



研究与开发

面向5G-Advanced与6G的智能无线电接入网： 关键标准技术与未来演进

赵喆^{1,2}, 陈嘉君^{1,2}, 高音^{1,2}

(1. 移动网络和移动多媒体技术国家重点实验室, 广东 深圳 518055;

2. 中兴通讯股份有限公司, 广东 深圳 518057)

摘要: 5G/5G-Advanced在持续提升关键性能指标方面被寄予厚望, 需要在时延、可靠性、连接数密度与用户体验等方面实现进一步突破。传统以人工操作为主的管理模式在效率、准确性与成本等方面的局限日益凸显。相较于传统优化方法, 人工智能技术凭借其预测性与前瞻性, 推动网络管理由被动应对转向主动感知与自优化, 实现从“监测-响应”到“预判-编排”的迁移。基于3GPP在无线电接入网(radio access network, RAN)智能化方向的关键技术与标准化路径, 结合典型用例场景, 分析了AI/ML模型管理、数据采集与交互机制。面向6G智能RAN, 进一步提出“意图驱动的协作任务”这一新型架构理念, 其关键是通过RAN对应用层信息的感知、任务级别的服务质量(quality of service, QoS)监控、动态组和资源管理等技术实现6G网络人机及碳硅生态系统的无缝交互。

关键词: 智能无线电接入网; 人工智能; 6G; 数据交互; 意图驱动网络

中图分类号: TP393; TN929.5

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2026029

Intelligent radio access network for 5G-Advanced and 6G: key standard technologies and future evolution

Zhao Zhe^{1,2}, Chen Jiajun^{1,2}, Gao Yin^{1,2}

1. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China

2. ZTE Corporation, Shenzhen 518057, China

Abstract: 5G/5G-Advanced is expected to continuously enhance key performance indicators, requiring further breakthroughs in aspects such as latency, reliability, connection density, and user experience. The limitations of traditional human-computer interactive network management, which relies primarily on manual operations, are becoming increasingly evident in terms of efficiency, accuracy, and cost. Compared to traditional optimization methods, the predictive and forward-looking capabilities of AI/ML enable the network to shift from passive response to active perception and self-optimization, achieving a transition from “monitoring-reaction” to “prediction-orchestration”. Based on the key technologies and standardization paths for radio access network (RAN) intelligence defined by 3GPP, the AI/



ML model management, data collection, and interaction mechanisms were analyzed in conjunction with typical use case scenarios. For intelligent RAN in 6G, a novel architectural concept termed the “intent-driven collaborative task” was proposed. Its key implementation relied on RAN's awareness of application-layer information, task-level quality of service (QoS) monitoring, dynamic grouping and resource management, and other technologies to achieve seamless interaction between human, machine, and the carbon-silicon ecosystem in 6G networks.

Key words: intelligent RAN, artificial intelligence, 6G, data exchange, intent-driven network

0 引言

从1G到5G,每一代移动通信技术都通过引入更高频段、更大带宽与更高速率以承载多样化业务需求,涵盖增强型移动宽带(enhanced mobile broadband, eMBB)、超可靠低时延通信(ultra-reliable low-latency communication, URLLC)与海量机器类通信(massive machine-type communication, mMTC)^[1]。相较于传统无线网络,业界对5G-Advanced(5G-A)和6G网络设定了更宏伟的愿景与更高的性能指标需求。此外,智能内生、智能无线环境、通信感知一体化、空天地海一体化等全新特征被认为是6G网络的潜在关键技术特性。然而,这些更高的设计目标与全新技术的引入也带来了诸多挑战。为满足未来多样化业务需求,亟须构建高速、智能的无线通信网络作为承载基础。为此,6G网络将会采用更大规模天线孔径,并将使用包括U6G频段(6.425~7.125 GHz)、中频段(7~24 GHz)、毫米波(millimeter wave, mmWave)等更高频段,这将使得近场传播特性更加显著^[2-3]。同时,智能超表面(reconfigurable intelligent surface, RIS)、超大规模多输入多输出(multiple-input multiple-output, MIMO)、去蜂窝(cell-free)等新技术的引入,也使得未来无线网络中的近场场景广泛存在^[4]。近场通信技术是满足未来6G网络更高的数据速率、高精度的感知,以及物联网无线能量传输等需求的关键使能技术之一,有望成为未来6G潜在无线空口的关键技术之一。然而,包括近场技术在内的多项新技

术也带来了全新的挑战。为此,需要引入高效的智能算法、智能化的物理层方案,以及探索更高频段的频谱资源,以满足下一代网络的性能需求。鉴于当前的无线系统高度依赖数学模型,且网络中产生的海量数据需要高效处理与分析,具备数据驱动预测与前瞻决策能力的人工智能(artificial intelligence, AI)技术将发挥重要作用,推动网络从以往的被动响应范式演进为主动调优的范式。

在无线电接入网(radio access network, RAN)中引入AI,在物理层,可利用AI或机器学习(machine learning, ML)技术应对近场通信引入的新需求,实现信道状态信息(channel state information, CSI)的测量与上报压缩^[5]、波束预测^[6]及高精度定位^[7]等功能;在高层,则可实时从海量测量数据与业务信息中挖掘规律与趋势,完成流量与小区负荷预测、智能资源编排及干扰抑制,从而缓解网络拥塞,提升系统整体性能、服务质量(quality of service, QoS)和体验质量(quality of experience, QoE)。面对连接设备的激增与行业应用的纵深发展,AI技术还可用于构建能耗建模与实现能效调度,在保障网络性能的同时降低系统能耗。进一步地,基于用户与终端的多源数据,AI能识别体验瓶颈并触发闭环优化机制,从而实现以用户体验为中心的精细化网络运营。目前,3GPP正在讨论智能化接入网的相关用例^[8-10]。所谓RAN智能化,是指通过运用AI/ML及大数据分析等技术,全面提升RAN在运营、管理与运维(OAM)方面的效率。借助这些智能技术,RAN将逐步具备自学

习、自优化与自愈等能力，从而更好地应对复杂的动态网络环境，并满足日益增长和多样化的用户需求。

智能RAN通常从两个维度加以刻画：面向网络的AI（AI for networks, AI4Net）^[11]与面向AI的网络（networks for AI, Net4AI）^[12]。AI4Net指利用AI技术来提升与优化无线网络的性能与管理能力，实现自动化运维、智能优化与网络自愈等功能^[13]。具体而言，AI4Net采用机器学习模型预测业务流量、动态分配网络资源，以及对网络状态进行实时监测并快速响应异常，从而提升网络运行效率、系统稳定性与用户体验。Net4AI则强调由网络为AI应用赋能，即通过网络侧保障AI应用所需的各项网络性能^[14]。其重点在于构建高带宽、低时延与高可靠的网络基础设施，以支持实时数据分析、边缘计算、分布式学习以及大规模智能设备互联等复杂AI应用场景。通过优化网络架构与传输机制，Net4AI能够为各类AI工作负载提供坚实支撑，促成更高效、更智能的服务。

现阶段，3GPP标准体系下的智能化RAN主要聚焦于单节点AI/ML模型的处理能力，以支撑各类优化用例，而在多节点协同智能方面尚缺乏系统性研究与标准化支撑。6G网络以“智能内生”为目标，即AI技术与通信协议、网络架构与算网资源实现协同设计，使模型的全生命周期（包括训练、更新、下发、存储）成为网络内建能力，空口与网元支持原生的学习与自治闭环，从而实现跨层、跨域、跨场景的持续优化与网络自我演进^[15-16]。为此，本文系统梳理了3GPP标准体系中智能化接入网的最新进展，涵盖智能接入网框架与高层用例。在此基础上，进一步提出面向6G的智能网络架构设想，并引入意图驱动的协作任务这一新场景，阐述其关键技术要点。

1 智能RAN关键技术

在3GPP标准化工作中，智能RAN主要聚焦于构建智能化的功能框架，并基于具体用例实现智能化解决方案的标准化^[5]。目前被标准化的AI/ML用例包含网络节能^[5]、网络负荷均衡^[5]、移动性优化^[5]、网络切片^[7]，以及覆盖与容量优化^[7]等。为支持RAN的智能化，基站需要收集来自邻区基站的数据，包括预测信息、反馈信息及来自UE的测量数据。针对AI/ML模型的部署场景，模型训练和模型推理可以部署在网络的各个节点上。

智能化RAN框架如图1所示。该框架展示了AI/ML在RAN中的功能模块及其交互关系。框架包括数据收集、模型训练、模型推理、执行等关键模块，以及它们之间的数据交互和反馈机制。

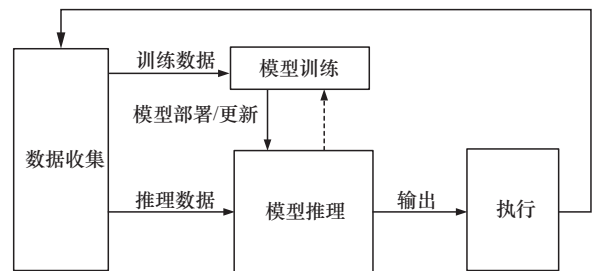


图1 智能化RAN框架

(1) 数据收集模块。该模块负责为AI/ML模型训练和推理提供所需的输入数据。需要指出的是，该功能本身不执行与具体AI/ML算法相关的数据准备工作，如数据预处理、清洗、格式化或转换等。其输入数据来源包括用户设备或网络实体的测量信息、执行模块的反馈信息，以及其他AI/ML模型的输出结果。

(2) 模型训练模块。该模块负责执行AI/ML模型的训练、验证与测试，并在测试过程中生成模型性能指标。在需要时，该模块还会对数据收集模块提供的训练数据进行数据准备处理，包括数据的预处理、清洗、格式化和转换等操作。



(3) 模型部署/更新。该机制用于将已训练、验证和测试的模型部署至模型推理模块，或用更新后的模型代替原有模型。

(4) 模型推理模块。其任务是根据推理数据生成模型预测结果或决策输出。在需要时，该模块也会对输入数据执行预处理、清洗、格式化或转换等操作。此外，在适用场景下，模型推理模块还可向模型训练模块反馈模型性能信息。这些推理输出的具体内容取决于所应用的具体用例。

(5) 模型性能反馈。模型性能反馈可用于监测 AI/ML 模型的性能表现。

(6) 执行模块。该模块负责接收模型推理模块的输出结果，并触发或执行相应动作的功能实体。执行器可触发针对其他实体或自身的动作。

(7) 反馈。执行过程中产生的反馈信息可用于生成训练数据或推理数据，也可通过更新关键性能指标来监测 AI/ML 模型性能及其对网络的影响。

1.1 AI功能部署方式

基于 AI/ML 的功能模块可灵活部署于 RAN 中的不同网络节点，具体的部署方案将影响相应的信令交互流程^[4]。根据网络架构和功能划分的不同，AI/ML 在 RAN 中可采用集中式或分布式的部署策略。现行标准支持的典型部署场景如下。

(1) 在集中式基站架构下：AI/ML 模型训练功能部署于 OAM，而模型推理功能部署于下一代无线电接入网（next generation radio access network, NG-RAN）节点；或 AI/ML 模型训练与推理功能均部署于 NG-RAN 节点内。

(2) 在分离式基站架构下：AI/ML 模型训练功能部署于 OAM，而模型推理功能部署于 gNB 中央单元（gNB central unit, gNB-CU）节点；或 AI/ML 模型训练与推理功能均部署于 gNB-CU 节点内。

AI/ML 模型训练功能部署于 OAM，模型推理功能部署在基站节点，其信令交互如图 2 所示。AI/ML 模型训练功能和模型推理功能均部署在基站节点，其信令交互如图 3 所示。

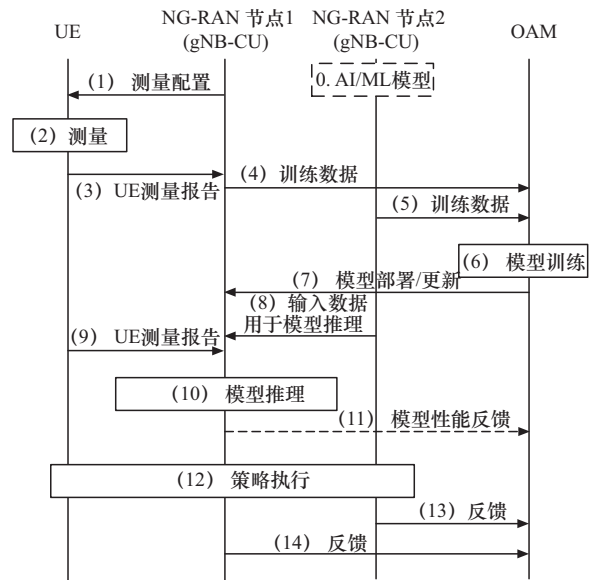


图2 AI/ML 模型训练功能部署于 OAM，模型推理功能部署在基站节点

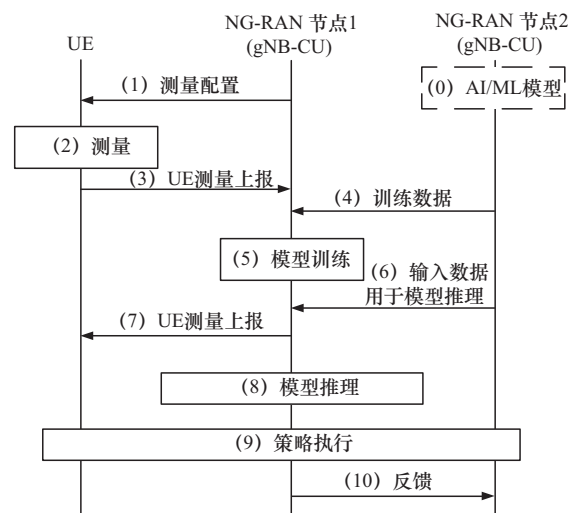


图3 AI/ML 模型训练功能和模型推理功能均部署在基站节点

在 OAM 侧训练完成的模型可部署在基站。基站通过收集的用户设备（user equipment, UE）侧数据或者邻区基站数据，根据具体用例需求执行模型推理，并基于模型推理结果生成相应的网络优化策略。此外，基站还可以进一步收集 UE

侧数据或者邻区基站数据，用于模型评估和性能评估。

基站也可基于收集的UE侧数据和邻区基站数据，进行在线训练或者离线训练。训练完成后，根据用例需求进行模型推理，生成相应的网络优化策略。在此部署场景中，模型评估都在基站侧完成。

1.2 AI/ML辅助网络优化用例

集成AI与5G的显著优势在于能够优化网络性能。AI算法可对海量数据进行实时分析，从而识别出可用于提高网络效率的模式和趋势。例如，AI可用于预测流量与小区负荷，并据此动态地分配资源，确保用户在需要时能够获得足够的带宽。这有助于缓解网络拥塞，提升整体服务质量。典型应用包括基于AI的负载均衡技术、覆盖和容量增强技术等。

将AI与5G集成的另一个优势是能够更有效地管理网络资源。随着连接设备数量的持续增长、交互应用的普及以及垂直行业业务的快速发展，网络需要在有效分配资源的同时，实现对网络能耗的有效管理，从而降低能耗。相关技术包括基于AI的节能技术与切片资源管理技术等。

此外，AI还可用于提升用户在无线网络中的体验。通过分析来自用户和设备的数据，AI能够识别可以改进的领域并实施改进，以提升用户整体体验。例如，基于AI的移动性优化技术便属于此类应用。下面将对上述具体用例进行阐述。

1.2.1 网络负荷均衡

随着业务流量的快速增长和多频段网络的并行使用，网络流量的均衡分布成为关键挑战。为此，在网络侧引入负载均衡策略，通过优化切换参数和用户行为，实现小区间及区域间的负载均衡分布、流量转移或用户卸载，从而提升网络性能、用户体验和系统容量，同时降低人工干预需

求。然而，该优化方法面临两大问题：一是其决策依赖当前或历史负载信息，容易引发乒乓切换、小区过载和服务质量下降；二是难以实现整体性能优化，如用户卸载可能导致目标小区负载激增。为应对上述问题，可引入AI/ML模型，通过收集UE、网络节点测量数据和历史反馈信息，实现对负载的预测与均衡策略的优化，从而提升系统容量和用户体验。

为实现基于预测的负载均衡决策，基站需要收集多类AI/ML输入数据，包括本地节点的当前及预测资源状态、UE轨迹预测、当前及预测UE业务流量，以及邻近基站节点的预测资源状态信息；UE信息，如位置坐标、服务小区ID、移动速度、移动历史信息，以及测量报告；邻近节点的当前和预测资源状态，以及在执行流量卸载时的UE性能测量数据。基于AI/ML的负载均衡模型可以输出如下信息：负载均衡所选定的目标小区、自身和邻近基站节点的预测资源状态信息，以及被预测选择切换到目标基站节点的UE。为持续优化模型性能，系统将从基站节点收集反馈信息，包括目标基站节点的UE性能数据（针对那些从源基站节点切换而来的UE）、资源状态信息的更新，以及系统关键性能指标（key performance index, KPI）。

标准化解决方案要求基站间进行多维度的预测性资源状态信息交换，如预测资源状态、活跃UE数量和RRC连接数，以支持模型的训练与推理过程，并为动态资源调度与策略优化提供数据基础。在完成负载均衡策略后，系统将收集反馈的数据用于模型性能评估与闭环优化，包括UE性能指标，如平均上下行吞吐量、时延和丢包率，用以反映卸载后用户的实际体验；当前资源状态信息，用于与预测值进行对比，以评估模型的预测准确性。

1.2.2 网络节能

为满足5G网络关键性能要求和移动用户增



长的需求，全球范围内正大规模部署基站，规模已达数百万级别。这一趋势也带来了高能耗、二氧化碳排放量增加和运营成本增加等问题，因此节能已成为5G网络的重要应用场景之一^[16]。小区激活/去激活是一种典型的空间域节能方案。其基本原理是在业务量低于阈值时关闭相应小区，并将UE卸载到目标小区。此外，还可通过降低负载、修改覆盖范围或调整RAN配置等方式实现能耗控制。最佳决策通常依赖于RAN节点负载、节点能力、KPI/QoS要求、活跃UE数量与移动性、小区利用率等因素。然而，识别高效的节能操作并非易事。错误地关闭小区可能影响网络性能，并增加额外的流量负担。当前节能机制易受负载预测不准、系统性能与能效目标冲突、参数配置僵化以及局部优化导致全局退化等问题影响。为解决上述问题，可引入机器学习技术，通过分析RAN数据，预测下一周

期的能效和负载状态，从而优化小区激活/去激活决策，实现对节能策略的动态配置，有效降低系统能耗。

基站之间交换能耗信息有助于了解邻近基站状态。当某一个基站实施如小区激活/去激活等节能策略时，不仅需要考虑自身能耗的降低，还应评估该策略对节能区域总能耗的影响。因此，系统需收集邻近基站的实际能耗数据，预测策略执行后的能耗变化，以支持全局优化的决策过程。能耗值及映射规则示意图如图4所示。在示例场景中，gNB 0负责容量管理。为通过关闭部分小区实现节能，gNB 0考虑将部分流量卸载到邻近节点gNB 1、gNB 2、gNB 3。在此过程中，gNB 0需要从相邻基站获取一致的能耗数据，并通过比较流量卸载前后区域的总体能耗，做出最终节能决策。

本文给出了基于AI/ML的节能方案仿真结

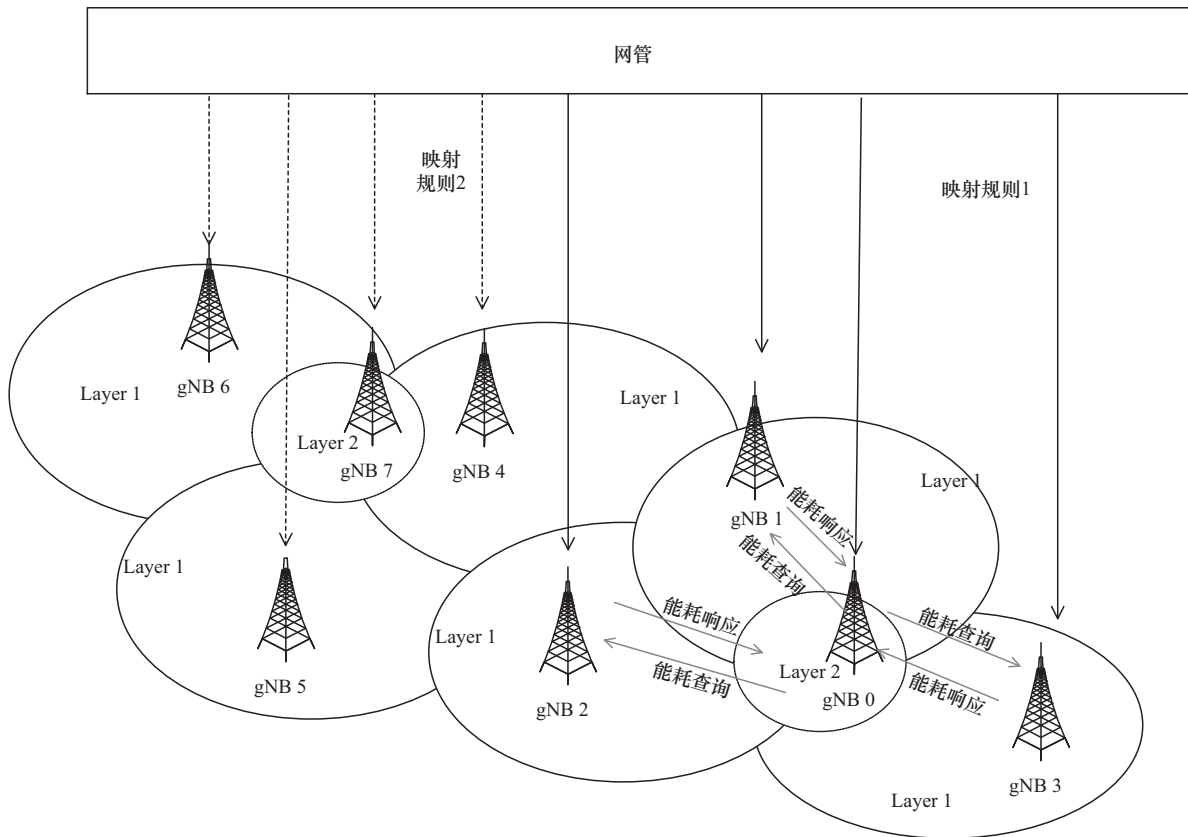


图4 能耗值及映射规则示意图

果。测试区域位于中国广州市番禺区，共包含 54 个有源天线单元 (active antenna unit, AAU)，节能策略基于负载预测阈值触发^[16]。为对比“无节能”“传统节能”与“基于 AI/ML 的节能”3 种方案的性能，设置了 3 个时间区间进行评估：T0 表示未启用任何节能策略的基准阶段，T1 表示采用传统节能策略的阶段，T2 表示采用 AI/ML 技术的节能策略阶段。节能配置参数见表 1。

表 1 节能配置参数

时间区间	类型	节能策略
T0 2022-06-09 至 2022-06-15	未启用	w/o 通道关断 w/o 符号关断 w/o 深度休眠
T1 2022-05-27 至 2022-06-22	传统节能	通道关断 符号关断 深度休眠
T2 2022-06-25 至 2022-07-01	AI/ML 辅助	通道关断 符号关断 深度休眠

3 种方案的节能策略能效统计见表 2。结果显示，深度休眠与通道关断的持续时间显著延长，会带来更高的节能收益。具体而言，引入 AI/ML 后，这些策略的关断时长进一步增加，使功耗降低至 452.18 W，相较传统方法能效提升 2.48%。上述结果表明，AI/ML 在节能场景中具有显著优势，尤其体现在延长节能策略持续时间、进而实现更可观的能耗降低方面。

3 种节能方案的电源平均功率如图 5 所示，5G 能效比如图 6 所示。由图 5 可知，基于 AI 的节能策略，可使电源平均功率下降 23.87%；图 6 则表明，在采用基于 AI 的节能策略时，5G 能效提升达到 23.40%。

表 2 节能策略能效统计

阶段	深度休眠时长/h	通道关断时长/h	符号关断时长/h	功耗/W	5G 能效比/ (GB·kW ⁻¹ ·h ⁻¹)	相较于无节能能效增益
T0-无节能	0	0	0	593.97	2.78	—
T1-基站节能	2.61	0.58	9.63	466.92	2.97	6.55%
T2-AI 节能	3.48	5.94	9.03	452.18	3.44	23.40%

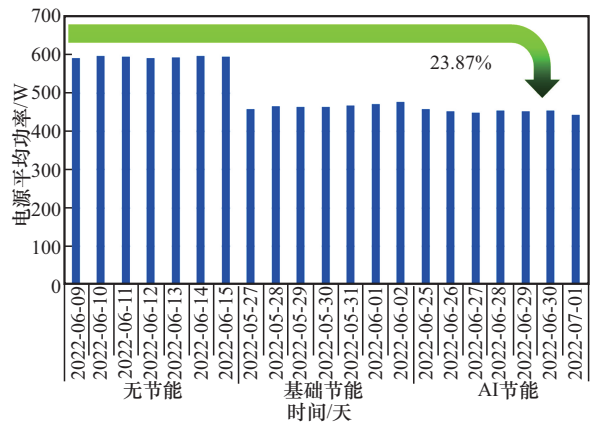


图 5 3 种节能方案的电源平均功率

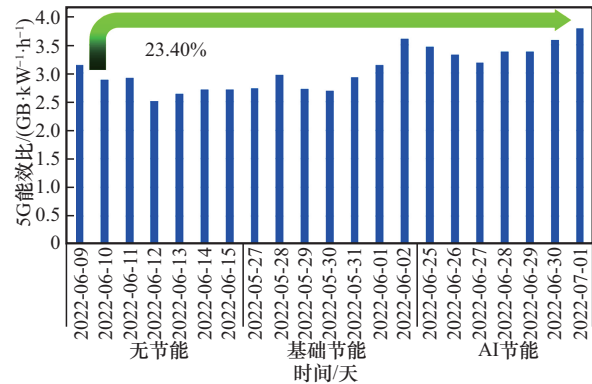


图 6 3 种节能方案的 5G 能效比

1.2.3 移动性优化

移动性管理旨在保障终端在移动过程中的业务连续性，其核心目标是降低掉话率、减少无线链路失败次数、避免不必要的切换和抑制乒乓切换。在未来高频段组网环境下，节点覆盖范围缩小将导致 UE 切换频率增加，这一问题在高速移动场景下尤为突出。同时，对 QoS 要求严苛的应用对切换性能较为敏感，任何失败或延迟都可能直接影响用户体验，因此需要尽可能避免切换失败并降低切换时延。然而，传统的基于试错的优



化方法难以实现零失败率的切换过程。切换失败常引起丢包或额外时延，这对时延敏感的应用来说是不可接受的。此外，依赖反馈调整的方式受环境随机性影响，优化效果有限。利用AI/ML技术增强移动性管理能力，可以从以下几个方向实现优化，包括降低意外事件概率，避免过早/过晚切换或错误小区选择；预测UE位置、移动性和性能表现。其中，基于位置轨迹的预测有助于优化资源管理和目标小区选择，从而实现效率提升与网络性能的优化。

在AI/ML驱动的移动性优化场景中，系统通过增强数据驱动能力实现精准的切换决策。与负载均衡等用例不同，此类场景特别引入UE历史位置信息作为关键输入特征。源基站结合UE历史位置和移动模式，利用AI/ML模型进行路径建模与趋势分析，进而推理其未来轨迹，包括UE未来驻留小区列表及在每个小区的预计驻留时长。该预测仅传至首跳目标基站节点。为准确评估轨迹预测模型的有效性，源基站通过新基站间的数据收集机制，从目标小区或后续服务小区获取UE的实际轨迹信息，包括已驻留的小区列表及实际驻留时间。源基站通过对比预测轨迹与实际轨迹，评估AI/ML模型的有效性和性能。

1.2.4 网络切片

网络切片是一种能够依据客户差异化需求，提供定制化服务保障的技术架构。借助该技术，移动网络运营商可以将客户划分为不同租户类型，为每类租户设定差异化的服务需求，并依据服务水平协议（service level agreement, SLA）和订阅策略，确定其可使用的切片类型。在此过程中，基站扮演着关键角色，需要在移动性管理、负载均衡及无线资源管理等方面做出相应决策，以满足各类切片所对应SLA的服务要求。AI/ML技术可通过分析网络侧及UE侧的各类性能指标，辅助实现网络切片场景下的智能资源管理与移动性管理策略的优化，从而更好地保障各类切片的服务质量，满足差异化需求。

为实现基于AI/ML的网络切片优化决策，基站需要获取以下输入信息：本地节点提供每条网络切片的无线资源状态和可用容量（测量或预测值），以及UE轨迹预测；相邻基站提供每条切片的无线资源状态和可用容量（测量值或预测值），包括小区UE性能的UE测量报告。基站中的AI/ML网络切片模型输出包括：每条切片的无线资源状态和可用容量预测、无线资源管理策略范围内的资源管理决策，以及具备切片感知的移动性决策。为优化模型性能，从基站采集反馈如每条切片的资源状态和可用容量测量值、针对从源基站切换UE的性能反馈，以及精细度UE性能反馈，以评估特定切片下的行为差异和体验质量。标准化解方案除小区级别的资源状态预测和UE性能测量外，还依赖切片级别的资源状态预测、切片级别的可用容量预测和UE性能测量指标，这些切片级数据可通过基站间的数据收集流程交互，类似于AI/ML负载均衡场景机制。

1.2.5 覆盖与容量优化

覆盖与容量优化（coverage and capacity optimization, CCO）功能旨在检测并解决（或缓解）CCO问题。基站可自主调整覆盖配置，并在变更时通知相邻节点，告知其涉及的小区及同步信号块（synchronization signal block, SSB）列表。传统CCO采用被动优化策略，仅在基站（在CU-DU架构下为gNB-CU）检测到问题影响性能后，才由gNB-DU生成覆盖配置。引入AI/ML技术后，CCO转向主动优化策略，通过模型预测潜在覆盖或容量问题，提前采取预防措施，从而避免问题发生或限制其进一步发展，减少对网络性能和用户体验的影响。

在基于AI/ML技术的CCO场景中，源基站通过模型预测未来可能出现的CCO问题，并生成小区或SSB粒度的覆盖调整方案，通过基站配置消息发送给邻区基站。邻区基站据此调整自身运行状态，实现协同优化。为支持预测与策略制

定，基站需要获取多源输入数据，包括本地基站的测量或预测无线资源状态，当前小区/SSB级别CCO状态，邻区基站的测量或预测资源状态，以及UE的测量报告。AI/ML模型输出包括预测的小区/波束级别CCO问题及未来小区/SSB级别CCO状态。为提升模型准确性，通过反馈已测资源状态和UE性能等形成闭环优化。标准化解决方案中，源基站通过AI/ML预测CCO问题，如覆盖盲区或容量瓶颈，生成未来配置状态，并通过消息发送给邻区基站；邻区基站确认并调整，实现联合优化。基站间协同交互支持AI/ML CCO场景如图7所示。在CU-DU分离架构中，gNB-CU基于全局信息预测问题及时间，通过消息发送给DU；DU生成未来CCO配置，通过消息反馈给CU，完成协同优化。基站内协同交互支持AI/ML CCO场景如图8所示。

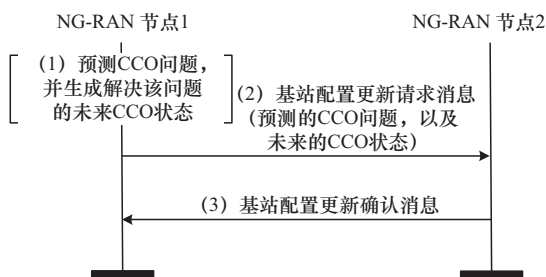


图7 基站间协同交互支持AI/ML CCO场景

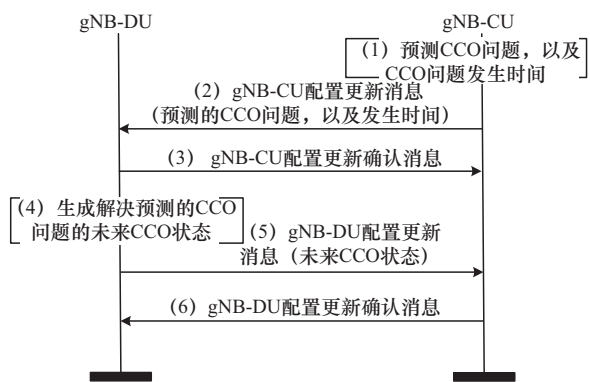


图8 基站内协同交互支持AI/ML CCO场景

2 6G 智能 RAN

2023年6月12日—22日，国际电信联盟无线

电通信部门5D工作组（ITU-R WP5D）第44次会议在瑞士日内瓦召开。会议审议通过了《IMT面向2030及未来发展的框架和总体目标建议书》（以下简称《建议书》）。作为6G发展的纲领性文件，《建议书》汇聚了全球对6G愿景的共识，明确了6G的发展目标与演进趋势，并系统提出了6G的典型应用场景及能力指标体系。《建议书》提出，面向2030及未来的6G系统将推动实现包容性、泛在连接、可持续性、创新、安全性与隐私性保障、标准化和互操作、互通性七大目标，支撑构建更加包容的信息社会，助力实现联合国可持续发展目标。

4G 改变生活，5G 改变社会，6G 将变革世界。6G 将持续深化人、机、物的互联互通，推动物理世界和虚拟世界的深度融合。6G 有望将感知和人工智能等能力融合到网络中，成为承载新用户、赋能新应用的新型数字基础设施。在5G三大场景基础上，6G进一步拓展与增强，形成六大典型场景，包括沉浸式通信、超大规模连接、极高可靠低时延、人工智能与通信融合、感知与通信融合、泛在连接等。与此同时，近年来AI技术飞速发展，特别是大模型技术的突破催生了AI代理（AI agent）、代理式AI（agentic AI）等新概念，相关方向正迅速成为全球信息通信领域的技术热点。

可以预见的是，6G网络的演进将展现出前所未有的自治性、适应性和智能水平。这一转变的核心在于代理式AI。这是一种新型范式，其中AI系统不再只是作为被动工具，而是作为主动、自治的代理，在复杂无线环境中具备感知、推理、决策和协作能力。与主要针对预定义任务进行推理的传统AI模型不同，6G中的代理式AI能够动态规划、与其他AI代理通信，并持续调整策略，以优化网络性能和用户体验。特别值得一提的是，这种代理式AI不仅适用于机器人，还可广泛部署于各类终端智能设备。



3GPP SA1 已就 AI 代理的定义及其相关要求展开讨论，具体涉及 AI/计算能力、AI 代理和 AI 相关协作任务等方面。

AI 代理指一种自动化智能实体，能够与环境进行交互，获取上下文信息，并具备推理、自主学习、决策及执行任务（可独立完成或与其他 AI 代理协作）的能力，以实现特定目标。

AI 服务作为一种 6G 服务（如 AI 模型推理），由 6G 核心网络向 6G 用户提供，旨在支持各类 AI 的相关活动，并考虑服务质量要求，如推理准确性与时延等。

基于 3GPP SA1 在技术报告 TR 22.870 中确定的用例，应用上的 AI 代理可通过意图调用 3GPP 服务（如通信、感知、AI 功能）或协调其他 AI 代理；而部署于 6G 网络中的 AI 代理，则可基于接收到的意图请求，向 UE 或第三方提供 3GPP 服务。代理式 AI 则更注重通过具体物理实体实现意图生成与任务执行、多代理任务协同和目标导向实时优化及反馈等功能。借助 6G 无线通信系统，AI 代理可真正实现用户意图的端到端实施。AI 代理意图驱动的协作任务示意图如图 9 所示。例如，分布于 RAN、核心和边缘域的 AI 代理可自主协商资源分配、预测用户移动模式或动态调整能耗，同时与运营商定义的政策和终端用户意图对齐。

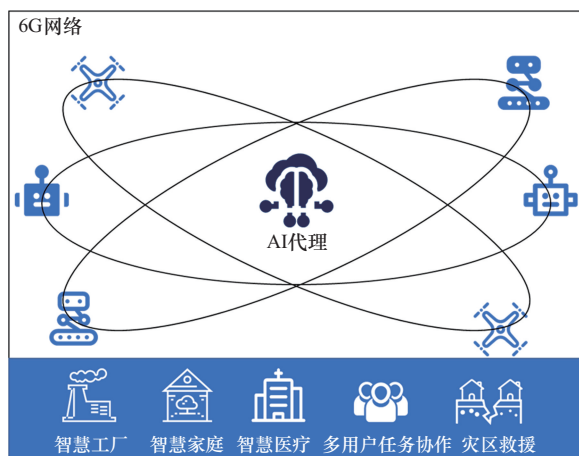


图9 AI代理意图驱动的协作任务示意图

基于代理式 AI 的理念，6G 系统中提出了“意图驱动的协作任务”作为新用例。意图可以为用户与网络交互提供一种新型方式。为确保对用户意图理解的一致性、实现更好的网络资源管理和保障用户体验，RAN 侧需要获取应用层提供的辅助信息。如果意图涉及多个设备，确保设备间的协调和一致性同样重要。6G 网络需要充分理解并解释意图，将其转化为每个参与设备的具体服务和资源分配，监控任务进度，并根据需要动态调整资源。为实现意图驱动的协作任务，6G 智能化接入网需要满足以下关键要求。

(1) 任务级别的 QoS 监控：在任务驱动的智能协作组服务中，应用侧可以向网络发送任务级别的 QoS 配置和需求信息，这可能包括针对特定任务的 QoS 配置、状态监控和性能预测。此外，系统还需要考虑涉及多个组成员的多个 QoS 流的同步和协调。

(2) RAN 对应用层信息的感知：为使 RAN 充分理解和解释用户意图，并将其转化为每个参与设备的具体服务和资源分配策略、监控任务进度并动态调整资源，RAN 侧需要获取应用层提供的辅助信息（如初始成员组信息、设备能力等），以确保对用户意图理解的一致性，并借助实时无线测量技术更好地保障基于用户意图的任务执行。

(3) 任务级别的动态组和资源管理：RAN 应支持动态管理任务组成员（如添加/移除成员、重新配置成员等），并能够选择最合适的授权成员以执行任务。此外，系统还需支持根据需求为不同成员配置差异化的传输路径和资源编排策略。

3 结束语

3GPP 在 AI/ML 高层标准化方面聚焦于 5G 及未来网络的优化，涵盖网络负载均衡、网络节能、移动性优化、网络切片以及 CCO 等典型场景。通过引入基于 AI/ML 的模型，网络能够预测负载、资源状态和 UE 轨迹，实现主动优化决策，

如均衡流量分布、动态小区激活/去激活、减少切换失败、切片级资源分配及预测潜在覆盖问题。标准化解决方案包括基站间数据交换、基于反馈的闭环模型优化,以及在CU-DU分离架构下的协同配置更新。这些机制有效减少了人工干预需求,提升了网络性能、能效和用户体验。在6G演进中,本文进一步提出将基于AI代理的意图驱动协作任务作为新用例,可广泛应用于智能工厂、智能家居和智能医院等场景。意图交互机制为用户与网络提供了新型协作方式,RAN通过获取应用层辅助信息,确保对用户意图的一致理解,进而实现精准的资源管理和可靠的性能保障。

参考文献:

- [1] Wang C X, You X H, Gao X Q, et al. On the road to 6G: visions, requirements, key technologies, and testbeds[J]. IEEE Communications Surveys & Tutorials, 2023, 25(2): 905-974.
- [2] FuTURE Forum. 6G near-field technologies white paper[R].2024.
- [3] Zhao Y J, Dai L L, Zhang J H, et al. Near-field communications: characteristics, technologies, and engineering[J]. Frontiers of Information Technology & Electronic Engineering, 2024, 25(12): 1580-1626.
- [4] 赵亚军. 基于智能超表面的6G近场网络构建[J]. 移动通信, 2024, 48(4): 2-11.
Zhao Y J. Reconfigurable intelligent surface constructing 6G near-field networks[J]. Mobile Communications, 2024, 48(4): 2-11.
- [5] Xiao H, Wang Z Q, Tian W Q, et al. AI enlightens wireless communication: analyses, solutions and opportunities on CSI feedback[J]. China Communications, 2021, 18(11): 104-116.
- [6] Sun C, Zhao L, Cui T, et al. AI model selection and monitoring for beam management in 5G-Advanced[J]. IEEE Open Journal of the Communications Society, 2024, 5: 38-50.
- [7] Alawieh M, Kontes G. 5G positioning advancements with AI/ML[PP]. arXiv preprint, 2023, arXiv: 2401.02427.
- [8] 3GPP. Study on enhancement for data collection for NR and EN-DC: TR 37.817[R]. 2022.
- [9] 3GPP. Study on artificial intelligence (AI)/machine learning (ML) for NR air interface: TR 38.843[R].2023.
- [10] 3GPP. Study on enhancements for artificial intelligence (AI)/machine learning (ML) for NG-RAN: TR 38.743[R].2023.
- [11] Gao Y, Chen J J, Li D P. Intelligence driven wireless networks in B5G and 6G era: a survey[J]. ZTE Communications, 2024, 22(3): 99-105.
- [12] Chen J J, Gao Y, Liu Y S, et al. Toward 6G technology: intent-driven autonomous intelligent wireless communication network[C]//Proceedings of the 2024 IEEE Globecom Workshops (GC Wkshps). Piscataway: IEEE Press, 2025: 1-5.
- [13] Chen M Z, Challita U, Saad W, et al. Artificial neural networks-based machine learning for wireless networks: a tutorial[J]. IEEE Communications Surveys & Tutorials, 2019, 21(4): 3039-3071.
- [14] Zhu G X, Liu D Z, Du Y Q, et al. Toward an intelligent edge: wireless communication meets machine learning[J]. IEEE Communications Magazine, 2020, 58(1): 19-25.
- [15] NGMN. Network architecture evolution towards 6G[R]. 2025.
- [16] Drampalou S F, Uzunidis D, Vetsos A, et al. A user-centric perspective of 6G networks: a survey[J]. IEEE Access, 2024, 12: 190255-190294.

[作者简介]



赵喆 (1974-), 男, 中兴通讯股份有限公司无线研究院副院长, 重点研究方向包括基带芯片总架构设计、通信基站IT化架构设计、无线关键算法和系统综合方案设计等。



陈嘉君 (1995-), 男, 中兴通讯股份有限公司工程师, 主要研究方向为人工智能与移动通信。



高音 (1980-), 女, 中兴通讯股份有限公司正高级工程师、标准预研总工程师, 3GPP RAN3 原主席, CCSA TC624 WG4 副组长, 主要从事5G/6G无线网络创新技术研究 and 标准推进工作。