



5G 网络 AI 应用典型场景技术 解决方案白皮书



中国移动通信集团有限公司

2024 年 6 月

前言

PREFACE

新一轮科技革命和产业变革深入发展，信息和能量融合创新纵深推进，数据、算力、AI 共同构成新质生产力的重要驱动因素。

随着业务应用及场景的多样化发展，5G 通信技术面临信道模型更复杂、数据处理更庞杂、网络需求差异大、跨域融合要求高等挑战。AI 技术具备的海量数据分析能力、复杂任务处理能力、灵活适配能力、智能决策能力，可以有效增强 5G 网络能力。

中国移动提出“5G 网络 AI 应用”技术方案，旨在将 5G 通信技术与 AI 应用相融合，构建“内生智能”的核心能力，解决 5G 网络发展的重大挑战，实现网络提质增效，推动网络向以智为先方向演进，目前已在基于业务的体验升级、智能网络优化、智能网络节能、5G+AI 用户画像、基于客户及特定场景的资源保障等方向进行重点布局，后续将驱动“+AI”向“AI+”跨越式发展，推动 5G 网络与 AI 技术双向赋能，助力经济社会数智化转型。

本白皮书由中国移动计划建设部委托中国移动研究院编制，联合华为、中兴、中信科移动、爱立信、诺基亚、高通、MTK、小米等公司共同撰写。

本白皮书版权归中国移动所有，未经授权，任何单位或个人不得复制或拷贝本建议之部分或全部内容。

目 录

1. 5G 网络 AI 应用简介	1
1.1 发展背景	1
1.2 技术挑战	1
1.3 技术原理	2
1.4 发展规划	5
2. 5G 网络 AI 应用高价值场景及解决方案	6
2.1 无线网络 AI 应用	6
2.1.1 无线网络 AI 应用技术架构	6
2.1.2 基于业务的体验升级场景	7
2.1.3 智能网络优化场景	11
2.1.4 智能网络节能场景	20
2.2 核心网智能控制面 AI 应用	22
2.2.1 核心网智能控制面 AI 应用技术架构	22
2.2.2 基于业务的体验升级场景	23
2.2.3 5G+AI 用户画像场景	24
2.2.4 基于客户及特定场景的资源保障场景	26
3. 5G 网络 AI 技术演进愿景	27
3.1 无线网络 AI 技术演进	28
3.2 核心网智能控制面 AI 演进	29
3.3 终端及芯片 AI 技术演进	31
4. 总结与展望	33
缩略语列表	35
参考文献	36
参编单位及人员	39

1. 5G 网络 AI 应用简介

1.1 发展背景

随着移动网络应用新业务的多样化和蓬勃发展，从传统语音、数据、视频等业务到新兴裸眼 3D、XR 等业务的高速率低时延要求，以及行业应用对大上行、精准定位等的新质要求，运营商对用户业务质量的多样性保障提出了更高的期望。与此同时，AI 技术的快速发展，AI 在数据分析、模型构建、智能决策等领域的能力持续增强，也将会加速 5G+AI 的融合，发展趋势将智能化由外挂叠加变成智能内生一体。网络中的 AI 应用将泛化升级，由用例为中心走向智能化能力为中心，实现业务、体验和网络规建维优的全面升级。

未来三年，全球 91% 的运营商计划将网络智能化纳入其战略并持续投资，5G+AI 将加速运营商智能化。在过去不久的 2024 年世界通信大会上，很多电信运营商也展示了他们在移动网络智能化方面的实践，如德国电信 Deutsche Telekom、法国电信 Orange 等运营商展示了关键智能化网络 AI 应用，涉及运维和网络优化等方面，聚焦网管层面实现 AI 应用，旨在提高人工效率，进而提升网络性能。

与此同时，作为技术领航者的国内外标准或行业组织，如 3GPP (3rd Generation Partnership Project, 第三代合作伙伴计划)、GTI (Global TD-LTE Initiative, TD-LTE 全球发展倡议)、TM Forum (TeleManagement Forum, 电信管理论坛)、ETSI (European Telecommunications Standards Institute, 欧洲电信标准化协会)、CCSA (China Communications Standards Association, 中国通信标准化协会) 等，都已经在各自负责的技术领域上启动了 AI 智能化的标准化工作，并且积极通过联合会议、联络函等形式进行了跨标准组织的技术分享。

1.2 技术挑战

5G 及 5G-A 可提供更高的数据传输速率、更低的延迟、更广泛的连接、更强大可靠的性能以及更加极效智能的网络，满足日益增长的数字化需求，对网络提出新的技术挑战：

1、信道模型更加复杂。在网络覆盖到工业封闭园区、低空等特殊场景后，传统覆盖地面的信道需要延展到密集大型封闭空间、低空等，造成信道模型复杂度提升；

2、数据处理更庞杂。除了传统的通信性能指标、基站及核心网数据等，还需持续引入业务体验数据、用户行为数据等，此外数据粒度也会从 15 分钟级统计细化到 1 分钟级，个别特殊场景到秒级数据，该变化会引来数据分析量成倍增加；

3、网络需求差异大。较之 4G 网络主要关注下行速率，5G 网络还需要关注时延、接入用户数、SLA 保障等多种网络要求，不同业务、场景对网络需求差异化大。

4、跨域融合要求高。从传统通信方案只关注单域能力，伴随网络演进水平维度需要考虑网络、终端及业务的协同，垂直维度需要考虑感知能力、业务特征等因素与通信能力的匹配。

基于以上挑战和发展趋势，引入 AI 技术，充分利用 AI 技术优势。在海量数据分析方面，AI 本身包含大量数据处理算法，从中提取有价值的信息和规律，提供决策支持和业务洞察。在智能决策方面，通过挖掘数据潜在关系并构建智能化模型，赋能网络在复杂和不确定的环境中做出更准确、更高效的决策。在算法模型方面，可以借助深度学习、强化学习、迁移学习等多种 AI 算法模型，实现不同任务类型的处理，如问题分类、时间序列预测等能力，以实现对复杂任务的学习和处理。

1.3 技术原理

通过分析网络场景化的问题，将网络问题转为数学问题，并有针对性的引入 AI 技术，典型应用场景技术原理及价值如下：

1、基于业务的体验升级

无线侧将 AI 技术应用在移动网络中将有助于更精准地预测业务特征、用户行为、移动性、信道环境等信息，最终通过更智能的资源管理与调度机制达成更好的服务质量和用户体验，实现更好的公平性和系统资源利用率，既可以针对某一模块算法提升性能，也可以优化整个端到端通信系统，促进移动通信网络性能的进一步提升。



图 1.3-1 基于业务的体验升级

核心网侧通过构建基于 NWDAF（网络数据分析功能）的核心网网络智能化架构，可定制化感知用户业务体验及网络状态实时数据，开展智能综合分析，并反馈实时保障策略给 PCF（策略控制功能）从而建立专载智能保障，有效提升特定用户/业务的按需体验，实现体验升格。

2、智能网络优化

当前运维系统中各个功能独立存在，人为编排组合系统功能需要进行多次交互，并且严重依赖于运维人员的技能，导致运维效率低，运维成本高。面向未来，借助 AI 技术，实现从以“功能”为中心到以“任务”为中心的运维模式转变，让系统去适应运维人员，为运维系统注入统一知识模型。运维人员通过自然语言与系统进行意图交互，系统能够根据人的意图，自动地编排组合使用功能来完成目标。此外通过 AI 技术，实现对网络资源差异化、精细化的编排，提高网络性能，实现通信资源合理化应用。



图 1.3-2 智能网络优化

3、智能网络节能

利用 AI 技术可对网络业务进行深度学习，识别出对体验敏感的业务（如 VoNR 语音通话及高清视频通话、实时游戏等）。基于业务识别结果，动态调整业务优先级，结合基于用户行为预测、网络负载预测等更多 AI 技术，可以提前为体验敏感业务预留足够的通信管道资源（如带宽、RB 等），防止节能措施引发的资源紧张对业务体验造成负面影响。同时，根据业务流量的时空分布特性，精确预测全时段内的网络负载和能耗需求，动态调整网络资源分配和节能策略，实现从闲时到忙时的平滑过渡。通过上述 AI 技术的实施，可以在移动网络实现体验与节能的高效协同，显著降低网络能耗，同时确保关键业务的优质体验不受影响。随着 AI 技术的不断发展和完善，网络的智能化节能将成为降低运营成本、实现绿色通信的重要手段。

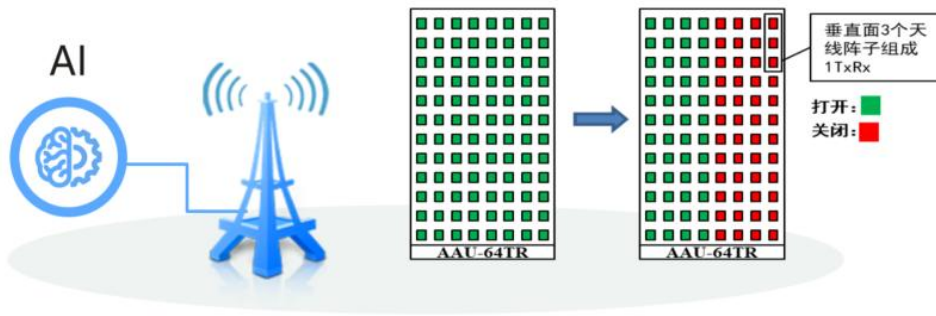


图 1.3-3 智能网络节能

4、5G+AI 用户画像

利用 AI 技术，结合用户访问位置、业务流量等网络数据生成用户画像，实现个性服务推荐和用户行为预测。现阶段运营商的网络资源分配策略不足以满足用户多样化、个性化需求，存在响应速度慢、难以快速识别和处理潜在安全风险等问题。利用核心网智能控制面的 AI 能力，并通过 NWDAF（网络数据分析功能）对数据的实时采集，可以实时分析用户行为，生成精准用户画像，形成个性化网络资源分配策略，动态优化不同用户服务优先级，不断提高个性服务推荐的精准度，进而提升用户满意度。同时，用户行为异常检测、预测等 AI 能力能够实时分析和识别用户行为数据中的异常情况，预测潜在风险，为用户提供实时的安全预警和风险防护方案，保障用户安全。

5、基于客户及特定场景的资源保障

高铁、景区、赛事等重点场景期间，用户会在相对集中地区，同一时间段同时使用网络，导致高流量冲击、资源分配不均等网络拥塞问题。利用 AI 技术可以精准预测流量需求，动态实时调度网络资源，实现重点场景高效保障。面向高

铁、景区等固定地理位置的场景，当前网络主要面临用户流量集中、需求高峰显著、网络资源分配不均等问题，核心网智能控制面的 AI 能力通过历史数据分析和实时数据采集可以提供精准的流量预测、智能化的资源分配和优化、以及实时的网络状态监测与调整，从而保障固定场景下的网络服务质量。面向重大活动、赛事等灵活地理位置的场景，当前网络主要面临临时高流量冲击、网络部署时间紧迫、网络资源难以灵活调配等问题，核心网智能控制面的 AI 能力可以基于实时流量监测和历史数据分析，预测重大活动和赛事期间的流量需求，并提前调度网络资源，通过自动化的网络规划和部署工具，快速完成网络部署和优化、以及智能化的网络状态监测与调整，实现对灵活场景的高效保障。

后章节会针对以上典型应用场景，关键技术及价值进行详细阐述。

1.4 发展规划

网络 AI 技术分两个大的分支，分别为 AI for Network 和 Network for AI。AI for Network 中 AI 用于满足移动网络新指标的关键路径，赋能网络并提升网络运行效率；另一方面 Network for AI 中移动网络是实现 AI 泛在普惠的基础平台，使能 AI 成为泛在化的社会级服务，即网络能提供什么能力服务于 AI。基于以上两个技术方向规划如下：

2024 年至 2026 年重心仍聚焦 AI for Network 方向，针对移动网络 AI 下价值应用的持续研发。AI 价值应用场景主要包含：基于业务的体验升级、智能网络优化、智能网络节能、5G+AI 用户画像和基于客户及特定场景的资源保障，以上技术方向预期在 2025 年相对成熟，结合验证效果推动规模商用。2024 年会同步启动新技术与 AI 技术的结合，包含通感一体、低空经济、数字孪生、空口 AI 等，如通过 AI 对低空飞行的目标进行更高精度识别，甚至判别为无人机或飞鸟。目前相关研究还处在理论研究阶段，预期 2025 年开展实验室验证。

2025 年同步会启动 Network for AI 方面技术探索，随着 AI 可以应用到工业、教育、等更多领域和场景，并且 AI 技术的飞跃发展，AI 模型增大，运行空间增加，AI 对底层资源要求也与日俱增。在该背景下，技术方向下需要研究网络在算力、存储、数据、模型管理等多个领域能为 AI 提供的关键技术。

此外，3GPP 标准也同步进行 AI 的相关研究。R18 标准基于分布式智能架构、新型分析机制、跨域智能协同、AI/ML 的网络节能、负荷均衡、以及移动性优化应用进行讨论；R19 聚焦业务质量智能决策、信令风暴预防、空口 AI 架构以及 CSI 反馈、波束管理、精准定位等应用，同步实现核心网与无线网的数据分析收集和训练模型。

2. 5G 网络 AI 应用高价值场景及解决方案

2.1 无线网 AI 应用

为了迎接网络发展遇到的挑战，通过把 AI 应用能力融入无线网络业务发展、性能提升、网络优化运维和绿色节能等方面，实现智能化业务感知及保障、智能网络优化、智慧节能等能力，推动网络向以智为先方向演进。

2.1.1 无线网 AI 应用技术架构

随着无线 AI 应用演进，国际国内通信标准组织也开展了积极的讨论，对无线网元以及网管的 AI 应用架构进行了定义，其中，无线 OMC 平台 (Operation and Maintenance Center) 负责无线单域内的网络智能，基站负责基站内的网元智能，整个网络通过分层处理来实现更高效的无线 AI 应用。

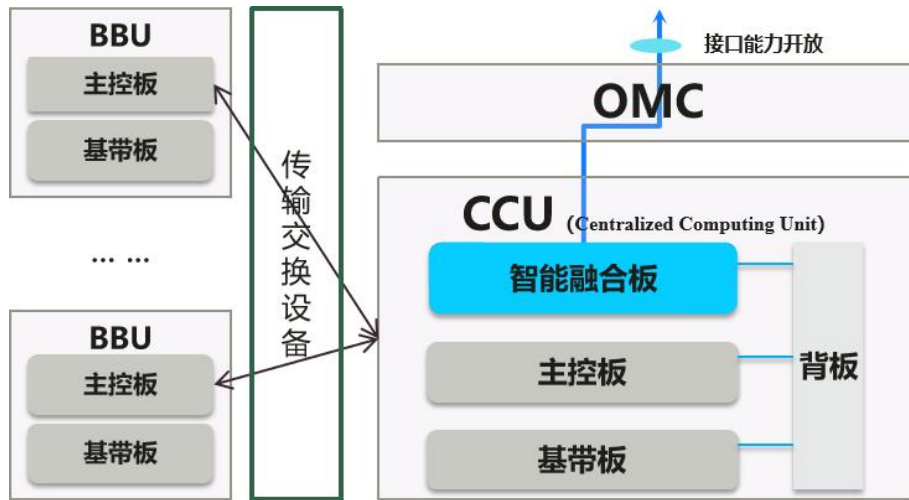


图 2.1.1 无线 AI 应用技术架构

基于上述技术架构，无线 AI 应用技术主要分为三个方面：

典型应用场景：科学合理实现通信与 AI 技术相结合，赋能网络新能力，进行智能业务实时保障，提升用户业务体验；在网络优化场景中，基于智能网络优化技术细耕网络，实现资源合理化使用；践行企业责任，基于智能化节能技术实现可持续发展，通过三个典型应用场景落实无线网络 AI 应用。

两级 BBU 架构：网络智能化技术对基站算力提出了新的挑战，需要基站为

AI 模型提供计算能力、存储能力等。通过在基站内部署新增算力，如新增智能融合板的方式，构建两级 BBU 部署架构，为无线网络 AI 应用提供硬件能力。

接口能力开放：基站内 AI 应用会产生丰富的价值数据，如业务感知数据等，此类数据及能力与传统的无线性能数据区别较大，属于 AI 特有数据。基于 OpenAPI 技术，实现网元的智能化能力标准化开放，提供细粒度数据及评估能力。向上为高层进阶的 AI 模块实现数据开放，向下让运营商一定程度实现与基站内 AI 能力的协同。

2.1.2 基于业务的体验升级场景

一、智能业务感知

移动互联网的迅猛发展催生了多种业务形态，包括视频流媒体、在线游戏、社交媒体、电子商务等。在这一过程中，无线网元肩负着重要的责任，不仅需要确保整体网络的稳定性和可靠性，还要保障各类型业务的高质量体验。这对于网元的业务感知能力提出了更高的要求。

智能业务感知技术是建设智能管道中较为重要的一环，基于业务感知结果对无线资源合理配置，使用户容量及业务感知能力最大化。智能业务感知技术方案通过基站 AI 算力自动进行传输数据采集，自动跟踪热门应用的数据传输特征变化，对热门应用的数据传输特征变化能迅速响应，从而确保了重点业务的高感知率。该技术保持对新业务持续追踪，并基于机器学习的数据传输特征感知系统，对海量经验数据进行智能分析，自动发现可能的数据传输特征，高效快速地构建特征库。业务指纹特征库在线的更新，无需断网就能完成系统特征库更新，并使用多种关键技术：

1、特征自动提取：深度学习能够自动从原始数据中提取特征，而无需依赖人工设计特征。网络流量复杂多样，传统方法难以准确提取所有关键特征，尤其是现在加密业务占比越来越高，能够利用的明文特征越来越少。通过深度学习模型，系统可以从流量数据自动提取有效特征，提高感知的准确性和效率。

2、注意力机制：引入 AI 中神经网络算法，通过注意力机制，能够有效捕捉输入序列中不同位置间的依赖关系，这对于业务体验感知任务非常重要，因为流量模式和特征往往跨越较长的时间跨度和范围，注意力机制可以动态调整每个输入位置的关注点，实现对全局信息的综合分析。

3、迁移学习：迁移学习通过将预训练模型应用于新的任务和数据集，可以大幅减少训练时间和计算资源。在业务体验感知中，迁移学习使 AI 模型能够在

大规模公开数据集上进行预训练，然后在特定网络上进行微调，快速达到高准确率和高性能，这种方式不仅提高了模型在不同网络中适应性和鲁棒性，还显著减少了对大规模标注数据的依赖。

4、在线监督：在线学习系统通过实时反馈机制，对模型的分类效果进行即时验证和校正。当模型对流量进行分类后，系统会根据实际情况对分类结果进行反馈。如果分类正确，模型参数得到正向强化，如果分类错误，模型会动态调整部分参数，以避免类似错误再次发生。这种基于反馈的更新方式，使模型能够持续学习和改进，逐步提高业务体验感知准确性。

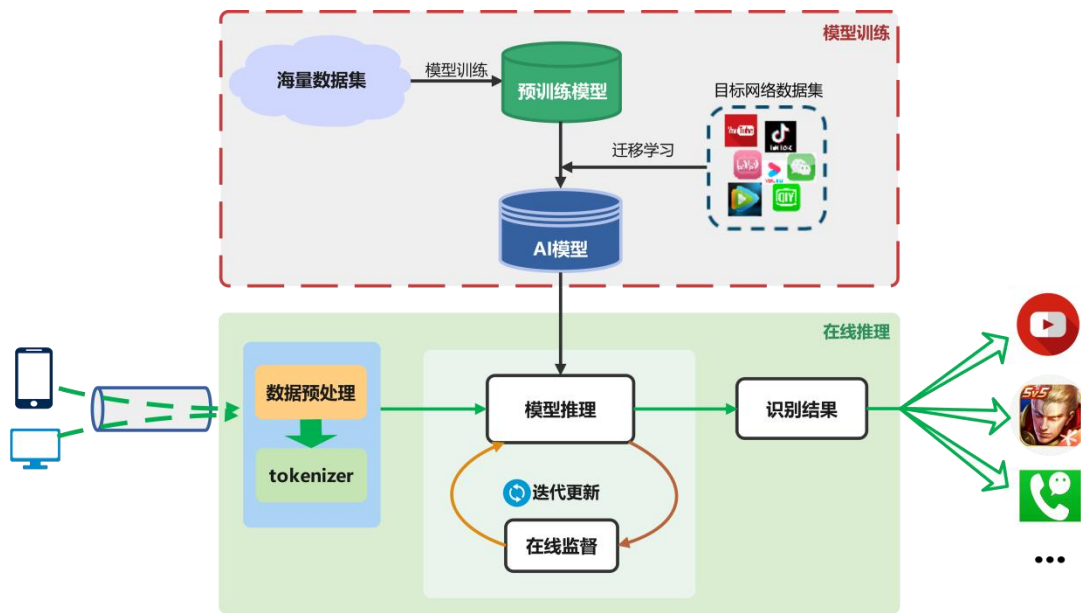


图 2.1.2-1 业务特征库的学习架构

智能业务体验感知通过建设自动化采样系统，实现了原始样本的自动化采集和样本标注的自动化。目前智能业务体验感知模块支持数以万计业务种类感知，包括国内外各种主流业务类型，并精细化感知应用中的各项子业务，如语音通话、视频通话、直播、视频会议、云游戏等，感知准确率达到 95%以上。

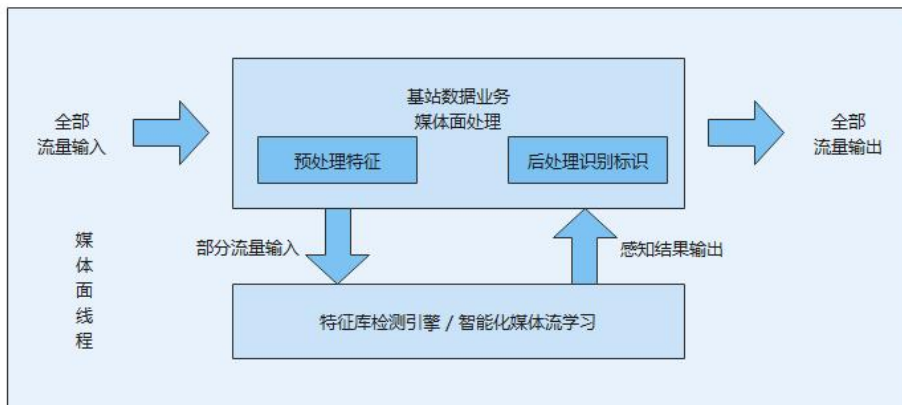


图 2.1.2-2 智能业务体验感知技术流程

智能业务体验感知对于现代网络管理具有重要价值，主要体现在以下几个方面：

1、优化用户体验。智能业务体验感知能够精确区分通话、视频流、语音流、游戏流等不同类型的流量。网元可以根据感知结果对相应的业务进行针对性优化和保障。例如，在高峰期优先保障视频通话和实时游戏流量，减少卡顿和延迟，极大地提升了用户的满意度和体验。

2、保障网络安全。通过业务体验感知，可以快速检测异常流量和潜在的安全威胁。例如，识别出恶意流量、欺诈流量、网络攻击等风险，网元可以及时采取相应的防护措施，这不仅保护了用户的隐私和数据安全，也增强了网络的整体稳定性和可靠性。

3、精细化业务和无线性能结合。智能业务体验感知为运营商提供了更为详尽的数据分析和洞察，帮助其更好地理解用户行为和需求。这些洞察可以用于开发新业务、优化现有服务和提高市场竞争力。

二、智能业务保障

5G 技术的推广和应用，使得网络不仅要满足更大的数据传输需求，还要满足不同行业和消费者对于服务质量（QoS）和服务可靠性（QoE）的高标准期望。而随着网络上新服务、新应用、新场景的不断出现，传统 5G 无线网络的保障方案越来越难以适应这些变化，因此无线智能化的在线业务体验感知，数据分析评估及智能化的精准保障就显得尤为重要。

智能化精准业务保障技术的演进，可以形成一套 AI+自优化的无线网络关键评估指标 KPI 和关键质量指标 KQI 监控及优化的智能解决方案，通过对网络 KPI/KQI 智能监控，及时有效地发现网络问题，进而通过分析网络配置、告警等多维度数据，找出导致网络质量异常的根因因素，再通过对参数调控、故障恢复等相关的优化手段，使得网络能够达到一个最佳的运行状态，精准提升高优先级用户的业务体验感知。

精准业务保障方案通过预测技术和保障策略的结合，形成一套闭环体系：

1、精准业务保障预测技术

智能服务层基于物理网络层的算力资源和数据资源，进行算力的建模和编排，以及数据的建模和治理，并基于多样化的算法模型进行智能化原子能力的编排，最终将算力、数据和算法以服务的形式向场景应用层开放。

在数据维度上，通过标准化的数据采集流程，以及统一的数据模型进行数据采集、清洗、关联和标注，无线网络系统内可以生成并维护用户画像、基站画像等模型化数据。分布式的数据服务再结合领域知识和模型知识，可高效支撑 RAN

内生智能的模型训练和推理。

在算法维度上，基站侧部署的实时智能引擎可支持极轻量级模型训练和近实时模型推理，通过精准预测和主动优化来提升网络效能和用户体验；网管侧部署的轻量化智能引擎可支持轻量级模型训练和非实时模型推理，通过质量洞察、智能排障等能力实现高效主动的网络运维。此外，该架构还支持在独立智能平台部署。

AI 应用服务层可将上述数据和算法原子能力以微服务的架构进行管理编排，并以 AI 服务（AI as a service, AIaaS）的形式向运营商开放，赋能场景应用层。运营商通过服务调用和低代码开发，可高效、低成本地将 AI 能力集成到自身的系统中，还可以实现灵活、按需的功能扩展以及资源弹缩。

基于以上技术，在该场景应用中可以借助 AI 服务实现网络可视化、建模及仿真能力按需调用、多场景精准网规、网络策略和参数寻优、网络故障精准诊断、预测性维护等网络规建维优全流程的高度自动化和智能化，还可以按需进行新业务的开通预演和性能评估，以及新技术研究和低成本试错等。

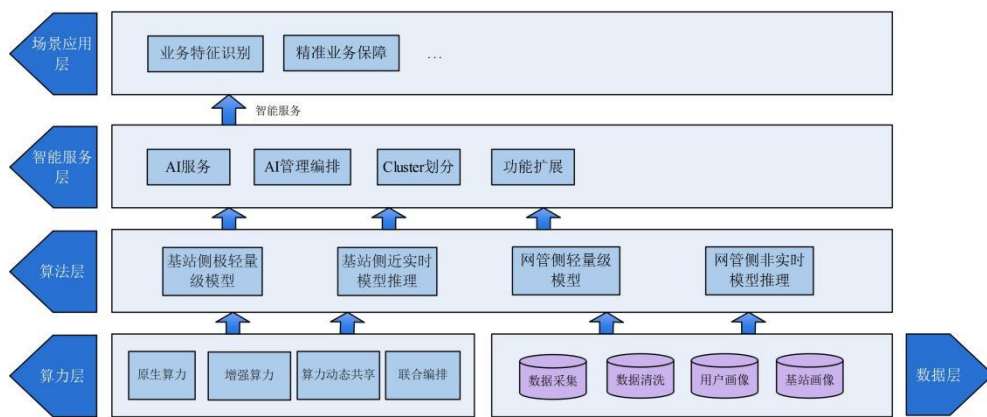


图 2.1.1 2-3 预测模型层级关系

2、精准业务保障策略

精准业务保障是在智能业务体验感知的基础上启动对应保障策略。保障策略是通过自定义业务保障类型配置不同的调度优先层级，在不同层级队列配置不同的保障速率从而实现对业务的精准保障。

为了确保无线网络能够提供优质的用户体验，根据不同业务特点，需要提出一套评估用户感知的体验指标体系。

基于智能精细化业务体验感知，及业务体验量化评估，对于有调度时延和速率需求的业务类型，在业务感知之后进行业务保障，可以优化业务的调度时延和业务速率。

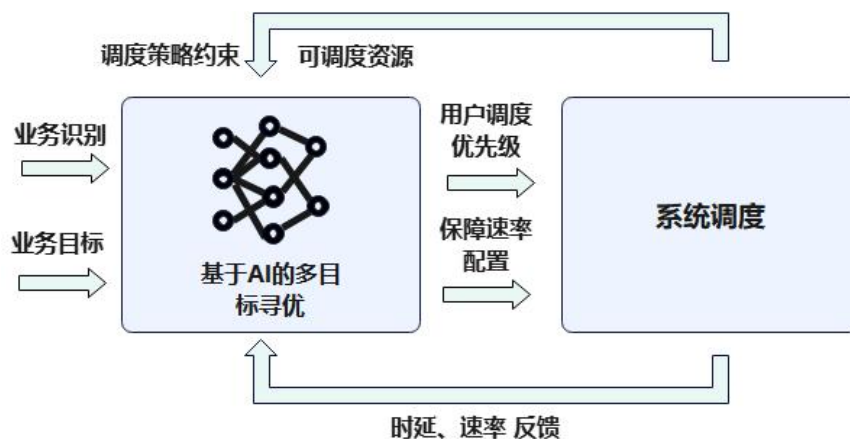


图 2.1.2-4 精准业务保障方案

如图 2.1.2-4 所示，业务体验感知后，根据业务的需求目标，如速率、时延等要求。并且根据当前业务目标的实际值、当前可调度的资源、调度策略约束，通过 AI 模型，实现多目标寻优，生成用户调度优先级以及保障速率配置等。同时根据调度后所产生的影响，如时延、速率等变化，将其反馈至基于 AI 的多目标寻优系统，再次完成多目标寻优，形成闭环。

基于 AI 的多目标优化系统根据实际的业务保障需求，可以分别针对业务优先级类保障、时延类保障、速率类保障、移动分流类保障等不同类型选择合适的调度优先级和调度配置参数，提升保障用户体验。

2.1.3 智能网络优化场景

一、智慧重保诊断

重大国际赛事、国际峰会、重要节假日等重大事件期间，终端用户会在相对集中地区，同一时间段同时使用网络，网络会承受比平时大几倍的业务量冲击。如此巨大的业务量冲击，很可能接近或超过为平常运营所设计的系统容量，导致网络运行下稳定发生突发问题，主要问题如下：

1、保前传统大话务预测方法主要面向初期场馆建网规划阶段宏观预测使用，且依赖专家固化规则，缺乏包括终端渗透率、用户数峰均比、上下行业务占空比等关键数据。造成难以精准评估不同时间段不同小区负荷、用户体验。于此同时，基于专家经验固化的预案，效果不存在不确定性，费时费力。

2、传统保障流程依赖工程师提取指标，筛选 TOP 异常小区，问题定界定因，再到闭环方案输出、执行时效性差，整个过程一般大于 45 分钟；尤其是性能异常与设备告警并存时，传统手段优先处理告警后再看性能是否恢复，未恢复再行空口问题定位，闭环问题约 60 分钟。无法满足保中问题快速发现和修复，进而

降低问题影响的诉求。

基于以上问题，将 AI 与重保场景结合，实现智慧重保。该 AI 应用在事前基于大话务风险仿真预测技术实现风险充分识别和预案准备，事后基于性能故障融合诊断技术准实时定界定因，以及方案智能生成：

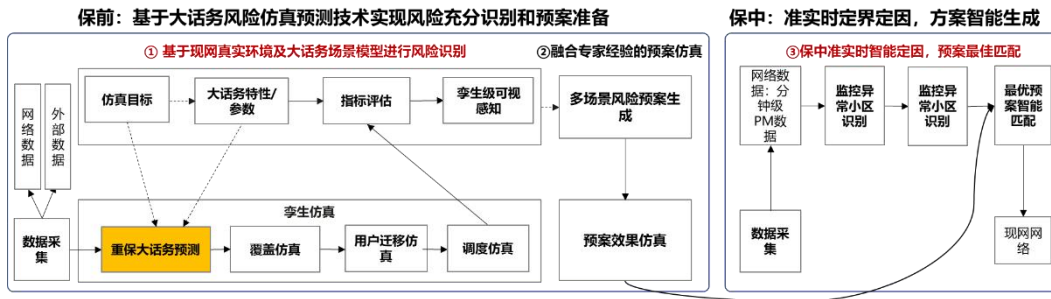


图 2.1.3-1 智慧重保调节业务架构

1、大话务风险仿真预测技术：基于现网数据、历史数据和实验室数据相结合，构建用户时空维度的大话务预测模型。针对大话务场景有效数据获取难的问题，提出“基础+增量”分层框架实现。基础模型：基于多次大话务赛事数据进行解析清洗后，提取赛事前、中、后不同细分场景、不同时间切片上的话务分布特征、话务变化特征。提取业务模型、时序变化、指标预测等共性特征。

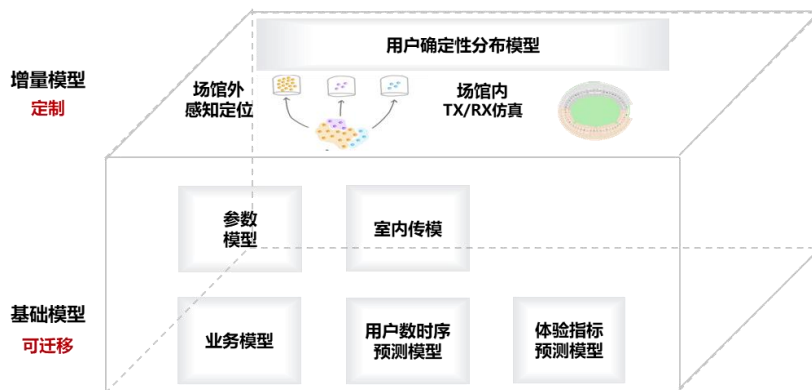


图 2.1.3-2 “基础+增量”分层框架

2、性能故障融合诊断技术：利用历史告警、性能 KPI、配置等数据，构建性能与告警影响统计模型；自动识别强、弱影响告警。针对强影响告警，采用告警发生时空口性能数据、告警数据基于神经网络注意力机制构建 KPI 与告警关联量化模型，准确识别告警影响。当出现性能异常事件，性能 KPI 属于时序连续数据、告警属于时序离散数据，传统基于时序相关性和因果算法无法匹配性能、告警异构数据量化分析要求，因此引入神经网络注意力机制技术实现性能与告警量化模型，准确评估告警对 KPI 的影响。

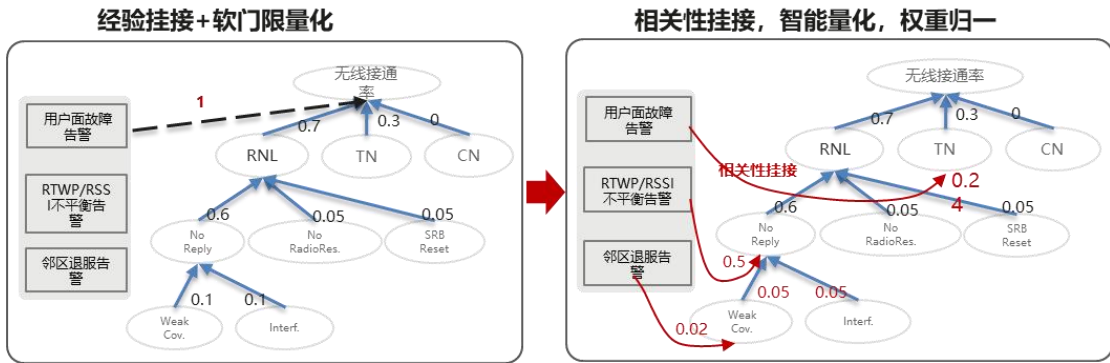


图 2.1.3-3 网络异常的诊断结果

智慧重保诊断可以精准预测网络风险和敏捷诊断故障，具有极高的应用价值：

- **面向保障前**，提供千站级规模的网络风险预测，可灵活设置仿真目标，智能识别网络风险隐患，辅助工程师精准定位网络问题，预防和消除网络潜在问题，实现网络加固。
- **面向保障中**，通过性能故障融合诊断、确保 1 分钟识别性能问题或设备故障，5 分钟解决突发问题，降低故障影响。

二、深度谱效提升

未来随着云游戏、XR、元宇宙等实时交互类、云渲染类业务发展，业务的实时性将原来的秒级发展为 10 毫秒级，对于网络时延和带宽的确定性带来新的诉求，同时应用对于网络质量的实时感知挑战也会加大，因此需要网络提供一定时延、带宽的确定性以高效保障 E2E 业务的体验。

从无线网络的频谱来看，频段越来越多，频段之间的差异也越来越大，从 Sub-1G 到毫米波，速率可以相差百倍、覆盖可以相差十倍，定位精度和时延等能力差异也很大。网络谱效的持续提升，满足大带宽、低时延的业务发展至关重要，基于以上背景，引入 AI 技术旨在分场景提升谱效。

1、多频场景的智能化协同技术

从用户视角看，需要满足随时随地一致性的体验。从网络视角看，一致性体验目标达成的关键在于如何进行一系列因素的测量、统计和计算。这些因素包括业务需求（上下行带宽、时延等）、频段类型（TDD/FDD/SUL/CA）、频段带宽、频段能效、终端能力，以及用户所处环境的覆盖、干扰、移动性等等。需要从多维度综合考虑多频资源的选择和组合，以达到整体的用户体验最优。

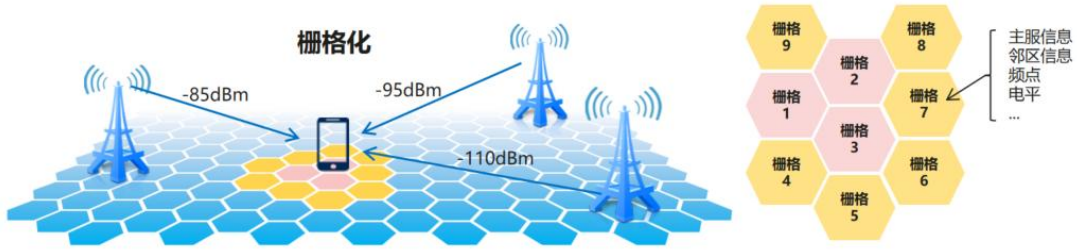


图 2.1.3-4 智能化虚拟栅格

智能化多频协同技术的关键价值在于提供多频段的空间预测能力，基于具有相同无线信号特性的终端上报 RSRP 信息特征作为输入，引用空口测量信息建立虚拟无线电环境特征库，并构建智能化虚拟栅格，基于该栅格可以快速预测出异频段的状态特征，从而计算出频段资源利用的最优解。在更优网络资源满足用户体验需求的同时，使得网络的整体谱效、能效最大化。

2、多波束场景智能化协同技术

大规模 MIMO 相比普通 MIMO，可提供 3~5 倍容量、10 倍边缘体验能力，以及立体覆盖、节能等方面的优势。在大规模 MIMO 场景下，基于复杂的多维度因子实时计算的多用户 MIMO 调度策略，是决定小区整体容量和性能的关键。但多用户 MIMO 调度策略的计算过程复杂，包括用户配对关系选择、无线资源分配、无线链路选择（RANK/MCS）三个环节，且每个环节涉及多个因素，包括波束分布、终端分布、业务特征、链路质量、干扰相关性等。

基于以上场景存在的问题，在选择 MU 调度策略时，不同维度调度量会影响 MU 配对效果和小区整体性能。全流程的调度联合寻优能找到最优的组合，但复杂度很高。因此，需要基于智能化的方法进行实时的调度寻优，这对于硬件的算力也有较高的要求。基于硬件能力的准备度，MU 调度能力的实现，从单环节少数因子的寻优，逐步走向多环节多因子的联合寻优。



图 2.1.3-5 多用户 MIMO 调度寻优

利用多频、多波束、多站场景的 AI 协同技术，有以下价值：

- 更加合理且高效利用基站频率资源，利用指纹库快速决策使用频率，达到最

优体验。

- 更加精确的波束调整，使得多用户 MIMO 配对比例增加，从而使得网络容量增加，谱效提升 3 到 5 倍，在不增加频谱的情况下实现多用户体验倍增。

三、智能空分配对

MU-MIMO，即多用户多输入多输出（Multi-User Multiple-Input Multiple-Output），是一种利用发射与接收端的多天线获取分集增益、充分利用空间资源，将多个终端的数据在相同频段收发，提升小区流量。

传统 MU-MIMO 按照相关性计算遍历所有终端，计算复杂度高，遍历周期长，多用户空分配对性能受限。智能 MU-MIMO 是传统 MU-MIMO 的增强技术，引入了基于数据驱动的智能模型包括基于用户的投影能量信息的空分配对模型、空分用户 MCS 智能适配模型。通过 AI 算法提升 MU-MIMO 的业务性能。

基于用户的投影能量信息的空分配对模型，是基于投影能量信息提炼用户空间特征、计算用户间干扰度量，并用干扰度量来指导用户配对和空分用户间流间干扰的计算。同时，根据用户空间特征，将空分资源向信道质量好、角度扩展小的空分用户倾斜。

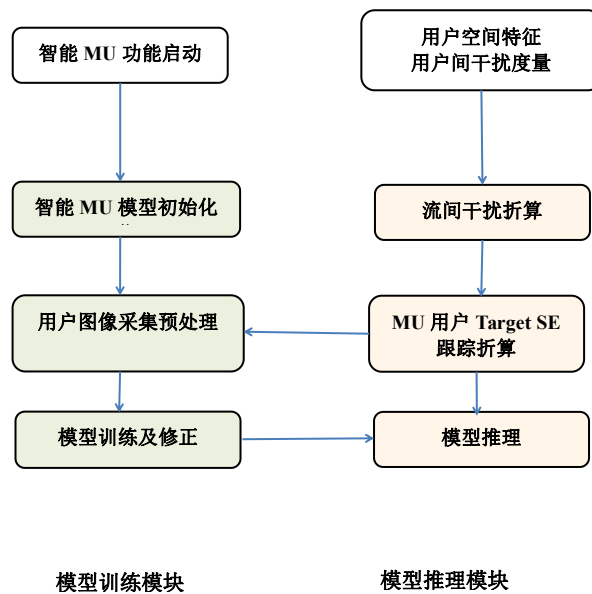


图 2.1.3-6 智能 MU-MIMO 模型

空分用户 MCS（Modulation and Coding Scheme，调制与编码策略）智能适配模型，在线推理空分配对组发生变化后的干扰度量，能够为配对组内用户提供流间干扰折算，综合提升小区 MU-MIMO 性能。

MCS 智能适配模型使用 CQI（Channel Quality Indication，信道质量指示）、空分 RI（rank indication，秩指示），空分配对的干扰度量信息，基于信道特

征划分构建逻辑栅格。同时基于 AI 模型定位到终端在当前时隙下归属栅格。为空分配对组内的 MU 用户提供时隙级的 MCS 跟踪和修正，适配模型计算最优 MCS。

相对传统 MU-MIMO 的配对 MCS 修正可以达到更加准确的 MCS 目标，快速收敛达到最佳频谱效率。

四、移动性 AI 增强

在现代移动通信网络中，多频协同部署日益普及，通常情况下，为了保证最优的性能与体验，终端会在不同的频率之间进行切换，以获得最佳的驻留频率。

频率切换需要进行频点测量，以获取信号强度数据作为参考，尤其是在覆盖较差的环境下，终端会对多个频率进行测量，甚至有些频率存在于距离终端很远的位置，多个频率测量会导致时延的增加，从而导致终端在弱覆盖场景下的时间延长；终端吞吐量的下降，也可能导致切换过晚，最终引起掉话。

为了解决上述问题，采用 AI 技术建立基于频率覆盖的模型通过预测频率覆盖，并结合源/目标站点负荷情况，进行有目的的移动性控制切换。同时结合平台级的移动性管理优化的 AI 功能，定期对切换参数进行优化，从多维度保证用户的最佳体验。

基于 AI/ML 的移动性增强功能，通过引入搜索区域创建 AI 模型，识别用于切换的候选频段来实现。模型使用监督学习算法，通过基于频率覆盖模型预测频率覆盖，再结合终端的测量信息来判断终端所处的位置，通过智能移动性管理模块，提取预测终端最有可能作为切换目标的频率，把终端的测量范围减少到该预测值，减少了终端的测量时间，增加测量的精准度，使得终端可以大大缩短在弱覆盖场景下的驻留时长。

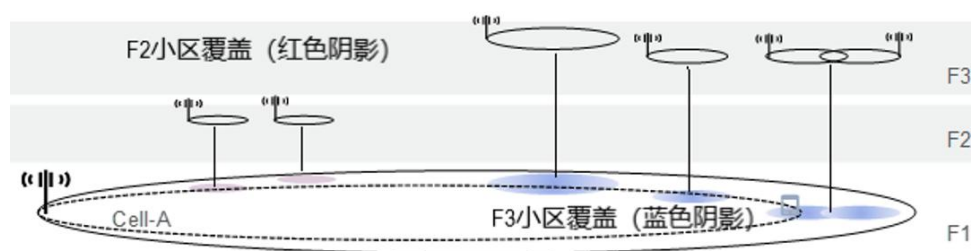


图 2.1.3-7 频率覆盖预测

基于平台级的移动性管理优化功能，利用 AI 从不同小区的配置、环境、邻区等中学习。当一个小区的性能比可比的移动场景差时，AI 管理优化器会自动提出最佳的切换参数更新，重点是改善小区边缘的性能。

将配置的优先级与预测覆盖，负载结合起来，可以更好的选择切换目标频率，减少终端在弱覆盖场景停留的时长，提升用户体验和速率。与传统的切换方

式相比，实验测试数据显示切换所需时间缩短 60%，提升用户的下行速率。

五、极致告警压缩

在通信领域，智能运维面临着由网络设备故障、链路拥塞、服务质量下降等引发的高频、多源告警冲击。尤其在 5G、物联网等新兴技术驱动下，海量数据传输与复杂网络架构加剧了告警管理难度。极致告警压缩旨在通过有效压缩海量告警，提升故障定位与恢复效率，确保通信服务的稳定与质量。面对海量、异构的通信告警数据，如何在短时间内从大量告警中精准识别出关键信息，同时保持对突发、重大事件的高度敏感性，是技术落地的关键难题。

针对以上问题，极致告警压缩引入 AI 技术中频繁项挖掘与时空关联分析能力予以解决。首先，运用基于频繁项挖掘算法和根告警识别算法，发现告警间的频繁模式和因果关系，构建关联规则库。其次，结合告警的时间戳与地理标签，利用时空数据分析方法揭示告警的时空分布特征、传播规律及潜在的区域性故障。再次，融合网络拓扑信息，运用图算法追踪故障根源，实现深层次的告警聚合与压缩。

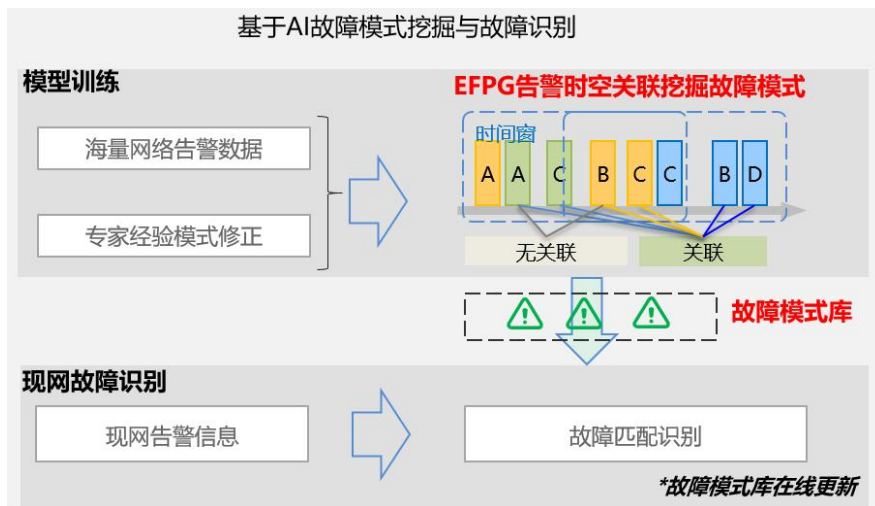


图 2.1.3-8 基于 AI 故障模式挖掘与故障识别

基于 AI 的故障模式挖掘相比人工经验准确性高，海量网络数据训练+专家经验修正确保基础故障库精准。极致告警压缩技术显著减轻通信运维的信息负担，提高告警处理效率与准确性，冗余工单减少 70%，有助于预测未来故障趋势，驱动预防性维护与网络优化策略，增强通信企业的服务可靠性与客户满意度。

六、小区性能监控

小区性能监控是保障用户体验、优化网络布局和提高运营效率的关键环节。然而，传统的小区性能监控方式往往受限于人工巡检和每个指标固定阈值的告警触发机制，难以实现实时、精准地监控网络性能。当前小区性能统计指标众多、

数据量巨大，使用传统的监控方式已经无法有效分析大量的性能数据，甚至可能导致关键信息的遗漏，从而无法精准有效地实现网络的管理和调整。

为了解决上述问题，在小区性能监控中引入 AI 技术，通过深度学习和大数据分析，AI 能够高效处理海量的监控数据，从中提取有价值的信息，为小区性能监控提供强大的支持。在小区性能监控系统中引入了 AI 技术后，可以在日常网络优化、网络故障的定位与排障以及重大节日、会议、赛事保障等场景中取得显著成果：

1、日常网络优化

由于统计指标众多，数据量巨大，传统人工分析的方式在日常网络优化中往往力不从心。而引入 AI 后，系统能够自动分析和处理大量的监控数据，为优化工作提供精准的数据支持。运维人员可以依据系统的建议，对网络布局、资源配置等进行精细化调整，从而显著提升网络的整体性能。

在现网的差小区优化工作中，借助异常检测算法和深度学习分类算法，初始数据筛查相较于传统人工分析，可以节省 90% 以上的的时间。

2、网络故障的迅速定位与快速排障

在面对海量数据时，传统的小区监控方式很难及时发现网络中的隐性故障。异常检测算法能够实时监控并分析海量网络数据，一旦发现异常指标或潜在故障，可以立即发出告警并提供故障定位建议。这使得运维人员能够迅速定位故障源并采取相应措施，大大缩短了故障处理时间。

在现网的实际应用中，借助 AI 系统进行的故障定位，相对于传统小区监控的网络故障定位，可以节约 50% 以上的定位时间。

3、重大节日、会议、赛事保障

在重大节日、会议和赛事期间，网络流量和性能需求激增，统计指标和数据量更是达到顶峰。通过预测和规划网络需求，提前进行资源扩容和优化调整，AI 确保了网络在关键时期的稳定性和安全性。

AI 在小区性能监控中的关键技术主要可以分为以下几个方面：

1、通过自动采集网络中的小区性能监控数据，为后续的分析 and 优化提供基础数据：基础的小区性能监控数据是后续分析和优化的基础，无论是传统的小区性能监控还是引入 AI 后的小区性能监控，都需要基于实时、准确的数据采集。

2、利用深度学习技术对采集的性能指标进行关联分析：目前小区性能统计指标众多，传统基于规则的监控方式基本上是基于固定阈值对单个基本指标进行门限判定。采用 AI 中深度学习可以在输入数据维度较大，数据组合复杂的情况下，也能够迅速关联出数据之间的内在联系，从而发现网络性能中的异常模式和潜在问题。

3、AI 系统根据分析结果，自动触发告警并提供优化建议：通过小区性能监控数据的分析，结合其它数据源，AI 系统可以自动触发告警，并提出调整优化建议，指导运维人员进行网络操作，提升网络的稳定性和性能。

AI 技术在小区性能监控中的应用，可以成功地将统计指标众多、数据量巨大带来的问题转化为巨大优势。通过高效处理和分析海量数据，AI 为日常网络优化、故障定位与排障以及重大活动保障提供了有力支持，在提高工作效率的同时，为用户带来更加优质、高效的通信体验。

七、干扰识别定位

网络优化中干扰问题分析主要依赖工程师的经验，且经验丰富的工程师很难短时间内完成干扰分析工作。同时，部分可能导致网络降级、用户投诉和其他不良影响的干扰类别，难以通过人工分析检测出来。

基于以上问题在干扰分析中引入 AI 技术。基于 AI 的干扰分析功能告别了依赖统一阈值的判断方法，通过多维数据，可自动识别网络中的各个干扰小区，尤其针对隐蔽的干扰小区，实现准确、全面的识别。同时可以将干扰问题自动划分为几种干扰类型，区分内部干扰和外部干扰。与传统的分步判定方法相比，干扰分析效率得到了大幅提升。

干扰智能分析的关键技术包括两个部分：离线训练模块和在线应用模块。

采用基于位置信息的 AI 聚类算法对原始干扰数据进行聚类，对聚类后的簇进行干扰类型的人工标注，以减少人工标注的工作量。然后采用 KNN（最近邻分类算法）算法建立模型，进行干扰类型识别。在智能干扰类型分析功能的基础上，通过结合图像特征向量距离算法与改进的最大连通子图算法，实现外部干扰源定位：

1、按相同干扰源对小区进行分组，结合小区位置特征使用三点定位法初步形成对干扰源的定位。

2、使用改进的最大连通子图法来改善干扰聚类。按位置形成相同干扰源的小区列表。

3、根据被同一干扰源影响的站点数目的不同，采用两种不同的定位策略（单站定位或多站定位）锁定干扰源。

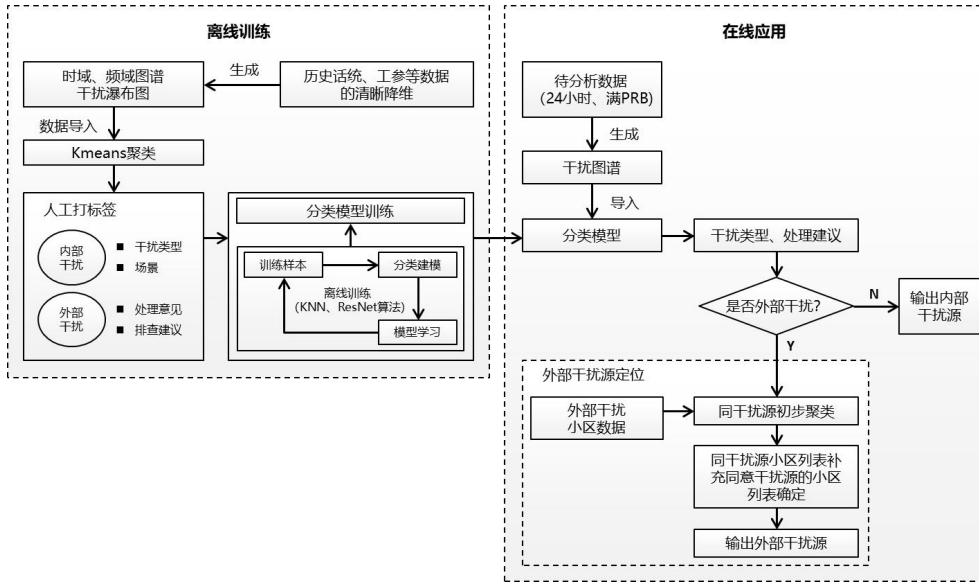


图 2.1.3-9 智能干扰类型识别及干扰源定位流程图

5G 网络智能干扰分析和定位的功能价值主要在于：

- 及时、准确和全面的干扰识别为提升网络性能和用户体验提供了帮助；
- 自动将干扰问题分类（区分内部干扰和外部干扰），极大地提高了干扰分析和定位的效率；
- 基于小区干扰数据，采用干扰类型簇集、最大连通子图计算、地理划分和单站/多站定位等方法，较为准确地完成干扰源定位。最终以地理拓扑和图形化方式直观展示干扰位置，大大节约了干扰定位工作的人力成本。

2.1.4 智能网络节能场景

随着无线网络规模的扩大，网络能耗问题日益突出。传统的通过固定时间段、人工设置参数的基础方式已不能满足节能降耗需求，急需向更精细化的智能节能方式升级。为此，运营商在重要的耗能设备和基础设施上，采用全面、智能的能源管理方法，使其成为一种必要和业务差异化的手段。

根据全球运营商的数据，RAN 在能耗支出方面占了 38%，空调和电源占了 7%。所以如果可以降低这部分能耗的支出，不仅可以减轻运营成本，而且可以减少运营商的碳排放量。

随着 AI 的发展，通过智能化节能同时兼顾节能与无线网络 KPI 的研究课题逐渐被业界重视。

基站能源互联，它是通过数字化技术实现基站能源生产、传输、储存和消费之间的高效连接和协同运行的系统。该系统通过采用物联网、大数据、AI 等技

术，结合无线设备的设备层、网络层和业务层的指标和数据，通过 AI/ML 的数据分析和决策，实现能源供需双方的信息共享和互动，促进清洁能源的应用和能源的高效利用。基站能源互联通过匹配无线设备的网管和 MR 等数据，形成无线大数据和能耗数据，由 AI/ML 对这些历史数据的分析、关联和预测，完成基站无线设备和配套设备全天 24 小时的联动节能遥控和遥调，实现由网络指标为导向的智能化节能。

通过收集到的网络指标和能耗数据，实现预测功能。通过使用历史时间序列数据来预测未来数据。预测功能是根据历史输入数据，如：业务指标和能耗指标，对未来数据进行单变量预测。目前收集的网络指标包含：ERAB 建立成功率、RLC、PRB 利用率、RRC 连接请求数等 46 个指标，通过采用梯度提升预测算法(Gradient Boosting Algorithm, GBA)，预测网络指标趋势。该算法可以使用历史目标值和时间戳数据拟合模型，也可以将可能影响目标值的外部数据（变量）包括在内，其中外部数据（变量），包括天气数据（温度、降雨量）或其他与能耗相关的外部因素。

通过上述的预测功能，能够根据业务或者无线网络指标的变化触发合适节能策略，控制无线和配套设备的功耗，解决节能和用户感知的平衡，实现节能开启期间的零投诉。

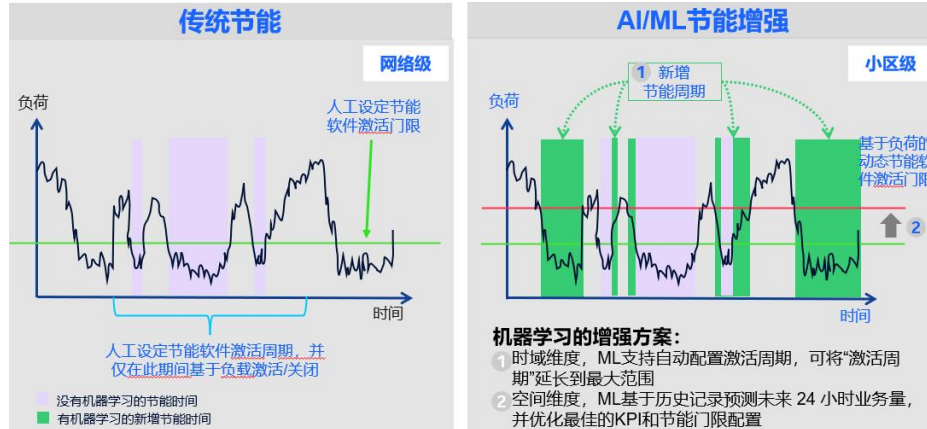


图 2.1.4-1 传统节能 VS. AI/ML 节能增强

通过智能化节能系统的 AI 算法预测每个基站的业务变化，可以比原来固定时间的节能手段，更加精准的确定节能时间段，增加 5%左右能够进入节能的基站数量，减少 37%因定时关闭站点造成的用户业务感知下降：

1、根据网络的运行特征及各种应用场景，调整主设备的运行状态，在对用户体验无感知的前提下，最大程度调节主设备功耗，降低闲时的能耗。

2、在原有主设备默认节能功能的基础上，根据跨 BBU 和跨系统业务负荷，采用 AI 算法选择最优节电方式，无人工干预的情况下实现跨 BBU、跨系统的协同节能。

2.2 核心网智能控制面 AI 应用

2.2.1 核心网 AI 应用技术架构

核心网作为业务连接中心，是业务使能的关键。为推进核心网智能化演进，3GPP 定义了 NWDAF（网络数据分析功能），将数据采集、模型训练、推理分析等 AI 基本能力引入网络运行机制。基于 3GPP 的定义，NWDAF 可与业务智能运行深入融合，实现逐用户、逐业务的实时数据采集、智能感知、分析推理、决策闭环，为智能业务体验感知与保障提供了新思路。

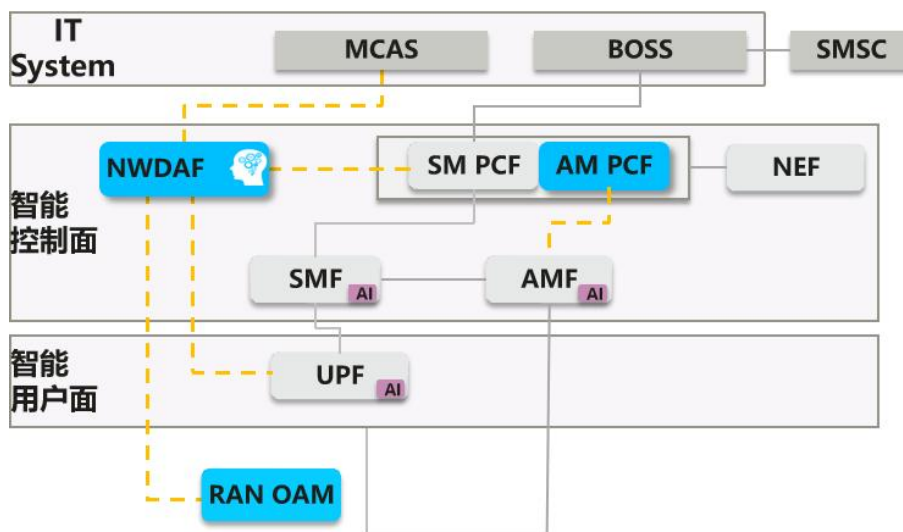


图 2.2.1-1 核心网 AI 应用技术架构

核心网 AI 应用技术架构包括三个方面：

1、**内生 AI**：从智能控制面通过用户行为分析、定制化 QoS 决策等能力与智能用户面体验感知、业务识别等能力，实现典型场景实现高价值应用，形成体验质差感知-保障优化-体验状态感知的实时监测、优化闭环的内生 AI。

2、**智能联合控制与决策**：为支持 AI 模型在网络中的部署应用，智能核心网硬件平台中的云化平台及基础设施所提供的 AI 算力可支持模型的训练及推理能力。同时新定义智能决策能力，融合网络资源预测、智能体验判断、智能业务识别等多 AI 原子能力有机协同，生成基于综合判断的策略建议，实现全局的策略智能推荐。

3、**智能感知及可视化**：网络侧可全方面实时感知应用类型、KPI/KQI 数据、网络负载状态（无线+网元功能）、用户移动数据，基于网络的能力开放架构，依据场景需求并结合网络隐私安全保护的要求，把网络侧部分数据及分析结果面向业务平台、特定行业应用开放，实现区域级、业务级、用户级体验可视化。

2.2.2 基于业务的体验升级场景

一、业务感知

随着 5G 的规模商用和 5G 流量消费快速增长，促进了直播、4K/8K 超清视频、云游戏、云手机、XR 等新兴业务的蓬勃发展。传统网络的流量管道能力，无法实时识别业务应用及特征变化，无法实时感知用户使用业务的体验。但实际上，要实现按需的精准体验保障、提高移动用户端到端业务体验感知，需要网络具备高效实时的业务感知能力。

基于 NWDAF 的核心网，可通过智能 UPF（内置 NWDAF）、NWDAF（独立）以及运营支撑系统的“AI+”能力升级和相互协同，实现面向指定用户/业务/切片的实时、定制化、智能业务感知分析及反馈，这表现在：

1、业务智能感知：改变传统 DPI 识别手段，智能 UPF 搭载业务识别模型，可通过对在网业务流的发包特征、流量特征等信息实时进行分析推理，实现实时智能业务感知。此外基于该技术持续业务流形态。结合智能 UPF 可在业务感知的基础上，进一步分析应用的业务流间关系、业务流关键指标需求等信息，推理业务形态关键信息，为每类/每个应用打造精准业务画像，为网络提供精准保障策略提供重要信息。

2、用户的业务体验感知及反馈：传统网络下发或执行体验保障策略后，因缺乏体验数据的实时采集和分析，网络无法实时感知用户使用业务的实际体验，亦无法为用户提供体验对账信息。为增强用户体验权益，打通智能 UPF-NWDAF-运营支撑系统，面向指定用户，NWDAF 向智能 UPF 定制化采集业务体验数据，进行统计分析后，实时反馈给运营支撑系统，进一步通过短信、微信、账单等方式为客户提供对账，客户感知升格。

二、体验升级

5G 应用呈现形态多样化、体验需求差异化、场景丰富化等特性，基于传统静态配置方式的策略控制机制，保障策略无法按需调整和适配，无法合理高效分配网络资源，无法实现客户的极致体验需求。

NWDAF 的引入，为 AI 内生于核心网策略控制和闭环优化提供了基础框架，可实现面向特定用户、特定业务、特定场景的定制化业务体验升级。基于 NWDAF 的体验升级关键技术包括：

1、定制化数据采集：根据签约要求，定制化实时采集指定用户/业务/位置区域/接入类型的体验数据、定制化采集业务开通区域的网络状态数据，实现面向业务体验保障相关数据的精准实时感知、分析、预判。

2、智能决策分析：根据采集的数据，开展综合分析，决策保障触发时机（如体验不好），并为 PCF 提供保障策略建议，真正实现了保障的按需触发和资源的按需分配。

3、智能控制闭环：NWDAF 实时反馈保障策略建议给 PCF，PCF 作为策略控制中心，进一步决策生成策略控制规则，建立专载保障。

基于 NWDAF 的体验升级技术，可面向用户签约需要，基于 AI 实现自感知、自优化，可有效提升特定用户/业务的按需体验，实现体验升格。

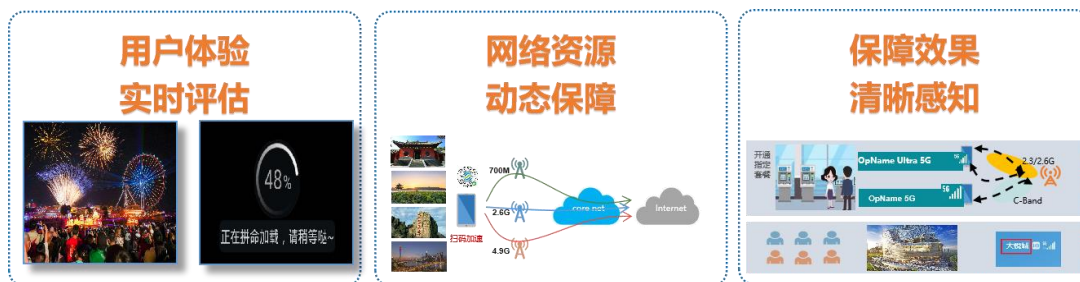


图 2.2.2-1 核心网智能控制面、用户体验可保障、感知升格示例

2.2.3 5G+AI 用户画像场景

一、个性服务推荐

当前移动通信网络环境，运营商面临着用户需求多样化、个性化需求无法精准满足等问题。这些问题主要体现在网络资源分配僵化，无法动态调整以优化用户体验，服务优先级设置固定，无法根据实际需求灵活调整。因此，利用 5G 核心网的 AI 能力，通过实时用户行为分析和动态资源分配，实现个性化推荐成为解决这些问题的关键。

在个性服务推荐场景中，网络实时获取用户的位置信息、业务流量等数据，这些数据通过 NWDAF 进行实时采集、分析。然后，利用机器学习等 AI 算法对这些数据进行分析，生成用户画像。基于用户画像，网络可以根据用户的实时需求和偏好，动态调整业务流量，优化服务优先级。例如，对于喜欢观看高清视频的用户，网络可以优先分配带宽资源，确保其观看体验的流畅性。

此外，个性服务推荐还可以根据用户的反馈，不断优化推荐算法和服务优先级设置。这一过程需要持续的数据采集和分析，以确保推荐系统能够准确反映用户的最新需求和偏好。通过这种闭环反馈机制，网络能够不断提高个性化服务的精准度和用户满意度。

个性服务推荐场景的具有重要价值，通过动态调整网络资源分配，满足用户的个性化需求，提升用户体验；根据实际需求灵活调整资源分配，优化网络利用

效率；通过个性化服务，提高用户对网络服务的依赖和满意度。

二、用户行为预测

随着 5G 网络的普及和应用，网络流量的波动性和不确定性给网络资源的分配和优化带来了挑战。用户的行为模式、使用习惯和偏好都在不断变化，这使得传统的基于规则和统计的预测方法难以适应。此外，网络环境的复杂性和不确定性也增加了预测的难度。核心网 AI 通过对用户行为数据的深度分析和建模，可以提供精准的用户行为预测能力，进而优化网络资源分配、提高服务质量，并促进个性化服务的发展。与此同时，安全风险监控是一个重要且具有挑战性的任务。现有的网络在安全风险监控方面存在诸多不足，主要表现为监控不及时、响应速度慢、难以及时识别和处理潜在威胁。

通过实时监控与分析用户行为数据，利用异常检测算法识别用户行为中的异常模式，预测潜在风险，并提供实时预警和应对方案，实现自动防护，保障用户安全。

首先，用户行为预测需要通过 NWDAF 实时采集用户行为数据。这些数据包括用户的登录行为、访问记录、操作习惯等，能够全面反映用户的行为模式。接下来，利用 AI 算法对采集的用户行为数据进行实时分析，识别出常规、异常等模式。通过对海量用户、数据的训练、建模，可准确识别出异常行为。例如，当用户的登录行为突然出现异常，或访问频率明显增加时，系统能够及时发现并识别这些异常行为。

其次用户行为预测还可以用于网络故障预测与预防。通过对网络运行数据的深度分析，核心网 AI 可以预测潜在的网络故障和性能瓶颈。例如，通过分析基站的运行状态、网络流量等信息，如 NWDAF 通过无线网管采集小区级 PRB 上行平均利用率、上行总流量等数据，基于周期性的时序预测 AI 算法可以预测某些基站可能会出现高负载情况，从而提前采取预防措施，避免网络服务中断。此外，通过对用户投诉和反馈数据的分析，AI 可以识别网络服务的薄弱环节，优化网络配置，提高网络服务的可靠性。

用户行为预测的优势在于通过对网络资源的精准配置，提升网络运营效率，减少资源浪费。同时，通过个性化服务推荐和动态业务流量调整，增强用户体验。预测网络故障和性能瓶颈，提前采取预防措施，确保网络服务的连续性和稳定性。此外，通过行为预测识别潜在的安全风险，提前采取措施，保障网络安全。



图 2.2.3-2 5G+AI 用户画像场景示例

2.2.4 基于客户及特定场景的资源保障场景

一、固定场景保障

在高铁、景区等固定场景下，当前网络面临的主要问题包括用户流量集中、业务需求高峰显著、网络资源分配不均、以及网络服务质量难以保障。这些问题常常导致用户体验不佳，特别是在高铁高速移动环境中问题尤为突出。为了解决这些问题，5G 核心网的 AI 能力可以提供精准的流量预测、智能化的资源分配和优化、以及实时的网络状态监测与调整，从而保障固定场景下的网络服务质量。

首先，固定场景保障需要通过 AI 技术实现精准的流量预测。核心网 AI 能够通过历史数据分析和实时数据采集，预测未来的用户流量趋势。例如，在高铁场景中，AI 可以分析历史乘客流量数据、天气状况、节假日等因素，预测未来一段时间内的网络流量需求；在景区场景中，AI 可以基于游客的历史数据、天气预报、景区活动安排等，预测未来的游客数量和网络使用情况。通过这些预测，运营商可以提前做好网络资源的准备和调整，避免突发流量高峰导致的网络拥堵。

其次，核心网 AI 可以实现智能化的资源分配和优化。在高铁、景区等固定场景下，网络资源需求会有明显的时间和空间分布特征。AI 技术可以根据流量预测结果，动态调整网络资源分配。

最后，实时的网络状态监测与调整是保障固定场景网络服务质量的关键。核心网 AI 可以通过实时监测网络的运行状态，及时发现和解决网络问题。例如在景区，当某个区域游客突然增多导致网络拥堵时，AI 可以快速响应，调整网络资源，缓解拥堵情况。

总体而言，固定场景保障通过 5G 核心网的 AI 能力，能够有效应对高流量固定场景下的网络挑战，提升网络服务质量和用户体验，优化网络资源利用，为运营商带来更高的商业价值。

二、灵活场景保障

在重大活动、赛事等灵活场景中，当前网络面临的主要问题包括临时高流量冲击、网络部署时间紧迫、服务质量难以保障、以及网络资源难以灵活调配。这些问题常常导致网络服务不稳定、用户体验差，尤其是在突发的高流量需求下，网络容易出现拥堵甚至瘫痪。为了应对这些挑战，5G 核心网的 AI 能力可以通过灵活的网络资源调度、快速的网络部署与优化、以及智能化的网络状态监测与调整，实现对灵活场景的高效保障。

首先，在灵活场景保障中，核心网 AI 能够通过灵活的网络资源调度，确保在突发高流量需求下的网络服务质量。AI 技术可以基于实时流量监测和历史数据分析，预测重大活动和赛事期间的流量需求，并提前调度网络资源。例如，在大型体育赛事或演唱会期间，AI 可以预测观众人数和网络使用情况，提前调度和分配更多的网络资源，确保现场观众的网络连接稳定、速度快。

其次，核心网 AI 可以实现快速的网络部署与优化。在重大活动和赛事等灵活场景中，网络部署时间通常非常紧迫。AI 技术可以通过自动化的网络规划和部署工具，快速完成网络部署和优化。例如，在大型活动现场，AI 可以根据现场环境和需求，自动生成最佳的网络部署方案。

此外，智能化的网络状态监测与调整是灵活场景保障的关键。核心网 AI 可以通过实时监测网络的运行状态，及时发现和解决网络问题。例如，在大型赛事期间，当比赛进入高潮时，观众的网络使用量会急剧增加，AI 可以实时调整网络资源，确保观众的网络体验。综合考虑，灵活场景保障的，有以下价值：

- **快速响应能力强：**通过灵活的资源调度和快速的网络部署，确保在紧急情况下的网络服务质量。
- **优化网络利用效率：**通过动态调整网络资源分配，优化网络利用效率，避免资源浪费。
- **提高网络服务可靠性：**通过实时监测和调整，及时解决网络问题，提高网络服务的可靠性。

3. 5G 网络 AI 技术演进愿景

2024 年，5G-A 已经从愿景走向现实，商用元年正式开启。面向未来，标准协议持续演进，赋予 5G-A 更强的生命力；业务应用不断扩展，带来更广阔的应用空间。在 5G-A 时代，无线网络首次明确构建“内生智能”的核心能力，针对多种业务需求如裸眼 3D、工业大上行、沉浸式 XR、通感一体、低空通信、高精定位等，引入通信大模型、意图驱动等智能化技术，此外随着 3GPP 标准的推进，终端及芯片也逐步关注内生 AI 在通信中的应用，旨在提升空口的性能，形成 AI

融合的立体智慧网络，推动无线网络数智化转型。AI 与网络融合作为 5G-A 的关键特征之一，将在网络规划、运行、优化、管理、运维等各个环节中发挥更大作用，推动网络向更加智能化、自动化、高效化的方向发展。核心网作为移动通信网络数据流向和控制的中枢，将在未来网络 AI 融合演进中占据重要位置。核心网网元将逐步内生 AI 能力，形成端到端 AI 服务自闭环，引入机器学习、大模型、数字孪生等 AI 前沿技术，对内赋能网络本身提质增效，为接入网、承载网以及其它通信网络提供 AI 基础能力底座，对外使能工业、医疗、教育、媒体等千行百业，在沉浸实时、智能上行、工业互联、通感一体等多个应用领域带来新的价值增长点和发展机遇。

未来随着网络持续发展，后续部分场景下网络问题需要跨域协同解决，伴随着未来 AI 能力泛在，届时需要多网元间 AI 能力协同，重点需要无线与核心网配合，协同实现网络问题的端到端解决。

3.1 无线网络 AI 技术演进

一、无线新技术智能化演进

伴随通感一体、低空智联、无源物联等 5G-A 新技术的发展，新技术新场景也对网络性能提出了更高的要求，在此背景下 AI 技术的结合将推动网络能力持续升级，进一步提升新场景下的网络性能，驱动 5G-A 智能化网络步入新阶段。

以通感一体为例，通感一体是 5G-A 网络的关键价值之一，利用移动通信网络已经广泛部署的站点资源，使用小部分时频资源，提供面向低空，水域和地面的广域泛在感知能力。但在城区中，高大建筑物多，站点环境复杂，导致反射杂波影响目标无人机的准确识别，同时城区各类移动物体较多，如飞鸟，易导致大量虚警产生。



图 3.1-1 通感一体识别示意图

因此如何通过 AI 自学习，消除杂波干扰，识别并排除非目标物，保障通感一体在实战顺利落地的关键研究方向。如利用 AI 手段识别静杂波（建筑、金属

牌等），提升检测率；通过微多普勒特征提取+AI 学习，提升目标识别能力等。

2、意图驱动节能技术

随着无线节能特性的增加以及功能的多样化，逐渐显现几类问题。首先，节能技术方案复杂繁多。为实现网络节能进一步想自治化的演进，意图驱动网络应运而生。意图驱动的节能根据用户的节能意图将现有的节能功能进行组合和动态寻优，以期尽量达到用户节能目标。意图输入包括节能的站点范围，时间范围以及节能策略要求等，通过意图理解和预测，将意图输入翻译为具体节能目标，不同节能目标对应不同节能特性的组合配置。利用 AI 和大数据分析，实现场景识别，将时、空、频、功率域的节能策略有机协同，对节能小区生成最优的节能策略。

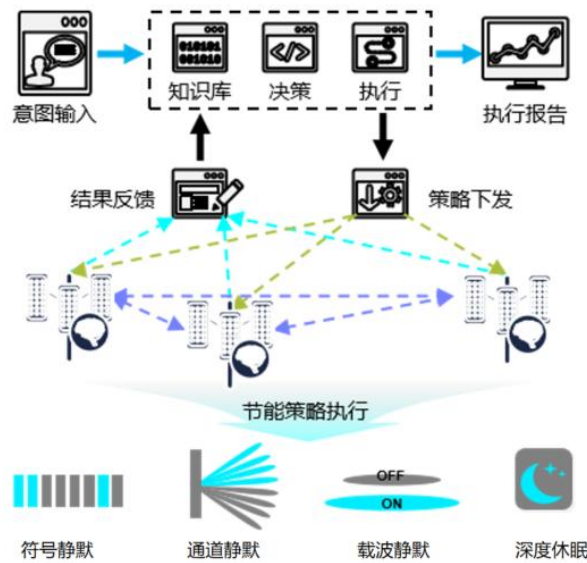


图 3.1-2 意图驱动的节能示意图

未来，如何在比特增长的同时尽量保证功耗无明显增加，需要挖掘 AI 技术结合绿色可持续的设计理念，不断寻求技术创新，全方位探索节能降碳手段，在保障极致性能和体验的同时，提升通信网络能效，积极推进经济数智化转型，助力净零碳目标达成。

3.2 核心网智能控制面 AI 演进

一、网元平台化能力构建

当前核心网 AI 场景业务需求大，且类型多变、迭代周期快，亟需构建高效、智能、灵活的平台，以支撑智能化技术与网络的深度融合和在网络的高效应用。在 AI 内生的设计理念下，核心网网元将迈向智能化演进之路，逐步引入 AI 软硬件系统，具备端到端 AI 服务能力，成为核心网 AI 的能力基石。因此，面向核心

网智能化网元，设计统一的平台化体系，是加速推动这一变革的核心。从 AI 产业生态来看，当前端到端体系较为封闭，软硬件耦合度较高，若采用异厂家 AI 能力完成系统构建，往往会导致性能下降、适配性差、开发复杂度高等一系列问题。为解决上述问题，需实现核心网智能网元平台化三项能力构建。

1、从 AI 黑盒化到 AI 插件化：通过平台化设计，解耦模型与运行环境，实现同平台对多类型 AI 模型的灵活高效加载，使得模型服务能力不依赖于特定硬件或框架，可实现核心网 AI 能力的快速迭代以及业务功能的快速上线，大大降低管理运营成本；

2、从软硬一体到跨架构迁移：通过平台化设计，实现同套 AI 代码在不同硬件上的无感迁移与高效部署，面向不同的硬件平台可高效、快速进行转换，无需针对性开发多套代码并进行多次集成、测试、优化工作；

3、从通用算力到异构智能算力：通过平台化设计，打破当前核心网算力以 CPU 为主，较为单一的现状，引入多厂家智算资源，实现算力的泛在融通，助力核心网 AI 生态的繁荣发展。

二、智能控制面数字孪生

数字孪生是一种采用数学建模仿真和数据驱动等方法、通过创建物理实体的虚拟副本以实现对现实世界精确映射和模拟的技术。它不仅能够实时映射和模拟实体系统的状态和行为，还能通过数据分析提供预测和决策支持。在通信核心网中，数字孪生技术可通过网络仿真模拟促进网络服务的持续优化，例如：

1、面向业务体验的数字孪生应用：针对业务体验优化，核心网 NWDAF 可根据历史流控参数及消息量的关系预测当下网络环境下适用的流控参数，但缺少闭环验证环境。数字孪生技术可建立虚拟网络环境即孪生网络，实时模拟真实网络流量和用户体验，结合流量数据和用户反馈，提供仿真测试平台，以提高策略准确性、降低调整风险。其核心在于验证 AI 生成的 QoS 策略，通过虚拟环境模拟网络操作，循环迭代测试和评估，不断优化策略。

2、面向信令风暴的数字孪生应用：信令风暴是网络故障或重大事件中短时间内大量信令请求导致的过载问题。数字孪生-信令风暴仿真架构通过构建网络拓扑和仿真参数设置，编排信令风暴场景，如 DC 故障等，并通过可视化展示结果。在某些网元故障发生时，网络的流量拓扑可能会发生改变，引入基于性能模型的快速仿真验证，可为网络拓扑发生改变后的流控阈值设置提供理论指导，增加流控阈值调整操作的准确性。

以上数字孪生技术在通信核心网中的应用，通过高度仿真的虚拟环境和先进算法，为网络管理提供预测、评估和解决方案，从而降低故障风险，提升稳定性和用户体验。

3.3 终端及芯片 AI 技术演进

3GPP 在 R18 标准化阶段中，定义了 AI 原生空口的通用功能框架，包括信道状态信息（Channel State Information, CSI）反馈增强、波束管理等。终端及芯片在无线网络 AI 领域的主要挑战是如何为 AI 使能的空口构建一个支持多种 AI 应用的统一功能框架。如下是终端内 AI 的典型应用。

1、终端信道信息反馈增强

大规模 MIMO 技术为基站和终端提供多天线的空间复用/分集增益，可以显著提升系统吞吐量，MIMO 的功率选择取决于信道状态信息（Channel State Information, CSI）的反馈。与时分双工（Time Division Duplex, TDD）系统不同，频分双工（Frequency Division Duplex, FDD）系统不具备上/下行信道互易性，基站通常依赖于终端反馈来获取下行 CSI。未来网络演进中，随着大量用户的 CSI 信息反馈，上行链路资源开销占比增大，给系统带来极大负担，因此，应用 AI/ML 技术（如 Transformer, Resnet, CNN 等）可以进行空域频域 CSI 压缩或预测 CSI 时域反馈，来减少开销并提高 CSI 反馈的准确性。

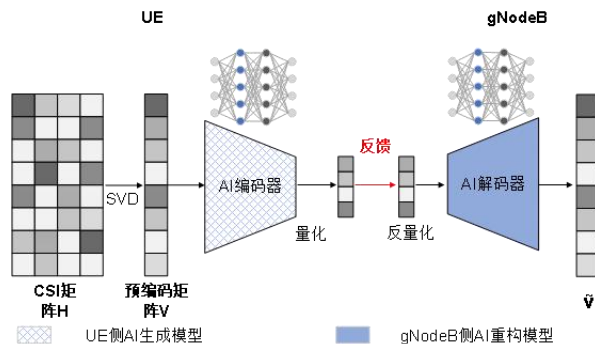


图 3.3-1 基于 AI 的 CSI 压缩推理过程示例

2、终端波束预测及管理

大规模 MIMO 技术通过特定的窄波束将发射能量集中在接收机上，从而提高系统性能。随着波束数量的增加，理论上可以提高性能增益。在实际系统部署中，为了获得性能增益，如何选择和映射最佳的窄波束是波束管理（Beam Management, BM）过程中最关键的挑战之一。3GPP 在 R18 标准化阶段启动了相关研究，利用 AI/ML 技术来减少波束管理开销，最小化延迟，并提高波束选择的准确性。空域下行波束预测技术通过利用 AI/ML 模型，结合另一组下行波束的测量结果，推断出一组下行波束中的最佳波束。时域下行波束预测技术通过利用 AI/ML 模型，结合另一组下行波束的历史测量结果，推断出在未来某个时刻的一组下行波束中的最佳波束。

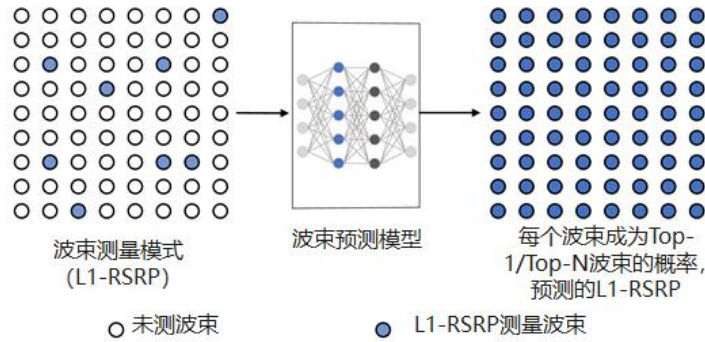


图 3.3-2 波束管理的推理过程示例

3、终端弱信号预测

对于终端用户，影响体验的重要问题是移动场景下弱信号/无信号/信号拥塞造成的链路卡顿。如果手机能够提前预测弱信号/无信号/链路卡顿区域，并且提前对视频做缓存或者提前规避异常小区，及时选择更优网络，就可以最大限度的保障视频流畅度，让用户有更好的观赏体验。

如图 3.3-3，终端基于 AI 的弱信号预测是通过机器学习方式，通过采集用户移动时的历史网络信号参数作为输入，用于模型推演预测出用户未来进入弱信号区域的时间，终端提前做出应对策略，以提升应用在弱网下的业务体验。

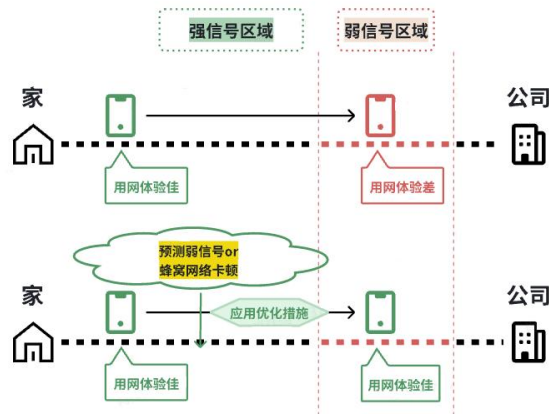


图 3.3-3 终端基于 AI 的弱信号预测示意图

4、终端 RRM 及移动性管理增强

3GPP 在 R19 标准化阶段启动在 NR 中应用 AI/ML 以增强移动性的研究，目标是评估 AI/ML 辅助移动性对于网络触发的基于 L3 的切换所带来的潜在优势和收益。基于 AI/ML 的 RRM 测量考虑 UE 端的模型和网络端模型，潜在的应用包括：

1、跨小区的空间/时间波束预测，将单一小区内波束预测扩展到了多个小区，以实现移动性的支持，从而减少参考信号（RS）的开销和 UE 测量工作。

2、在用户终端测量受限和需要测量间隙的情况下进行智能化的 UE 测量。一方面，这项应用可以减少对测量间隙的需求，从而减小数据传输中断，同时探索

基于一个频率的测量来预测另一个频率的 RRM 测量的可能性。另外一方面，通过 AI/ML 的智能化处理，用户终端可以对测量目标（频率/小区）进行优先级排序，帮助确定哪些测量目标（频率/小区）更为重要，从而优先分配资源，减少高优先级目标的测量延迟和保证其测量的要求。

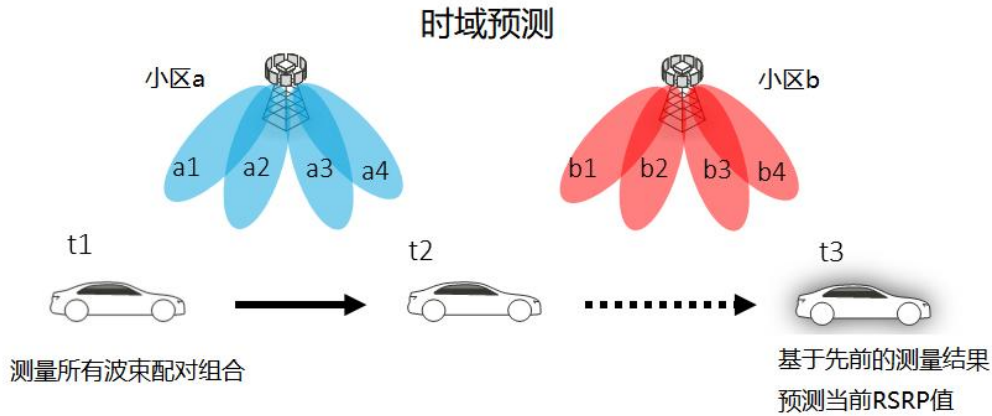


图 3.3-4 终端基于 AI 的 RRM 测量增强

4. 总结与展望

在 5G 移动通信系统中引入 AI 技术，不仅能够提升和增强传统网络业务和网络服务能力，更能进一步拓展 5G 应用新场景。在传统网络应用场景中，AI 技术对提升 5G 网络的运维能力和用户体验起到关键作用。本白皮书重点介绍了 5G 网络 AI 应用的发展趋势、总体架构、典型场景特点及解决方案和未来演进，为 5G 网络引入 AI 应用提供指导。2024 年中国移动重点对基于业务的体验升级、智能网络优化、智能网络节能、5G+AI 用户画像、基于客户及特定场景的资源保障等典型应用，基于网络 AI 应用开展技术攻关与验证，并将联合产业共同探索网络侧更多应用。同时，网络智能化能力也将逐步与终端协同，进一步提升网络效率。

本白皮书的发布，希望为 5G 网络 AI 应用的发展提供可参考的场景需求和技术方向，促进产业形成共识，共同推动 5G+AI 融合发展，加速经济社会发展从“+AI”向“AI+”迈进。

缩略语列表

缩略语	英文全名	中文解释
3GPP	3rd Generation Partnership Project	第三代合作伙伴计划
GTI	Global TD-LTE Initiative	TD-LTE全球发展倡议
ETSI	European Telecommunications Standards Institute	欧洲电信标准化协会
TMF	Tele Management Forum	电信管理论坛
CCSA	China Communications Standards Association	中国通信标准
KPI	Key Performance Indicator	关键性能指标
MIMO	Multiple-input multiple-output	多入多出技术
AI	Artificial Intelligence	人工智能
AN	Autonomous Network	自智网络
QoS	quality of service	服务质量
QoE	quality of experience	客户感知体验
XR	extended reality	扩展现实
eMBB	enhanced Mobile Broadband	增强移动宽带
SLA	Service Level Agreement	服务水平协议
OPEX	operational expenditure	运营支出
SDL	supplementary downlink	辅助下行
SUL	supplementary uplink	辅助上行
RSRP	reference signal received power	参考信号接收功率
MCS	modulation and coding scheme	调制和编码方案
eMBB	enhanced Mobile Broadband	增强移动宽带
MU-MIMO	Multi-User Multiple-Input -Multiple-Output	多用户多入多出技术
AI/ML	Artificial Intelligence/Machine Learning	人工智能/机器学习
OSS/MR	Operation support system/Measurement Report	网管/测量报告
ERAB	Evolved Radio Access Bearer	用户面的承载
RLC	Radio Link Control	无线链路层控制协议
PRB	Physical Resource Block	物理资源块
RRC	Radio Resource Control	无线资源控制层
BBU	BaseBand Unit	基带处理单元
PRB	Physical Resource Block	物理资源块
MR	Measurement Report	测量报告

KNN	K Nearest Neighbors	K- 近邻
Resnet	Residual Network	残差网络
Kmeans	K-means	K均值聚类

参考文献

- [1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30. preprint arXiv:2212.01020.
- [2] 3GPP TR 28.810: “Study on concept, requirements and solutions for levels of autonomous network”.
- [3] 3GPP TS 28.100: “Management and orchestration; Levels of autonomous network”.
- [4] 3GPP TR 28.812: “Telecommunication management; Study on scenarios for Intent driven management services for mobile networks”.
- [5] 3GPP TS 28.312: “Intent driven management services for mobile networks”.
- [6] 3GPP TS 28.535: “Management and orchestration; Management services for communication service assurance; Requirements”.
- [7] 3GPP TS 28.536: “Management and orchestration; Management services for communication service assurance; Stage 2 and stage 3”.
- [8] 3GPP TR 28.809: “Study on enhancement of Management Data Analytics (MDA)”.
- [9] 3GPP TS 28.313: “Self-Organizing Networks (SON) for 5G networks”.
- [10] 3GPP TR 28.861: “Study on the Self-Organizing Networks (SON) for 5G networks”.
- [11] 3GPP TR 37.816: “Study on RAN-centric data collection and utilization for LTE and NR”.
- [12] 3GPP TS 23.288: “Architecture enhancements for 5G System to support network data analytics services”.
- [13] 3GPP TR 23.700-91: “Study on Enablers for Network Automation for 5G – phase 2”.
- [14] TM Forum: IG1218, Autonomous Network Business Requirements and Framework. 2022

- [15] 左修玉. 一种基于模式识别的无线网络干扰自动排查方法 - 中国知网 [EB/OL]. /2023-07-05.
- [16] 李宗璋. 一种基于 MR 数据确定小区网络干扰分类的方法及装置 - 中国知网 [EB/OL]. /2023-07-05.
- [17] 左修玉. 一种用于干扰类型识别的低功耗深度学习网络方法 - 中国知网 [EB/OL]. /2023-07-05.
- [18] 张文婧. 一种多调制体制下基于深度学习的干扰类型识别方法 - 中国知网 [EB/OL]. /2023-07-10.
- [19] TM Forum. IG1252 Autonomous Network Levels Evaluation Methodology. 2021
- [20] TM Forum. IG1251 Autonomous Network Reference Architecture. 2021
- [21] TM Forum. IG1253 Intent in Autonomous Networks. 2022
- [22] TM Forum. IG1326 AN_Empowering_DT_evolution_towards_level4. 2023
- [23] ETSI GS ZSM 002: "Zero-touch Network and Service Management (ZSM); Reference Architecture".
- [24] ETSI GR ZSM 009-3: "Zero-touch Network and Service Management (ZSM); Closed-loop automation; Advanced topics".
- [25] ETSI GS ZSM 009-1: "Zero-touch Network and Service Management (ZSM); Closed-loop automation; Enabler".
- [26] ETSI GS ZSM 009-2: "Zero-touch Network and Service Management (ZSM); Closed-loop automation; Solutions".
- [27] ETSI, ETSI GR ENI 007 V1.1.1 (2019-11): "ENI; ENI Definition of Categories for AI Application to Networks".
- [28] IMT2020(5G) 推进组: 5G-Advanced 场景需求与关键技术

参编单位及人员

(排名不分先后)

中国移动计划建设部：张鹏、李继、毛剑慧、黎丹、刘莉莉

中国移动研究院：黄宇红、丁海煜、曹蕾、魏彬、王成、高精、阮航、张佳辰、
郭宣羽、江天明、李爱华、刘乐、史嫻嫻