综述

边缘算力发展态势分析

郭宇辉[1]，闫亚旗[1]，付韬[2]，张阔[1]，冉沛[1]，王哲[2]，刘文睿[1]
（1. 中国铁塔股份有限公司，北京 100089；
2. 中国信息通信研究院，北京 100191）

摘　要：边缘算力就近提供计算、智能等核心能力，已逐步成为垂直行业实现数字化、智能化的重要基础设施。我国高度重视边缘算力技术创新、产业发展和应用探索。分析了边缘算力整体发展态势，具体阐述了边缘算力概念及特征，梳理了边缘算力技术体系，总结了边缘算力典型应用场景，并提出边缘算力发展挑战及建议，为推进边缘算力技术产业、基础设施建设及应用发展提供参考。

关键词：边缘算力；云边协同；工业互联网；共建共享
中图分类号：TP393
文献标志码：A
doi: 10.11959/j.issn.1000−0801.2025195

Analysis on development trends of edge computing power

GUO Yuhui[1], YAN Yaqi[1], FU Tao[2], ZHANG Kuo[1], RAN Pei[1], WANG Zhe[2], LIU Wenrui[1]
1. China Tower Co., Ltd., Beijing 100089, China
2. China Academy of Information and Communications Technology, Beijing 100191, China

Abstract: Edge computing power, which delivers core capabilities such as computing and intelligence in proximity, has progressively evolved into critical infrastructure for vertical industries to achieve digital transformation and intelligentization. China has been placing high strategic emphasis on advancing technological innovation, fostering industrial development, and exploring practical applications in the field of edge computing power. A systematic analysis on the holistic development landscape of edge computing power was conducted. The conceptual definitions and distinctive characteristics of edge computing power were elucidated, its technical architecture was systematically organized, prototypical application scenarios were summarized, and developmental challenges alongside strategic recommendations were proposed. These research outcomes provides critical references for advancing technological innovation, infrastructure deployment, and application ecosystems in the domain of edge computing power.

Key words: edge computing power, cloud-edge synergy, industrial Internet, co-construction and co-sharing

0　引言

Machina Research研究报告显示，2025年全球物联网连接数将增至270亿个，同时由联网设备产生的数据达到847 ZB，给传统的集中式数据处理及网络传输带来挑战[1]。此外，随着AR/

收稿日期：2025−05−21；修回日期：2025−08−14

VR、云游戏、智能控制等新业务的蓬勃发展，大量低时延响应需求的出现也需要边缘算力完成数据的就近处理和分析，以满足用户的实时性要求。

当前，业界对边缘算力技术架构的研究主要集中于单一技术的突破。本文的创新之处在于构建涵盖边缘智能、边缘数据与算力网络协同的综合技术体系，通过梳理边缘训练、数据脱敏和边缘算网感知等关键技术模块，提出了一个层次分明的理论框架，为边缘算力在异构场景下的技术研究提供了新的理论支持和实践指导。

1 边缘算力整体态势

传统上，算力的部署主要依赖于数据中心等基础设施，而这些设施的地理位置需要经过慎重选择。然而，随着云计算、边缘算力等技术和业务的快速发展，算力部署方式正在发生深刻变革。集中式算力将计算、存储等能力封装为按需调用扩展的独立虚拟资源。边缘算力则将计算资源就近部署在数据源或用户边缘，并利用多种通信网络技术（如5G、Wi-Fi、光纤通信等）连接分布式算力节点，通过虚拟化和多层次的算力资源协同，实现对资源的灵活、按需和实时调度，提高边缘算力及网络的利用率，增强业务的服务质量和安全性，提升响应速率。

其中，边缘算力的“边缘”可以从多个角度去理解。一是地理位置的边缘，指的是数据源和用户设备所在的网络边缘位置，包括智能终端、工业传感器、自动驾驶汽车等物联网设备的所在地。上述设备生成大量数据，边缘算力就近处理，相较于集中式处理，降低了数据传输时延和网络带宽消耗。二是计算资源的边缘，指的是在用户侧部署的计算设备和资源。这些设备包括边缘服务器、网关、路由器等，它们可以在本地处理数据并做出决策。这种分布式计算模式使得数据处理更加高效和灵活。三是数据流的边缘，指的是数据在网络中传输路径上的关键节点。这些节点可以对数据流进行预处理、分析和决策，以降低需要传输的数据量。在网络边缘进行数据过滤和压缩，可以显著提高数据传输的效率和安全性。

1.1 边缘算力价值日益凸显

边缘算力已成为支撑信息技术（information technology，IT）、通信技术（communication technology，CT）、操作技术（operational technology，OT）创新演进的关键基础设施。

（1）IT领域。当前，计算架构的革新和硬件制程的进步推动算力部署模式从集中式向互联协同转变。多异构芯片集成、Chiplet技术以及新型封装技术等为边缘算力设备提供更强大的计算能力和灵活的定制化选项。此外，数据隐私和安全性受到各方高度关注，边缘算力可以实现数据本地处理分析，以降低数据传输风险，提高安全性。

（2）CT领域。新型网络架构将采用服务化设计，实时感知用户需求，支持资源按需调用，为不同垂直行业提供快速响应和灵活部署。实时感知能力需要边缘算力支撑响应，在5G等网络架构中明确定义边缘算力作为其重要组成部分，推进网络智能化演进。

（3）OT领域。随着算力总量的持续增长和算力类型的不断丰富，工业企业数字化转型加速，其中，边缘算力已成为工业数字化转型的关键基础设施之一。基于边缘算力部署的工业应用种类繁多且贯穿于各生产环节，目前已在设计模拟、订单排产、生产制造、运营维护、安环管理、质量追溯、物流库存等重要场景形成一批典型应用。

1.2 边缘算力产业关注度持续提升

边缘算力作为计算+网络融合交汇的新型基础设施，是打通算力普及的“最后一公里”，受

到产业各方的高度关注。

基础设施服务商持续加大边缘算力产业布局。浪潮、新华三均推出边缘超融合一体机，实现边缘算力敏捷部署、数据就近接入、本地智能处理，同时集成人工智能、物联网平台、数据平台等多项能力；中兴、华为积极探索边缘算力技术创新，提供基于边缘算力的行业解决方案，已在电力、制造、农业、医疗等行业广泛应用；阿里云扩展中心云至边缘，推出OpenYurt开源框架及边缘节点服务，实现云原生应用边缘部署。

随着5G规模化应用，5G+边缘算力产业进入快速发展期，基础设施运营商充分利用自身网络资源，将算力和网络接入端扩展成边缘算力节点，形成边缘算网一体化融合的独特优势。中国移动构建广泛边缘算力节点，并聚焦边缘应用生态构建，基于边缘算网融合能力提供泛在算力一体服务；中国电信自主研发边缘云，赋能低时延、大带宽及数据安全应用需求；中国联通融合5G、通信、大数据能力，构建5G边缘算力体系，面向工业、视频、车联网等领域推出定制化边缘服务；中国铁塔依托全国机房资源，积极推进“通信塔”变“数字塔”、“通信机房”变“数据机房”，按需网格化部署边缘算力，从边缘智算服务、算力共享服务等不同层面助力边缘算力建设。

工业企业基于行业经验积累，将工业软件部署至边缘算力节点，提高生产运营效率。三一重工、海尔、商飞等企业在边缘算力节点实现远程I/O毫秒级实时控制及逻辑控制器集中虚拟化部署，动态调配网络、控制、算力资源，降低设备固定资产投入，成功打造边缘算力应用于制造业的“样板间”。西门子、罗克韦尔等工业自动化企业推出基于边缘算力的强化学习服务，即基于强化学习智能算法创建大型设备预测性维护模型，实现工业生产高效运维管理。

2　边缘算力技术体系

随着边缘算力体系架构和技术体系的不断成熟，边缘算力技术的研究热点从基础设施构建向资源融合化、业务智能化发展。本文主要介绍的边缘算力技术体系如图1所示。边缘算力网络技术是实现边缘算力的信息基础设施，通过算网协同控制满足算力池化和自适应适配需求。边缘数

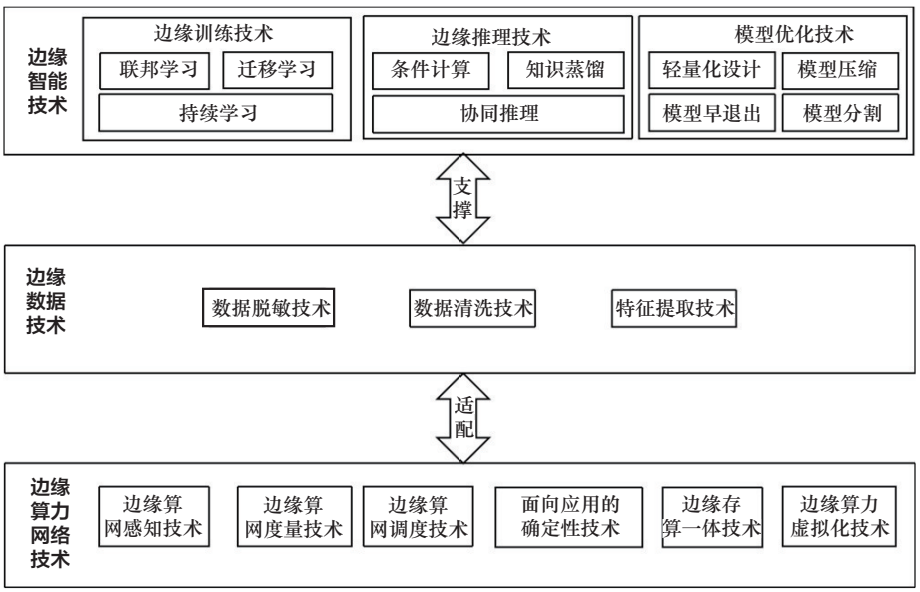

图1　边缘算力技术体系

据技术向边缘算力网络技术提供资源需求，通过数据挖掘调节带宽和算力的负载，平衡上层智能管控和下层资源供给，对智能应用起到支撑作用。人工智能和边缘算力相互赋能就催生了边缘智能技术，它将进一步提升边缘算力在工业互联网、车联网等领域的自主服务能力，开创端边协同模式新局面[2]。边缘智能技术为用户提供更快速、高效的智能训练和推断服务[3-4]。目前，边缘智能技术的难点在于边缘算力资源容量、大模型算力存储开销与决策准确度的折中[5]。边缘数据技术提供数据高质量挖掘能力，提升网边端协同效率。边缘算力网络技术是整合异构资源的重要发展方向，将边缘算力的分布式计算能力与网络资源深度融合，形成确定性的算网调度能力，以满足算网融合的需求。边缘算力将分布式算力资源和网络资源协同管控，实现面向业务的感知、度量、调度等。从图1可以看出，未来边缘算力将在资源层面呈现出异构性和确定性，在数据层面加强融合创新，在智能层面采用轻量化与协同化手段，提供灵活的智能应用。

2.1 边缘智能技术

2.1.1 边缘训练技术

边缘训练技术是向端设备就近提供训练服务的技术。利用该技术，数据传输量更少、距离更短，可以有效减少网络带宽、时延等方面的开销。联邦学习是允许模型在多个设备上进行训练的分布式机器学习方法。联邦学习适用于边缘设备多、数据分散且敏感的场景。文献[6]提出了一种利用多个边缘设备和云服务器协同训练剩余使用寿命（remaining useful life，RUL）预测模型的云边协同框架，通过变分自编码器提取特征并在边缘设备上进行本地训练，减少了数据传输时延。文献[7]针对海上通信环境差、通信带宽有限、通信链路易被干扰等问题，提出无人系统可以通过边缘人工智能模型协同实现数据采集与边缘处理任务，利用联邦互蒸馏的模型协同训练方法降低模型训练数据传输带宽要求。迁移学习通过复用源领域的知识，提升目标领域模型的泛化能力。持续学习是持续接收新数据并不断改进和调整模型参数的机器学习模式，使得模型在数据或任务不断变化时能够动态适应。

2.1.2 边缘推理技术

边缘推理技术利用边缘设备上的模型实现高效的数据处理和智能决策，该类技术利用特征提取技术得到特征数据，利用模型得到分析与决策结果。条件计算通过在推理过程中仅激活模型的最小计算路径来提高效率和减少资源消耗。知识蒸馏将知识从一个大型训练模型转移到一个较小的模型中进行部署。文献[8]提出一种基于掩码信息熵迁移的知识蒸馏策略，解决了文本检测任务中信息量不足的问题，有效提升了学生网络的性能和文本检测精度。协同推理利用节点间的协同，结合场景的资源特点等信息，组成用于协同推理的分布式模型[9]。文献[10]提出了一种结合联邦思想的分割学习方法，它将人工智能模型分割并在多个节点训练，有效降低端侧计算时延和资源消耗。

2.1.3 模型优化技术

模型优化技术旨在通过算法和策略降低模型规模、复杂度和资源开销[11]，解决边缘设备的资源限制问题。模型轻量化设计采用更小、更高效的模型结构，同时保持或提升模型的预测准确度和复用能力，结合深度可分离卷积、组卷积和通道分组等手段，减少深度神经网络的计算开销。模型分割利用大模型的访问局部性，将一个大模型拆分成多个部分并在不同设备上独立运行。模型早退出是指在深度神经网络的中间层嵌入多个早退出分支，并在这些分支上计算置信度，与设置的阈值进行比较，判断是否提前终止计算。常用的置信度计算方式包括计算最大类别概率（预测结果中最高概率值）和信息熵，当最大类别概

率超过阈值，或信息熵低于阈值时，模型即可在当前分支提前退出。文献[12]通过简化早退出口结构和优化其位置选择，提出了一种高效的模型早退出方法，在移动设备上实现了推理时间缩减。

2.2　边缘数据技术

2.2.1　数据脱敏技术

数据脱敏技术通过数据变形等方式防止隐私数据在采集、传输、使用等环节中暴露。根据处理时机和应用场景不同，脱敏技术主要可分为静态脱敏和动态脱敏。静态脱敏主要应用于数据存储等场景，对批量数据进行一次性脱敏处理。动态脱敏则侧重于实时数据访问场景，在数据被查询、调用或传输的瞬间，根据预定义的策略和访问者权限对敏感信息进行实时变形、替换或遮蔽，从而在保障数据可用性的同时，最小化隐私泄露风险。文献[13]为保障石油企业的数据资产，针对数据脱敏技术缺少图像数据脱敏能力的问题及油气田勘探中的曲线图像数据特点，借鉴深度学习图像修复方法，采用基于上下文编码器的脱敏技术对曲线图像进行脱敏。

2.2.2　数据清洗技术

数据清洗技术是确保数据质量和准确性的技术。该技术通过对数据进行去重、填补缺失值、处理异常值和转换格式等操作，去除或修复数据集中存在的错误、不一致、不完整和冗余等问题数据。文献[14]为了解决智能制造车间数据源庞杂繁复的问题，针对不同的业务筛选关键数据，提出了对应的数据清洗及数据降维的方法。

2.2.3　特征提取技术

特征提取技术是从原始数据中提取、选择、转换和创建对预测模型有用的特征。在数据层面，特征提取技术可进行数据的预处理，形成支撑上层应用的规范化特征。特征选择技术可以筛选出对模型训练和推理最具贡献的典型特征，减少冗余和相关性低的特征，防止模型过拟合。特征变换技术对现有边缘数据特征进行数学或统计变换，以提升特征的表现力。文献[15]针对故障诊断场景中变工况数据样本少、特征不明显等问题，通过有监督的模式分解算法从原始数据中分解成多频率的本征模态函数，采用联合最大平均差异特征迁移的方法使源域与目标域联合分布差异最小化。

2.3　边缘算力网络技术

2.3.1　边缘算网感知技术

边缘算网感知技术是针对边缘算力网络场景下的服务部署位置、边缘算力需求、资源实时状态和算网资源的感知，面向边缘算力、服务水平协议（service level agreement，SLA）需求的感知、分析技术。该技术聚焦于实时、动态地感知边缘场景下的多维关键信息，如用户服务请求及其SLA需求、边缘节点的实时算力资源状态，以及网络资源状态等。通过对上述感知数据的多源融合与智能分析，算网感知技术能够构建精确的全局资源视图，为后续资源调度与服务编排提供决策依据，从而支撑算力资源的动态匹配、按需调用，并最终实现面向用户需求的敏捷响应与业务的灵活部署。目前，算网感知技术已取得了一定的突破，能够有效实现对异构算力资源分布、算力节点状态以及网络负载等信息的监测收集。

2.3.2　边缘算网度量技术

边缘算网度量技术实现了对工业生产系统中算力与网络的统一描述，利用统一抽象方法提高感知精度和策略处理效率。在图形处理器（graphics processing unit，GPU）、张量处理器（tensor processing unit，TPU）等高性能模组中，存储能力和计算能力共同决定了计算任务的处理性能。网络能力度量包括网络带宽、时延、可用私网个数、可用公网IP地址数等。文献[16]提

出包括差异化服务层、异构融合网络层和智慧化适配层的总体架构。差异化服务层和异构融合网络层分别实现差异化确定性应用需求和异构化确定性网络要素的统一标识和描述，并通过标识解析映射实现确定性信息向智慧化适配层的统一封装和传递。智慧化适配层完成差异化确定性应用需求和异构化确定性网络要素的适配。

2.3.3 边缘算网调度技术

边缘算网调度技术对计算资源和网络资源进行动态管理和调度，将工作负载智能分配到边缘算力节点，以优化性能、降低时延、提升资源利用效率和降低能耗。在能耗和时延联合调度方面，文献[17]建立了一个以最小化时延和能量消耗为目标的联合计算卸载和资源分配问题，提出了一种协同任务拆分和资源分配优化算法。在负载均衡方面，文献[18]在移动边缘计算网络中分析用户表现出的各种运动模式，依据固有的移动性规律和数据连续时间段内随机到达，最大化网络系统的数据处理能力。在容器选择方面，文献[19]综合考虑响应时间和任务执行成本等多个因素，为具有紧迫截止期的任务确定任务优先级并决定对应的容器。

2.3.4 面向应用的确定性技术

围绕多应用协同需要统一调度算力和网络资源的问题，文献[20]提出面向应用的确定性技术。该技术是一种管理多个应用协同过程确定性的技术，可以依据边缘智能应用动态协同调度生产设备、边缘设备和网络设备的算网资源，使应用执行过程的时延、抖动、带宽等指标具有确定的极限值。在工业网络和边缘算力紧密协同后，端边云多个应用的通信环节和计算环节构成了实时性要求较高的闭环，一般具有固定的执行顺序拓扑和周期执行特征。面向应用的确定性技术开放工业应用到边缘算力网络的南北向接口，将底层边缘计算调度策略匹配工业应用的执行顺序，为应用之间选择确定性的最优路径，从而确保整个生产流程的确定性。因此，面向应用的确定性由网络传输的通信确定性、工业算力的计算确定性、应用流程的时序确定性和程序执行的逻辑确定性构成。该技术结合多实例GPU、vGPU等技术将计算任务划分到隔离的算力资源，实现了对算力和网络的细粒度管理。终端、边缘和云服务可以构成动态调整的功能链，从而在网络传输开销、计算开销和能耗之间取得折中。

3 边缘算力应用现状

3.1 工业互联网

传统工业企业依赖云计算中心集中式处理数据时，边缘算力通过构建分层式计算架构实现资源下沉。现场层部署的智能网关集成多协议转换引擎，实时采集PLC、传感器等设备数据并进行特征提取；边缘层采用容器化部署模式，通过轻量化运行环境实现工业软件的动态加载与版本管理。云边协同框架依托服务化通信中间件，实现控制策略的远程编排与增量更新，同时建立双向数据通道，以保障配置信息的安全同步。为解决异构设备接入难题，边缘节点内置工业协议栈解析模块，支持开放平台通信统一架构（open platform communications unified architecture，OPC UA）、Modbus等协议的语义化转换，并通过资源虚拟化技术实现CPU与现场可编程门阵列（field programmable gate array，FPGA）计算单元的混合调度。

云边协同架构通过在边缘部署代理程序，实现云端模型按需下发至本地执行，满足不同产品的定制化需求。基于边缘算力的工业生产系统架构如图2所示，该架构的优势在于利用边缘算力，提高本地响应速度，并能够满足企业多基地、多车间及外协工厂的管理需求，实现统一生产管控模式。

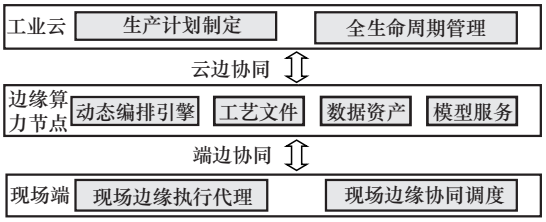

图2　基于边缘算力的工业生产系统架构

考虑到工业设备在不同工作位置可能需要执行不同的工种或任务，边缘算力服务应具备对现场端应用进行动态编排和部署的能力。同时，企业云化应用会随着产品生产的迭代而迭代演进，更新程序。在现场端侧，基于容器部署的云边协同应用框架，用户只需在云端开发脚本或进程，即可通过框架将应用派发到边缘网关执行。云端集中管理应用开发、测试和版本存储，实现应用共享。边缘侧设备可自动调度并执行不同任务，而用户只需在框架中编排不同脚本进程即可完成业务行为。

3.2　智慧社区

边缘算力作为连接感知终端与智慧应用的桥梁，通过边缘网关汇聚、清洗社区终端采集的视频、图片、结构化数据，并进行边缘智能分析与决策，最终将数据推送至统一管理平台，实现社区智慧化应用。目前，智慧社区建设通过部署边缘终端设备（如摄像头、传感器、监测设备等），实现对社区全方位覆盖和实时监控，并利用边缘算力节点进行数据智能分析。边缘算力可实现对采集数据的实时处理和解析，精准识别安全异常场景，及时发出告警通知，确保安全隐患得到有效处置。

基于边缘算力的智慧社区系统架构如图3所示，通过在社区内部搭建边缘算力设施，该系统构建了从数据采集到智能服务的闭环管理体系。核心架构包含4个主要部分：（1）本地化智能分析模块通过边缘算力实时处理智能摄像头、门禁及环境感知设备数据，实现多维特征（人/车/物）识别、异常行为解析以及安防/消防事故的即时监测；（2）感知终端智能化模块通过嵌入式边缘算力单元，支持实时视频结构化分析、动态人脸比对和行为模式识别，实现异常事件秒级报警联动；（3）基于模块化设计理念，系统采用标准化接口协议实现设备的统一运维管理，支持热插拔式扩容升级并优化机房空间利用率；（4）在数据安全层构筑本地加密存储与云端协同的双重保障机制，确保敏感数据的可信流通。

在技术实现层面，边缘算力节点搭载多模态数据融合引擎，对接入的RGB摄像机、红外传感器等多源异构数据进行关联分析。针对视频流处理需求，部署轻量化推理框架实现人脸特征向量提取与行为模式识别，通过层间张量压缩技术降低内存占用。在隐私保护方面，构建两级数据过滤机制，在终端侧实施基于区域掩码的敏感信息遮蔽，边缘算力节点采用基于动态脱敏的匿名化技术对结构化数据进行实时混淆处理，通过字段级令牌化与格式保留加密实现低延迟隐私保护。

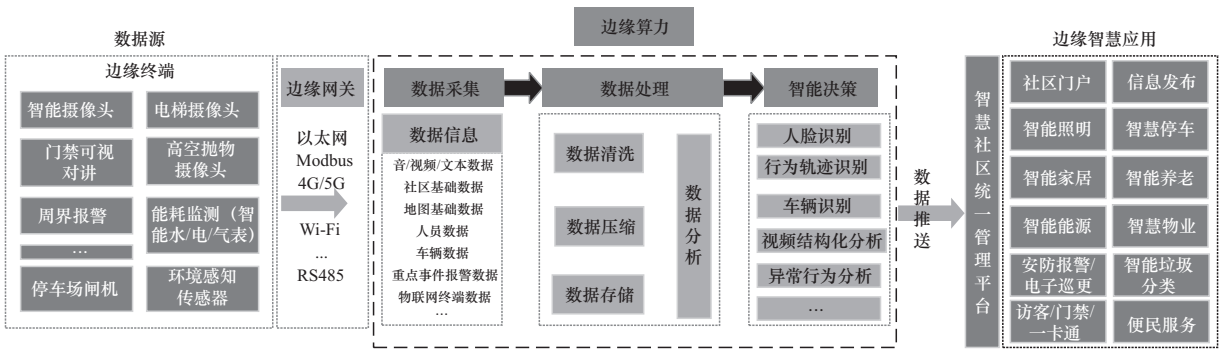

图3　基于边缘算力的智慧社区系统架构

在设备管理方面，开发自适应网络协议栈支持Zigbee、LoRa等不同通信制式设备的统一接入，并通过资源隔离技术保障关键安防应用的服务质量。

3.3 车联网

智能网联汽车渗透率不断提高，由此产生的车载信息娱乐、传感器、车路协同、地图等数据传输、处理、存储的需求极大地增加了网络负荷，并对网络时延、带宽、可靠性提出了更高的需求。将边缘算力应用于车联网之后，可以将业务部署在边缘节点，减少数据传输路由长度，从而降低端到端通信时延；还可以作为本地服务托管环境，提供强大的计算、存储资源，支持视频流的实时分析与处理、违章预警、危险驾驶处理等，以及部署本地更具地理和区域特色、更高吞吐量的车联网服务。

基于边缘算力的自主泊车系统架构如图4所示。该架构以5G通信网络为基础，融合人工智能、北斗定位等新技术，构建人、车、场、云四位一体的协同体系。该系统通过协同规划、感知、控制与定位的商业闭环，不仅支持车辆循迹安全行驶与场侧融合感知，更打造了覆盖泊车全流程的智慧运营服务，从行标V2X预警、双端运营监控到车位预约引导，从一键泊车、召车到多视角数字孪生监控，配合数据驱动的业务分析，显著提升停车场运营与用户出行双重效率。在技术实现层面，系统创新性地融合边缘算力节点与车载传感器的多源数据，突破盲区障碍物识别瓶颈；采用云端全局路径动态规划与车端实时感知导航的协同机制，实现车位资源最优分配；同时依托边缘算力节点毫秒级的运动趋势分析与碰撞风险评估，结合车端控制系统的快速响应能力，降低突发场景碰撞风险，在提升车位周转率的同时，全面保障停车场行车安全与通行效率。

3.4 轨道交通

传统列车运行状态监测系统依赖集中式数据处理，面临时延、带宽和可靠性挑战。为满

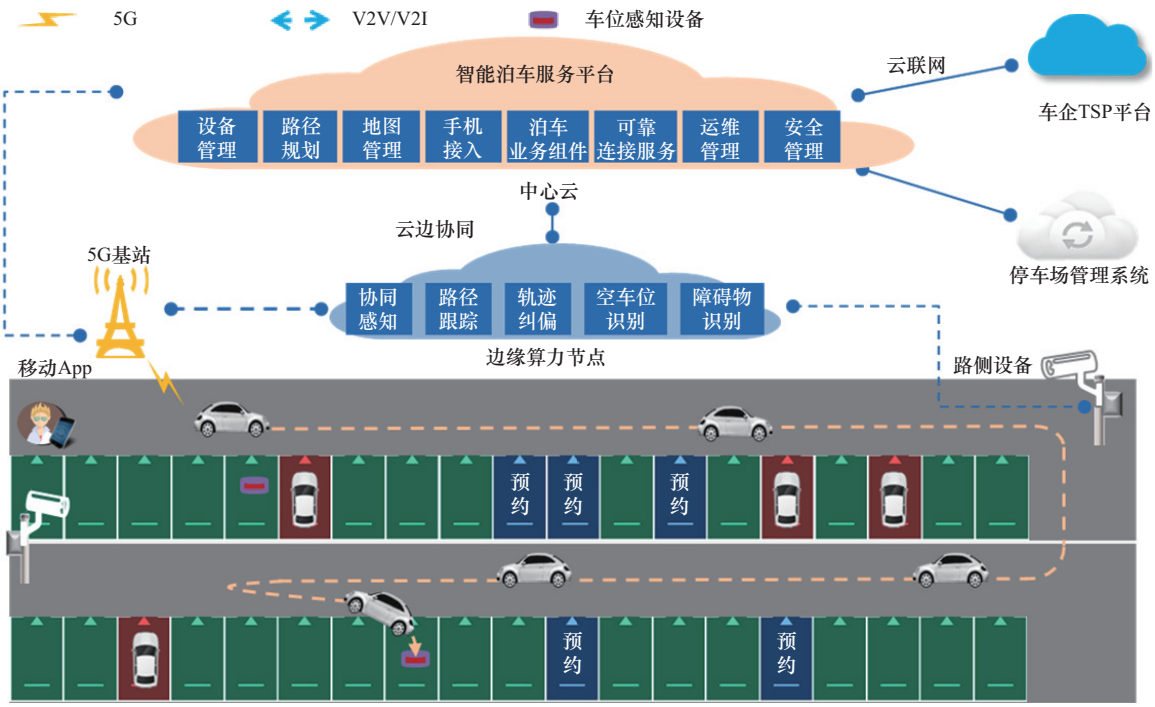

图4 基于边缘算力的自主泊车系统架构

足快速故障诊断等需求，亟须在列车中或轨道沿线部署边缘算力节点，对列车温度、制动压力、轨道环境等信息进行实时分析，并利用预训练模型在毫秒级时间尺度内识别潜在安全隐患，触发预警和应急处置，提升列车运行安全裕度。同时，边缘算力节点对数据进行压缩、特征提取和加密，降低车地通信负荷，提高数据传输安全性和效率。

基于边缘算力的轨道交通智能运维系统架构如图5所示。该架构采用分层部署策略，在列车、车站和指挥中心分别部署边缘算力节点，并与云端协同，构建高速、泛在、智能、安全的算力体系。边缘算力的实时性与云计算的强大算力结合，可为轨道交通行业数字化发展提供新动能。

通过构建“车–轨–站–控”全场景的轨道交通边缘智能体系，在关键环节部署多层次算力节点实现智能安全闭环。在列车车载设备中集成边缘算力模块，实时处理监测数据并通过5G回传云端，网络中断时启用本地模型保障运行安全；轨道沿线部署的智能感知单元不仅通过视频分析实现入侵预警和地质灾害预判，更内置故障预测模型，采用滑动窗口机制对振动频谱进行实时特征提取；车站及车辆段部署的运维系统构建分级存储策略，在本地完成原始波形数据的压缩归档，同时通过冗余编码传输关键特征值至中心云；指挥中心搭建的智能平台则通过抗噪声通信协议保障数据传输，物理层采用扩频调制技术提升信号鲁棒性，应用层实施重传优先级调度算法确保控制指令的可靠传输，最终整合多源数据构建数字孪生系统，实现列车调度、设备监控和应急指挥的全局协同。该体系通过融合实时计算与智能分析技术，在提升安全防护能力的同时，显著优化了轨道交通系统的运营效率与智能化水平。

3.5　云游戏

云游戏作为一种游戏服务模式，其核心在于将原本在本地终端进行的游戏渲染、逻辑运算等高负载任务卸载至边缘算力节点，用户终端仅需负责接收串流的音视频数据并进行简单的输入操作。这种架构对边缘算力节点的计算能力、图形处理能力以及网络传输性能提出了极高的要求，以确保流畅的游戏体验和低延迟的交互响应。

基于边缘算力的云游戏系统架构如图6所示，其核心思路是将云游戏的部分计算、渲染和网络

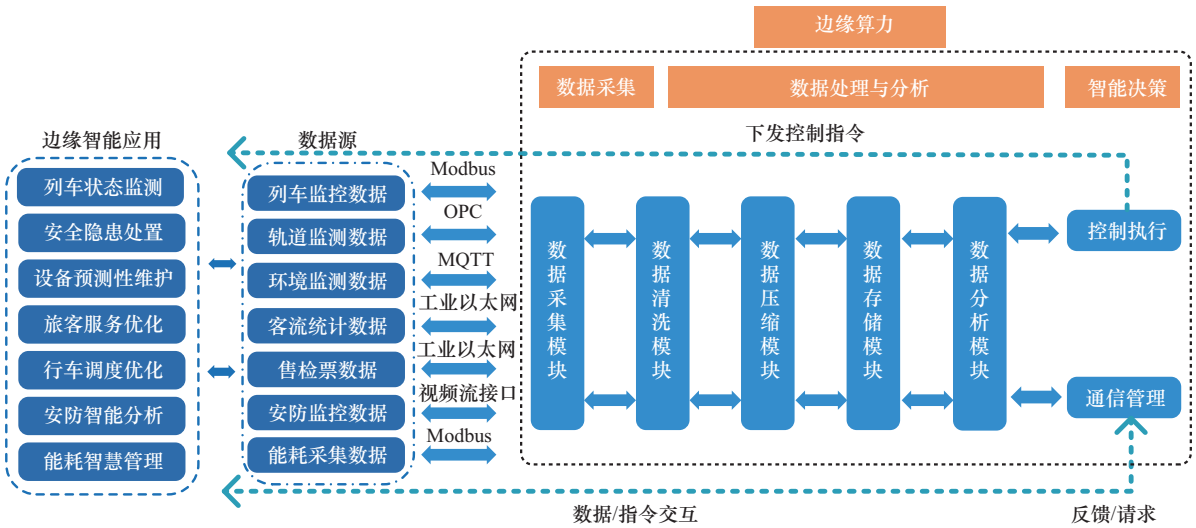

图5　基于边缘算力的轨道交通智能运维系统架构

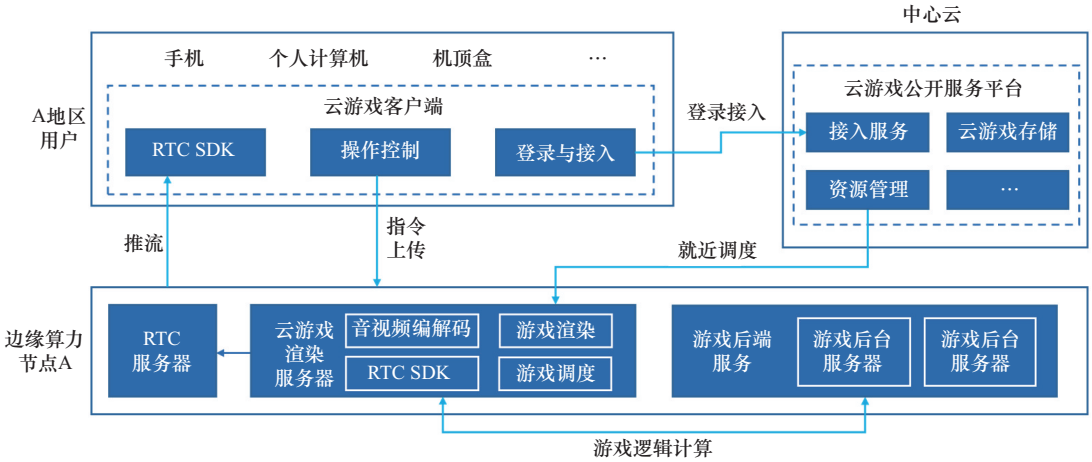

图6　基于边缘算力的云游戏系统架构

传输功能从中心云迁移至网络边缘，利用边缘算力节点的低时延、高带宽、分布式部署和丰富的云服务优势，打造更加流畅、高效、便捷的云游戏体验。

该云游戏架构采用分层设计实现端到端的低时延交互，主要由本地客户端、云游戏公共服务、后端服务及渲染服务器四大模块构成。本地客户端作为用户入口，集成解码与操控功能，既负责向云端发送操作指令，也承担音视频流解码呈现任务。云游戏公共服务部署于中心云，基于地理分布、网络质量、算力需求三重维度实施智能调度，通过动态匹配最优边缘节点实现游戏实例的全生命周期管理。云游戏后端服务和渲染服务器运行在分布式边缘算力节点，该节点构建了支持x86+GPU/ARM等架构的异构算力资源池，其中ARM架构通过精简指令集计算机（reduced instruction set computer，RISC）实现移动端游戏的低功耗处理，x86+GPU则依托复杂指令集计算机（complex instruction set computer，CISC）及GPU并行计算能力高效承载PC/主机级游戏渲染。异构架构的指令集兼容性问题通过硬件辅助转译技术（如Intel Houdini）与动态二进制翻译QEMU（quick EMUlator）协同解决，并基于硬件虚拟化技术实现GPU算力的弹性分配，具体执行指令解析、游戏逻辑运算、实时渲染及编解码处理等核心功能。在游戏渲染层面，采用分帧处理架构，将几何计算与光影渲染任务拆分至不同计算单元并行处理，显著提升渲染效率；网络传输层构建多路径冗余通道，基于用户位置智能选择最优接入点，并部署前向纠错机制，以有效降低画面卡顿率；游戏调度系统引入负载预测模型，依据历史访问规律预加载游戏实例，结合会话保持技术实现玩家状态的无缝迁移。同时通过实时时钟（real time clock，RTC）流媒体技术将渲染结果低延迟传输至终端，形成“输入–计算–渲染–传输”的完整闭环。在功耗控制方面，开发动态电压频率调节模块，根据实时负载，精细调整处理器工作状态，最终通过系统级优化实现全链路的高性能与低能耗平衡，确保玩家获得流畅稳定的沉浸式游戏体验。

4　结束语

随着大型语言模型、工业大模型等智能化应用热度不断高涨，边缘算力作为产业智能化发展的数字化底座将迎来战略机遇期。然而，由于目前标准化工作缺乏统一共识、建设应用模式尚不清晰等问题，边缘算力规模部署仍需产业各方协同推进。

（1）加速技术攻关，建立健全标准体系

以数字化转型需求为牵引，强化边缘算力与5G、人工智能等信息通信技术的协同攻关，聚焦边云协同、边缘算力网络及边缘智能为代表的边缘算力关键技术方向的基础理论研究攻关，完善边缘算力标准体系建设，制定实施核心设备、互通接口、测试规范、应用指南等急用先行标准，为边缘算力技术的应用转化夯实基础。

（2）加快基础设施统筹建设，提升综合服务水平

充分利用变电站、基站、通信机房等既有基础设施资源，以共建共享模式统筹规划边缘算力节点建设，推动边缘节点的规模化部署与异构资源整合，建立统一的边缘算力调度平台，盘活存量资源。有效支撑以GPT-4、Sora、DeepSeek等为代表的AI大模型向边缘侧延伸所带来的高吞吐量、低时延、泛在存储等新型算力需求，助力AI大模型在工业控制、智能交互、实时决策等复杂场景的规模化落地，并为企业提供低时延、低成本的智能算力服务。

（3）构建标杆案例，繁荣边缘算力应用生态

依托我国行业门类齐全、数字化转型需求旺盛等优势，支持各行业龙头企业按需灵活部署边缘算力基础设施，精准把握需求侧实际需求，聚焦共性应用场景，形成一批可复制、可推广的应用模板，“以用促建”，加快边缘算力应用推广落地。构建“边缘算力+*N*”应用体系，通过组织开展边缘算力应用促进活动激发创新活力，繁荣应用生态。

（4）深化产业主体间协作，构建开放共赢的产业生态

充分发挥联盟组织的平台汇聚作用，推动基础设施运营商、工业企业、互联网企业等各主体厂商交流合作，促进多方力量协同构建边缘算力产业协同，推动产业链上下游主体面向解决方案、生态构建等层面的深层次合作，加速推进规模化应用进程。

参考文献：

[1] 王哲. 边缘计算发展现状与趋势展望[J]. 自动化博览, 2021, 38(2): 22-29.
WANG Z. Development status and trend prospect of edge computing[J]. Automation Panorama, 2021, 38(2): 22-29.

[2] 中国铁塔股份有限公司. 边缘算力蓝皮书[R]. 2024.
China Tower Co., Ltd. Blue book on edge computing power[R]. 2024.

[3] 张晓东, 张朝昆, 赵继军. 边缘智能研究进展[J]. 计算机研究与发展, 2023, 60(12): 2749-2769.
ZHANG X D, ZHANG C K, ZHAO J J. State-of-the-art survey on edge intelligence[J]. Journal of Computer Research and Development, 2023, 60(12): 2749-2769.

[4] 莫梓嘉, 高志鹏, 苗东. 边缘智能: 人工智能向边缘分布式拓展的新触角[J]. 数据与计算发展前沿, 2020, 2(4): 16-27.
MO Z J, GAO Z P, MIAO D. Edge intelligence: a new exploration for artificial intelligence expanding to edge[J]. Frontiers of Data & Computing, 2020, 2(4): 16-27.

[5] 侯祥鹏, 兰兰, 陶长乐, 等. 边缘智能与协同计算: 前沿与进展[J]. 控制与决策, 2024, 39(7): 2385-2404.
HOU X P, LAN L, TAO C L, et al. Edge intelligence and collaborative computing: frontiers and advances[J]. Control and Decision, 2024, 39(7): 2385-2404.

[6] 于振军, 雷宁博, 莫语, 等. 基于联邦学习与云边协同的剩余寿命预测[J]. 清华大学学报(自然科学版), 2025, 65(5): 901-911.
YU Z J, LEI N B, MO Y, et al. Remaining useful life prediction based on federated learning and cloud-edge collaboration[J]. Journal of Tsinghua University (Science and Technology), 2025, 65(5): 901-911.

[7] 姜俊, 张家瑞, 潘吉龙, 等. 面向海上无人系统的边缘模型协同与数据压缩算法[J]. 系统工程与电子技术, 2025, 47(5): 1718-1727.
JIANG J, ZHANG J R, PAN J L, et al. Edge model collaboration and data compression algorithm for unmanned systems in sea[J]. Systems Engineering and Electronics, 2025, 47(5): 1718-1727.

[8] 陈建炜, 沈英龙, 杨帆, 等. 基于掩码信息熵迁移的场景文本

检测知识蒸馏[J]. 软件学报, 2025: 1-20.
CHEN J W, SHEN Y L, YANG F, et al. Knowledge distillation for scene text detection via mask information entropy transfer[J]. Journal of Software, 2025: 1-20.

[9] 赵婵婵, 吕飞, 石宝, 等. 面向边缘智能的协同推理方法研究综述[J]. 计算机工程与应用, 2025, 61(3): 1-20.
ZHAO C C, LYU F, SHI B, et al. Review of collaborative inference methods for edge intelligence[J]. Computer Engineering and Applications, 2025, 61(3): 1-20.

[10] 金博宇. 边缘网络中分割学习模型部署与资源调度研究[D]. 北京: 北京邮电大学, 2024.
JIN B Y. Research on deployment and resource scheduling of split learning models in edge networks[D]. Beijing: Beijing University of Posts and Telecommunications, 2024.

[11] 牛鑫, 吕现伟, 余辰. 边缘智能: 现状与挑战[J]. 武汉大学学报(理学版), 2023, 69(2): 270-282.
NIU X, LYU X W, YU C. Edge intelligence: state-of-the-art and challenges[J]. Journal of Wuhan University (Natural Science Edition), 2023, 69(2): 270-282.

[12] 成恩. 面向移动终端的模型早退网络结构优化及阈值设置研究[D]. 长沙: 中南大学, 2023.
CHENG E. Research on model early-exit network structure optimization and threshold setting for mobile terminals[D]. Changsha: Central South University, 2023.

[13] 刘烨, 李帆, 郑泽晨. 基于上下文编码器网络的油气曲线图像脱敏方法研究[J]. 计算机应用与软件, 2024, 41(2): 216-221.
LIU Y, LI F, ZHENG Z C. Desensitization method of oil and gas curve image based on context encoder network[J]. Computer Applications and Software, 2024, 41(2): 216-221.

[14] 郑良, 苑明海, 裴凤雀, 等. 智能制造车间数据采集和预处理方法研究[J]. 制造技术与机床, 2022(10): 85-90.
ZHENG L, YUAN M H, PEI F Q, et al. Research on data acquisition and preprocessing method based on intelligent manufacturing shop[J]. Manufacturing Technology & Machine Tool, 2022(10): 85-90.

[15] 尹洪申, 刘文峰, 俞啸, 等. 多尺度迁移学习的轴承故障诊断[J]. 机械设计与制造, 2025(1): 10-14.
YIN H S, LIU W F, YU X, et al. Bearing fault diagnosis based on multi-scale transfer learning[J]. Machinery Design & Manufacture, 2025(1): 10-14.

[16] 杨冬, 程宗荣, 田伟康, 等. 广义确定性标识网络[J]. 电子学报, 2024, 52(1): 1-18.
YANG D, CHENG Z R, TIAN W K, et al. Generalized deterministic identification networks[J]. Acta Electronica Sinica, 2024, 52(1): 1-18.

[17] LONG Y X, WANG Z Y, LAN S Z, et al. Energy-latency tradeoff for task offloading and resource allocation in vehicular edge computing[J]. Computer Networks, 2025, 258: 111026.

[18] PANG S C, CHENG Y Z, HE X, et al. Multi-user motion state task offloading strategy for load balancing in mobile edge computing networks[J]. Ad Hoc Networks, 2025, 169: 103759.

[19] MAHESAR A R, LI X P, SAJNANI D K. Enhancing task scheduling and QoS optimization in mobile edge computing via microservice-orientedcontainer selection[J]. Computing, 2025, 107(2): 60.

[20] FU T, ZHANG H S. The system architecture and decision mechanism of the deterministic application layer in the industrial internet[C]//Proceedings of the 2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE). Piscataway: IEEE Press, 2023: 18-25.

[作者简介]

郭宇辉（1972– ），男，中国铁塔股份有限公司通信技术研究院院长，中国通信企业协会低空经济专业委员会秘书长，主要研究方向为无线通信技术、低空经济、算力网络等。

闫亚旗（1988– ），男，中国铁塔股份有限公司高级工程师，主要研究方向为物联网、边缘计算、算力网络技术及产品创新。

付韬（1988– ），男，博士，中国信息通信研究院工程师，主要研究方向为工业互联网、确定性网络。

张阔（1988– ），男，中国铁塔股份有限公司高级工程师，主要研究方向为通信电子和人工智能。

王哲（1990– ），男，博士，中国信息通信研究院高级工程师，主要研究方向为工业互联网、边缘计算、边缘算力以及新型工业控制等。

冉沛（1981– ），男，中国铁塔股份有限公司高级工程师，主要研究方向为物联网及边缘计算。

刘文睿（1996– ），男，中国铁塔股份有限公司工程师，主要研究方向为边缘计算及算力网络。