

车-路-云协同自动驾驶安全技术研究进展

孙康康^{1,2}, 李建华^{1,2}, 陈秀真^{1,2}, 过敏意¹

(1. 上海交通大学计算机学院, 上海 200240;

2. 上海市信息安全综合管理技术研究重点实验室, 上海 200240)

摘要: 随着边缘智能技术的进步和城市化进程的加快, 智能交通系统 (intelligent transportation system, ITS) 得到快速发展。车-路-云协同实现车联网中车-车、车-路、车-云数据协同共享, 构建了更加高效的协作智能交通系统 (cooperative-intelligent transportation system, C-ITS)。然而, 车-路-云协同存在大量的安全威胁, 严重阻碍了协同自动驾驶的发展。首先总结了车-路-云协同的发展现状, 详细介绍了自动驾驶发展历史及车-路-云协同自动驾驶环境。然后系统性地将车-路-云协同中遇到的攻击及安全防御技术归纳成两类: 针对经典信息安全机制的威胁与防御技术, 主要从基于信息可用性、基于信息完整性、基于信息保密性、基于信息真实性、基于信息不可否认性5部分进行详细介绍; 针对基于机器学习技术的威胁与防御技术, 主要从集中式与分布式两部分进行分析。最后展望了车-路-云协同安全技术的发展方向和重点, 主要从联邦学习、区块链技术、安全多方计算、零知识证明、差分隐私技术展开介绍。

关键词: 车-路-云协同; 协作智能交通系统; 数据安全; 安全威胁

中图分类号: TP393.0

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2025058

Research progress on autonomous driving security technology for vehicle-road-cloud collaboration

SUN Kangkang^{1,2}, LI Jianhua^{1,2}, CHEN Xiuzhen^{1,2}, GUO Minyi¹

1. School of Computer Science, Shanghai Jiao Tong University, Shanghai 200240, China

2. Shanghai Municipal Key Lab of Integrated Management Technology for Information Security, Shanghai 200240, China

Abstract: With the advancement of edge intelligence technology and the acceleration of urbanization, intelligent transportation systems (ITS) have experienced rapid development. Vehicle-road-cloud (VRC) collaboration was enabled through the coordinated sharing of vehicle-to-vehicle (V2V), vehicle-to-road (V2R), and vehicle-to-cloud (V2C) data in the Internet of vehicles, thereby constructing a more efficient cooperative intelligent transportation

收稿日期: 2025-01-10; 修回日期: 2025-05-28

通信作者: 李建华, lijh888@sjtu.edu.cn

基金项目: 国家自然科学基金 (U2003206, 62501397, 62471301); 上海市科委“科技创新行动计划” (22511101202)

Foundation Items: The National Natural Science Foundation of China (U2003206, 62501397, 62471301), Science and Technology Commission of Shanghai Municipality “Innovation Action Plan” (22511101202)

引用格式: 孙康康, 李建华, 陈秀真, 等. 车-路-云协同自动驾驶安全技术研究进展[J]. 网络与信息安全学报, 2025, 11(5): 1-30.

Citation Format: SUN K K, LI J H, CHEN X Z, et al. Research progress on autonomous driving security technology for vehicle-road-cloud collaboration[J]. Chinese Journal of Network and Information Security, 2025, 11(5): 1-30.

system (C-ITS). However, numerous security threats in VRC collaboration were found to severely impede the development of cooperative autonomous driving. The development status of VRC collaboration was first summarized, and the history of autonomous driving and the VRC-based autonomous driving environment were elaborated. Subsequently, attacks and security defense technologies in VRC collaboration were systematically categorized into two types: classical information security mechanisms and defense technologies, which were detailed from five aspects—information availability, integrity, confidentiality, authenticity, and non-repudiation; and machine learning-based security threats and defense technologies, which were analyzed from both centralized and distributed perspectives. Finally, future development directions and research priorities of VRC collaborative security technologies were forecasted, primarily covering federated learning, blockchain technology, secure multi-party computation, zero-knowledge proof, and differential privacy technology.

Keywords: vehicle-road-cloud collaboration, collaborative intelligent transport system, data security, security threats

0 引言

随着车辆电动化、网联化、智能化的不断发展,智能网联汽车逐渐集成 5G/6G 通信、蜂窝车联网、人工智能、云计算等先进技术,形成新一代车-路-云智能交通系统。通过车-路-云之间的高效协同,智能网联汽车能够对交通环境进行全面感知,实现多个系统间、大范围、大容量的数据交互,对每辆智能网联汽车进行交通全程控制,对每一条道路进行交通全时空控制,以提供交通效率和交通安全为主的网络与应用服务。交通运输部数据^[1]显示,2023 年我国 L2 级及以上智能网联汽车渗透率达 30%,但伴随而来的通信时延、感知盲区等安全隐患导致交通事故率同比 2022 年上升 18%,其中因单车智能决策失误引发的事故占比高达 62%。因此,仅依靠单车智能已经无法满足动态复杂的交通环境,亟须构建更加安全可靠的车-路-云协同系统。

为应对上述挑战,近年来,我国出台了一系列促进智能网联汽车发展的措施,特别是构建车-路-云协同自动驾驶技术的发展措施。2015 年,国务院印发了《中国制造 2025》^[2],将智能网联汽车列入未来十年国家智能制造发展的重点领域。2018 年 1 月,国家发展和改革委员会等 11 个部门联合印发了《智能汽车创新发展战略》^[2],对智能网联汽车的发展进行了 3 个阶段的规划。到 2020 年,智能网联汽车新车规模占比达到 50%,中高级别智能网联汽车实现市场化应用,车用无线通信网络在大城市、高速公路的覆盖范围达到 90%,实现北斗高精度

卫星服务全面覆盖;到 2025 年,智能网联汽车基本实现智能化,车-路-云实现高度协同,新一代车用无线通信网络基本满足智能网联汽车发展需求;到 2035 年,中国将率先实现智能汽车强国。2023 年 11 月,工业和信息化部、公安部、住房和城乡建设部、交通运输部共同发布了《关于开展智能网联汽车准入和上路通行试点的通知》^[3],正式启动了智能网联汽车准入和上路通行试点,有利于我国在全球自动驾驶领域抢占技术制高点。2024 年 1 月 15 日,工业和信息化部等联合发布关于开展智能网联汽车车-路-云一体化应用试点,标志着我国智能网联汽车产业进入了全面协同、规模化应用的新阶段^[4]。

目前,智能网联汽车的研发已经成为全球汽车产业发展的战略方向。国际上多个国家纷纷发布多项政策以抢占智能网联协同发展的技术制高点。2023 年,美国发布《无人驾驶汽车乘客保护规定》,文件规定了无人驾驶汽车配置要求^[5]。同年,又发布了加速车联网部署计划草案,提出在 2024—2034 年期间推动 20 款智能网联汽车搭载蜂窝车联网 (cellular vehicle-to-everything, C-V2X) 通信技术,支持网联驾驶安全应用。欧盟在 2019 年推出《欧洲绿色协议》^[6],建议推动新能源的发展,从 2035 年起将禁止生产燃油车。2023 年,日本颁布《道路交通法》^[7],推动 L4 级自动驾驶车辆在出行和物流快递行业的应用。同年,韩国颁布了《汽车安全度测试和评价规定》^[8],规定了车联网 (vehicle-to-everything, V2X) 通信设备的试验方法和评价指标。

由于车辆单体智能的视野有限、模态缺失等引发的全局感知能力弱、群体协作能力差等特性,车-路-云高效协同的互联网环境是未来发展趋势。发展车-路-云协同自动驾驶,有利于融合智能交通系统、车载无线通信系统、智能网联汽车、云控平台、高精地图、高精度时空基准服务等领域交叉应用技术。然而,随着智能网联汽车规模不断增大、车辆移动性以及有限的自身计算资源都限制了车-路-云的高效协同。同时,在信息共享时,智能网联汽车极易受到网络攻击、敏感信息暴露等一系列威胁。因此,如何在车联网中实现高效安全的车-路-云协同共享,仍是一个亟待解决的问题。

目前,车-路-云协同技术处于发展初期阶段,相关文献较少,文献[9]对车-路-云协同架构与关键技术进行研究,并从智能网联V2X组网、融合定位和测试评价3个关键方面进行了分析。文献[10]梳理了混合交通系统的车-路-云协同框架,围绕混合交通流建模及优化控制的新兴方法,综述了基于车-路-云一体化的混合交通系统的研究进展。然而,目前尚缺乏全面的车-路-云协同自动驾驶安全技术研究综述。鉴于此,本文围绕车-路-云协同技术的发展历史进程、面临的安全威胁以及未来挑战进行整理,本文具体贡献如下。

1) 梳理了车-路-云协同技术发展历程,调研了车-车协同、车-路协同、车-云协同相关文献,归纳出车-路-云协同是实现全自动驾驶(L5级别自动驾驶)的必由之路。

2) 概括了车-路-云协同中遇到的安全威胁及安全防御技术,主要从两方面展开:其一,基于经典信息安全机制的车-路-云威胁与防御技术,包括基于信息可用性的攻击、基于信息完整性的攻击、基于信息保密性的攻击、基于信息真实性的攻击、基于信息不可否认性的攻击;其二,基于机器学习技术的威胁与防御技术,其中包括常见的黑盒攻击、灰盒攻击和白盒攻击集中式机器学习攻击,以及数据投毒攻击、模型投毒攻击、拜占庭式攻击和数据隐私攻击、模型隐私攻击;

3) 探讨了车-路-云协同技术目前遇到的机遇和挑战,并在此基础上提出了车-路-云协同安全领域未来的研究方向。

1 车-路-云协同技术发展现状

1.1 自动驾驶技术发展历程

随着我国汽车智能化水平不断提高,车辆通过安装不同的传感器完成对周围环境探测和定位,使得汽车具有环境感知、路径规划和自主控制的能力,从而实现不同等级的自动驾驶。根据最新国家标准(GB/T 40429-2021),将自动驾驶分为L0~L5共6个等级。简要说, L0级可实现无自动控制仅预警; L1级可实现单一功能辅助; L2级能实现组合控制,但驾驶员需全程监控; L3级在特定条件下自动驾驶系统主导驾驶,驾驶员需随时待命接管; L4级在限定场景下完全自动驾驶; L5级实现全场景自动驾驶,无须人类干预。目前, L0~L2级别的自动驾驶技术已经在智能网联汽车上广泛搭载,我国具备L2级组合驾驶辅助的乘用车新车渗透率达到42.4%, 2023年开始进入L3自动驾驶技术批量上车新阶段^[11]。L2级自动驾驶侧重于辅助驾驶员,减轻驾驶负担,驾驶员在驾驶过程中仍占据主导地位; L3级自动驾驶则在特定条件下,车辆驾驶权大部分移交至自动驾驶系统,驾驶员角色从操作者逐渐向监督者转变,但仍需随时待命接管车辆。2025年4月12日,工业和信息化部在智能网联汽车产品准入推进会上明确了相关政策,加强了车联网安全管理规范,需明确自动驾驶等级精准描述,严禁虚假宣传,明确事故车企责任划分,推动自动驾驶领域安全健康发展^[12]。

自动驾驶技术主要依靠车辆自身的摄像头、激光雷达等传感器和计算平台进行环境感知与决策控制,属于单车智能的范畴,但是单车智能个体感知的局限性,如视野有限、模态缺失、传感器数据系数等负面因素,导致智能全局感知能力弱、群体协作能力差。本文总结了单体智能主要面临的挑战如下。

1) 在感知层面上,单体智能对被遮挡的车辆或行人以及路口横向车辆感知较弱,如单体智能车辆在行驶到路口时,前向碰撞预警功能无法感知路口横向车辆,以及被遮挡的车辆。在极端环境的感知能力有限,包括极端天气、光线聚变、温差色差大、车辆高速移动等,如车辆在隧道环

境中行驶时, 自动紧急制动和高速公路辅助功能因光线变化, 导致车辆感知性能失效。

2) 在决策控制层面上, 单车智能对交通规则的感知和理解的准确性较低, 交通规则主要面向人类驾驶员设计, 单车智能感知到的标志标识在转换为机器可识别规则时存在误差。单车智能车辆之间无法实现协作, 每个车辆都以自身最优的策略来决策驾驶行为, 在安全和效率上无法实现最优驾驶。例如, 交通拥堵辅助功能仅能实现单车在拥堵条件下的安全驾驶, 无法从全局情况下对车辆行驶路况进行规划。

3) 在车辆资源与成本方面, 单车智能在计算资源上是有限的, 对于实现 L3/L4 级别自动驾驶依旧存在巨大挑战, 车辆需要对雷达、摄像头、高精度地图等多种感知数据进行融合, 以获取更全面的环境感知和更加精确的目标定位。车辆资源的有限性可以通过边缘计算、车-路-云协同等技术来解决, 通过合理使用车-路-云有效资源和算力共享, 保证自动驾驶算法的高效运行。

1.2 车-路-云协同自动驾驶环境

图 1 展示了车联网车-路-云协同自动驾驶环

境, 涵盖了智能网联汽车、智慧公路、云控平台及车载应用服务等, 基于蜂窝车联网技术, 形成以“聪明的车”“智慧的路”“融合的云”为基础的架构, 以数字孪生平台、高精度地图引擎、云控服务平台等产业为支撑, 融合信息安全、大数据、人工智能、边缘计算等关键技术, 提供安全、高效、便捷的出行服务, 形成车联网车-路-云协同自动驾驶系统, 该系统包含以下元素。

1) 车: 车端边缘感知节点, 包含车辆自身的传感器 (如摄像头、雷达等)、车载通信设备、无计算能力的路侧设备。收集车辆周边的环境信息, 如障碍物、交通标志、其他车辆状态等, 并通过车-车、车-路通信将信息发送出去, 同时接收来自路侧和其他车辆的信息, 为车辆的决策提供数据支持。

2) 路: 路侧边缘计算节点, 包含边缘云计算节点 (包括行业专网、LTE、5G/6G 移动通信网及卫星网络) 和有计算能力的路侧设备, 实现对路侧感知设备收集的交通数据 (如交通流量、路况信息等) 进行初步处理和分析, 将处理后的数据发送给城市数据融合中心, 同时接收中心的指

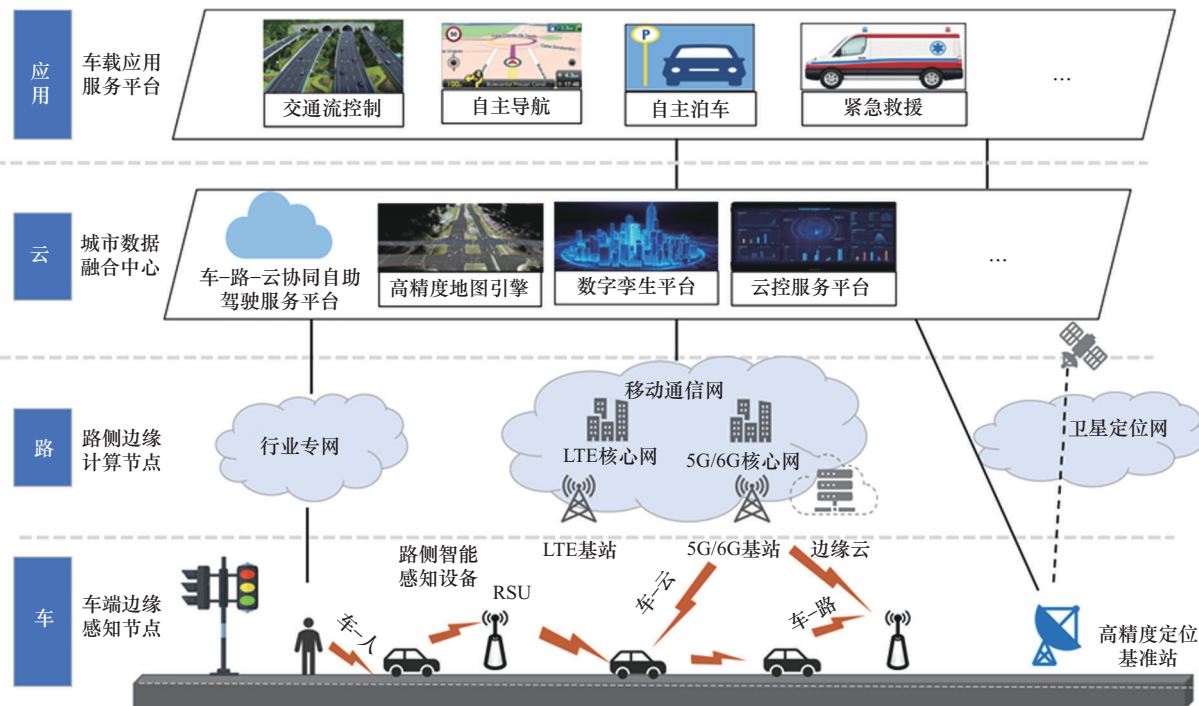


图 1 车-路-云系统交互环境

Figure 1 Interactive environment with vehicle-road-cloud system

令并转发给车端,实现路与车、路与云之间的信息交互和协同。

3) 云:城市数据融合中心,包含高精度地图引擎、数字孪生平台和云控服务平台,通过行业专网与路侧边缘计算节点相连,通过移动通信网与车载应用服务平台通信,还可接入卫星定位网获取定位信息。整合路侧和车端上传的数据,利用高精度地图引擎提供精确的地图信息,数字孪生平台构建虚拟交通场景辅助决策,云控服务平台进行交通资源的调配和管理,为车载应用服务平台提供数据支撑和决策指令。

4) 应用:车载应用服务平台,包含多种车载应用服务模块,例如交通流控制、自主导航、自主泊车、紧急救援等,通过移动通信网与城市数据融合中心通信,接收来自云中心的指令和数据,根据接收到的信息为驾驶员或自动驾驶系统提供具体的应用服务,如规划最优行驶路线、实现自动泊车、在紧急情况下触发救援,实现车辆的智能化应用和服务。

车-路-云协同自动驾驶技术可以分为车-车(vehicle-to-vehicle, V2V)协同、车-路(vehicle-to-infrastructure, V2I)协同、车-云(vehicle-to-cloud, V2C)协同和车-路-云协同。

1) 车-车协同:车辆之间通过通信和传感器技术进行信息交换和共享,以实现协同感知周围环境和交通情况的能力。车-车协同采用V2V通信,其本质是设备到设备(device-to-device, D2D)的通信方式,通过车辆之间的合作来增强交通和道路安全。车-车协同可以帮助车辆实现防碰撞、解决感知局限等。

2) 车-路协同:车辆与道路基础设施之间的信息交互和协同工作,以提高交通系统的安全性、效率和智能化水平。通过道路基础设施感知周围的交通环境和道路状态,以及车辆感知和通信,可以实现智能的交通控制、交通调节和交通安全警告。

3) 车-云协同:车辆与云端服务器之间通过互联网进行数据交换与互动,以实现车辆和驾驶者的智能化服务和管理。通过车辆与云端的连接,可以实现车辆远程监控、远程控制、远程诊断、远程更新等功能,提高车辆的智能化程度和

用户体验。

4) 车-路-云协同:车辆、道路基础设施和云端服务器之间通过互联网进行信息交互和协同工作,以实现智能交通管理和服务。这种技术结合了车-车协同、车-路协同和车-云协同的功能,通过数据共享和处理,实现智能的交通控制、路况监测、路线优化等功能,提高整个交通系统的智能化水平。

车-路-云协同存在以下特性。

1) 高度流动性。由于车辆之间的连接经常中断,传统的基于信誉的安全系统难以有效部署。网络流量会随着拓扑结构的变化而发生剧烈波动,影响安全数据包的路由及入侵检测。此外,高度流动性导致节点移动的不可预测性,从而导致实时安全信息传输困难。

2) 结构异构性。车-路-云协同系统由多种不同车辆和边缘基础设施组成,这些设备拥有不同的资源容量。如何在节点之间平衡这些资源,以确保安全策略的有效性,是一个巨大的挑战。在功耗、计算能力和存储容量受限的车载网络中,如何最大化安全性的适用性如同一场模糊博弈,很难找到最佳的解决方案。

3) 规模性。据文献[13]预测,到2030年,将有数百万辆汽车连接到网络中。然而,目前尚无全球统一的权威机构和标准来确保这一庞大网络的安全性。随着网络规模的不断增长,安全算法的计算和通信成本也在增加,导致响应时延和检测错误率随之增加。

4) 数据安全敏感性。车-路-云协同系统中,数据来源广泛多样,包括车辆的传感器数据(如摄像头拍摄的道路画面、雷达探测的周边物体距离和速度等)、用户的个人信息(如身份认证信息、行程偏好等)以及路侧设备采集的交通流量、路况等信息。这些数据涉及个人隐私、商业机密以及公共安全等方面,具有极高的敏感性。

5) 低时延性。在车-路-云协同系统中,对通信的实时性要求极高。特别是在求救信息广播、事故避免、紧急救援等特殊情景下,为保障车辆安全行驶、优化交通效率、快速响应突发状况,系统需在极短时间内完成信息传输与处理。从而确保车辆及时决策,交通资源动态调配,紧

急事件及时处置。

在车-路-云协同自动驾驶环境中, C-V2X 技术、协同自动驾驶技术以及数字孪生技术相互融合, 构建了智能化的交通生态系统。通过 C-V2X 技术, 车辆与道路基础设施及云端服务器实现高效通信并实时共享交通信息和路况数据; 协同自动驾驶技术借助这些信息, 实现车辆之间的协同行驶和智能决策, 提升行车安全性和交通流畅度。数字孪生技术可以为自动驾驶系统提供仿真环境和数据分析平台, 通过模拟现实场景、优化自动驾驶算法和系统性能。这些技术的有机结合, 将推动自动驾驶技术向更智能、更可靠的方向发展, 为未来交通运输领域带来革命性变革。表 1 调研了车-路-云协同自动驾驶主要的研究文献, 主要分为 V2V 协同、V2I 协同、V2C 协同和 V2X 协同。

2 车-路-云协同技术安全威胁

通过 V2X 通信技术、互联网技术、人工智能技术、大数据技术、边缘计算及云计算技术深度融合, 智能网联汽车逐步实现 GPS 协同感知、协同预测、协同决策、协同控制, 形成以“聪明的车+智慧的路+融合的云”一体化的协同自动驾驶架构, 满足不同等级的自动驾驶车辆行驶安全、高效节能的需求, 有效解决交通事故、交通拥堵、环境污染等问题^[38]。车-路-云协同自动驾驶技术带来了智能网联汽车和交通基础设施之间全方位协同的同时, 也带来了复杂的安全威胁。潜在的安全风险不仅涉及车辆自身的驾驶安全, 还关乎用户隐私和数据安全。

自动驾驶汽车通常配备多种高性能传感器, 例如摄像头、雷达和 GPS, 以感知周围环境。传感器捕获的数据被输入自动驾驶汽车的计算系统, 进行复杂的处理和计算, 实现车辆的自主控制。因此, 自动驾驶汽车严重依赖传感器数据来做出正确的驾驶决策, 这不可避免地扩大了潜在的威胁面, 并使传感器面临严重的安全风险。近年来, 机器学习, 特别是深度学习技术在自动驾驶领域取得了显著进展。通过强大的数据驱动能力, 机器学习方法可以显著提高车辆在感知、预测、规划和决策控制各个环节的性能, 推动车-路-云协同系统的智能化与安全性发展。本文在

传统车-路-云安全威胁与防御技术和基于机器学习的车-路-云安全威胁与防御技术方面展开相关调研。

表 2 总结了车-路-云协同中攻击种类与安全防御技术, 从车-路-云数据安全的角度, 分析了车辆、道路基础设施和云平台之间数据交互遇到的安全威胁及安全防御技术。基于经典信息安全机制的威胁与防御主要从基于信息的可用性、基于信息的完整性、基于信息的保密性、基于信息的真实性和基于信息的不可否认性分析。

2.1 基于经典信息安全机制的车-路-云威胁与防御

2.1.1 基于可用性的攻击

信息的可用性是指确保信息和系统在需要时可被授权用户访问和使用。将可用性威胁分为 3 种类型: 泛洪攻击、干扰攻击、共谋攻击。

1) 泛洪攻击: 攻击者发送大量信息, 导致被攻击者无法提供服务。由于车辆控制器局域网(controller area network, CAN) 作为常用的车载网络, 将 CAN 连接到外部网络时, CAN 报文泛洪是一种常见的拒绝服务(denial of service, DoS) 攻击。攻击者持续向目标车载 CAN 广播大量的 CAN 报文, 导致其功能瘫痪^[232]。车载 CAN 入侵检测系统(intrusion detection system, IDS) 常用来检测此类消息泛洪攻击, 例如基于消息频率的 IDS^[39], 基于时钟偏差的 IDS^[40], 基于电信号的 IDS^[41-43]和基于熵的 IDS^[44]。其中, 基于消息频率的 IDS 适用于高密度的交通环境以及特定车速下的车辆间通信, 尤其是在车辆与车辆之间、车辆与基础设施之间频繁交换信息时, 利用频率异常检测来识别异常行为。该方法的计算负载较轻, 因为它仅需简单统计和分析消息频率。基于时钟偏差的 IDS 适用于多个传感器和执行器进行高精度时间同步的系统, 例如文献[40]提出了一种基于时钟的异常的入侵检测系统, 来测量并利用周期性车载消息间隔对电子控制单元(electronic control unit, ECU) 进行指纹识别, 基于检测识别错误中的异常偏移判断是否发生入侵行为, 实验表明可以快速识别车载网络的入侵, 误报率低至 0.055%。虽然时钟偏差检测的精度较高, 但需要复杂的时钟同步机制和持续的时间

表 1 车-路-云协同自动驾驶主要研究文献

Table 1 Main research literature of vehicle-road-cloud collaborative autonomous driving					
类型	文献	时间	协同内容	协同技术	解决问题
V2V 协同	Su 等 ^[14]	2018 年	车辆间计算资源	DSRC/C-V2X 技术	利用软件定义网络（software-defined networking, SDN）架构，将车辆计算任务外包给周围自动驾驶车辆
	Li 等 ^[15]	2018 年	车辆间数据	DSRC/C-V2X 技术	提出了一种基于协作 3D LiDAR 点云配准和贝叶斯概率更新的新型半结构化多层停车库协作测绘和自主停车系统
	Li 等 ^[16]	2019 年	通信链路协议	DSRC/C-V2X 技术	利用车辆间相对间隙及速度差，实现自主车辆串行编队控制
	Deng 等 ^[17]	2019 年	车辆间数据	DSRC/C-V2X 技术	解决车辆超车防碰撞问题
	Chen 等 ^[18]	2019 年	车辆间数据	DSRC/C-V2X 技术	提出了一种基于点云的 3D 对象检测方法，用于处理多种对齐点云
	Chen 等 ^[19]	2020 年	车辆间数据	DSRC/C-V2X 技术	利用车辆间驾驶信息，构建深度模型预测附近车辆速度
	Yu 等 ^[20]	2020 年	车辆间数据	DSRC/C-V2X 技术	解决高速公路场景中自动驾驶车辆跟随和超车任务
	Liu 等 ^[21]	2020 年	通信链路协议	DSRC/C-V2X 技术	开发了可学习的握手通信进行协作感知
	Wang 等 ^[22]	2020 年	车辆间数据	DSRC/C-V2X 技术	通过车辆间共享数据，感知被遮挡物体
	Chen 等 ^[23]	2020 年	车辆间数据	DSRC/C-V2X 技术	车-车协同下交通地标检测的在线定位技术
	Chen 等 ^[24]	2021 年	车辆间数据	C-V2X 技术	基于 6G 通信技术，控制协同自动驾驶车辆车距
	Zhang 等 ^[25]	2022 年	通信链路协议	DSRC 技术	开发了一种 DSRC 辅助波束对准方法来提高 V2V 链路波束对准的效率
V2I 协同	Li 等 ^[26]	2022 年	车辆间数据	DSRC/C-V2X 技术	开发了公开的数据集（V2XSim），以推进 V2X 通信辅助自动驾驶的协作感知研究
	Lin 等 ^[27]	2024 年	车辆间数据	DSRC/C-V2X 技术	提出了一种新型车-车感知框架 V2VFormer，利用附近联网汽车的互补信息构建整体感知，扩大探测范围
	Kong 等 ^[28]	2022 年	车辆轨迹数据	DSRC/C-V2X 技术	提出了一种冲突证据融合算法，并考虑了出租车轨迹的地理约束以及车联网中出租车轨迹之间的协作
V2C 协同	Yang 等 ^[29]	2022 年	通信架构	DSRC/C-V2X 技术	设计了一个基于网络编码的 RSU 通信架构，实现 V2I 协同自动驾驶
	Yi 等 ^[30]	2024 年	车辆点云数据	C-V2X 技术	提出了一种时空异步融合方法，利用多层感知器生成后同步目标序列，这些序列随后用于融合以提高准确性
	Qiao 等 ^[31]	2018 年	协作边缘计算架构	DSRC/C-V2X 技术	引入了一种车辆边缘多接入网络，将车辆视为边缘计算资源，构建协作式分布式计算架构
V2X 协同	Xie 等 ^[32]	2019 年	协作边缘计算架构	DSRC/C-V2X 技术	探索了新颖的协作车辆边缘计算网络（CVECNet）架构
	Chen 等 ^[33]	2020 年	车辆数字孪生模型	DSRC/C-V2X 技术	基于数字孪生模型，帮助车辆在虚拟网络中协作驾驶
	He 等 ^[34]	2023 年	协同自动驾驶框架	C-V2X 技术	提出了基于 C-V2X 的协同自动驾驶框架（CCAD），可以从多个角度观察驾驶情况
	Qian 等 ^[35]	2024 年	智能交通系统	DSRC/C-V2X 技术	提出了多智能体强化学习算法，利用对比学习提高复杂驾驶场景中的策略稳定性
	Hu 等 ^[36]	2025 年	新型协作决策框架	DSRC/C-V2X 技术	提出了一种新型协作决策框架，借助大语言模型，提升车辆在多车道合流区域的安全性和效率
	Liu 等 ^[37]	2025 年	车联网辅助自动驾驶	DSRC/C-V2X 技术	提出了一种新型端到端协同驾驶系统，在整个自动驾驶流程中融入 V2X 通信，利用共享感知区域促进驾驶

表2 车-路-云协同中攻击种类与安全防御技术

Table 2 Attack types and security defense technologies of vehicle-road-cloud collaboration

分类	攻击种类	子类	介绍	防御方法
基于经典信息安全机制的车-路-云威胁与防御	基于信息可用性	泛洪攻击	攻击者发送大量信息, 使得被攻击者无法提供服务	基于消息频率的IDS ^[39] 、基于时钟偏差的IDS ^[40] 、基于电信号的IDS ^[41-43] 和基于熵的IDS ^[44] 等
		干扰攻击	通过产生干扰信号来阻碍无线通信系统正常工作的攻击手段	中继辅助 ^[45] 、基于MIMO ^[46] 、网络编码 ^[46-47] 、博弈论 ^[48] 、基于学习的干扰检测、功率控制和信道跳变技术 ^[48-55] 等
		共谋攻击	一组受感染的节点协作/串通发起恶意攻击的行为	声誉或信任方案 ^[56-59] 、身份认证和授权 ^[60] 、监控审计 ^[61] 、看门狗 ^[62-63] 、智能合约和共识机制 ^[64-65] 、机器学习 ^[66] 等
	基于信息完整性	改变或注入虚假信息攻击	通过篡改或注入恶意数据来破坏通信完整性的攻击手段	加密 ^[67] 、数据完整性校验 ^[68] 、身份认证 ^[68] 、IDS ^[69-70, 234] 、共识 ^[71-72] 、机器学习 ^[73] 等
		重放攻击	攻击者通过拦截并重复发送合法的数据包来进行攻击	数据水印 ^[74-76] 、机器学习 ^[77] 等
		GPS欺骗攻击	通过发送虚假的GPS信号来误导目标设备接收到错误位置、时间信息的攻击手段	加密认证 ^[78] 、信号监测 ^[79-82] 、多源导航(多传感器融合) ^[83-84] 等
	基于信息保密性	窃听攻击	未经授权的监听或获取通信内容的行为, 通常涉及通过技术手段获取他人的私密信息或机密信息	轻量级加密算法 ^[85] 、功率控制方案或定向天线 ^[86] 等
		位置追踪攻击	通过技术手段追踪某人或某物的物理位置	信号加密 ^[87] 、匿名性 ^[88-89] 、实时威胁检测 ^[90] 、频繁更新车辆位置共享策略 ^[91-92] 、差分隐私 ^[93-95] 等
	基于信息真实性	证书复制攻击	受感染的节点使用复制的证书, 通过删除添加到黑名单中的证书来隐藏自己	强证书管理 ^[96-97] 、证书吊销和更新 ^[98] 、证书透明度 ^[99-100] 、区块链 ^[101] 等
		女巫攻击	针对分布式网络的攻击方式, 攻击者通过伪造多个虚假的身份(或节点)来获得不成比例的影响力或控制权	基于资源测试 ^[102-103] 、基于可信认证 ^[104-106] 和基于位置验证的机制 ^[107-109] 等
		伪装攻击	攻击者通过冒充合法用户或设备的身份, 获取未授权的访问权限, 以进行恶意操作或窃取敏感信息	入侵检测系统(IDS) ^[110] 、机器学习 ^[111] 等
	基于信息不可否认性		不可否认性负责识别执行特定行为的真实节点的ID。它是一种通过数字签名或加密确保实体之间消息传输的方法	数字签名 ^[112-113] 、PKI技术 ^[114] 、区块链技术 ^[115-117] 等
基于机器学习技术的车-路-云威胁与防御	基于集中式机器学习攻击	白盒攻击	攻击者假设对底层的机器学习模型了解充分, 包括优化技术、已训练的模型、模型结构、激活函数、超参数、层权重和训练数据的信息	密码技术 ^[118-122] 、对抗训练 ^[123-126] 、正则化 ^[126-128] 、隐私保护方法 ^[129] 等
		灰盒攻击	攻击者假设对目标模型有一定了解, 而黑盒攻击则假设攻击者没有任何关于底层机器学习模型和训练数据的知识	密码技术 ^[118-122] 、对抗训练 ^[123-126] 、正则化 ^[126-128] 、隐私保护方法 ^[129] 等
		黑盒攻击	攻击者没有任何关于底层机器学习模型和训练数据的知识	密码技术 ^[118-122] 、对抗训练 ^[123-126] 、正则化 ^[126-128] 、隐私保护方法 ^[129] 等
	基于分布式机器学习攻击	数据/模型投毒攻击	攻击者通过操纵训练数据或修改模型参数破坏机器学习训练过程	基于数据的防御 ^[130-131] 和基于模型的防御 ^[132-134] 等
		拜占庭式对抗攻击	在分布式网络架构中, 网络中恶意节点发送错误信息, 使得网络中其他成员关于全体协作的策略得出不同的结论, 从而破坏了网络系统的一致性	基于中值的聚合器 ^[135] 、Krum ^[136] 和修剪平均值 ^[136] , 以及使用冗余和数据洗牌 ^[137] 等
		数据/模型隐私攻击	隐私攻击机器学习模型训练需要收集许多交通图像并将模型协同卸载到中央服务器, 这些图像包含许多用户的敏感信息, 从而引发潜在的隐私泄露	差分隐私 ^[138] 、区块链 ^[139] 、同态加密 ^[140] 、功能加密 ^[141] 和安全多方计算 ^[142] 等

截分析。基于电信号的IDS适用于需要对车辆内部硬件状态进行细粒度监控的系统,例如,文献[42]利用电CAN信号的特性来识别ECU,从而能够检测恶意ECU。该方法需要复杂的硬件支持和信号处理能力,尤其在高干扰环境下,会增加系统的计算复杂度和硬件要求。基于熵的IDS适用于大规模车联网中涉及复杂通信协议且数据包内容较为多样的场景,例如,文献[44]设计了一种基于熵的IDS系统,利用Bloom过滤来测试基于消息标识符和部分数据字段的帧周期性,从而有助于检测潜在的重放或修改攻击。熵计算涉及较为复杂的数学运算,尤其在高流量的网络中,会导致较高的计算开销。当前主流方案是基于网络流量异常检测,在车-路-云协同中,可通过跨域合作(车端、路端、云端)协同防御泛洪攻击。在对抗泛洪攻击时,基于消息频率的IDS实现简单但易受合法突发流量干扰且易被攻击者规避;基于时钟偏差的IDS对特定攻击检测有效,却依赖精确时钟同步且应用场景有限;基于电信号的IDS能检测隐蔽攻击,但硬件成本高、部署难且易受电磁干扰;基于熵的IDS对各类攻击检测敏感度高,对计算资源需求大且存在误判风险。因此,实际应用中常需多种方法结合,以提升对泛洪攻击的防御效能。

(2) 干扰攻击是一种通过产生干扰信号以阻碍无线通信系统正常工作的手段。此类攻击破坏通信链路,使合法用户无法有效地发送或接收数据。长期演进车联网(long term evolution vehicle-to-everything, LTE-V2X)技术物理层基于正交频分复用技术,实际上,干扰器需要识别数据包的存在才能发起干扰攻击。然而,当前的解决方案都无法有效防止干扰信号。使用定向天线可最大限度减少这种攻击的影响,并允许车辆避开干扰区域^[143]。常用的对抗干扰攻击的方法有中继辅助^[45],基于多输入多输出(multiple-input multiple-output, MIMO)^[46]技术,网络编码^[46-47],博弈论^[129],基于学习的干扰检测、功率控制和信道跳变技术^[48-55]。基于中继辅助的方法主要对来自车辆或基础设施的信号进行放大、编码或重新调制等处理,以增强对

抗干扰的能力。基于MIMO和网络编码的方法通过在信号发送端和接收端使用多副天线或者增加多个数据包编码以实现并行通信,主要是为了阻止持续干扰攻击,但需要复杂的硬件支持,并需在无线信道上处理大量的信号数据。基于博弈论的方法通过构建博弈模型,将干扰攻击者和防御机制视为博弈双方,分析双方策略选择和收益,主要用于应对低功率干扰攻击,但计算复杂度高,尤其在参与方众多、决策空间较大时,可能导致较高的计算开销。基于机器学习的方法通过大量干扰数据集训练高质量的对抗模型,对车-路-云系统中通信数据进行分析,以识别干扰信号,主要用于检测和定位周期性和随机性干扰攻击,但前期需要处理大量数据,可能带来较高的计算复杂性和较长的训练时间。文献[45]提出一种名为JammingBird的多天线802.11p接收器新架构,以确保车辆通信免受高功率持续攻击,并在停车场和高速公路场景中评估了接收器的抗干扰性。文献[46]研究了车联网环境中持续干扰攻击下的大规模MIMO路边单元(road side unit, RSU)的鲁棒性,其中RSU采用网络编码方案来服务多个车辆通信。针对6 GHz以下毫米波信道模型的类噪声干扰攻击,实验表明使用网络编码的RSU可在信道模型上实现双倍频谱效率。文献[49]提出了一种MAC层机制,能够区分车联网中的干扰攻击、干扰信号和碰撞,基于监控信道间隔内传输的冲突标帧信号,实现较低误报率下的干扰攻击检测。文献[50]提出了一种智能交通系统车辆队列抗干扰机制,结合通信的先验知识和监视冲突消息,利用数据挖掘技术检测网络中的异常传输。文献[51]采用数据包传输率检测车辆传输过程中的干扰攻击,可区分攻击类型和识别不良无线电链路的阈值。同时,文献[47]提出一种基于神经网络自动编码器的信息协调方案,以实现可靠性和保密性的动态平衡。文献[52]提出一种时域方法,分析干扰器位置对车辆编队协同的影响,实验结果表明,当干扰器接近跟随的第一辆车时,干扰器有更大的概率使车辆编队不稳定。文献[144]研究了相消干涉^[145]、压制干扰^[146]和欺骗性干扰^[47]3种攻击方

式对车辆编队协同稳定性的影响,实验结果表明,车辆编队协同极易受到干扰攻击。最近,机器学习方法被应用于车辆干扰攻击检测,文献[47]基于无监督机器学习方法,以干扰器车辆和目标车辆间的相对速度变化为指标进行监测,通过聚类的方式区分故意干扰情况,并提取每种攻击的具体特征。文献[55]提出一种基于半监督学习的车联网无线通信干扰攻击检测方法,通过干扰器和接收器之间的相对速度变换检测干扰攻击。对抗干扰攻击时,可分成:物理层的MIMO技术,能提升效率与抗干扰性,但硬件要求高^[46];网络层的中继辅助和网络编码技术,可增强传输可靠性但成本高且存在时延;协议层的抗干扰通信协议,通过差错控制、重传机制等保障数据的可靠性,但难以适应高度动态的网络拓扑结构^[52];智能化方法包括机器学习和博弈论,已有大量的工作基于该方法提升车-路-云协同的可靠性,但受限于高质量数据和复杂模型,仍存在收敛慢、对新干扰适应性差等问题。

(3) 共谋攻击是指一组受感染的节点协作/串通发起恶意攻击的行为。解决共谋攻击常见的方法是构建声誉或信任方案^[56-59]、身份认证和授权^[60]、监控审计^[61]、看门狗^[62-63]、智能合约和共识机制^[64-65]、机器学习^[66]等。基于声誉或信任方法通过对车辆过去行为的评分来计算其声誉评分,并通过车辆之间的直接交互经验和间接的推荐信息来建立信任关系。该方法主要用于分布式系统中多方协作的场景,特别是当参与者之间没有事先建立信任关系时,声誉机制能有效地防止共谋攻击。例如,文献[56]提出一种通用、轻量级、分布式的车辆串通攻击检测方法,主要针对欺诈信誉和欺诈信息识别不同的共谋攻击,以及一种抑制共谋节点的激励方法。文献[57]提出一种名为TWSN的信任机制,该机制评估通信信任和数据信任,可有效地防止共谋攻击。文献[58]提出一种基于集群的网络信任模型,以防御共谋攻击,但是该方法依赖于历史数据,可能无法应对突发的新型攻击,且信任管理也可能导致某些节点被错误标记为不可信。文献[59]提出了一种朋友分享系统,该系统包括声誉系统和加

权图系统,用于解决特定问题并应对共谋攻击和女巫攻击。基于身份认证和授权的方法通过数字证书、多因素验证等方式确认设备或用户的真实身份,利用基于角色的访问控制、动态权限管理等机制,仅赋予合法主体相互对应的权限,阻断攻击者通过伪造身份、越权操作等共谋手段非法获取资源或破坏系统。该方法能够防止伪造身份的攻击者加入系统,从根源上解决共谋攻击。例如,文献[60]提出了一个不可链接的经过身份验证的密钥协议,该协议对车联网具有抗串通性。在车-路-云协同中,采用多因素认证、加密技术和区块链技术进行身份验证,确保每个参与者在车-路-云协同中的身份唯一且可靠。文献[61]提出了一种监控验证方案,该方案能够验证监控节点的决策并准确检测网络中由共谋注入攻击引起的任何恶意行为。看门狗技术将车联网设备内置的硬件或软件定时器作为看门狗模块,来监测异常行为。该方法已经被证明能够检测静态非共谋邻域中的不当行为的节点,但是在模糊冲突、有限传输功率、共谋等情况下,仍存在一些问題。同时,监控过程可能导致系统的资源消耗增加,尤其在实时性要求高的场景中,可能会造成时延。文献[62]研究将通信节点分类为看门狗节点和路径评估器节点,分别用于识别行为不当的节点和帮助路由协议避开这些恶意节点。智能合约和共识机制通过预设规则自动执行车联网交互,并结合共识机制保障节点达成一致,抵御共谋攻击。该方法适用于去中心化的车联网平台,尤其是在多方共同决策和信息共享的场景中。例如,文献[64]设计了一种去中心化的安全协同方案,防止恶意车辆加入车辆协同社区。文献[65]探讨了以去中心化自治组织和运营为基础的去中心化车联网概念,促进车辆与其他实体之间的可信赖交互。然而,共识机制对计算资源要求较高,会增加系统负担,特别是在大规模网络中,可能对效率产生显著影响。一旦共识达成,智能合约能够快速执行,提升系统的自动化和效率。文献[66]提出一种基于长短期记忆网络(long short-term memory, LSTM)的IDS,以检测错误紧急消息。基于LSTM的分类器旨在根据来自正常攻击场景和共谋攻击场景的时间序列特征向量

来识别当前交通事件。在对抗共谋攻击中,构建声誉或信任方案凭借节点评价识别攻击,灵活性较好,但评价过程易受恶意操纵。身份认证的方法可以从源头上控制合法节点参与,然而认证管理成本较高,并且认证存在密钥泄露、侧信道攻击、DDoS等攻击^[147]。

2.1.2 基于完整性的攻击

信息完整性旨在确保数据的完整性,包括保证通过网络传播的数据准确性和一致性。将完整性威胁分成改变或注入虚假信息攻击、重放攻击、GPS欺骗攻击3种类型。

1) 改变或注入虚假信息攻击是一种通过篡改或注入恶意数据来破坏通信完整性的攻击手段。这种攻击可以在数据传输或存储过程中进行,导致接收方接收到错误或恶意信息。V2X通信过程中,内部攻击者可能尝试通过网络广播虚假的安全信息。广播消息适用于所有周围节点,但需要对其进行签名,以阻止外部攻击者生成虚假消息。由于内部攻击者是经过身份验证的节点,因此可以使用其数字证书来签署任意数量的虚假消息。然而,目前依旧缺少足够有效的方案来检测攻击者^[148]。对抗该攻击的主要方法包括加密^[67]、数据完整性校验^[68]、身份认证^[68]、IDS^[69-70]、共识^[71-72]、机器学习^[73]等。基于密码方法通过各种加密算法对传输信息进行加密,例如,利用对称/非对称加密确保传输信息的保密性,哈希函数确保数据完整性,数字签名验证发送方身份与信息真实性。该方法是对抗改变或注入虚假信息攻击最有用的方法之一,尤其对手是外部攻击者的情况,然而,当对手是内部攻击者时,拥有合法有效的识别证书,仅基于密码方法不足以解决车辆协同中的安全威胁。同时,加密操作需要较高的计算开销,尤其是在数据量大的情况下,可能导致系统性能下降。加密和解密操作也会增加通信时延,影响实时性。文献^[68]提出了一种基于身份认证技术的轻量级机制,用于建立车辆的数据完整性和检测恶意RSU的概率模型。IDS技术也是在车联网中应用广泛的技术,通过实时采集车联网中的通信数据,运用异常检测、误用检测等技术,基于预定义规则或机器学习构建的正常行为模型,识别数

据中改变或注入虚假信息的攻击行为,并及时发出警报以阻断攻击。文献^[69]提出一种基于反向传播机器学习神经网络的混合入侵检测机制,用于检测车联网中常见的入侵攻击,能够识别自动驾驶和半自动驾驶中的恶意车辆。文献^[70]提出了一种新型的滤波模型,可以快速过滤来自相邻车辆的信息,检测相邻车辆状态模式的变化,并预测相邻车辆的未来状态。此外,文献^[71]提出了一种新型的基于共识的V2V车辆控制策略,用于提高车辆编队协同稳定性,该策略仅将通过多数验证的数据记录到账本,能够防御消息伪造攻击,然而对车辆计算资源要求较高,给共识算法的部署带来挑战。文献^[73]提出一种基于人工智能技术的方法,用于检测车联网中欺骗性消息,如伪造仪表读数和禁用自动功能等安全威胁,该研究还探讨了消息身份认证和DoS预防之间的均衡点,该方法可以与IDS相结合,用于恶意车辆节点的监测和识别,然而,攻击者可以增加IDS误报率,从而掩盖真正的攻击行为,因此,受限于机器学习检测性能,在高动态、大流量场景中,IDS可能会过载,错过关键攻击检测点。

2) 重放攻击是一种网络攻击手段,攻击者通过拦截并重复发送合法的数据包来进行攻击,损害车辆数据完整性和新鲜度的攻击^[149]。在车联网中,重放攻击一般分为两步:首先,车辆的关键数据被攻击者获取,然后,攻击者对数据进行恶意修改后重新发给车辆。重放攻击通常利用通信协议中的漏洞,使攻击者能够冒充合法用户,重播合法的消息或命令,从而达到欺骗系统或窃取信息的目的^[235]。对抗该攻击的主要方法包括数据水印^[74-76]、加密^[150]和机器学习^[77]等。数据水印方法首先在文献^[74]中提出,是重放攻击检测常见的防御方法,其中关键是检查以输入噪声为特征的检测残差的分布特性,具有较强的隐蔽性和持久性,能够有效防止消息被篡改或重放。同时,数据水印在数据传输过程中不需要实时复杂的计算,且能在检测重放攻击时提供较高的准确性。后来,结合通用控制系统信号分布特性检查,提出了动态水印方法^[75,151],并设计了水印信号,如周期性水印^[152]和事件触发的水印^[153]。

文献[151]将动态水印方法应用到车联网中,用于对抗横摆角速度测量的不同类型的重放攻击。文献[75]设计了一种基于矩阵归一化因子动态水印方法,通过将私有激励注入控制输入以防止车联网中数据的重放攻击。然而,数据水印技术在高动态车联网环境中,水印的嵌入、提取和验证存在时延,攻击者可以利用时间差完成重放攻击^[154]。文献[150]设计了一种新颖的消毒算法来对抗车联网中的重放攻击,同时探索了加密数据重放攻击的安全模型,以支持安全清理。仅通过加密方法并不能从根本上解决重放攻击,例如,攻击者截获加密的车辆控制指令并直接重放,接收端通过解密依然会执行该指令。通过加密方法与设备物理层绑定,可确保重放数据无法跨设备使用,特斯拉车辆钥匙采用芯片设备物理 ID 对抗重放攻击。文献[77]提出了一种基于机器学习的方法来自动检测车联网中的重放攻击,其中从邻近车辆接收到的安全信息被恶意节点重播,且包含虚假的发送者信息。然而,现有机器学习模型多基于静态数据集训练,难以捕捉时空关联性,并且存在跨场景泛化性较弱等问题。例如,同一重放攻击在拥堵路段和高速路段中的特征差异性可能导致检测失败^[149]。

3) GPS 欺骗攻击是一种通过发送虚假的 GPS 信号来误导目标设备接收到错误位置和时间信息的攻击手段。此外, GPS 信号不仅受到欺骗攻击,还会受到干扰攻击和重放攻击^[155]。随着低成本无线电硬件平台和开源的 GPS 生成软件的普及,攻击者可以利用不到 700 元的硬件设备实现对 GPS 信号的欺骗攻击。攻击者可以欺骗车辆导航路线,并将车辆劫持到任意位置,而不会发出任何警报。对抗该攻击主要的方法包括加密认证^[78]、信号检测^[79-82]、多源导航(多传感器融合)^[83-84]等。文献[156]提出了一种基于导航消息认证的方法,通过将导航消息(GPS 消息等)进行加密或数字签名,以保证数据安全传输。其他加密防御方法,如隐私标记^[157]等,加密方法仍易受到重放攻击,攻击者可以通过记录合法信号并延时广播^[158],即可绕过认证。基于信号检测方法侧重于收集和处理 GPS 信号中的信息,如

准确的时钟信息、信号功率以及到达角度等,识别与真实 GPS 信号的差异,进而检测和抵御欺骗信号。文献[79]提出一种基于模型的信号滤波和异常检测方法,通过使用先前训练的一类支持向量机模型来检测传感器异常。通过机器学习方法可以提升 GPS 信号检测的性能,文献[80]提出一种数据驱动的方法来识别由受损车辆生成的伪造轨迹,计算每对轨迹之间的距离矩阵,并使用分层聚类算法识别伪造的轨迹。文献[81]将攻击检测描述为无监督的二元分类问题,基于 K 均值检测隐形 GPS 欺骗攻击,需要预先确定特征空间和距离函数来测量数据点之间的距离。文献[82]提出了一种基于支持向量机模型的方法来检测车辆传感器数据的异常,尽管在检测数据源的不一致上表现良好,但是会产生假阳性错误^[159]。多源导航方法融合了卫星导航与惯性、视觉、地磁等多种导航技术,通过数据融合交叉验证,识别并剔除欺骗信号。该方法可以通过部署多个传感器,将被攻击的传感器进行隔离,从而使系统能够依赖其他传感器进行导航^[83]。例如,若检测到 GPS 欺骗攻击,自动驾驶车辆可通过 LiDAR 进行定位。然而,判断哪个传感器被攻击并非易事,并且存在检测滞后性,在隔离传感器之前,被攻击的系统可能已经部分损害^[160]。文献[84]针对部署多个传感器车辆的场景,开发了一组攻击检测器来捕获每个传感器测量中的异常,该检测器是扩展卡尔曼滤波与累积总和鉴别器的组合,并采用基于规则的隔离方案来识别源异常传感器。还有一种对抗 GPS 攻击方式是基于其他实体,例如,通过基础设施安全数据进行 GPS 校正^[161]。文献[161]提出了一种通过路边单元作为独立、安全的数据源,为周边车辆广播安全可靠的 GPS 数据,设计了基于 RSU 安全数据的车辆与基础设施间全球位置测量方法,基于安全测量,开发和计算多个数据特征,并利用这些特征构建基于隔离森林的实时检测器检测 GPS 的欺骗攻击。然而,多源导航方法受限于传感器本身的误差,如 IMU 的漂移导致车辆导航存在误差累计,同时,在 GPS 欺骗攻击时,车辆视觉传感器因光照变化而失效,导致多源系统无法准确仲裁可信数据源^[162]。此外,多源融合算

法在资源消耗和实时性上同样存在瓶颈,处理多源异构数据需要车辆边缘设备具有高性能的计算资源。

2.1.3 基于保密性的攻击

信息保密性旨在确保信息只能被授权访问者读取或获取,防止未经授权的访问和信息泄露。保密性威胁主要分为窃听攻击和位置追踪攻击两类。

(1) 窃听攻击是指未经授权监听或获取通信内容的行为,通常涉及通过技术手段获取他人的私密信息或机密信息。广播消息通常包含被视为非机密信息的交通安全信息,任何经过认证的车辆均可接收该类信息。然而,内部攻击者在未经许可的情况下收集其他用户的信息,可被认为是一种窃听攻击^[163]。对抗窃听攻击常用的方法包括轻量级加密算法^[85]、功率控制方法、定向天线^[86]等。在LTE-V2X通信中,用户设备使用认证中心颁发的预共享密钥与基础设施进行安全通信,因此通信过程是加密的。在V2V通信中,UE对每个应用程序使用预共享密钥,但这会增加额外的计算成本。文献[86]提出并实现了一种共享的密钥生成协议,即利用现成的802.11n多天线设备的多天线Key生成器,通过从测量的接收信号强度指示器中提取的共享随机性来生成密钥,从而将比特生成率提高4倍以上。对于高速移动的车联网环境,定向天线需要实时调整方向以保持通信链路的稳定,这对天线的跟踪技术和系统的响应速度提出了很高的要求。

(2) 位置追踪攻击通常指通过技术手段追踪某人或某物的物理位置。这种攻击可以利用多种方式实施,包括利用GPS、蓝牙信号、Wi-Fi信号、移动网络信号或者其他传感器数据。位置追踪攻击可能会对个人隐私造成严重威胁,尤其是未经授权的情况。与相邻节点共享位置对于各种车辆应用非常重要。因此,攻击者可以收集并使用这些信息来跟踪用户。当车载单元广播消息以警告邻近车辆有关安全更新时,它都会使用自己的证书对该消息进行数字签名。接收节点可以识别发送车载单元及其当前位置,但车载环境中的无线接入无法支持匿名广播消息^[143]。在LTE-

V2X中,核心网为用户设备分配不同的临时标识符,并且这些临时标识符是不断变化的。临时标识符以纯文本形式发送给用户设备^[164]。因此,LTE协议并不是安全的,攻击者能够识别用户的位置。攻击者能够执行各种标识符与用户设备之间的映射,从而实现位置跟踪^[165]。因此,内部和外部攻击者都可以发起该攻击。例如,文献[166]通过连续捕捉用户的位置服务信息,从而实现对用户的位置追踪攻击。文献[167]评估了低功耗蓝牙协议中芯片组硬件设计的变化、传输功率水平、热条件的差异对准确识别指纹的限制程度,表明物理层识别是攻击者跟踪移动设备的一种可行方法。文献[168]针对Wi-Fi设备设计了一种新型室内位置推断方法,基于强化学习模拟最可能的设备移动路径,并利用这些建模轨迹来构建接收信号强度指示器值变化图,从而构建用于实时设备位置跟踪的系统。常用的对抗位置追踪攻击的方法包括信号加密^[87]、匿名性^[88-89]、实时威胁检测^[90]、频繁更新车辆位置共享策略^[91-92]、差分隐私^[93-95]等。文献[87]提出一种新的地理加密协议,该协议将允许移动节点通过限制在特定位置和时间对消息进行解码来安全地相互通信。只有从GPS接收机获取的坐标与目标坐标相匹配,接收机才能解密密文。匿名性方法通过假名替换真实身份、群签名隐藏发送者、混淆区域技术模糊车辆实际位置,结合时空混洗、多路径转发等策略,在车联网中实现匿名通信。文献[88-89]基于匿名技术保护用户位置隐私,以隐藏用户真实的位置数据。实时威胁检测方法通过对车辆的通信数据、位置数据及网络流量数据进行分析,及时识别位置追踪攻击行为并触发预警和阻断机制。文献[90]基于多点定位和信号强度进行实时威胁检测,提出了一种用于无线定位系统的攻击检测方案。然而,攻击者可以通过对抗样本攻击(例如GAN)、APT攻击等绕过检测模型,从而增加了实时威胁检测误报率^[169]。此外,通过频繁更新车辆位置信息,可以使攻击者难以依据规律追踪车辆真实位置。文献[91]建议经常更新LTE用户的物理层身份标识,以防被动窃听者获取位置信息。文献[92]提出了一种基于频繁更改的假名的隐私保护框架,使用户能够避免因

访问位置而被识别。但是, 频繁更新车辆位置信息会导致通信开销激增、影响服务连续性, 甚至可能引发信道堵塞^[170]。基于差分隐私方法通过在车辆位置数据中添加精心设计的随机噪声, 在不影响数据统计特性的前提下, 使单个车辆位置信息难以被精确识别和追踪。文献[93-95]基于差分隐私技术实现可证明的位置隐私保护。然而, 基于差分隐私的方法存在隐私-效用均衡问题: 在数据中添加随机噪声虽然可以保护隐私, 但不可避免地会降低数据的可用性和准确性, 而添加过多噪声可能会导致数据无法使用, 添加过少噪声则无法提供足够的隐私保护, 因此需要在隐私和效用之间进行权衡^[171]。

2.1.4 基于真实性的攻击

信息真实性旨在确保车-路-云通信真实性, 包括使用证书和数字签名, 根据节点的身份授予节点访问网络的过程。真实性威胁分为证书复制攻击、女巫攻击、伪装攻击 3 种类型。

1) 证书复制攻击是指受感染的节点使用复制的证书来隐藏自己, 通过删除添加到黑名单中的证书来逃避检测。现有的车联网证书管理主要基于公钥基础设施 (public key infrastructure, PKI) 体系和数字签名机制。使用公钥密码学 (public key cryptography, PKC) 保护 V2X 通信已被广泛应用, 由一组证书颁发机构 (certification authorities, CA) 构成车辆公钥基础设施 (vehicular public-key infrastructure, VPKI), 为合法车辆提供多个匿名凭证 (称为假名)。车辆从一种假名切换到以前未使用过的假名, 以实现数字签名消息的不可链接性, 并提高 V2X 消息的发送者隐私。常见的对抗证书复制攻击的方法包括强证书管理^[96-97]、证书吊销和更新^[98]、证书透明度^[99-100]、区块链^[101]等。然而, 针对车联网环境, 构建证书管理机构非常复杂, 包括证书创建、颁发、更新、撤销等, 管理较为烦琐, 需要构建强大的证书管理机构以适应大规模车联网环境, 整合车联网行业相关的车厂、通信运营商、交通管理部门、金融机构等, 实现证书信息的集中管理和共享。2025 年 4 月 1 日, 我国正式实施国家标准《基于 LTE 的车联网无线通信技术安全证书管理系统技术要求》, 该标准规定了基于

LTE 的车联网安全证书管理系统架构、证书管理要求、安全认证机制要求以及相关的证书格式^[172]。

2) 女巫攻击是一种针对分布式网络的攻击方式, 攻击者通过伪造多个虚假的身份或节点来获得不成比例的影响力或控制权。在女巫攻击中, 攻击者通常被称为恶意节点, 而模拟的虚拟节点被称为女巫节点。例如, 恶意车辆可能会制造大量具有虚假身份和虚假位置的虚拟车辆, 从而给附近其他车辆造成前方交通拥堵的错觉^[173]。随后, 邻近车辆可以选择其他路线, 而攻击者则可以获得良好的路况。恶意节点可能会通过多个女巫节点向目标车辆或 RSU 发送无用消息, 从而降低网络性能。由于女巫攻击造成的破坏非常严重, 研究人员提出了多种检测方法。这些技术可以分为 3 类: 基于资源测试^[102-103]、基于可信认证^[104-106]和基于位置验证的机制^[107-109]。基于资源测试假设每个物理节点具有有限的资源, 如计算、存储和通信资源。如果恶意节点伪造多个女巫节点, 则无法同时通过相应的测试。例如, 文献[102]使用计算谜题来测试每个身份的计算资源, 检测器生成大量随机谜题来同时挑战每个节点。如果一个节点在有限的时间内找不到结果, 那么它很可能是一个女巫节点。文献[103]提出了另一种资源测试方法, 采用通信通道来测试每个节点。研究表明, 女巫节点不能同时在不同的通道上传输消息。因此, 检测器将不同的通道分配给相邻节点进行广播, 然后选择一个随机通道进行监听。由于通道是随机选择的, 检测器将找到所有经过多轮测试的女巫节点。基于资源测试的方法在对抗女巫攻击时, 由于需要执行复杂测试任务, 会消耗大量系统资源, 在车联网中, 频繁测试会占用车辆计算与通信资源, 干扰其他功能的运行。并且, 这种方法难以适应不断变化的攻击手段, 攻击者容易找到绕过的方式, 面对分布式协同攻击时, 由于单个恶意节点资源占用少, 更难以识别^[174]。基于可信认证的方法是在所有节点之间建立信任关系的最流行的技术之一。此类方法通常使用证书颁发机构、公钥基础设施、数字签名和加密算法来确保每个身份的可信度。文献[104]设计了车辆公钥基础设施框架,

详细介绍了密钥管理、身份验证和密钥撤销机制。然而,该架构的部署仍然是一个不确定的问题,需要复杂的工作来完成。文献[105]提出了一种适合车联网初始阶段的解决方案,使用RSU发布的时间戳系列来检测女巫节点。每条消息应至少包含最后两个时间戳,并且两个女巫节点之间的这些时间戳序列总是非常接近。这种方法需要在每辆车上安装额外的硬件,即防篡改设备,以保证RSU授权的时间戳不能被修改。文献[106]提出了一种名为Footprint的方案来检测女巫攻击。当车辆接近RSU范围时,它需要包含RSU身份和时间戳的授权消息,以验证车辆在特定时间出现在某个RSU附近。然后,每辆车都可以生成一条轨迹,包含其沿路记录的所有消息。想要检测女巫节点周围的节点,则使用检查窗口和轨迹长度来排除不同的轨迹,因为两辆车具有完全相同轨迹的概率很小。文献[175]提出了一种基于Boneh-Shacham短群签名的方案,用于解决车联网中隐私和女巫攻击这两个相互冲突的目标。基于可信认证的方法能够在攻击初期识别女巫节点。然而,这通常需要一个中心化的信任实体来颁发数字签名或证书,这在车载自组网的初始阶段难以实施。此外,由于车载自组网拓扑结构变化快且密码算法复杂,轻量级、去中心化的检测方法更适合这种动态环境。基于位置验证方法通常通过物理测量手段来检测女巫攻击,如接收信号强度指示器(received signal strength indicator, RSSI)、到达角度和到达时间差。这些测量依赖于硬件和物理环境,恶意节点难以轻易伪造或篡改这些数据,能够快速验证节点位置,特别是在结合分布式验证机制时,可以提高系统效率^[107]。文献[108]提出了一种基于协作RSSI的女巫检测方案,不计算准确的位置,而是计算不同节点之间的距离。然后,每个节点将具有相似RSSI的身份分组在一起,并将该可疑分组结果广播给其他节点。最后,每个节点取所有接收到的组中最大的交集作为检测到的女巫节点。在动态场景中,如车辆快速移动,节点位置和信号状态变化迅速,物理测量的实时性和准确性难以保证。测量时延会导致位置验证滞后,无法及时发现攻击行为。而且,快速变化的信号

特征使基于静态模型的位置验证算法失效,攻击者可以利用动态场景的复杂性,频繁变换攻击策略。

3) 伪装攻击:攻击者通过冒充合法用户或设备的身份,获取未授权的访问权限,以进行恶意操作或窃取敏感信息^[176]。攻击者假装是一辆紧急车辆,迫使道路上的其他车辆改变车道或降低车速。为了确保VANET的安全,设计了不同的安全措施,其中最主流的是入侵检测系统^[110]。虽然IDS已被证明可以有效检测传统网络中的恶意节点,但由于车联网的特性,其在车联网中的应用存在不同之处,并且面临挑战。目前,机器学习已成为网络安全领域最广泛使用的技术之一,尤其是应用于IDS^[111]。这些智能模型提供了更大的灵活性,能够检测潜在威胁并提供有效的解决方案。

2.1.5 基于不可否认性的攻击

不可否认性用于识别执行特定行为的真实节点的ID。它是一种通过数字签名或加密确保实体之间消息传输的方法。在IEEE802.11p中,周期消息和警告消息使用发送者证书进行签名^[177]。此外,基于位置的服务和任何敏感信息都被加密,从而在节点发起恶意行为时能够识别其身份。由于信任概念的存在,协作式V2V通信在实现不可否认性时,对协作时效的要求极为严格。事实上,如果一个节点信任另一个节点,即使后者使用虚假身份,前者也会与其进行通信^[178]。对抗不可否认性常用的方法为数字签名^[112-113]、PKI技术^[114]、区块链技术^[115-117]等。每种防御手段都存在一定局限性,例如,数字签名依赖密钥管理,密钥泄露或丢失会削弱其安全性,且算法若被破解将导致签名失效;PKI技术以中心化认证机构为核心,存在单点故障风险,证书管理复杂易出漏洞,兼容性也较差;区块链技术则面临性能瓶颈,处理大量数据时效率低,存储成本高,还存在监管和合规方面的挑战,同时智能合约漏洞也可能影响不可否认性保障^[235]。因此,需要多种防御手段相结合来保证车-路-云系统的安全性。

2.2 基于机器学习技术的车-路-云威胁与防御

近年来,机器学习(machine learning, ML)

作为一种数据挖掘工具,可用于车-路-云协同系统,在车端实现感知^[179]、预测^[180]、规划^[181]及决策控制^[182]。例如,车辆通过摄像头、激光雷达等传感器设备扫描道路,检测交通标志、信号灯、人行道、障碍物等信息,并基于该信息训练 ML 模型,开发自动驾驶应用;在路端,智能基础设施集成来自车辆的 V2V 信号,根据实际交通负载优化速度限制、红绿灯计时和每个方向的车道数量。通过协作自适应巡航控制技术,进一步优化自动驾驶车辆的交通流量^[183]。另外,车辆可以利用巡航控制,以车队形式相互跟随,从而节省燃料;在云端,基于车辆和路边基础设施的数据,训练 ML 模型以指导该区域自动驾驶车辆的决策。

本文将传统机器学习系统、自动驾驶系统以及车-路-云协同系统进行对比,如图 2 所示。车-路-云协同系统相对于另外两个系统,更专注于车-路-云之间的协同数据以实现自动驾驶车辆的协同感知、协同预测、协同规划和协同决策。在数据收集环节,传统系统依赖传感器数据与知识库,自动驾驶系统借助传感器和摄像头,车-路-云协同系统则融合多源信息。在数据处理上,分别是数据预处理、特征提取与图像增强、网络和路况信息处理^[233]。模型输入阶段,分别基于特征、识别对象、协同环境进行训练。模型输出时,分别产生决策行为用于车辆控制和实现协同操作。图 2 中的攻击面涵盖数据篡改/投毒、模型

篡改/投毒、输出篡改。数据篡改/投毒是恶意改动或污染系统收集的数据,误导后续处理;模型篡改/投毒指对模型参数、算法等进行恶意修改,破坏模型正常学习;输出篡改是修改系统输出决策结果,使最终指令出错。这些攻击面会威胁传统机器学习系统、自动驾驶系统、车-路-云协同系统的安全性及可靠性,可能引发严重后果。本文将机器学习威胁分为集中式和分布式两种情况。

2.2.1 基于集中式机器学习的攻击

通常的机器学习攻击可分为白盒攻击^[184]、灰盒攻击^[185]和黑盒攻击^[186]。在白盒攻击中,攻击者假设对底层的机器学习模型了解充分,包括优化技术、已训练的模型、模型结构、激活函数、超参数、层权重以及训练数据的信息。在灰盒攻击中,攻击者假设对目标模型有一定了解,而在黑盒攻击中,则假设攻击者对底层机器学习模型和训练数据没有任何的知识。黑盒攻击模拟现实世界中的情况,攻击者扮演普通用户的角色,尝试从机器学习系统的输出中推断信息。黑盒对抗攻击利用对抗性示例的可转移性,即假设一个机器学习模型创建的对抗性示例会影响其他在相似数据分布上训练的模型。对抗这些攻击的常用方法包括加密技术^[118-122]、对抗训练^[123-126]、正则化^[126-128]、隐私保护方法^[129]等。

在车-路-云协同系统中,车辆利用搭载的 ML 模型来解析其接收到的来自周围车辆的时间

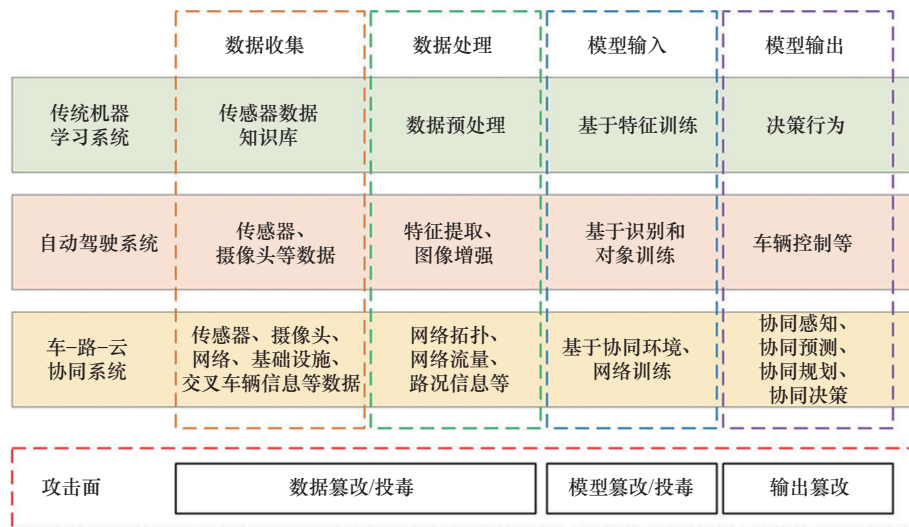


图2 传统机器学习系统、自动驾驶系统和车-路-云协同系统对比

Figure 2 Comparison of traditional machine learning system, autonomous driving system and vehicle-road-cloud collaborative system

序列驾驶信号，如速度和制动状态，并相应地采取行动以提高交通安全和道路效率。在这种情况下，攻击者可能会试图进行对抗性（白盒攻击、灰盒攻击或黑盒攻击）攻击，通过向实际的驾驶信号中添加微小的扰动，以欺骗其他协同车辆内部的ML模型，导致输出错误分类的结果，从而可能引发交通拥堵或事故。

本文调研了集中式机器学习安全威胁相关文献，见表3，并将其分为白盒攻击、灰盒攻击及黑盒攻击。

2.2.2 基于分布式机器学习的攻击

传统的训练方法通常在强大的云中心进行，需要高性能硬件，并依赖于集中式训练数据集来有效学习最佳模型参数，比如权重、偏差和隐藏模式。但是，这种基于云的训练方式面临一些挑战。首先，越来越大、越来越复杂的模型（如大型语言模型）增加了模型优化的复杂性^[187]。为

了在大型数据集上获得收敛的结果，单个计算机上训练这些模型可能需要数周甚至数月的时间，带来巨大的时间成本^[188]。其次，将大量数据传输到云中心会增加相当大的通信开销，可能导致网络拥塞。此外，收集数据集需要与中央服务器共享原始数据，引发了人们对隐私和安全的担忧，尤其是处理个人敏感信息。

分布式训练旨在通过利用车-路-云协同中各种分布式的数据源和计算资源来克服这些限制。该训练过程提供大量数据集，并利用专门的学习算法进行更新，生成可用的人工智能模型^[189]。在车联网环境中，车辆自组织网络通过V2X通信在车辆和基础设施之间建立通信链路。这种连接促进了训练数据和模型信息在不同位置之间的传输，使学习能够突破时间和空间限制^[187]。有两种基本不同的策略可用于扩展训练工作负载跨分布式实体：数据分区和模型分区。这两种合作

表3 集中式机器学习安全威胁文献

Table 3 Literature of centered machine learning security threats

文献	时间	类型	解决问题
Kumar等 ^[191]	2020年	黑盒攻击	通过最大限度地减少最混淆类的损失，解决了黑盒攻击中收敛较慢的问题
Won等 ^[192]	2020年	白盒攻击	提出了一种支持动态白盒加密的查找表改组机制。该机制使得攻击者很难确定表条目的位置，从而难以解密/加密密文/明文
Sarker等 ^[193]	2020年	黑盒攻击	提出了一种用于深度驾驶机动车分类模型的高级黑盒对抗攻击方法，基于二分搜索方法提高算法的求解效率
Sun等 ^[194]	2020年	黑盒攻击	根据已发现LiDAR点云的漏洞构建了第一个黑盒欺骗攻击，该攻击在所有目标模型上普遍实现了约80%的平均成功率
Deng等 ^[195]	2020年	黑盒攻击/白盒攻击	研究驾驶模型中的回归模型在对抗性攻击中的影响、现有防御技术的有效性以及对系统和中间件构建者的防御影响
Xiong等 ^[196]	2021年	黑盒攻击	提出了两种多源对抗样本攻击模型，包括并行攻击模型和融合攻击模型，以同时攻击自动驾驶汽车中的图像和激光雷达感知系统
Li等 ^[197]	2021年	白盒攻击	实现了两种对车载网络中使用基于LSTM检测模型有害的对抗性攻击模型
Xu等 ^[198]	2021年	白盒攻击	提出了一种基于投影梯度的迭代攻击方法，该方法能有效地骗过几种基于DNN的分割模型
Sarker等 ^[199]	2021年	黑盒攻击	提出了一种针对CAV场景中的时间序列DNN模型的上下文感知黑盒对抗攻击
Jia等 ^[200]	2021年	黑盒攻击	提出了针对IoU的黑盒攻击，根据当前帧和历史帧，预测IoU分数顺序生成扰动，以评估深度跟踪器的鲁棒性
Moukahal等 ^[185]	2021年	灰盒攻击	提出了一种混合模糊测试框架，将模糊测试的效率和一致性执行的精度结合起来，为汽车行业提供可靠的安全测试工具
Sun等 ^[201]	2022年	黑盒攻击	提出了用于轻量级攻击的错误变换器，将参数空间中的近似误差转化为特征空间中的扰动，并通过扰动特征来减少误差
Jia等 ^[202]	2024年	黑盒攻击	提出基于黑盒攻击的端到端框架，使商业交通标志识别系统在现实世界中以高成功率和强鲁棒性输出指定的错误分类结果
Xu等 ^[236]	2025年	白盒攻击/灰盒攻击	使用序贯卡尔曼滤波器构建多源状态估计系统，整合和估计前车通过多通道传输的冗余状态信息，测试白盒攻击和灰盒攻击下系统抵御攻击的韧性

策略的操作类似于协作推理模式。然而训练过程的不同之处在于，它侧重于对模型和数据集进行分区以学习最佳模型，而不是专注于特定的推理输入样本。

联邦学习 (federated learning, FL)^[184]是一种新的分布式机器学习 (distributed machine learning, DML) 架构，最初由 Google 提出，旨在让用户参与训练过程的同时保持数据私密性，从而解决隐私问题。在典型的 FL 框架中，所有参与者（如智能手机、车辆和服务器）都会使用本地数据对共享的 ML 模型进行迭代训练。例如，文献[190]研究了基于联邦学习的车联网入侵检测系统，该系统通过使用联邦学习加速模型训练和更新，并保护了车辆模型和数据隐私。在联邦学习初始阶段，服务器将模型分发给参与者。参与者在本地处理后，定期上传更新的模型权重以进行全局聚合。然后，更新后的全局模型会重新分发给参与者，进行下一轮迭代。这一迭代过程会持续进行，直到损失函数在全局模型优化中达到收敛。然而，在车-路-云协同过程中，联邦学习的训练数据和模型并不安全，面临多种安全威胁，本文将它们总结为以下 3 种安全威胁。

1) 数据/模型投毒攻击：投毒攻击是指攻击者通过操纵训练数据或修改模型参数，破坏机器学习训练过程的攻击方式。该攻击可分为数据投毒和模型投毒^[203]。自动驾驶车辆依靠基于深度学习模型的感知系统来识别物体并自动驾驶。然而，算法缺陷或模型错误，自动驾驶车辆中的感知系统可能会对物体进行错误分类，从而导致汽车事故^[204]。自动驾驶车辆数据投毒攻击包括目标类数据投毒攻击^[205]、干净标签投毒攻击^[206]和特殊类型攻击，如标签翻转攻击^[207]。文献[208]指出，攻击者可在路边设置有毒标志，欺骗自动驾驶车辆，使其深度学习模型对标志进行错误分类，从而导致鲁莽驾驶。文献[209]提出一种基于黑盒和清洁标签数据的数据投毒方法，通过混合粒子群算法和模拟退火算法优化扰动，降低目标模型性能。一个典型的案例是 Uber 自动驾驶汽车事件^[204]，一辆自动驾驶汽车将行人错误地分类为其他物体，未能及时刹车，导致碰撞。在对抗数据投毒方面，主要分为基于数据的

防御^[130-131]和基于模型的防御^[132-134]。基于数据的防御方法主要通过数据清理，分析训练数据在特征空间中的分布特征，并对比投毒攻击前后差异性。文献[130]提出一种在线分类的异常值检测方法，然而，该方法容易受到高维的影响，且仅适用于非负凸损失函数。文献[131]通过对比投毒样本和干净样本在模型较高层中的特征分布差异，筛选投毒样本，从而防御特征碰撞攻击^[131]和凸多胞体攻击^[210]。基于模型的防御方法主要在训练过程中设计额外步骤以检测投毒攻击行为。文献[132]利用最大熵原理和变分推理方法估计模型中间层的参数分布边界，发现嵌入层中模型参数边界会因投毒特征而增大，从而检测异常投毒数据。文献[133]提出一种基于认证的防御方案，利用随机平滑方法对抗标签翻转攻击，消除翻转标签对分类器的影响。在数据投毒检测架构方面，主要结合联邦学习架构来检测异常节点上传的梯度差异性。文献[134]提出了基于门控制循环单元的联邦持续学习异常检测架构，用于应对车联网中的女巫攻击数据投毒。文献[211]提出一种具有投毒识别功能的隐私保护对象检测架构，基于隐私保护机制对抗差分攻击威胁，防止模型逆向推理导致隐私泄露。

2) 拜占庭式对抗攻击：拜占庭攻击是分布式架构中常见的攻击方式^[212]。拜占庭攻击是指在分布式网络架构中，恶意节点通过发送错误信息，使其他成员对协作策略得出不同的结论，进而破坏网络系统的一致性。分布式机器学习方法，尤其是联邦学习，在协同自动驾驶中的应用保护了车辆本地数据的安全。然而，联邦学习非常容易受到拜占庭攻击，恶意用户可以伪造真实模型或梯度来破坏学习过程，或者直接毒害训练数据，使全局模型学习到错误信息^[213]。文献[214]提出了拜占庭式鲁棒聚合器能够有效抵御卷积神经网络上的标签翻转数据投毒攻击。目前对抗拜占庭攻击主要采用对平均值的稳健估计，如基于中值的聚合器^[135]、Krum^[136]和修剪平均值^[136]，以及通过冗余和数据洗牌^[137]的方式。虽然采用了多种方法对抗拜占庭攻击，文献[133]表明，在联邦学习中，多重拜占庭防

御方法对模型投毒攻击几乎没有。因此,有必要针对拜占庭攻击,在联邦学习中进行更多的有效性实证分析。

(3) 数据/模型隐私攻击:机器学习模型训练需要收集大量交通图像,并将模型协同卸载到中央服务器,由中央服务器挖掘这些图像以生成可用的预测模型^[215]。然而,这些图像往往包含许多用户的敏感信息,可能引发潜在的隐私泄露^[216-217],攻击者甚至可能窃取图像以挖掘用户的日常出行模式^[218]。文献[219]提出了智能网联汽车边缘辅助共享架构,基于秘密共享技术将原始数据分割为两个密文,并构造两类安全函数,以保护卷积神经网络隐私安全。文献[220]提出了一种基于同态加密的机器学习训练方案。然而,受目标检测速度的限制,基于加密的方法难以在准确性和隐私之间实现平衡,也无法满足自动驾驶车辆低时延需求。在车联网中,联邦学习可以解决车辆数据“孤岛”问题,通过本地训练机器学习模型,敏感输入不会离开自动驾驶车辆,只有模型参数更新才会发送到服务器进行模型收敛和更新。通过设置特定配置,可以最大限度地减少训练期间的数据存储,从而防止数据泄露^[221]。目前常用的隐私保护方法分为差分隐私(differential privacy, DP)^[138]、区块链^[139]、同态加密(homomorphic encryption, HE)^[140]、功能加密(functional encryption, FE)^[141]和安全多方计算(multi-party computation, MPC)^[142]。差分隐私技术通过在模型输入中引入噪声,有效保护数据隐私,防止个人敏感信息泄露。基于差分隐私的联邦学习既能保护个人隐私,又能实现分散的模型训练和知识共享,为隐私保护和合作学习之间的平衡提供了理想的解决方案。文献[138]提出了一种用于边缘计算的隐私保护联邦学习框架,旨在平衡隐私水平和低通信量,然而,由于每个客户端都会增加噪声,本地差分隐私严重影响了聚合模型的性能。为提高聚合模型的性能,有研究提出了基于洗牌模型的隐私保护框架,同时兼顾了隐私和高效通信。文献[222]提出了基于洗牌模型的局部差分隐私联邦学习框架,扩大了隐私预算,但未考虑通信时延。鉴于车联网对通信高可靠性、低时延

的需求^[223],需要设计一种兼顾通信效率、隐私保护和模型性能的方案。

3 车-路-云协同安全技术发展方向

本节总结基于联邦学习、区块链、安全多方计算、零知识证明及差分隐私相关安全技术在车-路-云协同中的应用及发展方向。各个安全技术并不是相互独立的,基于不同安全技术相互融合,为智能网联汽车和智慧交通安全性、隐私保护和协同效能提供坚实的技术保障。

3.1 基于联邦学习的车-路-云安全协同

尽管联邦学习的应用存在一些问题,但它保护数据隐私和分布式的特点使其能够解决跨多个互联设备或代理的大规模分布式训练问题。与传统的分布式机器学习系统相比,联邦学习具有更强的隐私保护功能。联邦学习依赖于本地训练的机器学习参数(如权重和偏差)的交换,而不是像分布式机器学习那样共享原始数据。与传统的大数据融合方法不同,联邦学习利用设备上的学习能力,并通过网络进行机器学习参数的密集交换。将联邦学习技术应用于第六代蜂窝车辆对基础设施范式中,有望实现零时延和高速率数据通信,为未来的车辆场景带来显著的好处。

由于车-路-云系统具有高度异构性,如何在计算能力有限的车载设备和高性能云端之间合理分配计算任务是关键问题。首先,通过轻量化模型、分层联邦学习以及动态节点管理等方法,提高训练效率和适应不同设备的需求。通过优化模型聚合机制,减少通信成本也是提升系统性能的重要手段。其次,隐私保护与安全性是联邦学习在车-路-云协同中的核心问题。由于车辆涉及大量个人敏感数据,如何保证数据的隐私性至关重要。差分隐私、同态加密以及多方安全计算等技术为保护数据隐私提供了有效手段。同时,针对恶意节点的防护也需要加强,可以采用鲁棒性增强的联邦学习算法,抵御数据中毒或模型篡改等攻击。文献[224]提出了一种半同步联邦学习协同协议,以提高车联网中机器学习的性能。跨域联邦学习成为未来研究的一个重要方向。车载数据、路侧单元数据和云端数据的异质性使得多模态数据融合成为必要。为此,

联邦学习框架需支持多源数据融合和跨区域模型迁移, 以实现不同区域或环境下自动驾驶系统的经验共享和优化决策。在通信与协同决策方面, 优化通信资源分配、减少数据传输量以及通过边缘计算提升响应速度, 将是未来车-路-云协同中的重点研究方向。此外, 基于深度强化学习在优化车-路-云协同效率和车联网资源分配方面展现出显著的优势, 尤其是结合事件触发控制等技术^[225-227]。

3.2 基于区块链的车-路-云安全协同

随着比特币的兴起, 区块链作为一种分布式账本开始受到关注。密码学和共识机制确保了数据的安全性和一致性, 其链式结构便于追溯数据来源并追究责任。智能合约作为一种不可篡改的代码, 通过点对点网络实现数据传播和验证, 推动了区块链的应用。区块链因其不可篡改性和一致性等安全特性, 在车载自组网等领域得到广泛应用。

在身份管理方面, 区块链通过分布式身份管理技术为车辆、RSU 和云端建立唯一可信的身份标识, 结合数字签名与加密技术, 能够高效抵御伪造身份和恶意节点的入侵。智能合约的引入为协同任务提供了自动化的规则执行能力, 例如, 在交通信号控制、车队协作、路径规划和资源分配等场景中, 通过动态调整规则提高系统效率, 并结合动态博弈模型提升对共谋攻击的防御能力。为应对车载设备和路侧单元资源受限的问题, 轻量化区块链技术成为重要的发展方向。例如, 利用侧链、分片技术以及高效共识算法, 可大幅降低计算与存储负担, 同时满足车联网的实时性和低时延需求。文献[228]提出了一种基于联邦学习和区块链的可追溯身份隐私保护方案, 并采用零知识证明的匿名身份方法。基于区块链的分布式入侵检测与溯源系统为车-路-云网络提供了全新的安全防御模式, 通过实时记录和广播攻击行为, 快速定位攻击来源并协调防御策略。

3.3 基于安全多方计算的车-路-云安全协同

安全多方计算在车-路-云协同系统中具有巨大的应用潜力。SMC 允许多个参与方在不公开私密数据的情况下进行计算, 以确保数据隐私和安全性。在车-路-云协同中, SMC 可用于处理

涉及车辆位置、行驶速度等敏感信息的任务。例如, 车辆可以通过 SMC 协议安全地共享其位置数据, 而无须暴露具体位置信息给其他车辆或云服务器。文献[229]设计了一种基于安全多方计算的车载移动众包感知隐私保护方案, 实现了更高的通信效率和更好的隐私保护。此外, SMC 还可用于实现联合决策和协同控制, 如在路口交通管理中协调车辆通过顺序。通过保护数据隐私并促进信息共享和合作, SMC 有助于提高车-路-云协同系统的安全和效率, 为智能交通和道路安全做出贡献。

轻量化 MPC 协议的设计是当前研究的重点。由于车载设备和路侧单元的计算资源有限, 传统 MPC 协议在计算复杂度和通信开销上难以满足车联网系统的实时性需求。首先, 通过优化协议流程、减少轮次交互以及引入高效密码学技术(如同态加密和秘密共享), 可以显著降低计算和通信成本。其次, 将 MPC 与差分隐私结合, 可为车-路-云协同提供双重隐私保护, 确保协同计算输出结果不会泄露个体信息, 满足严格的隐私保护需求。此外, 动态参与的 MPC 协议是面向车-路-云动态环境的重要方向。由于车辆和路侧单元的频繁加入与退出, 传统静态 MPC 协议难以适应, 需研究支持动态节点管理的灵活协议, 以确保计算的鲁棒性和参与方的安全性。

3.4 基于零知识证明的车-路-云安全协同

基于零知识证明(zero-knowledge proof, ZKP)的车-路-云安全协同系统为智能网联汽车和智慧交通提供了高度隐私保护和数据安全保障, 尤其是在敏感信息共享和身份认证方面具有显著优势, 其发展方向主要集中在性能优化、协议设计与实际应用的结合。文献[230]提出了一种结合区块链和零知识证明的电动汽车隐私保护认证框架, 旨在解决电动汽车充电过程中的隐私问题, 然而, 该框架会造成较大的计算和通信开销。轻量化零知识证明协议的设计是当前的重要研究方向。由于车载设备和路侧单元计算能力有限, 传统零知识证明协议计算复杂度高且通信开销大, 难以满足车-路-云协同的实时性需求。首先, 通过引入快速生成证明的算法和非交互式零知识证

明技术，可以显著降低资源占用并提高验证效率，从而适应资源受限的车联网环境。其次，零知识证明与分布式数据协同的结合为车-路-云环境中的隐私保护提供了创新解决方案。例如，利用零知识证明可在不泄露车辆位置、速度等敏感信息的前提下完成交通信号优化、路径规划和车辆协同任务，实现高效、安全的协作。此外，基于ZKP的分布式身份认证与访问控制是关键发展方向之一。通过零知识证明，车辆和路侧单元可在不暴露具体身份的情况下完成身份认证，防止敏感信息泄露，同时有效抵御伪造身份攻击和重放攻击。

3.5 基于差分隐私的车-路-云安全协同

基于差分隐私的车-路-云安全协同系统为智能网联汽车和智慧交通提供了一种有效的隐私保护方法，特别是在数据共享和协作计算中能够保护个体信息的隐私。差分隐私技术通过对数据进行噪声添加或处理，确保任何单个数据点的加入或移除都不会显著影响计算结果，从而防止泄露用户的敏感信息。研究人员正在探索通过不同类型的差分隐私机制（如局部差分隐私、集中式差分隐私）来平衡隐私保护和数据精度，满足车联网实时性的需求。此外，差分隐私与联邦学习的结合是当前的热门方向。在车-路-云系统中，车辆和路侧单元可以通过联邦学习协同训练机器学习模型，差分隐私则可为这些模型提供隐私保护。通过在每次训练过程中加入噪声，防止模型泄露用户的私密数据。文献[231]集成了联邦学习和本地差分隐私，以促进车联网中的众包应用，但本地差分隐私在计算资源和时间方面成本较高。

4 结束语

本文首先总结了车-路-云发展现状，并详细介绍了自动驾驶发展历程和车-路-云协同自动驾驶环境。其次，分析了车-路-云协同过程中遇到的安全威胁，并系统总结了相应的防御方法。最后，总结了对抗这些安全威胁的安全技术发展方向，包括联邦学习、区块链技术、安全多方计算、零知识证明、差分隐私技术。

参考文献：

- [1] 2023年交通运输行业发展统计公报[EB]. 2024.
Annual statistical bulletin of transportation development in China, 2023[EB]. 2024.
- [2] 国务院关于印发《中国制造2025》的通知[EB]. 2015.
Notice of the State Council on issuing the "Made in China 2025" [EB]. 2015.
- [3] 关于开展智能网联汽车准入和上路通行试点工作的通知[EB]. 2023.
Notice on the pilot work for the access and road passage of intelligent connected vehicles[EB]. 2023.
- [4] 关于开展智能网联汽车“车-路-云一体化”应用试点工作的通知[EB]. 2024.
Notice on the pilot work for the integrated application of intelligent connected vehicles in the "vehicle-road-cloud integration" system [EB]. 2024.
- [5] WANG W N, YING L, ZHANG J S. A game-theoretic approach to quality control for collecting privacy-preserving data[C]//Proceedings of the 2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton). Piscataway: IEEE Press, 2015: 474-479.
- [6] 王良民, 刘晓龙, 李春晓, 等. 5G车联网展望[J]. 网络与信息安全学报, 2016, 2(6): 1-12.
WANG L M, LIU X L, LI C X, et al. Overview of Internet of vehicles for 5G[J]. Chinese Journal of Network and Information Security, 2016, 2(6): 1-12.
- [7] 道路交通安全法[EB]. 2023.
Road traffic act[EB]. 2023.
- [8] 中国信息通信研究院. 车联网白皮书(2023年)[R]. 2023.
China Academy of Information and Communications Technology. White paper on internet of vehicles (2023) [R]. 2023.
- [9] 丁飞, 张楠, 李升波, 等. 智能网联车-路-云协同系统架构与关键技术研究综述[J]. 自动化学报, 2022, 48(12): 2863-2885.
DING F, ZHANG N, LI S B, et al. A survey of architecture and key technologies of intelligent connected vehicle-road-cloud cooperation system[J]. Acta Automatica Sinica, 2022, 48(12): 2863-2885.
- [10] 左志强, 刘正璇, 王一晶. 基于车-路-云一体化的混合交通系统优化控制综述[J]. 控制与决策, 2023, 38(3): 577-594.
ZUO Z Q, LIU Z X, WANG Y J. A survey of optimal control for mixed traffic system with vehicle-roadcloud integration[J]. Control and Decision, 2023, 38(3): 577-594.
- [11] 董晋之. 2024 智能互联网蓝皮书: 我国辅助自动驾驶乘用车市场渗透率达 47.3%[R]. 2024.
DONG J Z. 2024 smart internet blue book: the market penetration rate of China's assisted driving passenger cars reaches 47.3%[R]. 2024.
- [12] 中华人民共和国工业和信息化部. 装备工业一司召开智能网联汽车产品准入及软件在线升级管理工作推进会[EB]. 2025.
The Ministry of Industry and Information Technology of the People's Republic of China. Equipment Industry Department I held a promotion meeting on the access of intelligent and connected vehicle products and the management of software over-the-air up-

- grades[EB]. 2025.
- [13] DE LA FORTELLE A, QIAN X J, DIEMER S, et al. Network of automated vehicles: the AutoNet 2030 vision[C]//Proceedings of the 21st ITS World Congress. Detroit, USA, 2014.
 - [14] SU Z, HUI Y L, LUAN T H. Distributed task allocation to enable collaborative autonomous driving with network softwarization[J]. IEEE Journal on Selected Areas in Communications, 2018, 36(10): 2175-2189.
 - [15] LI B, YANG L, XIAO J Z, et al. Collaborative mapping and autonomous parking for multi-story parking garage[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(5): 1629-1639.
 - [16] LI Y F, TANG C C, LI K Z, et al. Consensus-based cooperative control for multi-platoon under the connected vehicles environment[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(6): 2220-2229.
 - [17] DENG R Q, DI B Y, SONG L Y. Cooperative collision avoidance for overtaking maneuvers in cellular V2X-based autonomous driving[J]. IEEE Transactions on Vehicular Technology, 2019, 68(5): 4434-4446.
 - [18] CHEN Q, TANG S H, YANG Q, et al. Cooper: cooperative perception for connected autonomous vehicles based on 3D point clouds[C]//Proceedings of the 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS). Piscataway: IEEE Press, 2019: 514-524.
 - [19] CHEN Y M, LU C, CHU W B. A cooperative driving strategy based on velocity prediction for connected vehicles with robust path-following control[J]. IEEE Internet of Things Journal, 2020, 7(5): 3822-3832.
 - [20] YU C, WANG X, XU X, et al. Distributed multiagent coordinated learning for autonomous driving in highways based on dynamic coordination graphs[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(2): 735-748.
 - [21] LIU Y C, TIAN J J, MA C Y, et al. Who2com: collaborative perception via learnable handshake communication[C]//Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE Press, 2020: 6876-6883.
 - [22] WANG T H, MANIVASAGAM S, LIANG M, et al. V2VNet: vehicle-to-vehicle communication for joint perception and prediction[M]//Computer Vision - ECCV 2020. Cham: Springer International Publishing, 2020: 605-621.
 - [23] CHEN S H, ZHANG N X, SUN H F. Collaborative localization based on traffic landmarks for autonomous driving[C]//Proceedings of the 2020 IEEE International Symposium on Circuits and Systems (ISCAS). Piscataway: IEEE Press, 2020: 1-5.
 - [24] CHEN X S, LENG S P, HE J H, et al. Deep-learning-based intelligent intervehicle distance control for 6G-enabled cooperative autonomous driving[J]. IEEE Internet of Things Journal, 2021, 8(20): 15180-15190.
 - [25] ZHANG Y, LI C L, LUAN T H, et al. Collaborative driving: learning-aided joint topology formulation and beamforming[J]. IEEE Vehicular Technology Magazine, 2022, 17(2): 103-111.
 - [26] LI Y M, MA D K, AN Z Y, et al. V2X-sim: multi-agent collaborative perception dataset and benchmark for autonomous driving[J]. IEEE Robotics and Automation Letters, 2022, 7(4): 10914-10921.
 - [27] LIN C M, TIAN D X, DUAN X T, et al. V2VFormer: vehicle-to-vehicle cooperative perception with spatial-channel transformer[J]. IEEE Transactions on Intelligent Vehicles, 2024, 9(2): 3384-3395.
 - [28] KONG X J, ZHU B, SHEN G J, et al. Spatial-temporal-cost combination based taxi driving fraud detection for collaborative Internet of vehicles[J]. IEEE Transactions on Industrial Informatics, 2022, 18(5): 3426-3436.
 - [29] YANG S H, YIN H H F, YEUNG R W, et al. On scalable network communication for infrastructure-vehicle collaborative autonomous driving[J]. IEEE Open Journal of Vehicular Technology, 2022, 4: 310-324.
 - [30] YI S, ZHANG H, LIU K. V2IViewer: towards efficient collaborative perception via point cloud data fusion and vehicle-to-infrastructure communications[J]. IEEE Transactions on Network Science and Engineering, 2024, 11(6): 6219-6230.
 - [31] QIAO G H, LENG S P, ZHANG K, et al. Collaborative task offloading in vehicular edge multi-access networks[J]. IEEE Communications Magazine, 2018, 56(8): 48-54.
 - [32] XIE R C, TANG Q Q, WANG Q N, et al. Collaborative vehicular edge computing networks: architecture design and research challenges[J]. IEEE Access, 2019, 7: 178942-178952.
 - [33] CHEN Y M, LU C, CHU W B. A cooperative driving strategy based on velocity prediction for connected vehicles with robust path-following control[J]. IEEE Internet of Things Journal, 2020, 7(5): 3822-3832.
 - [34] HE Y K, WU B F, DONG Z, et al. Towards C-V2X enabled collaborative autonomous driving[J]. IEEE Transactions on Vehicular Technology, 2023, 72(12): 15450-15462.
 - [35] QIAN H, ZHAO L, HAWBANI A, et al. Collaborative overtaking strategy for enhancing overall effectiveness of mixed connected and connectionless vehicles[J]. IEEE Transactions on Mobile Computing, 2024, 23(12): 13556-13572.
 - [36] HU S K, FANG Z R, FANG Z H, et al. AgentsCoMerge: large language model empowered collaborative decision making for ramp merging[J]. IEEE Transactions on Mobile Computing, 2025, (99): 1-15.
 - [37] LIU G J, HU Y, XU C X, et al. Towards collaborative autonomous driving: simulation platform and end-to-end system[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, (99): 1-18.
 - [38] LIU S S, LIU L K, TANG J, et al. Edge computing for autonomous driving: opportunities and challenges[J]. Proceedings of the IEEE, 2019, 107(8): 1697-1716.
 - [39] MILLER C, VALASEK C. Adventures in automotive networks and control units[J]. Def Con, 2013, 21(260-264): 15-31.
 - [40] CHO K T, SHIN K G. Fingerprinting electronic control units for vehicle intrusion detection[C]//25th USENIX Security Symposium. 2016: 911-927.
 - [41] CHO K T, SHIN K G. Viden: attacker identification on in-vehicle networks[C]//Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2017: 1109-1123.

- [42] CHOI W, JO H J, WOO S, et al. Identifying ECUs using inimitable characteristics of signals in controller area networks[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(6): 4757-4770.
- [43] CHOI W, JOO K, JO H J, et al. VoltageIDS: low-level communication characteristics for automotive intrusion detection system[J]. *IEEE Transactions on Information Forensics and Security*, 2018, 13(8): 2114-2129.
- [44] GROZA B, MURVAY P S. Efficient intrusion detection with bloom filtering in controller area networks[J]. *IEEE Transactions on Information Forensics and Security*, 2019, 14(4): 1037-1051.
- [45] PIRAYESH H, SANGDEH P K, ZHANG S C, et al. JammingBird: jamming-resilient communications for vehicular ad hoc networks[C]// *Proceedings of the 2021 18th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. Piscataway: IEEE Press, 2021: 1-9.
- [46] OKYERE B, MUSAVIAN L, ÖZBEK B, et al. The resilience of massive MIMO PNC to jamming attacks in vehicular networks[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 22(7): 4110-4117.
- [47] 汪涵, 陈立全, 王忠民, 等. 基于深度学习的车联网无线密钥生成系统[J]. *网络与信息安全学报*, 2024, 10(1): 102-111.
WANG H, CHEN L Q, WANG Z M, et al. Wireless key generation system for Internet of vehicles based on deep learning[J]. *Chinese Journal of Network and Information Security*, 2024, 10(1): 102-111.
- [48] XIAO L, LU X Z, XU D J, et al. UAV relay in VANETs against smart jamming with reinforcement learning[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(5): 4087-4097.
- [49] BENSLIMANE A, NGUYEN-MINH H. Jamming attack model and detection method for beacons under multichannel operation in vehicular networks[J]. *IEEE Transactions on Vehicular Technology*, 2017, 66(7): 6475-6488.
- [50] LYAMIN N, KLEYKO D, DELOOZ Q, et al. Real-time jamming DoS detection in safety-critical V2V C-ITS using data mining[J]. *IEEE Communications Letters*, 2019, 23(3): 442-445.
- [51] MOKDAD L, BEN-OTHTMAN J, NGUYEN A T. DJAVAN: Detecting jamming attacks in vehicle ad hoc networks[J]. *Performance Evaluation*, 2015, 87: 47-59.
- [52] ALIPOUR-FANID A, DABAGHCHIAN M, ZENG K. Impact of jamming attacks on vehicular cooperative adaptive cruise control systems[J]. *IEEE Transactions on Vehicular Technology*, 2020, 69(11): 12679-12693.
- [53] PUÑAL O, AKTAŞ I, SCHNELKE C J, et al. Machine learning-based jamming detection for IEEE 802.11: design and experimental evaluation[C]// *Proceeding of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks 2014*. Piscataway: IEEE Press, 2014: 1-10.
- [54] PAWLAK J, LI Y C, PRICE J, et al. A machine learning approach for detecting and classifying jamming attacks against OFDM-based UAVs[C]// *Proceedings of the 3rd ACM Workshop on Wireless Security and Machine Learning*. New York: ACM, 2021: 1-6.
- [55] KARAGIANNIS D, ARGYRIOU A. Jamming attack detection in a pair of RF communicating vehicles using unsupervised machine learning[J]. *Vehicular Communications*, 2018, 13: 56-63.
- [56] XU Z S, ZHANG R R, HUANG P, et al. An aggregation protocol resisting collusion attacks in the Internet of vehicles environment[C]// *Proceedings of the 2023 IEEE 10th International Conference on Cyber Security and Cloud Computing (CSCloud)/2023 IEEE 9th International Conference on Edge Computing and Scalable Cloud (EdgeCom)*. Piscataway: IEEE Press, 2023: 24-29.
- [57] BUSI REDDY V, VENKATARAMAN S, NEGI A. Communication and data trust for wireless sensor networks using D-S theory[J]. *IEEE Sensors Journal*, 2017, 17(12): 3921-3929.
- [58] GAUTAM A K, KUMAR R. A robust trust model for wireless sensor networks[C]// *Proceedings of the 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*. Piscataway: IEEE Press, 2018: 1-5.
- [59] JAIN A, KUMAR S. FriendShare: a secure and reliable framework for file sharing on network[J]. *Journal of Network and Computer Applications*, 2018, 120: 1-16.
- [60] LI X H, LIU J, OBAIDAT M S, et al. An unlinkable authenticated key agreement with collusion resistant for VANETs[J]. *IEEE Transactions on Vehicular Technology*, 2021, 70(8): 7992-8006.
- [61] KANDAH F, SINGH Y, ZHANG W Y, et al. Mitigating colluding injected attack using monitoring verification in mobile ad-hoc networks[J]. *Security and Communication Networks*, 2013, 6(4): 539-547.
- [62] MARTI S, GIULI T J, LAI K, et al. Mitigating routing misbehavior in mobile ad hoc networks[C]// *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*. New York: ACM, 2000: 255-265.
- [63] JEDARI B, XIA F, CHEN H L, et al. A social-based watchdog system to detect selfish nodes in opportunistic mobile networks[J]. *Future Generation Computer Systems*, 2019, 92: 777-788.
- [64] HALABI T, ZULKERNINE M. Trust-based cooperative game model for secure collaboration in the Internet of vehicles[C]// *Proceedings of the ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*. Piscataway: IEEE Press, 2019: 1-6.
- [65] YANG J, NI Q H, LUO G Y, et al. A trustworthy Internet of vehicles: the DAO to safe, secure, and collaborative autonomous driving[J]. *IEEE Transactions on Intelligent Vehicles*, 2023, 8(12): 4678-4681.
- [66] FENG J Y, LIU N, CAO J, et al. Securing traffic-related messages exchange against inside-and-outside collusive attack in vehicular networks[J]. *IEEE Internet of Things Journal*, 2019, 6(6): 9979-9992.
- [67] SONG H M, KIM H R, KIM H K. Intrusion detection system based on the analysis of time intervals of CAN messages for in-vehicle network[C]// *Proceedings of the 2016 International Conference on Information Networking (ICOIN)*. Piscataway: IEEE Press, 2016: 63-68.
- [68] ABHISHEK N V, AMAN M N, LIM T J, et al. DRiVe: detecting malicious roadside units in the Internet of vehicles with low latency data integrity[J]. *IEEE Internet of Things Journal*, 2022, 9(5): 3270-3281.
- [69] ALI ALHEETI K M, MCDONALD-MAIER K. Hybrid intrusion

- detection in connected self-driving vehicles[C]//Proceedings of the 2016 22nd International Conference on Automation and Computing (ICAC). Piscataway: IEEE Press, 2016: 456-461.
- [70] LIANG J W, LIN Q Z, CHEN J Y, et al. A filter model based on hidden generalized mixture transition distribution model for intrusion detection system in vehicle ad hoc networks[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(7): 2707-2722.
- [71] PETRILLO A, PESCAPÉ A, SANTINI S. A collaborative control strategy for platoons of autonomous vehicles in the presence of message falsification attacks[C]//Proceedings of the 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS). Piscataway: IEEE Press, 2017: 110-115.
- [72] PETRILLO A, PESCAPÉ A, SANTINI S. A collaborative approach for improving the security of vehicular scenarios: The case of platooning[J]. Computer Communications, 2018, 122: 59-75.
- [73] SHARMA P, LIU H, WANG H G, et al. Securing wireless communications of connected vehicles with artificial intelligence[C]//Proceedings of the 2017 IEEE International Symposium on Technologies for Homeland Security (HST). Piscataway: IEEE Press, 2017: 1-7.
- [74] SATCHIDANANDAN B, KUMAR P R. Dynamic watermarking: active defense of networked cyber - physical systems[J]. Proceedings of the IEEE, 2017, 105(2): 219-240.
- [75] PORTER M, HESPAHOL P, ASWANI A, et al. Detecting generalized replay attacks via time-varying dynamic watermarking[J]. IEEE Transactions on Automatic Control, 2021, 66(8): 3502-3517.
- [76] NAHA A, TEIXEIRA A, AHLÉN A, et al. Sequential detection of replay attacks[J]. IEEE Transactions on Automatic Control, 2023, 68(3): 1941-1948.
- [77] KUMAR A, SHAHID M A, JAEKEL A, et al. Machine learning based detection of replay attacks in VANET[C]//Proceedings of the NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium. Piscataway: IEEE Press, 2023: 1-6.
- [78] BRECHT B, THERRIAULT D, WEIMERSKIRCHA, et al. A security credential management system for V2X communications[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(12): 3850-3871.
- [79] WANG Y Y, MASOUD N, KHOJANDI A. Real-time sensor anomaly detection and recovery in connected automated vehicle sensors[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(3): 1411-1421.
- [80] HUANG S E, FENG Y H, LIU H X. A data-driven method for falsified vehicle trajectory identification by anomaly detection[J]. Transportation Research Part C: Emerging Technologies, 2021, 128: 103196.
- [81] HOANG T M, NGUYEN N M, DUONG T Q. Detection of eavesdropping attack in UAV-aided wireless systems: unsupervised learning with one-class SVM and K-means clustering[J]. IEEE Wireless Communications Letters, 2020, 9(2): 139-142.
- [82] WANG Y Y, MASOUD N, KHOJANDI A. Real-time sensor anomaly detection and recovery in connected automated vehicle sensors[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(3): 1411-1421.
- [83] WANG Y Z, LIU Q P, MIHANKHAH E, et al. Detection and isolation of sensor attacks for autonomous vehicles: framework, algorithms, and validation[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(7): 8247-8259.
- [84] FANG Y, MIN H, WU X, et al. Anomaly diagnosis of connected autonomous vehicles: A survey[J]. Information Fusion, 2024, 105: 102223.
- [85] MATHUR S, TRAPPE W, MANDAYAM N, et al. Radio-telepathy: extracting a secret key from an unauthenticated wireless channel[C]//Proceedings of the 14th ACM International Conference on Mobile Computing and Networking. New York: ACM, 2008: 128-139.
- [86] ZENG K, WU D, CHAN A, et al. Exploiting multiple-antenna diversity for shared secret key generation in wireless networks[C]//2010 Proceedings IEEE INFOCOM. Piscataway: IEEE Press, 2010: 1-9.
- [87] KARIMI R, KALANTARI M. Enhancing security and confidentiality in location-based data encryption algorithms[C]//Proceedings of the Fourth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2011). Piscataway: IEEE Press, 2011: 30-35.
- [88] KIDO H, YANAGISAWA Y, SATOH T. An anonymous communication technique using dummies for location-based services[C]//Proceedings. International Conference on Pervasive Services, 2005. Piscataway: IEEE Press, 2005: 88-97.
- [89] NIU B, LI Q H, ZHU X Y, et al. Achieving k-anonymity in privacy-aware location-based services[C]//Proceedings of the IEEE INFOCOM 2014 - IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2014: 754-762.
- [90] CHEN Y, TRAPPE W, MARTIN R P. Attack detection in wireless localization[C]//Proceedings of the IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications. Piscataway: IEEE Press, 2007: 1964-1972.
- [91] JOVER R P. LTE security and protocol exploits[C]//Proceeding of Shmoocon 2016. Washington, DC: Shmoo Group, 2016: 1-15.
- [92] BERESFORD A R, STAJANO F. Location privacy in pervasive computing[J]. IEEE Pervasive Computing, 2003, 2(1): 46-55.
- [93] HO S S, RUAN S H. Differential privacy for location pattern mining[C]//Proceedings of the 4th ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS. New York: ACM, 2011: 17-24.
- [94] DEWRI R. Local differential perturbations: location privacy under approximate knowledge attackers[J]. IEEE Transactions on Mobile Computing, 2013, 12(12): 2360-2372.
- [95] ZHAO P, JIANG H B, LUI J C S, et al. P3-LOC: a privacy-preserving paradigm-driven framework for indoor localization[J]. IEEE/ACM Transactions on Networking, 2018, 26(6): 2856-2869.
- [96] ALICHERY M, KEROMYTIS A D. DoubleCheck: Multi-path verification against man-in-the-middle attacks[C]//Proceedings of the 2009 IEEE Symposium on Computers and Communications. Piscataway: IEEE Press, 2009: 557-563.
- [97] ABADI M, BIRRELL A, MIRONOV I, et al. Global authentication in an untrustworthy world[C]//14th Workshop on Hot Topics in Op-

- erating Systems (HotOS XIV). 2013.
- [98] ALI M, NELSON J, SHEA R, et al. Blockstack: a global naming and storage system secured by blockchains[C]//2016 USENIX Annual Technical Conference (USENIX ATC 16). 2016: 181-194.
- [99] RYAN M D. Enhanced certificate transparency and end-to-end encrypted mail[C]//Proceedings 2014 Network and Distributed System Security Symposium. Internet Society, 2014.
- [100] LI B Y, LIN J Q, LI F J, et al. Certificate transparency in the wild: exploring the reliability of monitors[C]//Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2019: 2505-2520.
- [101] WANG Z, LIN J Q, CAI Q W, et al. Blockchain-based certificate transparency and revocation transparency[C]//Financial Cryptography and Data Security. Berlin, Heidelberg: Springer, 2019: 144-162.
- [102] DOUCEUR J R. The sybil attack[C]// the 1st International Workshop on Peer-to-Peer Systems (IPTPS'02), Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 2022, 2429: 51-260.
- [103] NEWSOME J, SHI E, SONG D, et al. The sybil attack in sensor networks: analysis & defenses[C]//Proceedings of the Third International Symposium on Information Processing in Sensor Networks, 2004. Piscataway: IEEE Press, 2004: 259-268.
- [104] RAYA M, PAPADIMITRATOS P, HUBAUX J. Securing vehicular communication [J]. IEEE Wireless Communications, 2006, 13 (5): 8-15.
- [105] PARK S, ASLAM B, TURGUT D, et al. Defense against Sybil attack in vehicular ad hoc network based on roadside unit support[C]// Proceedings of the MILCOM 2009 - 2009 IEEE Military Communications Conference. Piscataway: IEEE Press, 2009: 1-7.
- [106] CHANG S, QI Y, ZHU H Z, et al. Footprint: detecting sybil attacks in urban vehicular networks[J]. IEEE Transactions on Parallel and Distributed Systems, 2012, 23(6): 1103-1114.
- [107] DEMIRBAS M, SONG Y. An RSSI-based scheme for sybil attack detection in wireless sensor networks[C]//Proceedings of the 2006 International Symposium on a World of Wireless, Mobile and Multimedia Networks(WoWMoM'06). Piscataway: IEEE Press, 2006: 1-5.
- [108] LV S H, WANG X D, ZHAO X, et al. Detecting the sybil attack cooperatively in wireless sensor networks[C]//Proceedings of the 2008 International Conference on Computational Intelligence and Security. Piscataway: IEEE Press, 2008: 442-446.
- [109] BOUASSIDA M S, GUETTE G, SHAWKY M, et al. Sybil nodes detection based on received signal strength variations within VANET[J]. International Journal of Network Security, 2009, 9(1): 22-33.
- [110] SHARMA S, KAUL A. A survey on Intrusion Detection Systems and Honeypot based proactive security mechanisms in VANETs and VANET Cloud[J]. Vehicular Communications, 2018, 12: 138-164.
- [111] KHATRI S, VACHHANI H, SHAH S, et al. Machine learning models and techniques for VANET based traffic management: Implementation issues and challenges[J]. Peer-to-Peer Networking and Applications, 2021, 14(3): 1778-1805.
- [112] CHAUHAN S S, JAIN N, PANDEY S C. Digital signature with message security process[C]//Proceedings of the 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). Piscataway: IEEE Press, 2022: 182-187.
- [113] HEGDE N, MANVI S S. Distributed integrity and non-repudiation scheme in the dynamic vehicular cloud environment[J]. International Journal of Information and Computer Security, 2023, 20(3/4): 315.
- [114] ARM J, FIEDLER P, BASTAN O. Offline access to a vehicle via PKI-based authentication[M]//Computer Safety, Reliability, and Security. SAFECOMP 2021 Workshops. Cham: Springer International Publishing, 2021: 76-88.
- [115] CHEN F, WANG J H, LI J Q, et al. TrustBuilder: a non-repudiation scheme for IoT cloud applications[J]. Computers & Security, 2022, 116: 102664.
- [116] CHEN F, WANG J H, JIANG C K, et al. Blockchain based non-repudiable IoT data trading: simpler, faster, and cheaper[C]//Proceedings of the IEEE INFOCOM 2022 - IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2022: 1958-1967.
- [117] MOUSSAOUI D, KADRI B, FEHAM M, et al. A Distributed Blockchain Based PKI (BCPKI) architecture to enhance privacy in VANET[C]//Proceedings of the 2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH). Piscataway: IEEE Press, 2021: 75-79.
- [118] CHOW S, EISEN P, JOHNSON H, et al. A white-box DES implementation for DRM applications[M]//Digital Rights Management. Berlin: Springer, 2003: 1-15.
- [119] BILLET O, GILBERT H, ECH-CHATBI C. Cryptanalysis of a white box AES implementation[C]//Selected Areas in Cryptography. Berlin: Springer, 2004: 227-240.
- [120] BIRYUKOV A, BOUILLAGUET C, KHOVRATOVICH D. Cryptographic schemes based on the ASASA structure: black-box, white-box, and public-key (extended abstract) [M]//Advances in Cryptology-ASIACRYPT 2014. Berlin: Springer, 2014: 63-84.
- [121] BOGDANOV A, ISOBE T. White-box cryptography revisited: space-hard ciphers[C]//Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2015: 1058-1069.
- [122] BOGDANOV A, ISOBE T, TISCHHAUSER E. Towards practical whitebox cryptography: optimizing efficiency and space hardness[M]//Advances in Cryptology-ASIACRYPT 2016. Berlin: Springer, 2016: 126-158.
- [123] WANG J Y, ZHANG H C. Bilateral adversarial training: towards fast training of more robust models against adversarial attacks[C]// Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2019: 6628-6637.
- [124] NAZEMI A, FIEGUTH P. Potential adversarial samples for white-box attacks[EB]. 2019.

- [125] WANG Y X, LIU J Q, CHANG X L, et al. DI-AA: an interpretable white-box attack for fooling deep neural networks[J]. *Information Sciences*, 2022, 610: 14-32.
- [126] LIU H Y, GE Z J, ZHOU Z Y, et al. Gradient correction for white-box adversarial attacks[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 35(12): 18419-18430.
- [127] KAYA Y, HONG S, DUMITRAS T. On the effectiveness of regularization against membership inference attacks[EB]. 2020.
- [128] LI T J, LI M S, YANG Y H, et al. Frequency domain regularization for iterative adversarial attacks[J]. *Pattern Recognition*, 2023, 134: 109075.
- [129] 易杰, 曹腾飞, 郜帅, 等. 车联网中基于攻防博弈的蜜罐防御及传输策略[J]. *网络与信息安全学报*, 2022, 8(4): 157-167.
- YI J, CAO T F, GAO S, et al. Honeypot defense and transmission strategy based on offensive and defensive games in vehicular networks[J]. *Chinese Journal of Network and Information Security*, 2022, 8(4): 157-167.
- [130] STEINHARDT J, KOH P W, LIANG P. Certified defenses for data poisoning attacks[EB]. 2017.
- [131] SHAFABI A, HUANG W R, NAJIBI M, et al. Poison frogs targeted clean-label poisoning attacks on neural networks[EB]. 2018.
- [132] CHACON H, SILVA S, RAD P. Deep learning poison data attack detection[C]//*Proceedings of the 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*. Piscataway: IEEE Press, 2019: 971-978.
- [133] ROSENFELD E, WINSTON E, RAVIKUMAR P, et al. Certified robustness to label-flipping attacks via randomized smoothing[C]//*International Conference on Machine Learning*. PMLR, 2020: 8230-8241.
- [134] TALPUR A, GURUSAMY M. GFCL: a GRU-based federated continual learning framework against data poisoning attacks in IoV[EB]. 2022.
- [135] LI X, HUANG K X, YANG W H, et al. On the convergence of FedAvg on non-IID data[EB]. 2019.
- [136] AVENT B, DUBEY Y, KOROLOVA A. The power of the hybrid model for mean estimation[EB]. 2018.
- [137] DWORK C, ROTHBLUM G N, VADHAN S. Boosting and differential privacy[C]//*Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science*. Piscataway: IEEE Press, 2010: 51-60.
- [138] ZHOU H, YANG G, DAI H, et al. PFLF: privacy-preserving federated learning framework for edge computing[J]. *IEEE Transactions on Information Forensics and Security*, 2022, 17: 1905-1918.
- [139] WANG N Y, YANG W T, WANG X D, et al. A blockchain based privacy-preserving federated learning scheme for Internet of Vehicles[J]. *Digital Communications and Networks*, 2024, 10(1): 126-134.
- [140] 王瑞锦, 唐榆程, 张巍琦, 等. 基于同态加密和区块链技术的车联网隐私保护方案[J]. *网络与信息安全学报*, 2020, 6(1): 46-53.
- WANG R J, TANG Y C, ZHANG W Q, et al. Privacy protection scheme for Internet of vehicles based on homomorphic encryption and block chain technology[J]. *Chinese Journal of Network and Information Security*, 2020, 6(1): 46-53.
- [141] CHANG Y S, ZHANG K, GONG J Q, et al. Privacy-preserving federated learning via functional encryption, revisited[J]. *IEEE Transactions on Information Forensics and Security*, 2023, 18: 1855-1869.
- [142] KANAGAVELU R, LI Z X, SAMSUDIN J, et al. Two-phase multi-party computation enabled privacy-preserving federated learning[C]//*Proceedings of the 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID)*. Piscataway: IEEE Press, 2020: 410-419.
- [143] LAURENDEAU C, BARBEAU M. Threats to security in DSRC/WAVE[C]//*Ad-Hoc, Mobile, and Wireless Networks: 5th International Conference*. Ottawa, Canada, 2006, 4104: 266-279.
- [144] MALEKI M, MALIK M, FOLKESSON P, et al. Modeling and evaluating the effects of jamming attacks on connected automated road vehicles[C]//*Proceedings of the 2022 IEEE 27th Pacific Rim International Symposium on Dependable Computing (PRDC)*. Piscataway: IEEE Press, 2022: 12-23.
- [145] MOSER D, LENDERS V, CAPKUN S. Digital radio signal cancellation attacks: an experimental evaluation[C]//*Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks*. New York: ACM, 2019: 23-33.
- [146] LICHTMAN M, POSTON J D, AMURU S, et al. A communications jamming taxonomy[J]. *IEEE Security & Privacy*, 2016, 14(1): 47-54.
- [147] SIKARWAR H, DAS D. A novel MAC-based authentication scheme (NoMAS) for Internet of vehicles (IoV)[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023, 24(5): 4904-4916.
- [148] JEONG S, PARK E, SEO K U, et al. MUVIDS: false MAVLink injection attack detection in communication for unmanned vehicles[C]//*Proceedings Third International Workshop on Automotive and Autonomous Vehicle Security*. Internet Society, 2021.
- [149] PAN D F, DING D R, GE X H, et al. Edge-based event-triggered platooning control of heterogeneous connected automated vehicles under replay attacks[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(11): 17443-17456.
- [150] PENG Z F, PENG C G, TIAN Y L, et al. Bilateral control for secure communication against replay attack in ORAN-based vehicular networks[J]. *IEEE Transactions on Vehicular Technology*, 2025, 74(5): 7565-7579.
- [151] SHANGGUAN L T, CHOUR K, KO W H, et al. Dynamic watermarking for cybersecurity of autonomous vehicles[J]. *IEEE Transactions on Industrial Electronics*, 2023, 70(11): 11735-11743.
- [152] FANG C R, QI Y F, CHENG P, et al. Optimal periodic watermarking schedule for replay attack detection in cyber-physical systems[J]. *Automatica*, 2020, 112: 108698.
- [153] DU D J, ZHANG C D, LI X, et al. Attack detection for networked control systems using event-triggered dynamic watermarking[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(1): 351-361.

- [154] 陈明奇, 钮心忻, 杨义先. 数字水印的研究进展和应用[J]. 通信学报, 2001, 22(5): 71-79.
CHEN M Q, NIU X X, YANG Y X. The research developments and applications of digital watermarking[J]. Journal of China Institute of Communications, 2001, 22(5): 71-79.
- [155] PETIT J, SHLADOVER S E. Potential cyberattacks on automated vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(2): 546-556.
- [156] SCOTT L. Anti-spoofing & authenticated signal architectures for civil navigation systems[C]//Proceedings of the 16th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS/GNSS 2003). 2003: 1543-1552.
- [157] KUHN M G. An asymmetric security mechanism for navigation signals[M]. Information Hiding. Berlin: Springer, 2004: 239-252.
- [158] PAPADIMITRATOS P, JOVANOVIĆ A. GNSS-based positioning: attacks and countermeasures[C]//Proceedings of the MILCOM 2008 IEEE Military Communications Conference. Piscataway: IEEE Press, 2008: 1-7.
- [159] YIN S, ZHU X P, JING C. Fault detection based on a robust one class support vector machine[J]. Neurocomputing, 2014, 145: 263-268.
- [160] SHEN J J, WON J Y, CHEN Z Y, et al. Drift with devil: Security of multi-sensor fusion based localization in high-level autonomous driving under GPS spoofing[C]//29th USENIX Security Symposium (USENIX Security 20). 2020: 931-948.
- [161] WANG F L, HONG Y, BAN X G. Infrastructure-enabled GPS spoofing detection and correction[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(12): 13878-13892.
- [162] VAN WYK F, WANG Y Y, KHOJANDI A, et al. Real-time sensor anomaly detection and identification in automated vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(3): 1264-1276.
- [163] MOKHTAR B, AZAB M. Survey on security issues in vehicular ad hoc networks[J]. Alexandria Engineering Journal, 2015, 54(4): 1115-1126.
- [164] HOLTMANN S, RAO S P, OLIVER I. User location tracking attacks for LTE networks using the interworking functionality[C]//Proceedings of the 2016 IFIP Networking Conference (IFIP Networking) and Workshops. Piscataway: IEEE Press, 2016: 315-322.
- [165] SEDDIGH N, NANDY B, MAKKAR R, et al. Security advances and challenges in 4G wireless networks[C]//Proceedings of the 2010 Eighth International Conference on Privacy, Security and Trust. Piscataway: IEEE Press, 2010: 62-71.
- [166] CHEN X H, MIZERA A, PANG J. Activity tracking: a new attack on location privacy[C]//Proceedings of the 2015 IEEE Conference on Communications and Network Security (CNS). Piscataway: IEEE Press, 2015: 22-30.
- [167] GIVEHCHIAN H, BHASKAR N, HERRERA E R, et al. Evaluating physical-layer BLE location tracking attacks on mobile devices[C]//Proceedings of the 2022 IEEE Symposium on Security and Privacy (SP). Piscataway: IEEE Press, 2022: 1690-1704.
- [168] LI R H, HU H B, YE Q Q. RFTrack: stealthy location inference and tracking attack on Wi-Fi devices[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 5925-5939.
- [169] DE ARAUJO-FILHO P F, NAILI M, KADDOUM G, et al. Unsupervised GAN-based intrusion detection system using temporal convolutional networks and self-attention[J]. IEEE Transactions on Network and Service Management, 2023, 20(4): 4951-4963.
- [170] JIANG H B, LI J, ZHAO P, et al. Location privacy-preserving mechanisms in location-based services[J]. ACM Computing Surveys, 2022, 54(1): 1-36.
- [171] SUN K K, XU H S, HUA K, et al. Joint top-k sparsification and shuffle model for communication-privacy-accuracy tradeoffs in federated-learning-based IoV[J]. IEEE Internet of Things Journal, 2024, 11(11): 19721-19735.
- [172] 梁应敞, 谭俊杰, NIYATO D. 智能无线通信技术研究概况[J]. 通信学报, 2020, 41(7): 1-17.
LIANG Y C, TAN J J, NIYATO D. Overview on intelligent wireless communication technology[J]. Journal on Communications, 2020, 41(7): 1-17.
- [173] YAO Y, XIAO B, WU G F, et al. Multi-channel based sybil attack detection in vehicular ad hoc networks using RSSI[J]. IEEE Transactions on Mobile Computing, 2019, 18(2): 362-375.
- [174] TASLIMASA H, DADKHAH S, NETO E C P, et al. Security issues in Internet of vehicles (IoV): a comprehensive survey[J]. Internet of Things, 2023, 22: 100809.
- [175] ALIMOHAMMADI M, POUYAN A A. Sybil attack detection using a low cost short group signature in VANET[C]//Proceedings of the 2015 12th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC). Piscataway: IEEE Press, 2015: 23-28.
- [176] GHOSAL A, CONTI M. Security issues and challenges in V2X: a survey[J]. Computer Networks, 2020, 169: 107093.
- [177] KERRACHE C A, CALAFATE C T, CANO J C, et al. Trust management for vehicular networks: an adversary-oriented overview[J]. IEEE Access, 2016, 4: 9293-9307.
- [178] 翟宝琴, 王健, 韩磊, 等. 基于信任值的车联网分层共识优化协议[J]. 网络与信息安全学报, 2022, 8(3): 142-153.
ZHAI B Q, WANG J, HAN L, et al. Hierarchical proxy consensus optimization for IoV based on blockchain and trust value[J]. Chinese Journal of Network and Information Security, 2022, 8(3): 142-153.
- [179] YAN C G, XIE H T, YANG D B, et al. Supervised hash coding with deep neural network for environment perception of intelligent vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(1): 284-295.
- [180] RAMOS S, GEHRIG S, PINGGERA P, et al. Detecting unexpected obstacles for self-driving cars: Fusing deep learning and geometric modeling[C]//Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV). Piscataway: IEEE Press, 2017: 1025-1032.
- [181] KATRAKAZAS C, QUDDUS M, CHEN W H, et al. Real-time motion planning methods for autonomous on-road driving: State-of-the-art and future research directions[J]. Transportation Research Part C: Emerging Technologies, 2015, 60: 416-442.

- [182] ZHANG Y, SUN P, YIN Y H, et al. Human-like autonomous vehicle speed control by deep reinforcement learning with double Q-learning[C]//Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV). Piscataway: IEEE Press, 2018: 1251-1256.
- [183] MILANÉS V, SHLADOVER S E, SPRING J, et al. Cooperative adaptive cruise control in real traffic situations[J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15(1): 296-305.
- [184] BALUJA S, FISCHER I. Learning to attack: adversarial transformation networks[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1).
- [185] MOUKAHAL L J, ZULKERNINE M, SOUKUP M. Boosting grey-box fuzzing for connected autonomous vehicle systems[C]//Proceedings of the 2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion (QRS-C). Piscataway: IEEE Press, 2021: 516-527.
- [186] SU J W, VARGAS D V, SAKURAI K. One pixel attack for fooling deep neural networks[J]. IEEE Transactions on Evolutionary Computation, 2019, 23(5): 828-841.
- [187] XU D L, LI T, LI Y, et al. Edge intelligence: empowering intelligence to the edge of network[J]. Proceedings of the IEEE, 2021, 109(11): 1778-1837.
- [188] LANGER M, HE Z, RAHAYU W, et al. Distributed training of deep learning models: a taxonomic perspective[J]. IEEE Transactions on Parallel and Distributed Systems, 2020, 31(12): 2802-2818.
- [189] VERBRAEKEN J, WOLTING M, KATZY J, et al. A survey on distributed machine learning[J]. ACM Computing Surveys, 2021, 53(2): 1-33.
- [190] CHEN X Z, QIU W C, CHEN L X, et al. Fast and practical intrusion detection system based on federated learning for VANET[J]. Computers & Security, 2024, 142: 103881.
- [191] KUMAR K N, VISHNU C, MITRA R, et al. Black-box adversarial attacks in autonomous vehicle technology[C]//Proceedings of the 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR). Piscataway: IEEE Press, 2020: 1-7.
- [192] WON J, SEO S H, BERTINO E. A secure shuffling mechanism for white-box attack-resistant unmanned vehicles[J]. IEEE Transactions on Mobile Computing, 2020, 19(5): 1023-1039.
- [193] SARKER A, SHEN H Y, SEN T, et al. An advanced black-box adversarial attack for deep driving maneuver classification models[C]//Proceedings of the 2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS). Piscataway: IEEE Press, 2020: 184-192.
- [194] SUN J C, CAO Y L, CHEN Q A, et al. Towards robust LiDAR-based perception in autonomous driving: General black-box adversarial sensor attack and countermeasures[C]//29th USENIX Security Symposium (USENIX Security 20). 2020: 877-894.
- [195] DENG Y, ZHENG X, ZHANG T Y, et al. An analysis of adversarial attacks and defenses on autonomous driving models[C]//Proceedings of the 2020 IEEE International Conference on Pervasive Computing and Communications (PerCom). Piscataway: IEEE Press, 2020: 1-10.
- [196] XIONG Z B, XU H H, LI W, et al. Multi-source adversarial sample attack on autonomous vehicles[J]. IEEE Transactions on Vehicular Technology, 2021, 70(3): 2822-2835.
- [197] LI Y, LIN J, XIONG K Q. An adversarial attack defending system for securing in-vehicle networks[C]//Proceedings of the 2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC). Piscataway: IEEE Press, 2021: 1-6.
- [198] XU X, ZHANG J R, LI Y J, et al. Adversarial attack against urban scene segmentation for autonomous vehicles[J]. IEEE Transactions on Industrial Informatics, 2021, 17(6): 4117-4126.
- [199] SARKER A, SHEN H Y, SEN T. A context-aware black-box adversarial attack for deep driving maneuver classification models[C]//Proceedings of the 2021 18th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON). Piscataway: IEEE Press, 2021: 1-9.
- [200] JIA S, SONG Y B, MA C, et al. IoU attack: towards temporally coherent black-box adversarial attack for visual object tracking[C]//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2021: 6705-6714.
- [201] SUN C, ZHANG Y, CHAOQUN W, et al. Towards lightweight black-box attack against deep neural networks[J]. Advances in Neural Information Processing Systems, 2022, 35: 19319-19331.
- [202] JIA W, LU Z J, YU R Z, et al. Fooling decision-based black-box automotive vision perception systems in physical world[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(7): 7081-7092.
- [203] BIGGIO B, NELSON B, LASKOV P. Poisoning attacks against support vector machines[C]//Proceedings of the 29th International Conference on Machine Learning. 2012: 1807-1814.
- [204] BALAKRISHNAN A, PURANIC A G, QIN X, et al. Specifying and evaluating quality metrics for vision-based perception systems[C]//Proceedings of the 2019 Design, Automation & Test in Europe Conference & Exhibition (DATE). Piscataway: IEEE Press, 2019: 1433-1438.
- [205] MUÑOZ-GONZÁLEZ L, BIGGIO B, DEMONTIS A, et al. Towards poisoning of deep learning algorithms with back-gradient optimization[C]//Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security. New York: ACM, 2017: 27-38.
- [206] KOH P W, LIANG P. Understanding black-box predictions via influence functions[C]//International conference on machine learning. 2017: 1885-1894..
- [207] HAN X, HUANG X, CLAUDIA E. Adversarial label flips attack on support vector machines[C]//ECAI 2012. IOS Press, 2012.
- [208] SITAWARIN C, BHAGOJI A N, MOSENIA A, et al. DARTS: deceiving autonomous cars with toxic signs[EB]. 2024.
- [209] CUI C, DU H P, JIA Z J, et al. Data poisoning attack using hybrid particle swarm optimization in connected and autonomous vehicles[C]//Proceedings of the 2022 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE). Piscataway: IEEE Press, 2022: 1-5.
- [210] ZHU C, HUANG W R, LI H, et al. Transferable clean-label poisoning attacks on deep neural nets[C]//International Conference on Machine Learning. 2019: 7614-7623.

- [211] LI J Y, GUO W Z, XIE L H, et al. Privacy-preserving object detection with poisoning recognition for autonomous vehicles[J]. *IEEE Transactions on Network Science and Engineering*, 2023, 10(3): 1487-1500.
- [212] YIN D, CHEN Y D, RAMCHANDRAN K, et al. Byzantine-robust distributed learning: Towards optimal statistical rates[C]// *International Conference on Machine Learning*. 2018: 5650-5659.
- [213] BLANCHARD P, EL MHAMDI E M, GUERRAOUI R, et al. Machine learning with adversaries: byzantine tolerant gradient descent[J]. *Advances in Neural Information Processing Systems*, 2017, 30.
- [214] CHANG K, BALACHANDAR N, LAM C, et al. Distributed deep learning networks among institutions for medical imaging open access[J]. *Journal of the American Medical Informatics Association*, 2018, 25(8): 945-954.
- [215] GUO J, SONG B, CHEN S Q, et al. Context-aware object detection for vehicular networks based on edge-cloud cooperation[J]. *IEEE Internet of Things Journal*, 2020, 7(7): 5783-5791.
- [216] SUN K K, WU J, LI J H. Reputation-aware incentive mechanism of federated learning: a mean field game approach[C]// *Proceedings of the 2024 9th IEEE International Conference on Smart Cloud (SmartCloud)*. Piscataway: IEEE Press, 2024: 48-53.
- [217] SUN K K, XU H S, ZHANG X J, et al. Time-sensitive local differential privacy-based federated learning for vehicular digital twin networks[C]// *Proceedings of the International Symposium on Intelligent Computing and Networking 2024*. Cham: Springer Nature Switzerland, 2024: 105-118.
- [218] RAZZAQ W, ARIF U, DIN Z M U. Visual perception deep drive model for self-driving car[J]. *Pakistan Journal of Scientific Research*, 2022, 1(1): 18-21.
- [219] XIONG J B, BI R W, ZHAO M F, et al. Edge-assisted privacy-preserving raw data sharing framework for connected autonomous vehicles[J]. *IEEE Wireless Communications*, 2020, 27(3): 24-30.
- [220] ZHANG Q C, YANG L T, CHEN Z K. Privacy preserving deep computation model on cloud for big data feature learning[J]. *IEEE Transactions on Computers*, 2016, 65(5): 1351-1362.
- [221] SUN K K, LIU L Z, PAN Q Q, et al. Large-scale mean-field federated learning for detection and defense: a Byzantine robustness approach in IoT[J]. *IEEE Internet of Things Journal*, 2024, 11(22): 36370-36383.
- [222] SUN L C, QIAN J W, CHEN X. LDP-FL: Practical private aggregation in federated learning with local differential privacy[C]// *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*. 2021: 1571-1578.
- [223] 陈秀真, 吴越, 李建华. 车载信息系统的安全测评体系及方法[J]. *信息安全学报*, 2017, 2(2): 15-23.
- CHEN X Z, WU Y, LI J H. Security evaluation system and method of in-vehicle information system[J]. *Journal of Information Security*, 2017, 2(2): 15-23.
- [224] SUN K K, WU J, BASHIR A K, et al. Personalized privacy-preserving distributed artificial intelligence for digital-twin-driven vehicle road cooperation[J]. *IEEE Internet of Things Journal*, 2024, 11(22): 35902-35916.
- [225] LIANG F Y, YANG Q L, LIU R Q, et al. Semi-synchronous federated learning protocol with dynamic aggregation in Internet of vehicles[J]. *IEEE Transactions on Vehicular Technology*, 2022, 71(5): 4677-4691.
- [226] PAN D F, DING D R, GE X H, et al. Edge-based event-triggered platooning control of heterogeneous connected automated vehicles under replay attacks[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(11): 17443-17456.
- [227] KHAN N, COLERI S. Event-triggered reinforcement learning based joint resource allocation for ultra-reliable low-latency V2X communications[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(11): 16991-17006.
- [228] SUN K K, WU J, PAN Q Q, et al. Leveraging digital twin and DRL for collaborative context offloading in C-V2X autonomous driving[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(4): 5020-5035.
- [229] LI Y J, TAO X F, ZHANG X F, et al. Privacy-preserved federated learning for autonomous driving[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(7): 8423-8434.
- [230] PENG T, ZHONG W T, WANG G J, et al. Privacy-preserving truth discovery based on secure multi-party computation in vehicle-based mobile crowdsensing[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2024, 25(7): 7767-7779.
- [231] GABAY D, AKKAYA K, CEBE M. Privacy-preserving authentication scheme for connected electric vehicles using blockchain and zero knowledge proofs[J]. *IEEE Transactions on Vehicular Technology*, 2020, 69(6): 5760-5772.
- [232] ZHAO Y, ZHAO J, YANG M M, et al. Local differential privacy-based federated learning for Internet of things[J]. *IEEE Internet of Things Journal*, 2021, 8(11): 8836-8853.
- [233] HOU S, CHEN X Z, MA J, et al. An ontology-based dynamic attack graph generation approach for the internet of vehicles[J]. *Frontiers in energy research* 2022, 10: 928919.
- [234] QIU W, MA Y, CHEN X Z, et al. Hybrid intrusion detection system based on Dempster-Shafer evidence theory[J]. *Computers & Security* 2022, 117: 102709.
- [235] 于海洋, 陈秀真, 马进, 等. 面向智能汽车的信息安全漏洞评分模型[J]. *网络与信息安全学报*, 2022, 8(1): 167-179.
- YU H Y, CHEN X Z, MA J, et al. Information security vulnerability scoring model for intelligent vehicles[J]. *Journal of Network and Information Security*, 2022, 8(1): 167-179.
- [236] XU C, ZHENG F, GUO G. Stealthy false data injection attacks in multi-channel vehicular communication: white-box and gray-box strategies[J]. *IEEE Transactions on Vehicular Technology*, 2025.

[作者简介]



孙康康 (1993-), 男, 山东滨州人, 博士, 上海交通大学助理研究员, 主要研究方向为车联网及物联网安全、隐私保护、博弈论、分布式机器学习安全。



陈秀真 (1977-), 女, 山东聊城人, 博士, 上海交通大学副教授、博士生导师, 主要研究方向为社交网络数据分析、车联网信息安全、安全检测与评估。



李建华 (1965-), 男, 江西九江人, 博士, 上海交通大学教授、博士生导师, 主要研究方向为信息内容安全技术、态势感知、网络安全。



过敏意 (1963-), 男, 江苏无锡人, 博士, 上海交通大学教授、博士生导师, 主要研究方向为并行与分布式处理、云计算。