

基于AI的通信网络资源分配策略研究综述

堵 颢¹ 安平省² 张 晨¹

1. 河南省信息咨询设计研究有限公司 河南 郑州 450000

2. 中国联合网络通信有限公司驻马店市分公司 河南 驻马店 463000

摘要: 随着通信网络规模持续扩张与业务需求复杂化,传统资源分配方法在实时性、动态性和效率方面面临严峻挑战。人工智能(AI)技术的引入为通信网络资源分配提供了全新范式,通过机器学习、深度强化学习、群体智能算法等手段,实现了资源分配的智能化与自适应优化。本文系统梳理了AI在通信网络资源分配中的核心应用场景,涵盖动态资源分配、安全多方计算、网络切片、算力交易等关键领域,深入分析了粒子群优化、强化学习、联邦学习等主流算法的技术原理与实现路径,结合工业通信网络、边缘计算等典型案例验证了AI驱动策略的有效性。研究指出,AI技术可显著提升网络吞吐量、降低延迟并优化资源利用率,但需进一步突破计算资源限制、数据隐私保护等瓶颈。未来发展趋势包括多算法融合、边缘智能与6G原生AI架构的深度集成。

关键词: 人工智能; 通信网络; 资源分配; 动态优化; 强化学习; 网络切片

引言

通信网络作为现代社会信息交互的基础设施,其资源分配效率直接影响用户体验与系统性能。传统资源分配方法多基于静态规则或经验模型,难以应对网络规模指数级增长、业务需求动态变化及异构设备共存等挑战。例如,在工业物联网场景中,传感器节点需实时传输高精度数据至边缘服务器,而传统轮询调度易导致关键任务延迟;在5G毫米波通信中,高频段信号衰减特性要求资源分配需具备动态信道适配能力。人工智能(AI)凭借其强大的数据处理能力与自主学习特性,为解决上述问题提供了新思路。通过构建数据驱动的决策模型,AI可实时感知网络状态、预测流量变化并动态优化资源分配策略,从而提升网络吞吐量、降低延迟并保障服务质量(QoS)。

本文系统综述了AI在通信网络资源分配中的研究进展,从核心应用场景、主流算法原理、案例分析及未来发展趋势四个维度展开论述,旨在为通信网络智能化升级提供理论参考与实践指导。

1 AI在通信网络资源分配中的核心应用场景

1.1 动态资源分配与流量优化

动态资源分配是AI在通信网络中的基础应用场景,其核心目标是根据实时流量需求调整带宽、功率等资源分配比例。传统方法如最大信噪比(Max-SNR)调度易导致边缘用户体验下降,而AI技术通过构建流量预测模型与动态优化算法,实现了资源分配的全局均衡。例如,GoogleDeepMind研发的深度强化学习模型可基于历史流量数据预测基站负载峰值,动态调整频谱分配策略,使基站

能耗降低15%-20%;诺基亚实验室提出的AI驱动型接收机采用深度神经网络替代传统均衡器,在5G毫米波场景中将误码率降低40%,显著提升了高频段信号覆盖半径^[1]。在工业通信网络中,AI驱动的动态资源分配已展现显著优势。某大型制造企业部署的AI优化系统通过实时监测生产线传感器数据,利用LSTM网络预测设备故障风险,动态调整网络带宽分配优先级。测试数据显示,该系统使关键任务传输延迟降低20%,同时将非关键数据丢包率控制在0.5%以内,有效保障了工业控制的实时性与可靠性。

1.2 安全多方计算(SMPC)与隐私保护

在分布式AI计算场景中,数据隐私保护是资源分配的核心挑战。安全多方计算(SMPC)通过密码学协议允许多方在不泄露原始数据的前提下协同完成计算任务,而AI技术可进一步优化SMPC的资源分配效率。例如,在医疗影像分析场景中,多家医院需联合训练诊断模型,但受限于数据隐私法规无法直接共享患者影像。基于AI的SMPC资源分配策略通过加法秘密分享技术将影像数据分割为多个份额,仅在计算节点间传输加密份额,同时利用强化学习算法动态分配计算任务至空闲节点。实验表明,该策略在保证数据隐私的前提下,将模型训练时间缩短30%,资源利用率提升25%。

1.3 网络切片智能切分与资源分配

网络切片技术通过逻辑隔离实现资源按需分配,是5G/6G网络支持多样化业务的关键。AI技术可基于历史流量数据与用户行为模式,动态优化切片划分与资源分配策略。例如,华为提出的AI驱动型网络切片方案采用卷积神经网络(CNN)对网络拓扑进行建模,通过遗传算

法优化切片路径规划,在多约束条件下实现负载均衡。测试数据显示,该方案使切片间资源冲突率降低40%,关键业务(如自动驾驶)的传输延迟稳定在5ms以内。

1.4 算力交易与异构资源调度

随着边缘计算与云计算的融合,算力交易成为AI算力网络的核心商业模式。AI技术通过构建算力资源定价模型与动态调度算法,实现了异构资源的高效匹配。例如,某云服务商部署的AI调度系统采用粒子群优化

$$v_i(t+1) = \omega \cdot v_i(t) + c_1 \cdot r_1 \cdot (pbest_i - x_i(t)) + c_2 \cdot r_2 \cdot x_i(t+1) = x_i(t) + v_i(t+1)$$

其中, ω 为惯性权重,控制粒子保持原有速度的能力; c_1 , c_2 为加速系数,分别调节个体最优与全局最优的吸引力; r_1 , r_2 为随机因子,增加搜索随机性。在通信资

$$F = \sum_{i=1}^N R_i \cdot \log_2(1 + SNR_i) \quad \text{或} \quad F = \sum_{i=1}^N \left(\frac{D_i}{R_i} + \frac{C_i}{F_i} \right)$$

PSO算法的优势在于无需复杂数学推导、易实现且适合嵌入式设备,但其收敛速度受参数设置影响较大,需结合模拟退火等算法进行改进。

2.2 强化学习(RL)算法

强化学习通过智能体与环境的交互学习最优策略,适用于动态资源分配场景。以深度Q网络(DQN)为例,其核心流程包括状态感知、动作选择与奖励反馈:(1)状态感知:智能体实时采集网络状态数据(如节点负载、链路延迟);(2)动作选择:基于Q值函数选择最优动作(如增加某节点带宽分配);(3)奖励反馈:根据动作执行结果更新Q值表,优化后续决策。在AI算力网络中,DQN算法可动态调整任务调度策略以平衡负载^[3]。例如,某边缘计算平台部署的DQN模型通过定义多维度奖励函数(包括任务完成时间、资源利用率与能耗),实现了算力资源的自适应分配。实验表明,该模型在突发流量场景下使任务排队延迟降低35%,同时将服务器空转率控制在10%以内。

2.3 联邦学习(FL)算法

联邦学习通过分布式模型训练保护数据隐私,其核心思想是在本地设备上训练模型参数,仅上传梯度信息至中央服务器聚合。在通信网络资源分配中,FL可联合多个基站共享模型参数,提升全网资源调度效率。例如,爱立信提出的联邦学习框架通过加密梯度传输技术,在保护用户数据的前提下实现了基站间负载均衡策略的协同优化。测试数据显示,该框架使跨基站资源冲突率降低50%,同时将模型训练时间缩短40%。

3 典型案例分析

3.1 工业通信网络动态资源分配

(PSO)算法,根据任务优先级、节点算力容量及网络延迟等参数,动态分配计算任务至最优节点^[2]。实际应用中,该系统使算力资源利用率提升35%,任务完成时间缩短22%,同时降低运营商能耗成本18%。

2 AI驱动资源分配的主流算法原理与实现

2.1 粒子群优化(PSO)算法

PSO算法通过模拟鸟群觅食行为实现资源分配的全局优化。其核心公式包括速度更新与位置更新:

源分配中,粒子位置 x_i 代表资源分配方案(如带宽分配比例),适应度函数 F 可定义为总吞吐量最大化或总延迟最小化:

某汽车制造企业部署的AI优化系统采用LSTM网络预测生产线传感器数据流量,结合PSO算法动态调整网络带宽分配。系统架构包括数据采集层(传感器节点)、AI推理层(边缘服务器)与决策执行层(网络交换机)。实际应用中,该系统使关键任务(如焊接机器人控制)的传输延迟稳定在2ms以内,同时将非关键数据(如环境监测)的带宽占用率控制在15%以下,显著提升了生产线整体效率。

3.2 边缘计算场景下的算力交易

某云服务商推出的AI算力交易平台采用DQN算法实现异构资源调度。平台通过定义多维度奖励函数(包括任务优先级、节点算力容量与网络延迟),动态分配计算任务至最优节点。测试数据显示,在1000个并发任务场景下,该平台使任务完成时间标准差降低60%,资源利用率提升至90%以上,同时将运营商能耗成本降低20%。

4 挑战与未来发展方向

4.1 现有挑战

4.1.1 计算资源限制:算力瓶颈与实时性需求的矛盾

AI模型在通信网络资源分配中的应用高度依赖计算资源,尤其是深度强化学习(DRL)和大规模神经网络模型。然而,边缘设备(如基站、终端设备)的算力有限,难以支撑复杂模型的实时推理。例如,DRL算法中的策略网络和价值网络需频繁更新参数,导致推理延迟增加,无法满足低时延通信需求(如工业互联网、车联网场景)。现有解决方案包括模型压缩(如量化、知识蒸馏)和剪枝技术,通过减少模型参数量或计算复杂度降低延迟,但可能牺牲部分精度。此外,分布式计算框架(如边缘-云协同)虽能分担计算压力,但需解决数据

同步与通信开销问题。

4.1.2 数据隐私保护：隐私计算技术的安全性漏洞

通信网络数据涉及用户位置、行为模式等敏感信息，AI驱动的资源分配需在保护隐私的前提下实现数据共享。联邦学习（FL）通过本地训练、全局聚合的方式避免原始数据泄露，但梯度信息仍可能被逆向攻击（如差分隐私攻击）。同态加密（HE）虽能实现加密数据上的计算，但计算开销大，难以直接应用于实时资源分配^[4]。现有研究尝试结合多方安全计算（MPC）和差分隐私技术，通过添加噪声或分割数据片段增强安全性，但需在隐私保护与模型性能间权衡。此外，跨运营商数据共享场景下的隐私合规性（如GDPR）仍是待解难题。

4.1.3 动态性与稳定性平衡：强化学习算法的适应性挑战

通信网络具有高动态性（如用户移动、流量突发），要求资源分配策略快速适应环境变化。DRL通过试错学习动态调整策略，但在快速变化场景中易出现策略震荡（如频繁切换信道或功率分配），导致资源分配效率下降。经验回放机制通过存储历史经验缓解数据相关性问题的，目标网络则通过固定部分参数稳定训练目标，但二者需结合具体场景调整超参数（如回放缓冲区大小、目标网络更新频率）。此外，模型泛化能力不足（如对未见过网络拓扑的适应性差）也是制约DRL应用的关键因素。

4.2 未来发展方向

4.2.1 多算法融合：构建混合优化框架

单一算法（如DRL、粒子群优化PSO）在全局搜索或局部收敛上存在局限，未来可通过融合多算法优势提升资源分配性能。例如，结合PSO的全局搜索能力与DRL的动态适应能力，设计“PSO初始化+DRL微调”的混合框架：PSO快速定位近似最优解区域，DRL在该区域内精细化优化策略。此外，图神经网络（GNN）可建模通信网络的拓扑结构，与DRL结合实现空间-时间资源的联合分配，适用于大规模异构网络场景。

4.2.2 边缘智能原生设计：硬件级优化资源分配

现有AI模型与通信硬件分离设计导致计算效率低

下。未来6G基站芯片将集成AI加速单元（如NPU），支持模型推理的硬件级优化。例如，通过定制化指令集加速矩阵运算，或采用存算一体架构减少数据搬运延迟。此外，模型量化感知训练（QAT）可在训练阶段模拟量化效果，直接生成适配硬件的低位宽模型，进一步降低推理能耗。边缘智能原生设计将推动资源分配从“软件定义”向“硬件定义”演进，满足超低时延（如亚毫秒级）需求。

4.2.3 数字孪生与强化学习集成：降低现实部署风险

通信网络数字孪生体通过虚拟映射实时模拟网络状态，为DRL提供低成本训练环境。未来可构建“数字孪生预训练+现实环境微调”的两阶段框架：在孪生体中预训练DRL策略，利用其无限试错能力加速收敛；在现实网络中通过少量交互微调策略，避免直接部署导致的性能崩溃。此外，数字孪生可结合数字线程技术，实现物理网络与虚拟模型的双向同步，支持资源分配策略的动态验证与优化。该方向需解决孪生体建模精度（如信道模型、用户行为模拟）与计算复杂度的平衡问题。

5 结语

AI技术为通信网络资源分配提供了从静态规则到动态优化的范式转变。通过粒子群优化、强化学习与联邦学习等算法，AI实现了资源分配的全局均衡、隐私保护与自适应调整。典型案例验证了AI驱动策略在工业通信、边缘计算等场景中的有效性，但需进一步突破计算资源限制与数据隐私保护等瓶颈。未来，随着6G原生AI架构与边缘智能的深度融合，通信网络资源分配将迈向更高水平的智能化与自动化。

参考文献

- [1]张明洁,王丽萍.智能通信网络中的资源分配[J].电脑知识与技术,2025,21(17):90-92.
- [2]张冬冬.基于深度强化学习的通信网络资源分配方法[J].信息与电脑(理论版),2023,35(18):167-169.
- [3]刘晗.基于智能算法的移动通信网络干扰预测与资源分配研究[D].北京邮电大学,2022.
- [4]赵燕萍.基于改进粒子群算法的无线通信网络资源分配研究[J].长江信息通信,2024,37(08):15-17.