汽车大数据应用流程	50
3.1	数据采集	50
3.1.1	数据采集方法	50
3.1.2	新能源汽车数据采集	51
3.1.3	小结	54
3.2	数据预处理	54
3.2.1	数据预处理目标	54
3.2.2	数据预处理方法	55
3.2.3	新能源汽车数据预处理	66
3.2.4	小结	68
3.3	数据存储	69
3.3.1	数据存储技术路线	69
3.3.2	新能源汽车数据存储	70
3.3.3	小结	72
3.4	数据探索与分析	72
3.4.1	数据探索与统计	73
3.4.2	数据挖掘与应用	75
3.4.3	新能源汽车数据探索与分析	80
3.4.4	小结	82
3.5	数据可视化	83
3.5.1	数据可视化工具	84
3.5.2	新能源汽车数据可视化	90
3.5.3	小结	93

第4章 数据分析的基础理论	94
4.1 相关与回归分析	94
4.1.1 相关分析	94
4.1.2 一元线性回归分析	99
4.1.3 多元线性回归分析	109
4.1.4 非线性回归模型	114
4.1.5 小结	117
4.2 聚类方法	117
4.2.1 聚类方法概要	118
4.2.2 K-means 方法	121
4.2.3 层次聚类	122
4.2.4 类别数的确定方法	125
4.2.5 小结	127
4.3 分类方法	127
4.3.1 分类方法概要	127
4.3.2 K-近邻	129
4.3.3 贝叶斯分类	132
4.3.4 分类的评判	135
4.3.5 小结	138
4.4 诊断方法	139
4.4.1 离群点诊断概要	139
4.4.2 基于统计的离群点诊断	140
4.4.3 基于距离的离群点诊断	142
4.4.4 基于密度的离群点挖掘	144
4.4.5 基于聚类的离群点挖掘	145
4.4.6 小结	146
4.5 时间序列数据分析与预测	146
4.5.1 时间序列概述	146
4.5.2 时间序列的描述分析	148
4.5.3 长期趋势分析	151
4.5.4 季节变动分析	153
4.5.5 循环变动分析	154
4.5.6 时间序列的预测	155
4.5.7 小结	157

第 5 章 新能源汽车的运行大数据统计分析与应用实例	158
5.1 新能源汽车大数据的应用概述	158
5.1.1 大数据标准化	158
5.1.2 大数据应用领域	159
5.2 新能源汽车技术分析与应用	160
5.2.1 动力电池系统故障分析	160
5.2.2 动力电池系统健康状态评估	170
5.2.3 动力电池系统梯次利用分析	175
5.3 新能源汽车使用行为分析与应用	180
5.3.1 驾驶行为分析与应用	180
5.3.2 充电行为分析	185
5.3.3 行驶里程分析	189
5.4 新能源汽车宏观经济分析与应用	195
5.4.1 分时租赁应用	195
5.4.2 对城市交通运行的影响分析	204
第 6 章 大数据分析在未来交通出行中的应用及发展前景	213
6.1 未来的交通出行	213
6.1.1 未来的汽车出行	213
6.1.2 未来的交通网络	214
6.2 未来交通出行中大数据的分析与应用	215
6.2.1 未来汽车行业——以人为本	216
6.2.2 未来交通系统——智慧出行网络	218
6.2.3 未来社会发展——国计民生	220
6.3 未来新挑战	222
6.3.1 数据的安全性	222
6.3.2 数据的复杂性	223
6.3.3 计算的复杂性	223
6.3.4 系统的复杂性	224
参考文献	225



汽车作为人们日常出行中重要的交通工具，自 1886 年发明以来经过了 100 多年的发展历程。这不仅仅是汽车的发展历史，也代表了现代工业文明的发展轨迹。可以说汽车是人类技术发展方向的一个重要载体，人类每一次的技术进步都会在车辆制造上有相应的体现。进入 21 世纪以来，人们对于汽车改造的主要方向是让汽车在满足原有的运载能力的基础上实现节能减排以及车辆网联化的目的。本章将对目前汽车的发展趋势进行介绍，并着重介绍目前在汽车工业领域研究发展十分迅猛的车联网技术和车辆大数据分析与应用技术。

1.1 汽车电气化与智能化

1.1.1 电气化

随着社会的迅速发展，能源危机已成为世界上绝大多数国家都必须面对的问题。化石能源作为不可再生能源，将在可预见的未来成为稀缺资源。而汽车是消耗化石能源的主要工业品之一，其尾气排放引起的环境污染问题也日益严峻。在节能减排的迫切需求下，新能源汽车凭借其能源经济性与环境友好性在汽车产业中得到了企业和消费者的一致青睐。

新能源汽车在解决能源与环境问题上有着巨大的潜力。首先，作为电气化核心部件，驱动电机的效率非常高，最高能达到 97%，相对于发动机大约 30% 的热效率有着巨大的优势。此外，电机的全工况效率很高，有着“双 80”的说法，即 80% 以上的工况下的效率都是在 80% 以上的，这是传统内燃机远达不到的。而且在驱动电机的配合下，能够实现传统动力总成的效率最大化。以君越 30H 为例，在驱动电机的配合下，发动机可以采用米勒循环（或称阿特金森循环），发动机最高效率可达 40%，而通过双电机与双行星排的配合，可以让发动机一直工作在最高效区域。

目前,全球几乎所有主流车企都在大力推广新能源汽车。在美国、日本、欧洲等发达国家,政府纷纷出台各种扶持和优惠政策,在技术研发、税收、补贴等方面提供支持来大力推动新能源汽车发展,引导新能源汽车产业从由政府主导到市场消费推动。虽然新能源汽车目前还处于产业化和商业化的初期阶段,但发展新能源汽车已成为全球汽车业未来发展的重点方向。

1.1.2 智能化

随着互联网技术、通信技术、人工智能、计算机技术的快速发展,智能化已经成为一种潮流和趋势。从智能手机、智能家电,到企业的智能制造、智能物流等,智能化已经渗透到整个社会的各行各业。在“工业4.0”“智能交通”“智慧城市”和“互联网+”的大背景下,汽车智能化已经成为汽车产业发展的重要潮流和趋势。

智能汽车在解决能源、安全和环境问题上具有巨大的潜力,例如:通过采用自动驾驶技术能够减少90%由于人为操作引起的交通事故;通过车-车通信和智能速度规划,在智能化发展的前期可以将道路通行率提高10%以上,在高度自动化阶段可以将道路通行率提高50%~90%;在节能减排方面,通过经济性驾驶和整体智能交通规划,能源消耗至少能降低15%~20%。由于智能汽车存在的巨大潜力,汽车的智能化已经成为行业发展的热点,并且正在引发行业的巨大变革。可以预见,汽车的电动化与智能化将会对传统的汽车行业格局产生很大的冲击。如果说汽车的电气化只是将汽车的动力由内燃机换为电动机,对传统汽车行业格局的改变有限,那么汽车的智能化就是把一辆汽车变化成一辆有着自己智慧的机器。

在车辆智能化的背景下,世界各国纷纷制定相应的汽车智能化研究计划,欧盟、美国和日本均发布政策法规来推动智能网联汽车发展。中国在《中国制造2025》中也明确给出了汽车智能化技术的总体目标,即制定中国自动驾驶标准:基于多源信息融合、多网融合,利用人工智能、深度挖掘及自动控制技术,配合智能环境和辅助设施实现自动驾驶;可改变出行模式、消除拥堵、提高道路利用率;装备自动驾驶系统的汽车,综合能耗较常规汽车降低10%,排放减少20%,交通事故次数减少80%,基本消除交通死亡。在《中国制造2025》后,国家工业和信息化部、发改委、测绘局等相关部委出台多部政策,从汽车智能化、网联化、智能制造、地图信息采集、大数据等多个方面促进智能汽车的发展。

1.2 信息化的车联网

物联网被称为是继计算机、互联网之后世界信息产业发展的第三次浪潮。在中国，物联网已经被正式列为国家五大新兴战略性产业之一，而车联网是战略新兴产业中物联网与智能汽车两大领域的重要交集。通过车联网技术，汽车厂商能够为消费者提供全方位的个性化服务，提升消费者的使用体验；此外，通过汽车、道路和基础设施的相互联通（图 1-1 和图 1-2），能够有效提高道路通行效率，减少交通碰撞事故的发生；另外，车联网技术还能够降低交通对环境的影响，在环境保护方面也发挥着重要的作用。

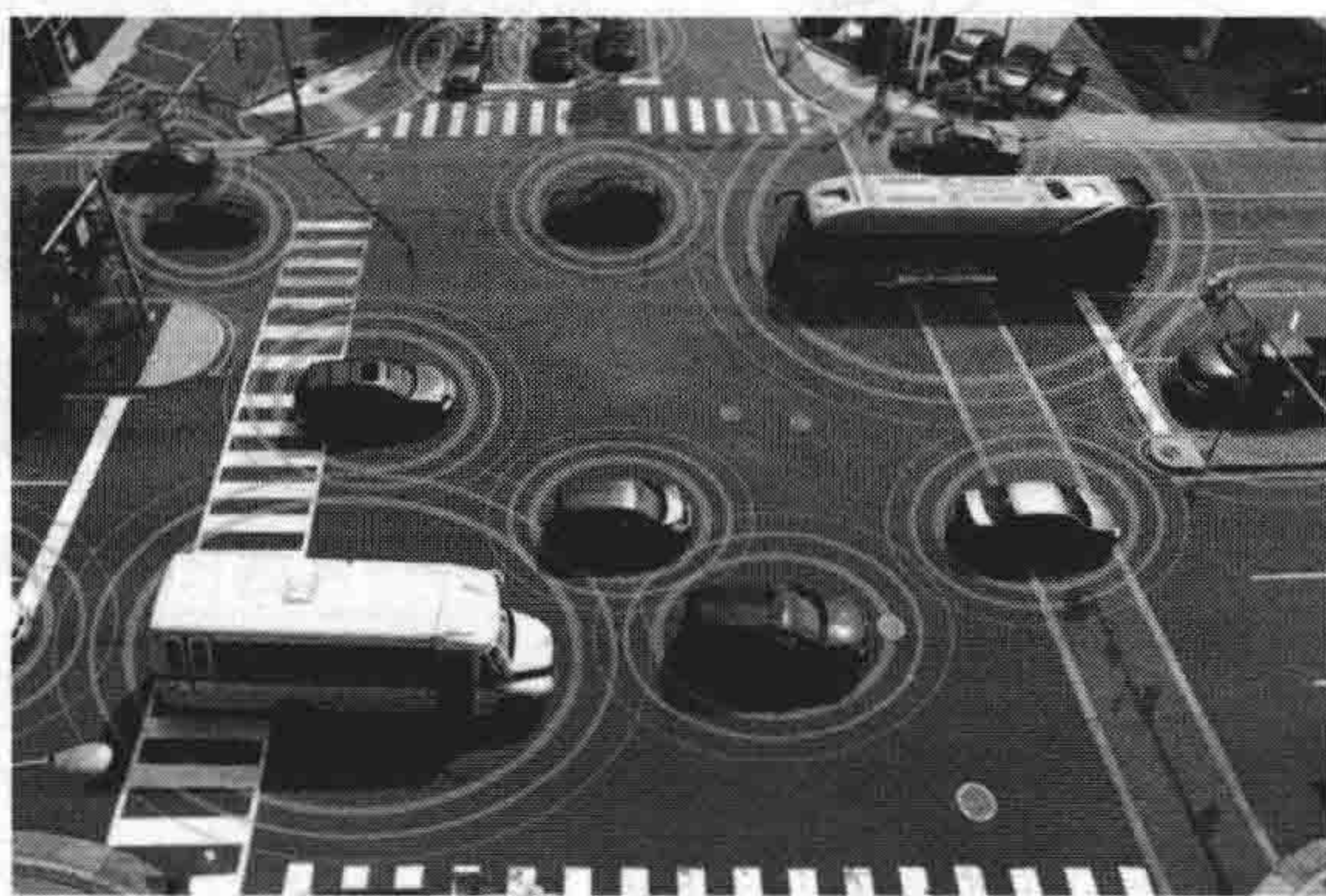


图 1-1 汽车与汽车的相互联通

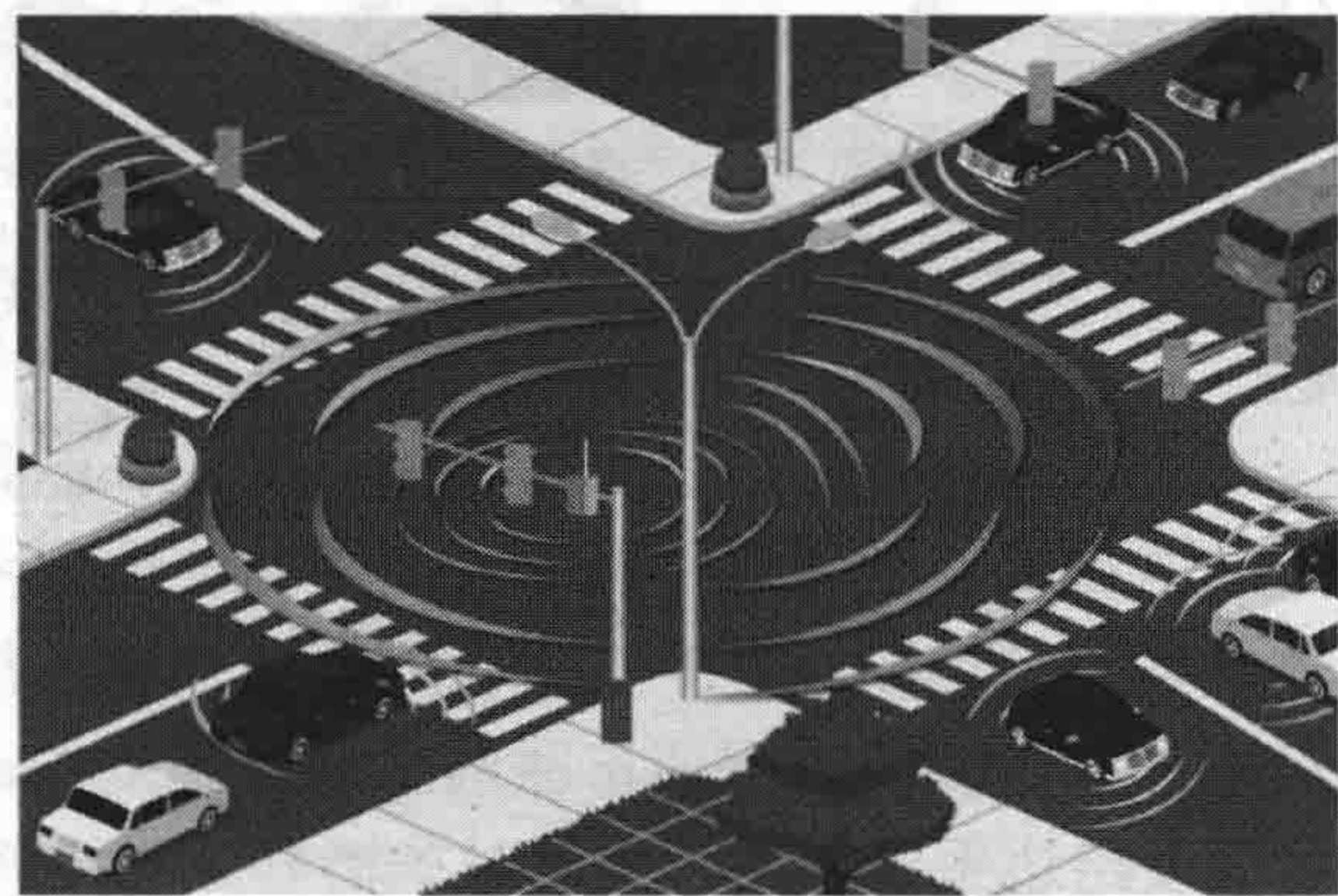


图 1-2 汽车与基础设施的相互联通

1.2.1 车联网技术

根据中国物联网校企联盟的定义，车联网（Internet of Vehicles）是由车辆位

置、速度和路线等信息构成的巨大交互网络。通过 GPS、RFID、传感器、摄像头图像处理等装置，车辆可以完成自身环境和状态信息的采集；通过互联网技术，所有车辆可以将自身的各种信息传输汇聚到中央处理器；通过计算机技术，这些车辆的信息可以被分析和处理，从而计算出不同车辆的最佳路线，并及时汇报路况、安排信号灯周期。简言之，车联网是以车、路以及道路的基本设施为节点组成网络，用以实现车与车、车与人、车与路的信息交换，利用先进的技术（包括网络技术、传感器技术、控制技术、计算技术、智能技术等）实现安全防护、智能驾驶、车辆售后服务、位置服务，最终达到提高交通效率、提升道路通行能力、降低交通事故等目的。

车联网的基本架构如图 1-3 所示。

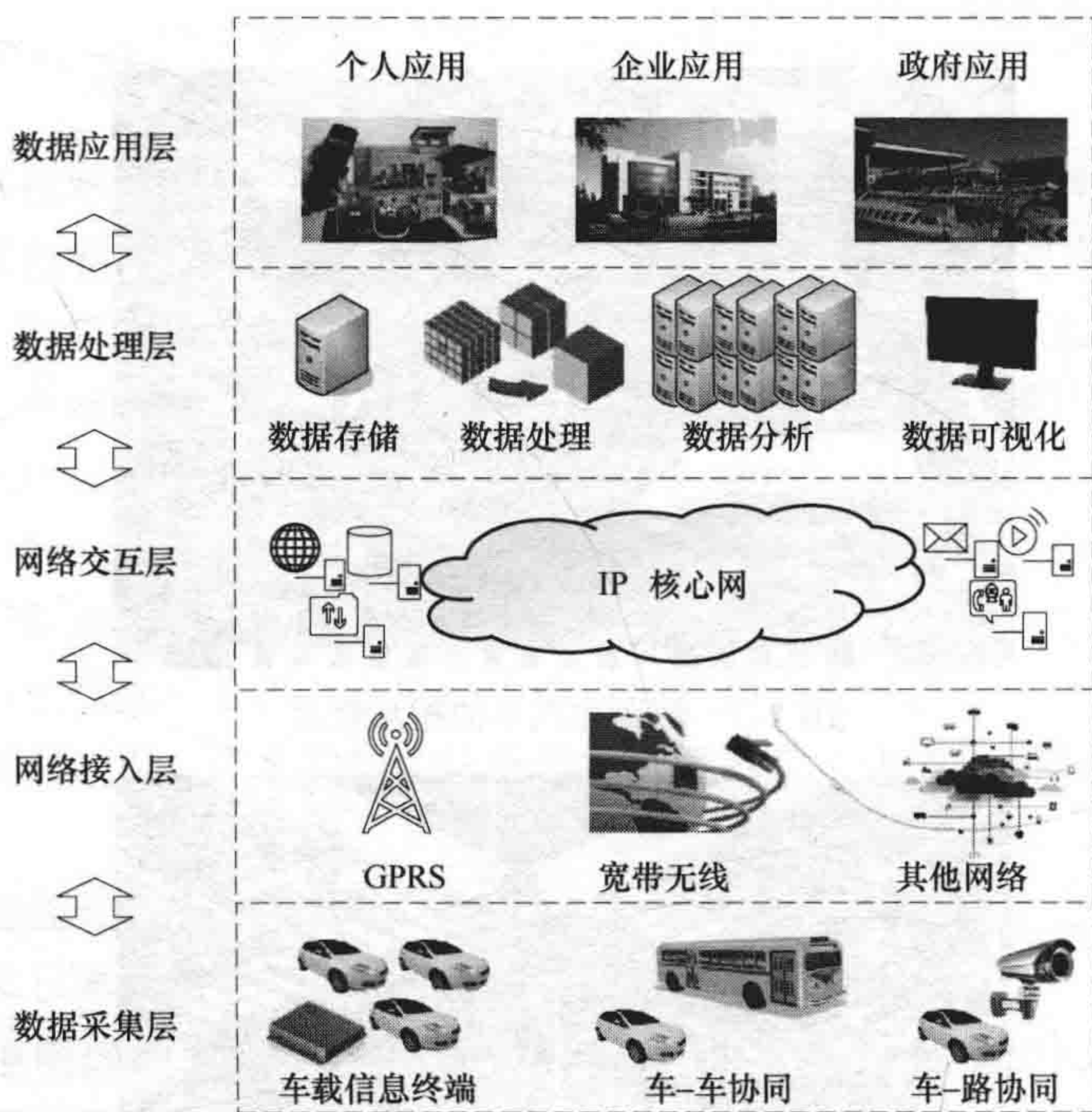


图 1-3 车联网的基本架构

根据车联网的基本框架结构，为保证车联网系统顺利工作，首先要通过感知技术、车载信息终端以及路边系统设备，实现对车辆自身的位置、速度、加速度、行进方向等行驶和运行信息以及车辆外在属性（如道路、人和环境）等信息的提取，通过轻量级的车载设备完成车辆相关信息的收集和处理，同时接收和执行来自上层的智能交通和信息服务等交互控制指令。在该过程中，汽车既是数据的收集和感应器，也是实时信息的发布者。

车联网基于 GPRS、3G、4G 以及未来通信网络（5G）等移动通信网络和宽带无线城域网络基础设施，实现运行系统（车辆信息系统、路网信息、信息采集基站系统和运行管控服务中心系统）和运营系统（运营管控平台系统、关键服务子系统）之间的数据传输。然后通过移动无线网和专用核心网实现汽车信息源与数据中心之间的信息传输，提供用户终端连接和对用户终端的管理，完成对业务的承载。作为承载网络提供到外部网络的接口，从而实现汽车各种服务、管理和交互过程的控制。

最后，数据平台能够对在网车辆和设施产生的海量数据的存储和处理提供支撑，同时集成其他服务基础数据，为智能交通管控和车载信息服务提供支撑。智能交通管理中心拥有超大的数据库和数据分析能力，用以存储、分析从路边设施传来的数据，并根据分析结果发送相应指令。车载信息服务与运营中心负责面向不同类型用户提供开放多样的车载信息服务，同时提供安全可靠的运营支撑环境，支持具有新型服务形态和商业模式的车联网应用的开展。

为实现以上过程及服务，有一些关键的技术需要解决，包括异构无线网络的融合、全面的感知、智能化信息处理、与新能源汽车的整合。具体来说，车联网需要解决的关键性技术问题可总结为以下 4 条。

（1）异构无线网络的融合

在车联网中将有多种不同的无线通信技术并存，包括 WLAN（如 IEEE 802.11a/b/g/n/p 协议）、WIMAX（IEEE802.16a/e）、超宽带通信 UWB（IEEE 802.15.3a）、2G/3G/4G/5G 蜂窝通信、LTE 以及卫星通信等网络。不同的网络有不同的通信方式和特点，适用于不同的场景。为了达到信息共享的目的，车联网中的很多信息需要在不同的网络中传递。同时，车辆作为一个移动单元，在移动过程中将发生水平切换和垂直切换，也需要进行移动性管理。因此，需要在车联网环境下考虑异构无线网络的融合，实现无缝的信息交换和无缝的网联切换需求。

（2）全面的感知

车联网想要为地面交通提供极限通行能力，首先必须依赖于全面的感知，包括对整个道路的感知和对车辆的感知，从而分别结合道路和车辆获取相应的状态信息。如今，各种不同类型的感知节点已经大量应用于地面交通。如何将这多元的感知节点进行有效的利用是一个非常关键的问题。它涉及感知节点的选择、功能定位（如汇聚节点）、布局、特征提取与分析以及多元信息的融合。车内感知和车外感知考虑的重点不一样，而道路的感知与车辆状态的感知关注的重点也不一样。比如，道路感知对路面是否结冰很关心，但车辆感知可能更关心车辆的行驶速度和当前的位置。