

5G-AxAI 新技术， 新案例，新模型 白皮书

<https://gtigroup.org/>

GTI

5G - AxAI 新技术，新 案例，新型模型白皮书



版本：	V1.0
交付类型	程序性文件 ☒ 工作文档
保密等级	☒ 对GTI运营商成员开放 ☒ 愿意与GTI伙伴合作 ☐ 面向公众
项目	5G-AxAI
工作组	N/A
项目	项目1：网络智能 项目2：数字孪生网络智能 项目3：应用智能 项目4：可持续性洞察力
任务	N/A
源成员	中国移动，日本电信运营商北京实验室，华为，爱立信，中兴通讯，诺基亚、英特尔、高通、联发科、紫光展锐、小米、OPPO、vivo、荣耀
支持成员	
编者	
最后编辑日期	
批准日期	

保密： 此文件可能包含机密信息，且访问权限仅限于保密级别中列出的个人。未经GTI事先书面授权，不得使用、披露或部分或全部复制此文件，且获授权者只能为授权目的使用此文件。GTI对于此文件中信息的准确性、完整性或及时性不承担任何责任。此文件中的信息可能会在未事先通知的情况下发生变化。

文档历史

日期	会议 #	版本 n #	修订内容

目录表

GTI 5G-AxAI 新技术、新案例、新模型 白皮书	2
文档历史	3
目录	4
1 高管总结	5
2 5G-AxAI：新能力满足新需求，释放新价值	5
2.1 5G-A 商业进步	5
2.2 新的人工智能能力	6
2.3 跨领域融合的机遇	7
2.4 5G-A与AI的乘数效应	11
3 5G-AxAI 新兴技术	13
3.1 网络智能	14
3.2 数字孪生网络智能	20
3.3 应用智能	23
3.4 可持续智能	29
4 5G-AxAI 实现新案例	31
4.1 区分化的体验保障	31
4.2 新的呼叫服务	33
4.3 工业确定性服务	37
4.4 绿色能源节约	39
4.5 高可靠性网络	41
4.6 多模态个人助理	45
4.7 内在人工智能	47
4.8 沉浸式体验	48
4.9 网络智能驾驶	53
4.10 低空智能连接	55
第五次新模型产业革命	59
5.1 三个ToC/BtoC的创新商业模式	59
5.2 三种针对ToB的创新商业模式	60
6 全球产业合作提案	62
7 词汇表	62
8 参考文献	64

1 摘要

在当前技术创新的时代，人工智能（AI）技术正迅速发展，为网络带来了新的机遇。5G-Advanced（5G-A）与人工智能的融合将成为一种必然的产业趋势，推动电信及其他产业产生乘数效应。一方面，5G-A×AI将满足网络需求，提升网络性能和效率。另一方面，5G-A×AI还将为产业提供新的服务，加速产业智能化革命。5G-A×AI的创新技术已在4个主要领域取得显著成效：在网络智能领域，它使网络表现出极高的质量；在数字孪生网络智能领域，它实现了低成本试验和高效率创新；在应用智能领域，它扩大了网络服务范围；在可持续智能领域，它实现了低碳目标。所有这些技术都为新应用提供了创新引擎，涉及服务保障、个人AI代理、具身人工智能、自动驾驶等。同时，它也引领了新的商业模式，标志着网络服务从连接向连接、计算和智能融合的转变。

2 5G-AxAI：新能力满足新需求， 释放新价值

前沿的人工智能技术为社会带来新的能力，使人们的生产和生活方式发生创新。^[1] 人工智能发展进入新阶段，变得轻量化、通用化和具体化。优化算法极大地降低了人工智能成本，并降低了小型企业的进入门槛；多模态基础模型统一了异构的人体数据，并扩大了应用范围；机器人集成了理解和思考能力，使人工智能在现实世界中具体化。

在同一阶段，通信网络正经历大规模部署和创新的时期。5G网络数量达到398个，全球5G SA运营商达到154家。下一阶段的5G-A系统开始成形，并由超过15家运营商实施。

一部新的故事在两个先进且快速发展的行业相遇时展开。随着5G-A网络的加入，人工智能的效果和价值进一步扩大。5G-A与人工智能的融合成为一种不可避免的趋势，并在技术、应用、产业等多个方面带来变革。基于5G-AxAI技术的发展，将满足网络和产业的新需求，提高社会效率并创造新价值。

2.1 5G-A 商业进步

2024年中，3GPP已经完成了Release 18的相关工作，预计其Release 19的工作将在2025年底完成。5G已正式进入5G-A阶段，并为社会提供多重价值。^[2-4] 首先，高速、低延迟和几乎无处不在的连接能力。其次，各种创新信息的整合。

技术方面。第三，行业服务的能力基础。总之，5G-A在速率和智能能力方面实现了显著的升级。根据数据，实际的5G-A速率已跃升至3-5Gbps，预计未来将超过10Gbps，与传统的5G网络相比增加了10倍。极大规模天线阵列（ELAA）等技术被用于提高高频带的覆盖范围，多载波协调技术被用于高效聚合多个频带的带宽，从而大大提高了频谱效率。在终端侧，2024年发布了新型高端芯片，以支持超过6组件载体（CC）的大带宽能力。此外，5G-A的持续能力升级也开辟了行业的新蓝海。例如，5G-A车联网（IoV）支持在车辆道路云上进行大规模试点项目，并加速了智能网联汽车行业的发展。

从全球视角来看，世界上超过60家运营商和合作伙伴已宣布5G-A商用计划。5G-A已成为业内热点话题。2024年12月10日，以“世界首个5G-A区域启航”为主题的特殊研讨会启动。来自权威行业组织、监管机构、领先运营商和设备供应商（如GSMA、TDRA、DU、e & Vodafone、Ooredoo、华为、爱立信和诺基亚）的专家分享了5G-A的创新。在中国，5G-A网络已在300多个城市启动，中国为世界树立了一个良好的商用典范。在其他地区，许多运营商已开展5G-A验证，例如亚太地区HKT、CTM、马来西亚Maxis、欧洲DNA、Vodafone、拉丁美洲VIVO和TIM。

在中国，中国移动积极推动5G-A网络的实施，在采用创新技术方面走在前列。例如，它部署了3CC聚合网络和5G-A智能控制平面（NWDAF），提高了传输效率，并为关键客户提供独特的保障能力。中国移动还为旅行、游戏和直播发行了5G-A收费套餐，已有超过50万5G-A用户订阅。例如，5G-A商务旅行计划在下行链路提供高达3Gbps的高速体验，在上行链路提供高达200Mbps的体验。

在中东，运营商推出了最高300Mbps的无线宽带套餐。无线家庭宽带利用5G-A高速网络提供与光纤相似的高质量体验，并解决了光纤部署困难和高建设成本的问题。

总体而言，5G-A 在发展和商业化方面具有巨大潜力。它目前处于持续发展阶段和大规模部署阶段，期待新驱动力的到来以实现进一步的突破。

2.2 新的人工智能能力

人工智能是使机器模拟人类智能的技术，使他们能够处理诸如学习、理解、问题解决、感知、决策和自主性等任务。最新的大型基础模型（例如，GPT-4o、Deekseek）已展现出其在理解自然语言和思考问题上的非凡能力。人工智能算法包括监督学习、无监督学习和强化学习。其中，深度学习是一种有效且强大的方法，已在各个领域得到广泛应用。具有多种不同层（例如，MLP、CNN、RNN、Transformers）的神经网络

展现了他们在感知、预测、决策和优化方面的强大能力。这些能力已经在医疗保健、金融、交通、治理和其他许多行业中显示出其价值。

1) AI感知

人工智能能够通过处理从各种传感器收集到的数据来感知和理解周围环境。它使机器能够模拟人类感知世界的方式，包括视觉、听觉、触觉和其他感官。它甚至超越了人类，因为它还可以直接从网络空间获取信息。例如，在医院中，人工智能可以监测患者的生命体征并检测早期恶化的迹象，这样医疗人员就能及时发出警报，提供紧急援助。

2) 人工智能预测

人工智能也能根据历史数据和当前信息预测未来的事件、趋势或结果。它具有处理复杂数据并识别人类可能无法察觉的模式的能力。在天气预报中，AI模型可以通过分析大量气象数据（包括温度、湿度、风速和气压）来准确预测未来的天气状况。在金融领域，AI可以根据历史股票数据、经济指标和新闻情绪预测股价和市场趋势。

3) 人工智能决策

人工智能可以分析数据，评估各种选项，并提出建议或做出决策以实现特定目标。它具有将多模态输入条件归一化到统一矢量空间的能力，提取共同特征，并对它们进行加权以得到最终决策。在自动驾驶领域，AI系统可以帮助分析来自不同传感器的数据，如摄像头、激光雷达和雷达，以检测障碍物、其他车辆和行人。然后，它根据何时加速、制动或转向来避免碰撞并遵守交通规则做出决策。

4) 人工智能生成

人工智能能够通过学习和分析大量相关数据来创造新的内容。这种能力可以适应并扩展到不同的领域（文本、音频、图像、视频）。在创意领域，它为艺术家、设计师、作家、音乐家提供创意灵感，帮助他们快速生成创意内容。在教育领域，它生成教学资源，如教学动画和个性化学习材料，有助于提高教学效果。

所有这些人工智能能力都可以在网络领域被利用和改进。结合低延迟、高可靠性、普遍连接等网络能力，通信智能融合服务将出现并推动行业进入下一阶段。

2.3 跨领域融合的机遇

2.3.1 满足演进的网络需求

5G-A和AI的融合在满足网络演进需求和实现卓越网络性能方面提供了变革性潜力。AI技术可以以数据驱动的方式优化整个链路或系统中的多个功能模块，从而提高整个端到端通信系统。通过结合5G-A和AI的先进能力，网络可以实现前所未有的效率、适应性和可靠性，满足现代连接不断增长的需求。以下是关于如何的分析。

人工智能技术在网络领域解决关键挑战。

1) 效率资源利用

高效资源利用在5G-A中至关重要，以满足对连接性的日益增长需求，尤其是在密集城市地区、工业物联网（IoT）应用和高流量环境中。人工智能在优化关键网络资源，如频谱、带宽和基础设施方面发挥着关键作用。

动态资源管理：人工智能算法分析实时交通状况并预测需求模式，使5G-A网络能够动态分配诸如带宽、电力和频谱等资源到最需要的区域。这确保了网络资产的效率利用和流畅适应不断变化的交通负载。

网络切片优化：通过5G-A的网络切片能力，虚拟网络可以根据特定应用（例如，物联网、增强型移动宽带和低延迟服务）进行定制。人工智能通过实时优化每个切片的资源分配来增强这一点，以适应网络性能指标和用户需求。

预测交通分析：AI驱动的预测模型利用历史数据和使用趋势来预测网络拥堵或瓶颈。这种主动方法允许网络提前重新分配资源，确保即使在高峰使用期间也能保持不间断运行。

AI增强频谱效率：AI算法根据实时需求动态分配可用频率，最小化干扰并最大化频谱效率。AI还能实现运营商和技术之间的有效频谱共享，提高网络容量。

优化MIMO技术：5G-A的多输入多输出（MIMO）技术显著提升了数据吞吐量。人工智能动态调整MIMO参数，在维持高数据速率的同时，以最少的功耗优化系统性能。

2) 以用户为中心的网络运营

用户中心化的运营对于确保用户获得一致可靠的服务水平至关重要，包括关键指标如带宽、延迟、抖动和丢包率。对于5G-A，由于网络支持的服务范围广泛，服务质量（QoS）优化尤为重要，这保证了用户体验并平衡了网络负载。

实时监控与调整：人工智能使网络性能和用户体验的持续监控成为可能。通过实时分析延迟、吞吐量和数据包丢失等指标，人工智能可以动态调整网络参数，以确保最佳的QoS。例如，它可以将时间敏感型流量，如自动驾驶汽车或远程手术，优先处理，以最大限度地减少延迟和数据包丢失，同时为视频流等其他服务分配充足资源。

智能交通整形：人工智能算法可以根据服务类型（如视频、语音或物联网）对流量进行分类。通过了解每个服务的具体需求，人工智能可以优先处理高优先级流量，并高效地路由低优先级流量，确保流畅的性能并避免拥堵。

上下文感知交通优先级：人工智能利用上下文数据，如一天中的时间、用户位置和网络条件，动态调整服务质量设置。例如，在高峰时段，人工智能可以优先处理与紧急服务或医疗保健相关的应用程序，而将一般浏览或流媒体应用程序放在次要位置，确保关键服务不受影响。

端到端服务质量保证：人工智能可以实现对端到端服务质量的控制，确保服务质量标准不仅在网络层面得到满足，同时也在应用层面得到满足。人工智能可以分析

整个网络路径——从边缘到核心——用于检测可能影响服务交付的任何潜在问题。这种整体方法确保了更好的用户体验，尤其是在智能城市、自动驾驶和工业自动化等关键应用场景中。

智能交通卸载：人工智能预测交通模式并主动将流量卸载到替代的、高效的路径中。这包括将数据重新路由到边缘计算节点、相邻小区或其他优化网络路径。这些前瞻性措施确保核心网络不拥塞，从而提高整体网络的稳定性性能。

3) 网络自动化

智能算法可以显著提升网络管道的性能，降低运维（O&M）成本，并提高新服务推出的效率。此外，人工智能将推动运维实践向L4/L5级别发展，协助电信运营商提升服务质量并增加收入。

敏捷网络：在当前的操作维护系统中，各个功能独立运行。为了实现任务目标，操作维护人员必须手动编排和整合这些功能，这需要频繁与系统交互，导致操作维护效率降低。

人工智能技术正在将运维模式从以功能为中心转变为以任务为中心，使系统更好地满足人类需求。通过将统一的知识模型纳入运维系统，运维人员可以使用自然语言与之进行交流。这使系统能够理解用户意图，并自主编排和集成功能，无缝地完成任务目标。

绿色网络：网络节能与体验保证之间常常存在矛盾。节能方面最大的矛盾在于我们不知道何时关闭以及关闭到什么程度。

利用人工智能算法，网络可以动态平衡网络资源分配和节能策略，以实现从非高峰时段到高峰时段的平稳过渡。通过实施上述智能策略，可以实现经验保证与节能之间的有效协作，显著降低网络能耗，同时保证用户体验。

服务创新赋能：电信行业正在积极探索新的网络能力和服务。为了满足各种需求，有必要摆脱基于人工经验和操作的传统的提供服务工作模式。

在新服务提供阶段，运营商可以直接以自然语言的形式传达服务需求。借助人工智能方法，网络能够通过将意图翻译成机器代码来理解意图，在网络孪生环境中进行在线模拟，并自动生成网络配置以满足新的需求。

2.3.2 满足多元行业需求

5G-AxAI的行业需求因5G-A和AI两种颠覆性技术的融合而迅速增长。这种组合预计将能够在多个行业中实现各种新的应用和服务。

1) 远程医疗和远程医疗服务

5G-A提供低延迟和高速数据传输能力，这对于实现实时视频咨询甚至远程手术操作至关重要。AI增强了这些功能。

通过支持分析医学图像、传感器数据和患者病史的诊断工具，提升能力，协助医疗专业人士做出明智的治疗决策。

远程医疗和远程手术：5G-A提供了实时远程会诊和手术所需的宽带和低延迟。人工智能可以协助医生进行诊断或通过机器人辅助进行手术，利用人工智能驱动的成像和数据分析提高准确性。

健康监测：由5G-A网络增强的人工智能设备能够实时监测患者健康数据。这些系统通过分析可穿戴设备、传感器和其他医疗设备的数据，能够在问题变得严重之前预测潜在的健康问题。

人工智能辅助诊断：AI驱动的模式可以以显著的准确性分析医学扫描，包括X射线和磁共振成像（MRI）。当与5G-A快速传输大数据集的能力相结合时，这些工具能够实现更快、更精确的诊断，尤其是在偏远或欠发达地区，从而提高获得优质医疗保健的途径。

2) 自动驾驶车辆与交通

5G-A在促进自动驾驶车辆发展中扮演着至关重要的角色，它提供了令人印象深刻的数据传输并支持大规模的连接。随着网络中部署的计算系统，车辆能够对周围环境拥有更全面、更详细的感知，并做出更准确的驾驶决策。

自动驾驶汽车：5G-A提供了自动驾驶汽车与其他车辆及基础设施（如交通灯和道路传感器）无缝通信所必需的高速、超低延迟的连接性。人工智能处理来自多个传感器（包括激光雷达、摄像头和雷达）的数据，使得汽车能够进行实时决策，实现安全高效的驾驶。

车辆与一切（V2X）：5G-A支持V2X通信，促进车辆、基础设施甚至行人间数据的交换。人工智能利用这些数据进行高级交通管理、预测性路线优化和实时交通更新，提升道路安全与出行效率。

车队管理和物流：5G-A使得在大规模车辆队伍中实现实时监控和无缝通信。通过人工智能算法分析数据来预测维护需求，优化配送路线并提高燃油效率，确保成本效益和可持续的物流运营。

3) 制造业与工业4.0

5G-A网络连接工厂中大量的设备和传感器，并将大量数据（例如，生产过程参数、设备运行状态）传输至网络或应用服务器中的中央控制系统进行分析和处理，从而实现生产过程的实时监控和优化。

智能制造：5G-A提供所需的宽带，以便实时监控机器、生产线和机器人系统。人工智能分析传感器数据以优化生产流程，在它们导致故障之前识别出低效或潜在问题。

预测性维护：人工智能算法利用传感器数据在设备发生故障之前预测故障。凭借5G-A的持续数据传输能力，系统可以发送实时维护警报，减少停机时间并提高运营效率。

机器人自动化：5G-A的超低延迟使得AI驱动的机器人能够实时执行高度同步的任务，如组装和质量控制，显著提高制造环境中的精度和生产力。

4) 零售与电子商务

5G-A和AI对零售和电子商务产生了巨大影响，带来了购物体验、物流以及供应链管理等方面的变革。实时数据共享和通信使需求预测更加准确，客户体验更佳，从而提高了整体效率、灵活性和吸引力。

个性化客户体验：AI分析客户行为，提供个性化推荐和针对性促销。5G-A通过实现实时数据交换和在零售店中使用增强现实（AR）等沉浸式体验来支持这一功能。

库存管理与物流：人工智能可以实时预测产品需求并优化库存水平。结合5G-A追踪运输和仓库操作的能力，零售商可以提高效率并减少缺货情况。

智能商店：5G-A技术使得物联网设备（例如智能货架、传感器）在零售环境中的集成成为可能。人工智能分析这些设备的数据，以优化产品陈列、检测盗窃和优化店内体验。

5) 智慧城市与城市基础设施

5G-A 和人工智能是构建智能城市的关键推动因素，促进了智能治理、智能安全、智能交通等领域的重要进步。通过技术发展，城市能够实时监控状态，做出全面决策，并管理城市运营。

交通管理：5G-A网络非常适合智能城市基础设施，它能够实现从城市各处的传感器进行实时数据分析，以优化交通流量、减少事故并提高整体城市效率。在处理大规模数据以做出智能决策方面，人工智能是关键。

公共安全和监控：基于人工智能的安全系统，包括人脸识别和异常检测，得益于5G-A网络的速率和连接性。这些系统可以处理监控摄像头的实时数据，以增强公共安全。

5G-A和人工智能技术的需求覆盖了广泛的行业，每个行业都在利用高速连接、实时数据处理和智能决策的结合。医疗保健、汽车、制造、电信、零售和智能城市等行业正在利用这些技术来提高效率、提升客户体验并解锁新的应用。5G-A和人工智能之间的协同作用不仅增强了当前的使用案例，还促使开发出创新的新一代应用，这些应用曾经是无法想象的，从而推动这些行业的数字化转型。

2.4 5G-A 和 AI 的乘数效应

5G-A与AI的结合产生强大的乘数效应，放大了每项技术在多个领域的影响力。这种结合加速了创新，提高了效率，提升了客户体验，并促进了新商业模式的开发。 [5] .

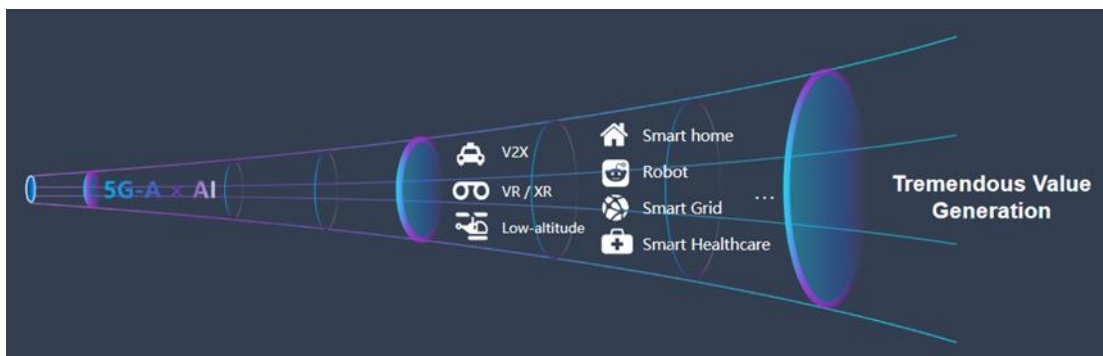


图2-1 5G-A与AI的融合产生巨大价值

5G-A与AI碰撞的成果是融合效应，而非简单的相加。两个平行宇宙（5G-A领域和AI领域）碰撞并扩展到更高的维度，其内在粒子之间产生相互作用力，释放出巨大能量。这种能量扩散到不同的应用领域，并为人类社会展开巨大价值。这种价值主要体现在三个方面：

1) 提升网络性能与效率

网络优化：人工智能通过促进自主网络管理、预测性资源分配和故障检测，增强了5G-A技术。这导致网络资源（如带宽和电力）的更有效利用，降低服务提供商的运营成本，同时提升用户体验。

网络自动化：人工智能（AI）使得网络自动化成为可能，允许5G-A网络进行自我优化和自我修复。这减少了人工干预的需求，加快了网络部署速度，并提高了可扩展性，以满足物联网（IoT）和连接设备的不断增长需求。

低延迟应用：AI模型帮助优化交通路由和预测网络拥堵，支持自动驾驶汽车、增强现实和远程手术等超低延迟服务。5G-A提供高数据速率和海量连接的能力确保了这些AI驱动服务的平稳且不间断运行。

预测性分析：由AI驱动的分析性维护和交通管理允许5G-A网络预测需求变化、检测瓶颈和进行主动维护。这最小化了服务中断、降低了运营成本并提高了网络可靠性。

2) 解锁新的服务和应用

边缘计算与实时处理：5G-A支持边缘计算，即在数据源附近进行处理，例如物联网设备或传感器。这降低了延迟和带宽使用，使得需要实时处理的AI应用成为可能，包括自动驾驶汽车、智能制造和智慧城市。

大规模连接性：5G-A能够容纳庞大的物联网生态系统，数百万个设备传输小数据包。当与AI处理和基于这些数据进行决策的能力相结合时，行业可以优化运营，改善客户服务，并开发新的商业模式。例如，在智能农业领域，AI可以分析传感器数据以优化灌溉和作物管理，而5G-A确保农场内无缝实时数据传输。

沉浸式通信：在教育、娱乐和零售等领域的AI驱动AR/VR应用通过5G-A能力提供高质量、低延迟的沉浸式体验得到增强。

体验。这为虚拟培训、远程协作和虚拟产品试用开辟了创新机会。

3) 加速产业革命

自动化与优化：人工智能推动智能化在各个行业的应用，包括制造业（工业4.0）、医疗保健、交通和智能城市。在制造业中，由人工智能驱动的机器人和预测性维护系统，得益于5G-A的超低延迟，提高了生产效率并最小化了停机时间。在医疗保健领域，人工智能算法实时分析患者数据，以创建个性化的治疗方案，并得益于5G-A的快速数据传输。

实时分析和个性化：借助5G-A的连接能力，人工智能使行业能够根据实时数据实现个性化服务。例如，在零售领域，人工智能分析客户行为以提供定制化推荐，而5G-A确保从物联网设备、传感器和移动应用程序无缝传输数据，以提升购物体验。

改进决策：AI处理复杂数据集的能力为企业提供市场趋势、客户偏好和运营低效率的实时洞察。这使得行业能够做出快速、基于数据的决策，提高企业敏捷性和竞争力。

预测性和规范性分析：在物流领域，人工智能驱动的预测模型优化了路线和库存管理，而5G-A的连接性促进了货物运输中的实时监控。这导致了成本节约、提升客户满意度以及更高的运营效率。

5G-A与人工智能融合的乘数效应深远，创造了性能提升、新的商业机会和创新的巨大价值。人工智能驱动智能决策和自动化，而5G-A提供了支持这些进步所需的速度和可靠性的连接基础设施。共同作用下，它们推动了行业的数字化转型，提高了运营效率，促进了新的商业模式，并带来了社会效益。结果是，一个动态的跨行业生态系统，增长和创新得到放大，使企业、消费者和社会受益。

3 5G-AxAI 培育新技术

5G-A与人工智能的融合带来了革命性的创新。核心创新集中在四个主要领域：1) 网络智能，旨在通过智能优化提升网络质量；2) 数字孪生网络智能，旨在实现低成本试验和高效率创新；3) 应用智能，旨在通过增强功能扩展网络服务范围；4) 可持续智能，旨在通过智能节能实现重要的环境目标。这些创新共同推动了覆盖服务保障、个人AI代理、具身AI和智能网络基础设施生命周期管理等多元领域的应用新浪潮。通过在这些技术领域的持续进步，5G-AxAI正在积极重塑网络范式，加速行业转型，并推动从简单的连接到连接、计算和智能融合的一体化进化。

3.1 网络智能

3.1.1 智能实时网络闭环

通过引入智能控制平面和用户平面，5G-A能够实现实时应用识别、体验感知以及动态服务控制，从而支持运营商持续优化网络资源调度，并实现差异化服务保证。

1) 实时体验感知

通过网络数据分析功能（NWDAF）智能分析能力和用户平面功能（UPF）流量识别能力的协调，智能网络能够根据大量网络数据提供精确且实时的用户体验评估。该过程包括四个阶段：数据处理、AI模型结构设计、模型预训练和模型微调。

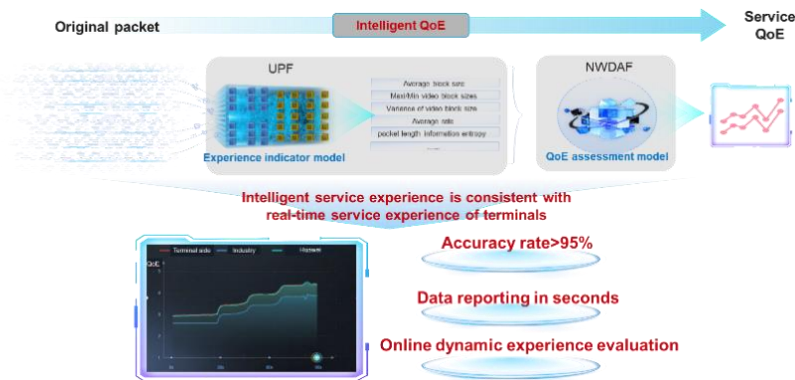


图3-1 实时体验感知框架

数据处理旨在为模型训练提供高质量、易于使用的语料库，为后续AI模型的准确性和泛化能力奠定基础。数据处理考虑了多个方面，包括隐私安全、特征选择、编码、分词。

人工智能模型结构设计涉及定义和设计适用于网络操作状态数据压缩和建模的基本模型。它基于Transformer和Mamba结构设计领域模型，集成多模态学习以融合信令、流量和其他多源数据。通过结合知识蒸馏和模型量化技术，复杂模型被压缩成轻量级版本，以满足UPF和NWDAF的实时数据分析和应用需求。

模型预训练旨在通过在大规模数据上全球训练来构建具有一般特性的基础模型，同时优化训练过程以提高训练效率和效果。通过自适应损失加权算法优化多任务学习性能，动态调整任务优先级；此外，通过结合分布式训练框架，它还提高了大规模模型训练的效率。

该模型微调通过基于特定服务场景的训练为模型提供高精度任务能力。在微调阶段，可以使用迁移学习来...

快速适应不同的场景，并且在小样本学习中可以使用，以在数据不足的场景中提高模型性能。基于微调技术，应用感知和质量体验（QoE）感知任务可以共享相同的基本模型，并且可以独立升级。

2) 区分化的体验保证

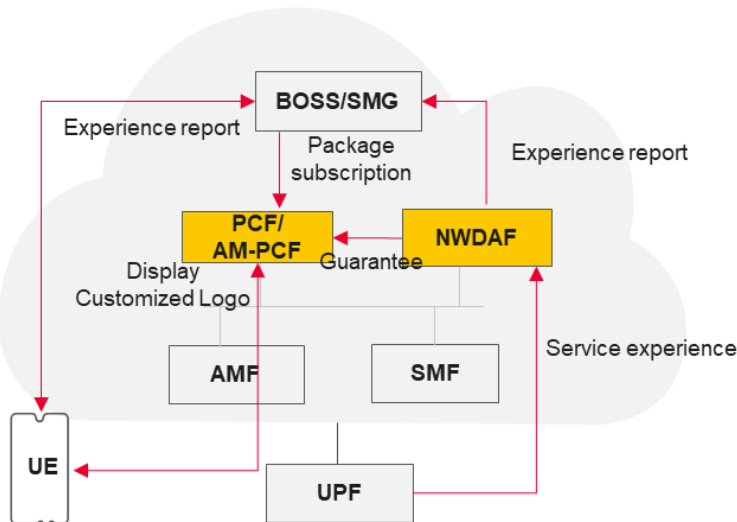


图3-2 分层体验保障框架

差异化的体验保证技术有助于运营商最大化资源利用并提高盈利能力。具体而言，当网络处于低负载状态时，差异化的体验保证技术可以刺激用户流量，使用户愿意将直播等高流量服务迁移到移动网络；当网络处于高负载状态时，差异化的体验技术可以为关键客户的关键服务提供充分的保证，从而确保网络状态不会降低他们的体验。

在服务保障过程中，UPF（用户平面的功能）持续分析用户体验并向NWDAF（网络数据访问功能）报告分析结果。NWDAF根据不良体验触发保障策略。在政策控制功能（PCF）交付保障策略后，它会在用户设备（UE）上触发专用的品牌标识显示。保障完成后，主动推送体验报告，以使用户感知保障结果并简化端到端体验闭环。

3.1.2 网络基础模型

通信网络是所有流量的聚合节点，可以获取大量有价值的数据。借助大型语言模型（LLM）技术，构建融合网络知识的基座模型成为可能，从而实现精确的网络操作和全面的用户分析。与传统的AI方法相比，基座模型在通用性和鲁棒性方面得到了显著提升，提供了更强的感知和优化能力。此外，它使得任务适应更加容易，避免了网络功能的重复开发。

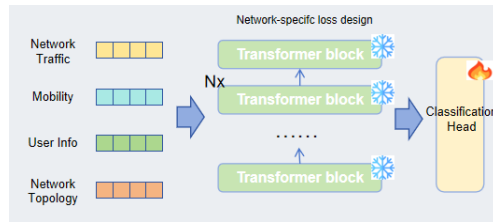


图3-3 智能网络的基础模型设计

总体而言，大型语言模型（LLM）主要关注处理语音、文本和图像，而不是网络数据。为了充分发挥LLM在网络领域的潜力，需要从以下方面重新设计基础模型。首先，重要的一点是定制网络数据的标记化方法。网络数据需要根据其特性进行标记化，这会将多维网络数据（例如，交通、用户信息、拓扑）映射到向量空间，并帮助模型提取高级数据特征。其次，应重新构建模型结构和优化目标。适当的层、参数、损失函数和训练策略的设置将确保模型在整合如此多的异构网络数据时能够收敛。第三，鉴于当前网络状况尚不具备高效计算能力，但需要实时处理，应降低模型适应和部署的成本。这可以通过参数高效微调（PEFT）和模型压缩实现。

在网络流量识别领域，与典型方法不同，基础模型不依赖于纯文本规则，并且能够提取高级隐含特征以区分不同类型的包和流量。特别是在加密流量、私有协议和未注册服务类型的情况下，它带来了更高的准确性。此外，网络本身支持跟踪与用户相关的数据，如位置变化、网络接入和流量使用，这使得网络能够感知用户的偏好、生活方式和职业。鉴于基础模型在语义理解、上下文信息处理和跨域适应能力方面的出色表现，它可以通过整合多源异构的用户相关数据，成为更精确的移动用户画像的重要工具，从而进一步促进更个性化营销策略的发展。

3.1.3 智能网络基础设施生命周期

网络基础设施涉及硬件设备、云平台、网络功能和网络管理系统。为了更好地支持日益复杂的上层应用，网络基础设施需要高效且高质量的构建和交付。在研究和开发、测试、部署和集成等阶段需要进行广泛的重复验证，以确保基础设施的稳定性。因此，将人工智能能力引入网络基础设施是迫切需要的，以取代大量的劳动力成本。

为了使网络更加智能化，AI代理可以在基础设施集成和验证的全生命周期中得到应用和增强，例如在5G-A基础设施的测试、集成、交付、运营和维护过程中。在测试和验证阶段，AI代理可以帮助进行需求管理、解决方案制定、验收验证和试点验证。在建设和交付阶段，AI代理可以自动完成环境设置、交付验证、配置、部署和问题定位等任务。在

运营和维护评估阶段，AI智能体可以执行质量评估、问题预测、运营维护监控和灾难恢复演练。因此，智能基础设施集成和验证技术将大大提高网络基础设施测试自动化和管理效率的水平。

智能网络基础设施的架构分为四层，类似于人的感官、大脑、神经信号和手。每一层在5G-A网络中都有其相应的功能和应用。

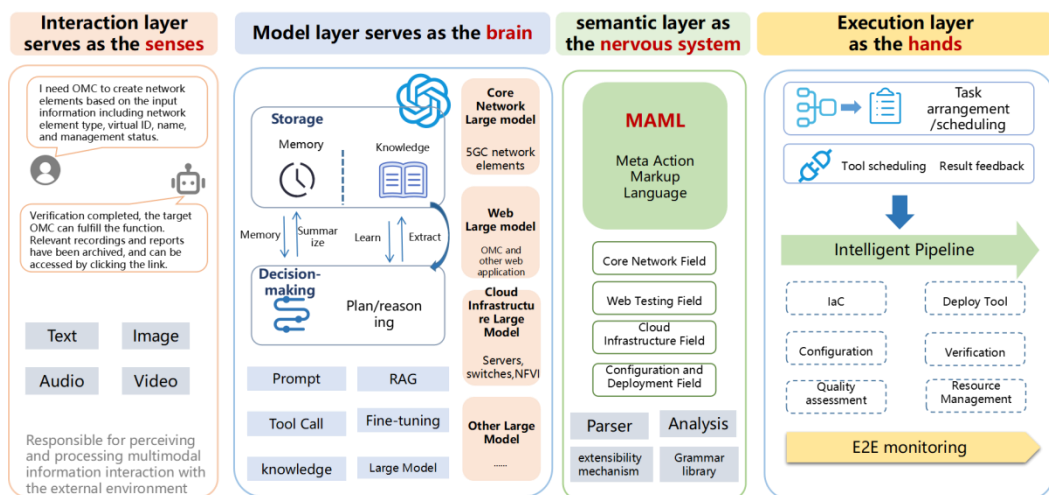


图3-4 智能网络基础设施解决方案架构

交互层： 作为5G-A人工智能代理与人类交互的“感官”，这一层负责感知和处理与外部环境的多模态信息交互，包括文本、图像、音频和视频。它还通过交互层将任务处理的结果反馈给用户，并启用迭代交互过程。

模型层： 作为AI代理的大脑，这一层负责任务分解、规划、推理、记忆、反思以及工具使用。在5G-A基础设施的不同场景中，大模型被划分为核心网络模型、网页模型、云平台模型等，以完成不同领域的特定任务。

语义层： 为了弥合任务执行与自然语言之间的差距，建立了一个语义层，作为AI代理的神经信号层，使大型模型能够理解和执行各种任务。这个语义层设定了统一的MAML（元动作标记语言），并为网络验证、网页验证、云平台验证、部署和配置等子领域定义了子语言。

执行层： 作为AI代理的“手和脚”，这一层负责执行、协调和调度特定任务。利用智能管道，执行层根据自动化工具或脚本完成环境部署、测试用例验证、质量评估和资源管理等工作，然后将结果反馈给模型层，以便进行迭代、反思和后续任务的规划。

通过这项技术，5G-A网络基础设施可以实现更高效和智能的管理与运营，以满足未来网络技术快速发展的需求。

3.1.4 网络移动性中的AI

人工智能通信网络的应用代表了5G-A领域的一项重大突破。其中一项关键创新是网络移动性方面的AI。为应对网络的动态和复杂需求，AI提供了新的解决方案，以优化切换、预测可能的故障，并减少资源和测量开销。

1) RRM测量预测

无线电资源管理（RRM）测量预测在促进高效的移动性管理中发挥着关键作用。^[6] 基于人工智能的方法利用跨细胞簇的历史和实时测量数据来生成预测，以改善切换决策并减少频繁测量和报告的需求。这些人工智能模型可以被训练来预测时间、频率和空间域中的测量值。如图3-5所示的框架图说明了整体预测结构，其中人工智能模型可以在网络的各个位置部署。在所提出的方法中，我们强调基于簇细胞的组合方法与案例3（直接预测）相结合，显示出利用人工智能预测收益的巨大潜力。这是通过结合更丰富的空间信息并在多个模块间进行联合优化来实现的，从而提高了预测准确性和网络性能。

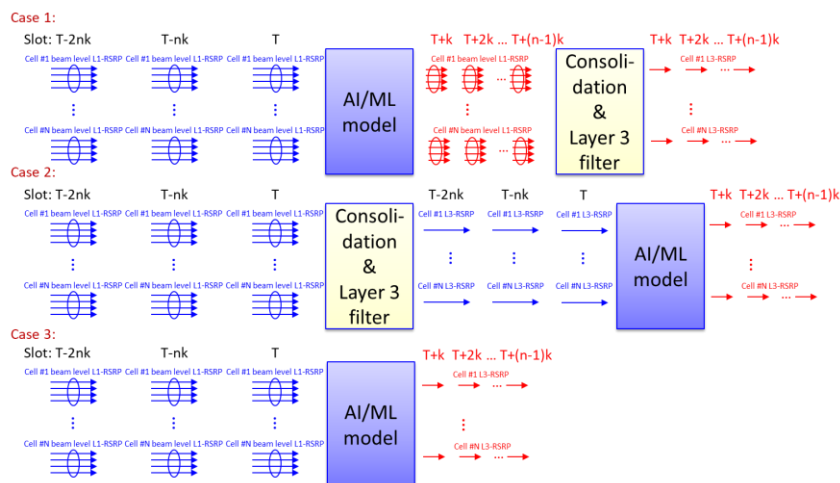


图3-5 RRM测量预测框架

2) 故障预测

预测切换失败（HOF）和无线链路失败（RLF）对于在用户移动性期间保持无缝连接至关重要。^[7] AI模型分析信号质量随时间的变化趋势，以识别可能发生故障的规律，从而实现主动式网络调整，减轻服务中断。与T310过期相关的RLF场景——作为故障预测的代表性案例——可以通过短期和长期方法来解决。

在短期场景中，当T310已经被触发时，AI模型估计剩余计时器期间的故障可能性。这些预测提供了对故障风险的即时洞察，支持实时决策和恢复行动。在T310激活之前进行的长期预测，侧重于估计潜在故障的概率和时间。通过结合扩展的RRM测量趋势

在多个细胞中，长期方法允许进行先发制人的规划和优化流动性策略。

直接预测模型，直接输出故障概率，由于其能够提供精确且可操作的信息，特别适用于短期场景。相反，依赖时间相关RRM测量预测的间接预测方法在长期场景中更有效。这些方法能够捕捉在较长时间窗口内服务小区和相邻小区的演变条件，为资源分配和移动管理提供全面的视角。

在具有较高失败风险的环境中，例如FR2部署，这些预测能力尤其有价值。虽然直接预测解决了即时风险，但间接预测则利用多小区空间和时间数据来支持战略规划，确保网络性能的可靠性。这些方法共同提高了降低高掉话率（HOF）和重连失败率（RLF）的能力，有助于构建更稳健和适应性强的网络运营。

通过整合这些预测方法，人工智能模型在确保无缝服务交付方面发挥着变革性作用，即使在最动态和复杂的移动场景中也是如此。

3) 事件预测

通过利用人工智能模型，对测量事件（如A3事件）的预测有助于优化网络操作，降低测量开销，并通过更精确和及时的事件触发提升切换性能。^[8]与故障预测类似，事件预测也可以通过间接和直接两种方法实现。间接预测利用RRM测量预测作为输入，允许网络根据服务小区和相邻小区的时空数据趋势推断事件的发生。此方法与现有的移动框架无缝集成，利用多小区相关性来维持预测精度，同时降低操作复杂性。另一方面，直接预测绕过中间步骤，AI模型直接输出事件触发器。这种方法特别适合于实时应用，可以实现更快的响应时间和简化处理。

在专注于减少测量开销的场景中，如FR1部署，预测测量事件可以替代实际测量，显著降低信令负载同时保持等效性能。例如，时域预测允许网络在不牺牲事件准确性的情况下减少测量报告的频率。同时，在FR2环境中，由于移动性需求更高，这些预测通过主动调整如滞后和触发时间等参数，增强了切换性能。这减少了晚切换的可能性，并改善了最佳目标小区的选择，确保网络节点之间的过渡更加平滑。

评估这些预测的准确性对其有效性至关重要。指标如漏报预测比率、误报预测比率和时间准确性对于评估可靠性是必不可少的。准确的预测确保事件触发在正确的时间发生，从而实现及时的手续转移并最小化中断。这些能力有助于构建一个更具适应性和高效的移动管理框架，解决下一代网络的动态挑战。

4) 确定前进道路

AI在5G-A中的移动性代表了向前迈出的一个变革性步伐，正如测量精度、故障预测可靠性和事件预测方面的进步所展示的那样。

能力。开发将预测结果整合到网络决策过程中的稳健框架对于实现这些益处至关重要。未来的研究将继续提高人工智能模型对不同部署条件的适应性，确保一致性和可扩展性。

通过将AI驱动的洞察嵌入到5G-A的核心，行业正逐步接近实现其构建更智能、更灵活、更高效的网络的愿景。这些创新为未来在移动管理方面的能力奠定了坚实基础，为真正互联的未来铺平了道路。

3.2 数字孪生网络智能

数字孪生网络（DTN）代表着我们对5G网络以及更高级网络的管里和优化方法的一种范式转变。通过创建物理网络基础设施的虚拟副本，DTN使得运营商能够以前所未有的精确度模拟、分析和优化网络操作。这个数字表示不仅是当前网络状态的反映，而且也是一个测试未来情景的沙箱，为创新和网络优化提供了一个强大的平台。

3.2.1 网络数据映射技术

数据映射是网络数字孪生实施的基础，为孪生建模提供必要输入。收集数据的品质与全面性直接决定了数字孪生如何忠实地反映其物理对应物。然而，由于现代网络的规模庞大、设备生态系统多样化、复杂的接口配置和动态流量模式，这些挑战使得高精度数据收集既资源密集又可能对网络性能造成破坏。

为了应对这些挑战，重点已经转向开发一种高效的网络数据收集机制，以满足网络孪生的特定需求。该机制采用按需数据收集方法，精心选择收集目标、精度水平、协议和传输方式。目标是实现一种平衡的解决方案，在保持效率和节能的同时提供全面的数据。

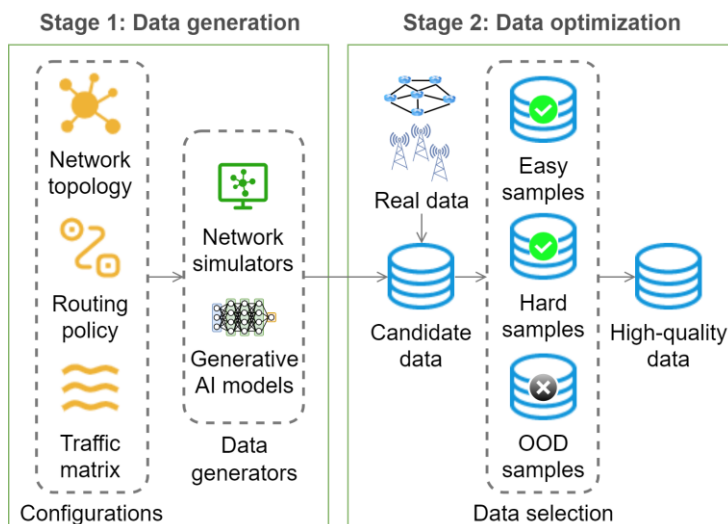


图3-6 AutoOPT框架

AutoOPT框架通过利用生成式AI模型进行数据生成和优化，代表了对这一挑战的创新性解决方法。如图3-6所示，此框架通过两个不同的阶段运作。在数据生成阶段，它使用规模无关性指标从小型网络中创建合成数据，确保DTN AI模型^[9]可以有效推广到更大的网络环境。后续优化阶段通过种子样本选择和增量细化自动识别和筛选高质量数据，最终提高DTN人工智能模型的准确性和泛化能力。

这种方法不仅解决了数据收集的即时挑战，还为数字孪生精度持续改进奠定了基础。通过将数字孪生本身作为模拟数据生成实体，我们可以丰富机器学习训练数据集，并为网络分析和优化创建更稳健的模型。

3.2.2 数字孪生建模技术

NDT建模技术在实现复杂5G-Advanced和未来网络中的有效网络智能方面至关重要。它侧重于构建物理网络基础设施的高保真虚拟副本，在模型精度与计算效率之间取得平衡。在日益复杂和动态的网络架构背景下，一个稳健的NDT模型对于准确状态复制、预测模拟和历史事件分析等功能至关重要，从而实现主动式网络管理和优化。

NDT建模围绕三个基本维度构建，协同贡献于全面的数字表示。首先，网络状态数字孪生解决了网络实体、属性及其相互关系的实时和历史表示问题。这一维度包括不变的网络配置和动态操作遥测。采用的技术包括：

- **网络模拟器**：利用模拟不同层次抽象的网络环境的平台，从详细的包级仿真到更高层次的行为模型。
- **数据抽象表达式**：采用正式化的数据模型和架构来

高效表示网络状态数据并促进语义互操作性。这可能涉及使用数据序列化格式和标准化信息模型。

- **知识图谱**：构建基于图的表示以捕捉网络拓扑、实体关系和状态依赖，从而在数字孪生中实现知识驱动的推理和推断。

其次，服务与环境匹配主要关注服务层及其运营环境的建模。这对于理解网络内部的服务性能和用户体验至关重要。关键方法包括：

- **生成对抗网络 (GANs)**：利用生成对抗网络 (GAN) 架构来合成真实且动态的服务行为模式并模拟用户交互配置文件。GAN在捕捉复杂、非线性服务动态以及生成用于模型训练和验证的合成数据集方面特别有效。
- **基于深度学习的环境重建**：利用深度学习算法，尤其是卷积神经网络 (CNN) 和点云处理技术，对传感器数据 (例如，激光雷达、相机图像、部署蓝图) 进行处理，以实现物理网络部署环境的自动识别、提取和三维重建。这为网络模拟和可视化生成空间精确的上下文。

第三，网络行为孪生模型模拟网络元素的功能特性。这一维度采用一系列基于所需精度和计算成本的建模方法：

- **白盒建模**：在网络虚拟环境内部直接实施网络协议和设备功能，通常通过软件定义网络 (SDN) 和网络功能虚拟化 (NFV) 原则。这提供了对模拟网络操作的较高透明度和控制，适用于详细的协议分析和功能验证。
- **黑盒建模**：利用数据驱动的建模技术，如机器学习统计建模，根据观察到的历史数据预测网络状态转换和性能指标。此方法抽象于内部协议细节，聚焦于输入输出关系和预测准确性，适用于性能预测和异常检测。
- **灰盒建模**：结合抽象网络机制表示与选定的实际网络功能。这种混合方法在建模精度和计算复杂性之间提供了一种实用的平衡。在当前语境中，形式化方法如Petri网特别有价值，它们为分布式工业物联网网络数字孪生中并发和异步事件的建模提供了一个数学上严谨的框架，如先前研究所展示。^[10]

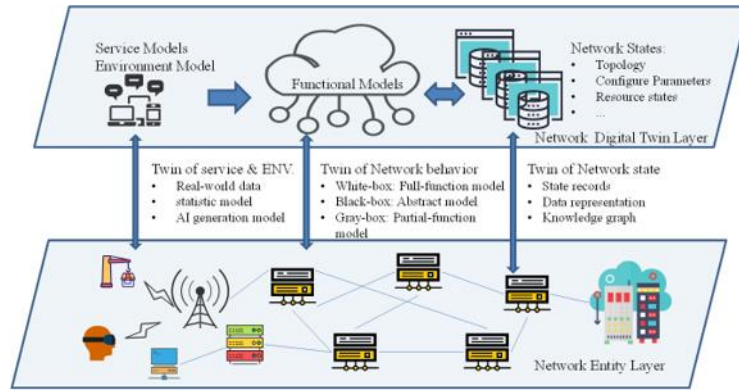


图3-7 NDT建模方法

这三个维度，如图3-7所示，在构建全面无损检测技术（NDT）方面至关重要。这种集成建模范例赋予了网络运营商对网络理解、预测分析和主动优化的增强能力，最终有助于实现向智能、自主网络管理的演变，并支持现代网络基础设施的先进服务需求。

3.3 应用智能

通信网络正在从基本连接服务向智能应用服务演变。通过人工智能技术的深度融合，网络不仅提供传统的语音和数据服务，还支持多模态交互、实时翻译和数字人等新型智能服务。为了支持这些创新应用，网络经历了技术演进。这些增强不仅扩展了网络的应用边界，还为运营商创造了新的商业模式，从流量运营转向基于价值的运营。

3.3.1 IMS 数据通道

IP多媒体子系统的数据通道（IMS DC）是一项先进功能，它可在通信会话期间实现非语音数据的高速、实时传输，利用基于IMS的语音和视频服务的基础设施。其设计旨在满足与延迟、带宽和可靠性相关的特定需求，这些对于应用如增强现实、实时多媒体和物联网交互至关重要。本节详细探讨了IMS DC的架构、功能和优点。

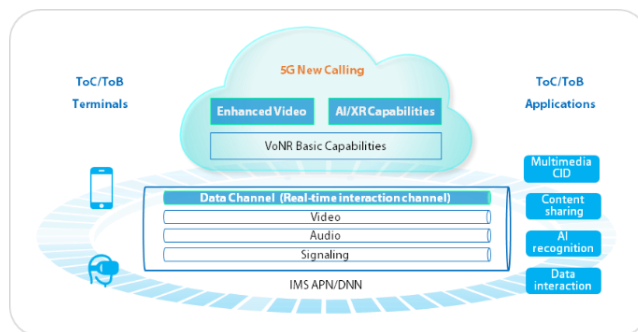


图3-8 IMS DC实现实时多媒体交互

如图3-8所示，IMS DC集成在IMS框架内，基于语音和视频通道创建一个无缝、多功能的通信渠道。这种集成允许在传统语音或视频通话的同时同步传输数据类型，确保数据流与多媒体内容保持同步。IMS DC在现有的IMS架构下高效运行，利用电信网络固有的优势，如通过电话号码的全球连接、统一的认证和强大的会话管理。

IMS DC 被分为两个主要组件：

- **引导数据通道 (BDC)：** 该通道促进了IMS DC应用程序在终端上的下载和执行。这些应用程序可能包括网页内容（例如，HTML5页面）、媒体元素以及用于实时交互的控制脚本（例如，JavaScript）。
- **应用数据通道 (ADC)：** 这是数据通道，终端上的应用程序通过它传输数据到其他设备或网络。模数转换器 (ADC) 确保在语音/视频流量并行传输的同时，实时传输特定于应用程序的数据，从而保持一致的用户体验和最小延迟。

IMS DC在终端和网络之间建立，以有效地管理数据交换。当通信会话开始时，DC服务器建立BDC，将DC应用程序推送到主叫和接收终端。一旦这些应用程序激活，ADC用于处理数据传输，确保所有多媒体和应用数据实时同步。

在IMS DC架构中，控制和媒体功能是分离的。该方法使用数据通道信号功能 (DCSF) 来处理信号控制，并使用媒体功能 (MF) 来管理媒体资源，支持如AR渲染和媒体处理等功能，这大大提高了未来应用的灵活性和可扩展性。

IMS DC在现代通信中提供基本的好处，包括：

- **实时、多模式交互：** 通过结合视频、语音和数据流，它能够在IMS框架内直接实现沉浸式和交互式通信体验，例如AR和虚拟会议。
- **服务质量 (QoS) 和安全性增强：** 利用IMS强大的功能，IMS DC确保了QoS管理和强大的安全协议，保护数据完整性和最小化延迟。
- **灵活的应用集成：** 通过BDC流动态引入新应用的能力简化了扩展IMS功能的流程，确保未来创新可以迅速集成到网络中，而无需进行重大调整。

基础设施变更。

总之，IMS DC代表了电信服务演变中的一个关键步骤，提供了一个统一的数据、语音和视频通信通道，优化了跨不同网络环境的实时、多媒体交互。IMS DC的持续进步，尤其是与增强现实（AR）和未来通信服务的集成，凸显了其在5G-A及其以后通信系统基石作用。

3.3.2 交互式新呼叫

5G-A新呼叫服务为用户提供交互式、智能化和沉浸式的创新新服务和新场景，例如数字人、智能翻译、意图通信等。

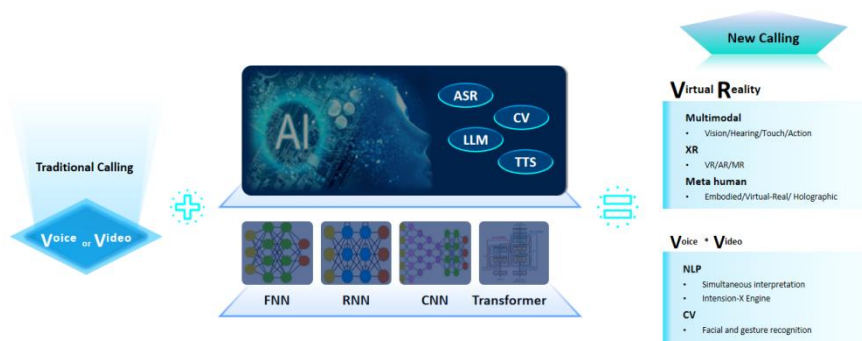


图 3-9 为新兴职业引入的新人工智能技术。

新的AI技术刷新了传统的电话服务，并促进了业务模式从语音或视频拓展到多模态通信。随着AI技术的快速发展，AI已经成为推动通信技术革新的不可或缺的力量。通过AI赋能网络，持续的创新被注入到电话服务中。以下列举了许多用于新叫通中使用的AI新技术：

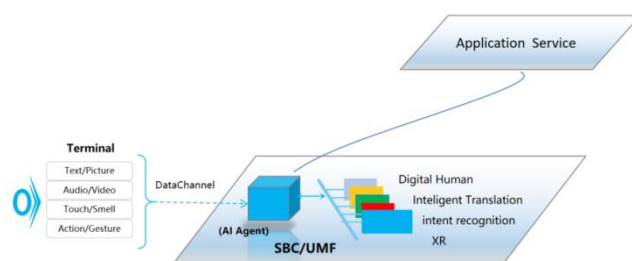


图 3-10 交互式新呼叫

- **意图识别**：意图识别指的是通过自然语言处理技术（如大型语言模型、自然语言处理NLP）对人类语言中表达出的意图或目的进行识别和理解。基于意图的通信极大地简化了新呼叫的交互流程。系统

自动根据文本和图片识别并执行意图。例如，如果您订购1000个披萨，系统将自动执行意图识别并生成订单。

- **XR + AGI技术：** 扩展现实（XR）交互技术指的是通过计算机技术和可穿戴设备生成的真实与虚拟结合、人机交互的环境。它将虚拟信息与真实场景相结合，创造出一个人机交互的虚拟环境。该技术可用于娱乐、教育、医疗和工业领域，以提供更加丰富和沉浸式的体验。通用人工智能（AGI）技术可用于快速生成XR视频和数字世界。

- **数字人：** 数字人技术是一种高科技产品，它将计算机图形、动作捕捉、图像渲染和人工智能结合在一起。它允许创建具有人类外观、行为和特征的虚拟角色。这些虚拟角色可以存在于数字空间中，并与现实世界互动。

- **智能翻译：** 结合实时通信，智能翻译将用户的语音转换为文本信息，并在用户的手机上显示这些文本信息。智能翻译可以根据人工智能技术如自动语音识别（ASR）、大型语言模型（LLM）等实现不同语言或方言之间的翻译。它还适用于在无法进行语音通话的情况下进行实时语音/文本转换，例如，正常人与听力障碍人士之间的交流。

通过应用人工智能技术如语音识别（ASR）、意图识别、智能翻译和数字人技术，新通话将受益于以下方面：

- **新商业模式：** 在将来，通信将从交通和时间操作发展到价值操作，并提供价值服务，例如数字人、个人助理和化身/自我沟通。
- **缩短服务上线时间：** 通过整合人工智能技术，服务推出时间可以简化。通过插件智能AI平台，新服务组件可以实现即插即用。

3.3.3 云-边-端协同

云边端协同是优化通信网络的关键，尤其是在人工智能驱动应用方面。随着人工智能模型复杂性的增加，在云、边缘和终端层之间高效分配任务变得至关重要。像计算卸载、资源调度和数据协作这样的关键策略对于实现这些层之间的无缝协作至关重要。

1) 计算卸载

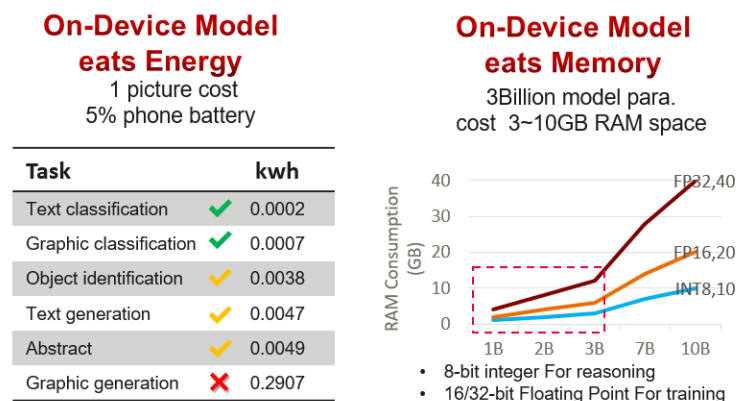


图3-11 设备端AI：能耗与内存消耗

随着人工智能模型复杂性的增加，终端设备通常无法独立处理其计算负荷。将这些任务卸载到云端或边缘设备，可以更好地分配计算工作负载，确保在终端设备上的实时性能。如图3-11所示，随着人工智能模型复杂性的增加，终端设备上的能量和内存使用量显著上升。这凸显了将更多资源密集型任务卸载到云端或边缘设备的重要性，从而使终端设备能专注于更轻的计算任务。

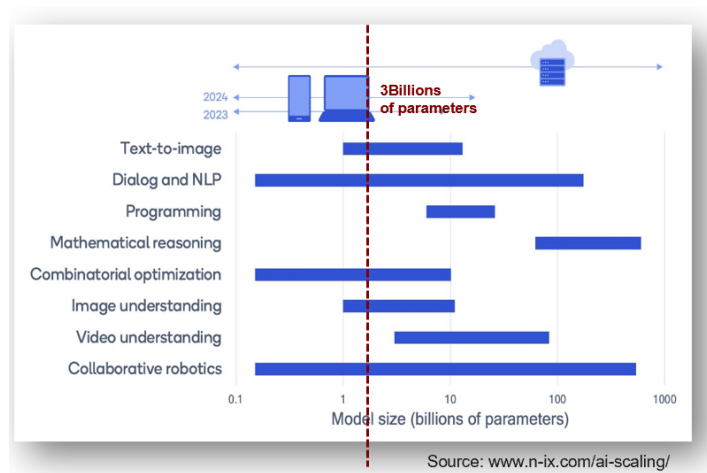


图3-12 不同任务对AI模型大小的需求

同样，图3-12显示了不同AI模型的计算资源需求变化。大型模型，如深度学习使用的模型，通常被卸载到云端，而需要实时推理的小型模型则在边缘或终端处理。这种分层方法确保AI应用可以根据手头任务的特定计算需求有效扩展。

2) 以上行链路为中心的宽带通信

随着人工智能应用如AI视频通话和AI助手的普及，人工智能与人类之间的交互变得多模态，例如图像和视频，以及从终端到云端的图像和视频上传变得必要。根据相关研究，要确保80%常用应用的交互体验，需要20Mbps的上传速率，而要确保60%高品质应用（如增强型应用）的交互体验，则需要30Mbps的上传速率。

补充上行链路 (SUL)。

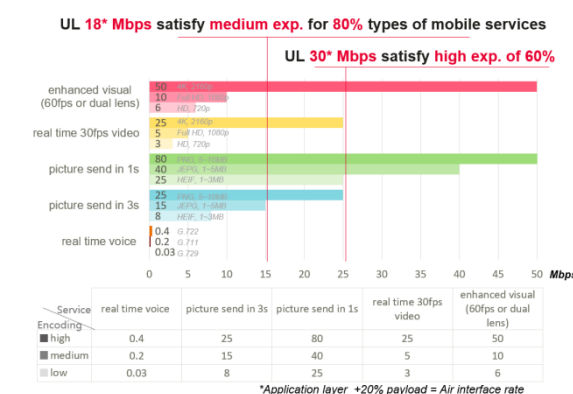


图 3-13 多模态交互的上行链路体验需求

上行为中心的宽带通信 (UCBC) 将上行带宽能力提升10倍，满足多模态交互、机器视觉和不同AI应用场景下的海量宽带物联网上传需求，加速数千个行业的智能化升级。目前，更多频谱和带宽正基于当前的下行通道进行聚合，以提升上行能力。主要技术包括灵活频谱接入、SUL增强和上行载波聚合。FA SUL网络技术可应用于实现大上行带宽并满足多模态交互需求。此外，TDD频段的上下行时隙分配进行了调整，以增加上行时间槽资源的调度，从而提升上行能力。

3) 资源调度

高效资源调度对于确保任务在云、边缘和终端节点之间被最优分配至关重要。资源的动态分配确保计算能力、内存和带宽得到有效利用，最小化延迟并最大化吞吐量。AI应用通常需要大量的上行带宽，尤其是在处理大量数据集时。满足这一需求对于确保终端、边缘和云层之间的数据传输平滑至关重要。

在云计算边缘协作的背景下，资源调度的作用变得更加重要。核心焦点是依据当前网络负载和资源可用性来管理任务，确保计算资源尽可能地高效使用。这种动态调度有助于维持最佳性能，减少瓶颈，并实现任务执行的加速。

4) 数据协作

数据协作确保了云、边缘和终端之间信息的顺畅流动。在终端收集的数据首先在边缘进行初步处理和缓存，从而减少需要传输到云的大量数据。这种本地处理最小化了延迟，使得实时决策成为可能，尤其是在对时间敏感的应用中尤为重要。

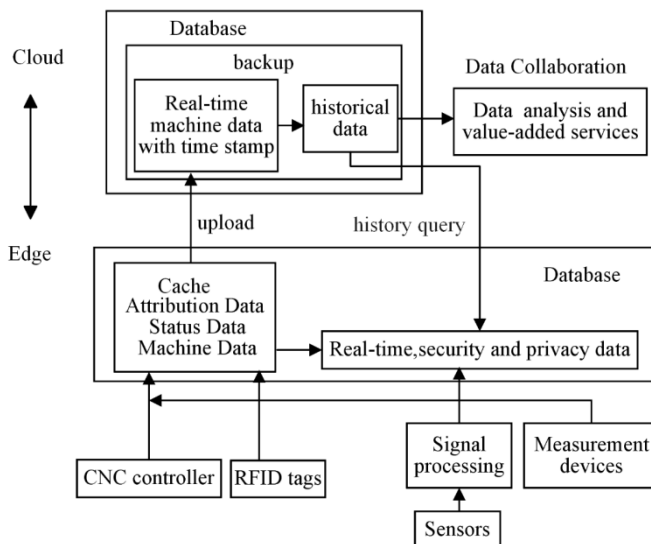


图3-14 云边数据协作方案

图3-14展示了数据在这些层之间是如何被处理和管理的。数据在上传到云端之前，在边缘进行预处理，然后云存储并进一步分析数据。这种协作确保只有最相关的数据被传输到云端，优化带宽使用并确保高效的模型训练。它还使得敏感数据的处理更加安全，尤其是在医疗和金融等领域。

计算卸载、资源调度和数据协作之间的协同作用，使云端-边缘-终端协作能够高效运行。通过利用这些策略，人工智能应用的总体性能得到提升，计算能力得到合理分配，延迟最小化，数据流得到优化。

随着人工智能驱动应用的持续发展，动态分配资源以及管理云、边缘和终端数据的能力将变得愈发重要。这一合作不仅提高了人工智能处理的效率，而且允许应用程序有效扩展，以满足下一代服务的日益增长需求。

3.4 可持续智能

随着全球电价的上涨，节能和能源效率已成为电信行业的核心优先事项。对于移动网络而言，降低能耗以实现净零排放已成为运营商的核心目标，超过85%的移动网络运营商承诺达成此目标。鉴于无线接入网络（RAN）和核心网络对整体能耗的贡献重大，提高这两个领域的能源效率至关重要。

人工智能的集成在提高能源效率方面发挥着关键作用。通过分析实时交通模式和资源利用，人工智能使资源调度和电力管理最优化，显著减少能源消耗，同时保持服务质量。

3.4.1 提高设备级能效

在设备层面，无线单元中的主要能量消耗者是功率放大器（PA），它用于放大传输信号。优化PA的效率对于提高整体节能至关重要。现有技术能够在低负载条件下改善PA的线性和效率；然而，大多数解决方案在低负载期间（在网络操作中很常见）的优化效果有限。

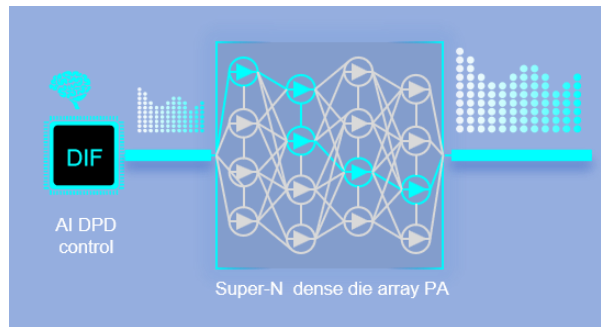


图3-15 智能功率放大器控制

为了解决这一问题，采用了AI驱动的数字预失真（DPD）技术来提高功率放大器（PA）的线性度。传统的DPD算法已经通过机器学习模型进行升级，这些模型能够动态适应不同的环境和信号条件。深度神经网络（DNNs）模拟PA的非线性特性，并在实时调整输入信号，以实现更好的功率效率和信号质量。

AI驱动的DPD系统在广泛的真实世界信号数据上训练，考虑到各种频段、功率水平和环境条件。这些系统可以在复杂的网络环境中自主调整参数，减少信号失真并提高性能。

3.4.2 优化网络级能源效率

在网络层面上，通过分析历史交通模式、天气数据和地方事件，AI预测交通负载波动，从而允许动态调整网络资源以减少不必要的能耗。AI驱动的关闭机制可以预测低交通时段并自动关闭包括功率放大器和收发器在内的组件，在不影响服务质量的情况下减少能耗。

人工智能模型通过分析来自终端设备和网络组件的实时反馈，也允许对网络设备进行智能控制。这使得网络能够动态调整电力分配和传输策略。例如，基站中的功率放大器（PA）和基带单元（BBU）在低负载期间可以进入深度睡眠状态，只有在流量增加时才会被重新激活。

此外，人工智能优化不仅限于无线接入网（RAN），还包括核心网络的能效改进。在核心网络中，智能调度和资源管理机制分析流量模式、用户需求和网络拓扑，以动态路由数据流量并最小化不必要的能耗。

通过多维度的节能策略，包括时间、空间、频率和功率域的优化，人工智能能够实现精确的能源优化，而不会影响性能。预计这种能力将在R18及以后得到进一步发展，结合网络数字孪生技术进行离线节能预测，为未来网络的部署实现更精确的节能策略。

4 5G-AxAI 推动新应用场景

4.1 差异化经验保证

随着移动互联网的发展，用户对网络的需求正变得越来越个性化。从4G到5G，运营商主要通过流量管理来管理数据服务。在5G-A时代，利用人工智能技术动态调整网络资源分配对于满足不同用户级别（如钻石、白金、黄金、银色和标准卡）的不同类型服务（如短视频、游戏和直播）、中高速移动场景（如高速铁路和地铁）以及高容量场景（如音乐厅和旅游景点）的差异化体验需求至关重要。网络感知、体验保证和体验评估等技术可以显著提升用户满意度。

网络感知：目前，智能网络感知模块支持识别数千个服务类别，包括主流国内与国际服务流量。此外，它能精确识别超级应用内的各种子服务，例如语音通话、视频通话、直播、视频会议以及云游戏，整体感知率超过95%。

经验保证：基于网络感知的结果，为对应服务提供有针对性的保障策略。在无线侧采用的经验保证的关键技术包括精确服务保证预测、多目标优化服务保证策略和智能多频率协调。在核心网络侧，基于NWDAF的网络安全智能被引入，以满足用户需求提供动态保证比特率（GBR）保证。

经验评估：用户体验评估涉及基于不同服务特点制定一套全面的QoE（用户体验质量）指标。对于视频服务，这包括通过多个维度如内容质量、传输质量和交互质量等方面评估清晰度、流畅性和及时交互，这些维度随后将被整合成一个统一的标准分数[1, 5]。对于即时通讯服务，重点在于服务类型和响应及时性等因素，指标包括文本、语音、图像和视频传输类型，以及服务量、持续时间、速度和TCP/UDP协议下的性能指标（如丢包率、延迟、抖动和TCP重传率），所有这些均用于计算即时通讯服务的平均意见分数（MOS）。这种全面的方法确保了评估能够准确反映用户在全网服务中的体验质量。

现有网络的测试结果如下展示：

- **在无线网络方面：**在部署精确的服务保障策略后，各种类型的服务（如短视频、二维码支付和网页）的延迟均得到降低。

浏览量减少了大约20%。

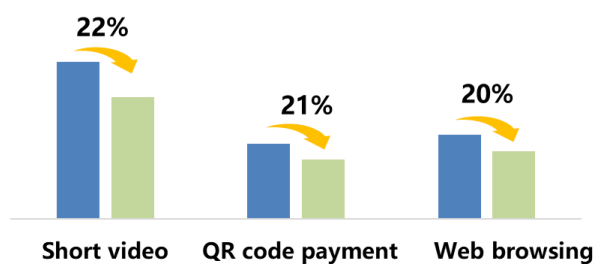


图4-1 短视频、二维码支付和网页浏览服务的延迟降低。

- 在核心网络方面：在引入基于 NWDAF 的智能后，当检测到用户体验下降时，随后采用 GBR 确保机制以提升用户体验。



图4-2 对于VIP用户，在拥塞前上行速度为2Mbps，拥塞后降至255Kbps，在建立GBR后恢复至2Mbps。

此外，在服务保证过程中，用户移动设备上的标志可以实时动态显示“VIP保证激活”文字，与增强的体验相匹配，提升用户的感知。服务保证过程结束后，用户会收到一份体验报告，提供对可测量保证效果的实时反馈。这完成了闭环体验管理，实现了对客户更高水平的关怀。



图4-3 浙江移动测试结果

4.2 新呼叫服务

近年来，多媒体大型模型、AI代理和DC技术的发展正推动通信行业从传统的音频和视频向多模态、互动通信转变。这为用户提供了智能视频通话、互动视频通话和呼叫智能助手等一系列丰富多样的服务，显著提升了个人和企业通信体验和效率。它促进了通信行业的繁荣发展，并为运营商业务开辟了新的增长空间。2023年，中国移动与合作伙伴共同推出了5G新通话服务，并在年底前完成了覆盖全国31个省的新通话网络第一阶段建设。相关新通话网络元素已部署在八大区域的资源池中，包括现有IMS网络元素的升级和新通话能力元素的部署。这支持了六种新通话服务场景的发布：点亮屏幕、趣味通话、名人通话、AI速记、语音转文本和实时翻译。

1) 点亮屏幕：从语音到视频再到内容，开辟了一个新的领域 内容运营

2023年，中国移动江苏分公司与华为及其他合作伙伴共同开发了一项新的通话点亮屏幕服务，以满足消费者情感表达需求。该服务于当年第四季度发布。



图 4-4 新呼叫点亮屏幕服务

点亮屏幕是一款以用户沟通社交属性为核心的产品。用户可以使用该产品设置个人虚拟形象。在语音通话过程中，用户可以传输和

在通话过程中无需切换视频通话即可向通话方展示预设的图片、视频或虚拟数字人图像，以传达情感和表达个性。企业用户可以定制企业形象、传播企业价值或进行产品营销。在通话期间，他们可以根据不同的关键词触发介绍视频或图片，以提高沟通效率和效果。运营商可以公布反欺诈和安全知识，承担更多社会责任。不到一年时间，江苏移动的屏幕照明服务已有超过500万用户，每月内容展示量超过1亿次。它收到了良好的市场反馈，并成为新的呼叫服务，受到消费者的欢迎。中国移动广东为所有安装和维护人员启用ToB应用。在与客户的沟通中，使用关键词触发“光纤到户”（FTTR）品牌推广和套餐推广视频，促进服务推广和高效沟通。2024年下半年，中国移动升级了其屏幕照明业务，植入人工智能生成内容（AIGC）的自创作能力，使用户能够在通话中创建自己的风格化图像并相互展示。江苏移动还探索了数字人通话的“另一个我”功能。基于用户的外观特征，AI生成风格化的数字人图像。在通话过程中，数字人根据用户的语音由网络AI驱动。我的声音与数字人的表情和唇形实时同步，为通话注入更多新鲜体验。截至2024年底，已有超过1500万用户订阅了屏幕照明服务。开启屏幕将使运营商将通话持续时间操作提升至平均90秒，为通话带来更大的商业和社会价值，成为新的移动媒体。

2) 实时翻译与语音转文字：科技助力善举，构建实时沟通之桥。

在2023年9月，中国杭州举行的第19届亚洲运动会上，中国移动推出了“智能翻译”服务。在需要跨语言沟通的情况下，“智能翻译”服务可以自动识别并翻译在原生视频通话中的语音内容，无需依赖翻译软件。同时显示双语文本字幕。在亚洲运动会期间，可以支持包括英语、韩语、日语、阿拉伯语等多种语言的实时翻译，使沟通不受语言障碍的限制。呼叫字幕服务可识别通话中的语音内容，并在移动电话屏幕上以大字体显示语音内容，从而减少沟通障碍，大幅提高听力受损者的通话体验。



Real-Time Translation Speech-to-Text Conversion

图4-5 实时翻译和电话字幕服务

在2024年，中国移动在所有省份启动了实时翻译和通话字幕服务。不到一年的时间里，超过500万用户订阅了通话字幕服务或

调用字幕服务。这项新的服务真正实现了“科技向善”的宗旨。帮助用户朝着“无障碍沟通”的目标迈进。

到2024年底，在网新5G通话用户数量已达4000万，极大地提升了用户通话体验和通信效率，为运营商带来了相应的商业和社会价值。

除了原有的语音和视频通道外，DCs（数据中继）被添加到音频和视频通话中传输数据。引入了交互式通话，以实现图像和视频共享、消息通信、屏幕共享、AR标注和文件传输等多媒体内容。这提高了远程通信的体验和效率，并扩展了行业价值边界。在2024年第四季度，中国移动在直播网络上启动了交互式视频用户的试验性商业应用。它推出了六个DC应用，即内容共享、消息箱、趣味通话+、实时翻译+、屏幕共享和数字人，以实现通话中的快速设置、共享和便捷互动。



图 4-6 互动视频用户

与新的通信行业合作伙伴合作，GSMA于2024年6月发布了TS.66，该标准定义了设备侧的DC API标准。来自高通、MediaTek和UNISOC等芯片供应商的主流芯片也支持DC功能。来自vivo、OPPO、小米、三星和华为的一些主流终端机型也支持DC功能。DC交互通话行业生态初步建立。预计2025年将成为DC系列应用大规模商业使用的第一年。

3) 人工智能代理启用新通话功能，帮助运营商构建人工智能服务入口。

人工智能全面升级原生通话体验。作为智能通话助手，AI提供智能接听和智能聊天等服务，并巩固了个人通话入口。升级企业客户服务热线并构建企业智能应用门户。2024年下半年，中国移动开始逐步在网络构建呼叫代理能力。在浙江、广东和江苏三省，中国移动开展了智能接听、AI辅助聊天和星级通话等场景的服务测试和试点。目标是构建每位用户的数字助手。打造个人AI服务入口。

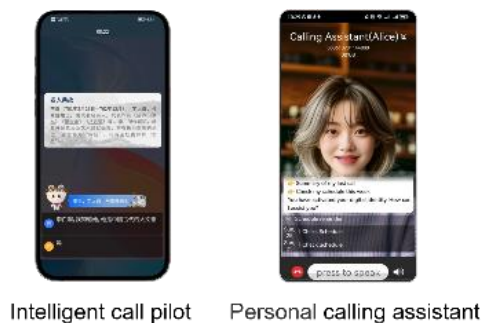


图4-7 智能呼叫助手

在无法断电、快递和促销等场景中，呼叫智能助手帮助用户智能接听电话。基于通话内容分析，拦截条件可以灵活设置以减少误判。拦截的内容也会通知用户，并可回拨。为了实现反骚扰、反泄露和反欺诈的目的；它还可以帮助用户在通话中整理大量信息，根据通话场景实时提醒重要信息，减轻用户的记忆负担，并实现智能聊天。呼叫代理还可以调用外部垂直领域代理和搜索工具，帮助用户完成餐饮预订和预约，提高通话体验和效率，使电话成为个人AI服务的统一入口。到2024年6月，中国移动江苏分公司简化了公司A客服热线车辆保险索赔和贷款客户服务场景的业务流程。AI被用于帮助企业客户服务智能化升级，从传统语音交互发展到多模态的自然语言、视频和数据渠道代理。企业数字人客服人员通过自然语言交互一次性获取关键信息。在大规模商用数据通道原生终端之前，多模态代理能力解决了复杂的双音多频（DTMF）交互、多层嵌套和耗时等问题。企业在通话中可以快速关闭交易，降低交易成本，实现用户、企业和运营商的共赢，并帮助运营商拓展企业市场份额。



Intelligent customer service

图4-8 智能客户服务

用户可以根据实际需求，通过代码编号调用第三方垂直领域代理。运营商可以使用代码编号构建通话应用统一的入口，实现一号通，一位代理，并构建企业应用生态系统。在2025年，呼叫代理技术将快速发展，逐渐重塑呼叫行业生态系统。正如俗语所说：“呼叫使AI无处不在，AI使呼叫无所不能。”

4.3 工业确定性服务

随着5G技术的快速发展以及网络基础设施的逐步改善，利用5G推动产业数字化转型取得了显著进展。5G和人工智能等信息技术已经在生产管理和辅助生产过程中得到广泛应用。数据收集、视频监控和自动导引车（AGV）等解决方案已实现大规模部署。然而，对于生产控制场景，由于短包传输周期（毫秒级）需要极高程度的确定性，以及严格的包级要求（连续三个包异常可能导致系统停机）和多种具有显著参数配置差异的协议，目前5G技术尚未完全应用于这一领域。存在两大主要挑战：

- 传统的5G服务质量标识符（5QI）机制仅能保证平均网络性能，无法提供所需的包级确定性。这使得满足工业控制系统长期稳定运行的需求变得困难。
- 即使采用先进的网络技术，如网络切片、差异化服务策略（如DS帧结构和PDCP乱序交付）以及链路冗余（双重传输与选择性接收），一些高要求的场景，尤其是需要极低网络抖动的场景，仍难以满足服务要求。现有的网络解决方案不足以满足这些严格的需求。

为确保确定性的服务体验，通过利用人工智能技术，推出了一种“1+2+3”确定性保障解决方案，提升了网络的确定性能力。

“1”指的是一种基于人工智能的服务感知算法，专注于工业场景，其中人工智能算法用于服务特征的自我学习。这种持续迭代的优化可以准确识别工业控制包的包级服务类型、消息长度、传输间隔、到达时间和关键词。

“2”指的是两种智能计算基础设施解决方案，包括基站用工业智能卡和边缘智能UPF。

1) **基站用工业智能卡**：这些卡可以通过插件方式快速将传统基站升级为智能工业基站，为人工智能能力如智能服务感知提供基础计算平台。采用向前兼容的设计，它们可以灵活适应现有的室外和室内主流基站模型，无需完全更换基站，有效降低成本。

2) **边缘智能UPF**：基于此基础设施，构建了一个用于确定性服务的智能“感知-推理-编排”闭环保证技术系统。它实现了对确定性协议数据包的智能检测、对确定性服务场景的智能推广以及对确定性编排策略的智能管理。这解决了诸如系统间协调困难、编排复杂度高、部署周期长等挑战。此外，它将5G计算能力与工业控制和工业人工智能能力相结合，实现了混合实时虚拟化操作系统和基本计算能力编排，推动了5G与工业的深度融合。

应用生态系统，为企业提供更高效、智能和确定性的生产和管理工作方法。

“3”代表三种基于服务的调度增强解决方案，通过服务感知优化基于特征的数据包级调度，以实现网络资源与服务需求的精确匹配。这提高了确定性的性能，同时增加了确定性用户的容量，确保服务的长期稳定运行。三种保证级别包括：

- 1) **基于服务类型的固定配置保证：** 调整用户级协议参数，例如禁用省电功能、不活跃计时器、BWP切换和UM模式，以在工业控制终端和普通终端之间提供差异化的保障。
- 2) **基于服务特征的精确调度保证：** 优化服务流量的资源分配策略。一方面，根据数据包的周期和大小预留资源，使资源分配更加精确，减少传输等待延迟。另一方面，协调多个工业控制流进行错峰传输，避免冲突，显著增强在多用户并发情况下保证低延迟边界的能力。
- 3) **基于服务逻辑的关键分组调度保证：** 优化服务流中关键数据包的调度队列策略和可靠性保障策略，以确保及时且准确的传输，防止看门狗复位和系统停机，从而实现工业控制服务的长期稳定运行。



图 4-9 智能工业示例

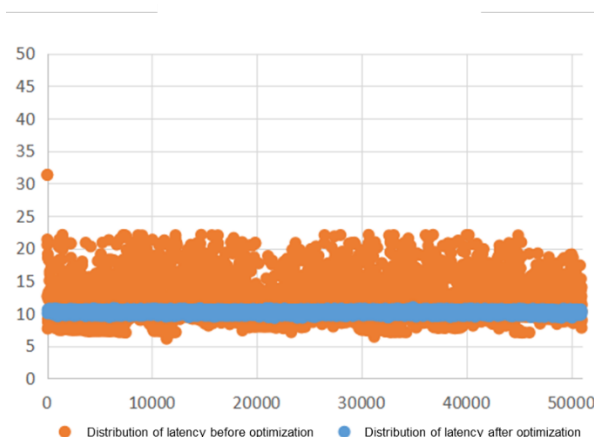


图 4-10 终端到终端延迟分布

目前，确定性保障技术解决方案和产品已在林克汽车、安钢和熊猫电子等工厂的40多个试点项目中得到验证。网络性能：确定性网络延迟已达到12ms@99.99%，确定性延迟抖动已达到8ms@99.99%。确定性保障的有效性提高了30%以上，显著提升了延迟和抖动性能以及一次性成功传输的可靠性。实际应用：这项技术广泛应用于各种场景，包括全向无线可编程逻辑控制器（PLC）、起重机的远程控制以及AGV的协同控制。通过实现无线工业控制，这些解决方案促进了工厂的智能化生产，从而提高了生产效率、改善了生产质量、降低了劳动力成本、降低了维护成本，并加速了行业的数字化和智能化转型。

4.4 绿色能源节约

在全球绿色发展趋势和可持续发展目标推动下，低碳移动通信网络的部署已成为行业共识。随着5G基站快速扩张，电费现在构成了运营商整体运营费用的很大一部分，使得节能和效率提升对于提高运营效率至关重要。数字技术在实现碳中和中扮演着至关重要的角色，并在帮助世界应对气候变化方面必不可少。积极探索节能技术的创新与应用，利用数字和智能化方法提升网络能源效率，不仅是一种有效降低运营成本的方式，也是推动行业可持续发展的关键步骤。

基站节能技术在时间域、频域、功率域和空间域已经实现了多维度的综合能耗管理。然而，鉴于多标准和多频段网络、多样化的服务场景以及不断变化的用户感知需求，传统的粗放式节能策略难以满足这些多样化的需求。关键挑战在于精确预测服务需求、动态触发节能机制以及在复杂网络条件下区分节能策略。实现这种精确性和效率对于节能技术的成功应用至关重要。

为解决单一目标基站节能策略的局限性，这些策略无法适应复杂网络结构、多样场景和服务需求，引入了大数据分析和AI算法。这些技术增强了节能策略在时间、频率、空间和功率域内的实时性能和精确度，从而扩大了节能潜力。方案如下所述。

1) 精确识别节能电池

通过进行联合区域分析以识别覆盖重叠区域，自动配置节能补偿细胞，以及发现更多节能机会，我们可以提高网络能效。为了实现这一目标，AI算法基于不同类型基站的不同特征，分析大量基站数据。这允许自动识别和分类基站。利用基站能效与流量之间的数据关系模型，识别问题基站，并采用节能技术来提升整个网络的能效。



图4-9 能耗细胞识别

2) 能源节约持续时间准确预测

使用规律分析来分析历史网络运营商/蜂窝组数据，并使用时间序列模型训练最优负荷预测算法。识别在低负荷条件下符合触发运营商关闭或深度睡眠功能的期限。利用人工智能的自我学习能力，在监视性能指标（基本KPIs，服务KPIs，感知KPIs）波动的时候迭代调整阈值参数。通过优化节能阈值，这有助于找到节能和系统性能之间的最佳平衡。

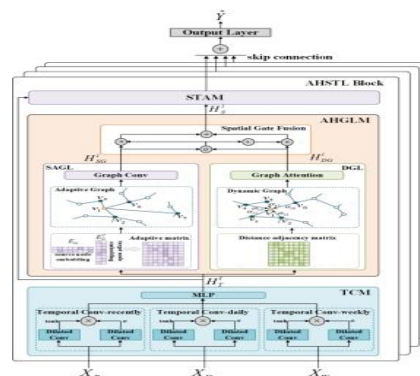


图4-10 能耗持续时间预测模型

3) 能源节约策略的智能优化

对于多场景、多标准、多频段和多节能技术，制定协调的节能策略。收集网络状态数据，如测量报告（MR）、性能测量（PM）和工程参数。使用高斯

基于混合密度的带有噪声的应与应用空间聚类（DBSCN）算法，用于确定蜂窝能量节省状态。为不同的蜂窝训练模型，并在各种参数配置下优化多维度网络性能指标。通过区域内的联合自学习，持续迭代和优化，生成最佳能量节省技术与网络参数配置的最佳组合，以适应当前的覆盖场景和业务需求。从而实现针对“一站点、一时段、一策略”的智能动态能量节省策略，旨在实现最大的能源节省。

4) 基于服务感知的网络安全节能

引入服务感知能力以基于不同服务类型和需求优化节能策略。例如，对于对延迟敏感的服务，开发差异化的集中式调度策略以平衡节能策略对网络性能的影响。这使网络能够在不影响服务质量的情况下激活节能功能，并延长这些功能的有效期。此外，根据服务特性和网络监控指标的变化动态调整和优化节能策略，确保网络性能并满足对延迟敏感的服务需求。

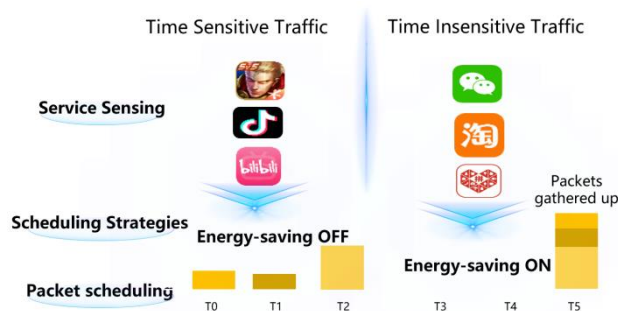


图 4-11 基于服务感知的差异化集中调度

智能方案，如服务预测、节能小区识别和节能优化策略，已在中国移动的网络中覆盖超过25个省份，服务约450万个基站。服务差异化的节能方案也已在全国范围内试点。在传统的节能功能之上，智能节能策略可以实现平均每天超过5%的节能收益，确保节能与性能之间的平衡。这显著促进了无线网络的绿色和可持续发展。

4.5 高可靠性网络

网络稳定性涉及多个方面，包括站点级和数据中心的运营。本白皮书从站点和数据中心的角度讨论了网络运营保证的挑战、解决方案和实际成果。

1) 网站运营稳定性保障

随着5G的大规模部署和发展，网络结构正变得越来越复杂，应用场景和服务需求也在多样化，这对网络运维的效率和效果提出了更高的要求。传统的站点运维方法面临以下挑战：

- **延迟故障响应**：依赖于人工检查和被动应对，故障检测和处理被延误，从而导致问题升级。
- **预测不足的隐蔽风险**：缺乏对设备状态的实时监控和数据分析能力，使得提前预测潜在隐藏隐患变得困难。
- **数据孤岛和信息不透明**：设备、板、电源等相关数据分散，难以整合和共享信息。缺乏统一的可视化平台使快速定位问题变得复杂。
- **低运维效率**：人工检查和故障排除耗时且劳动密集，需要高超的技术专长，并且高度依赖专家经验。

5G和AI的深度融合，以及大数据、机器学习、数字孪生和其他技术的采用，使得智能数字化场所的建设成为可能，显著提升了网络运维的自动化水平。这不仅满足了当前通信网络的需求，也为未来网络的智能化发展提供了强有力的支持。关键技术在以下方面得到阐述：

- **构建数字站点**：通过数字技术，基站设备、组件、运行状态和故障信息被整合到一个统一模型中，实现了对站点的精确理解，提高了站点感知的准确性和可靠性。
- **自动故障诊断**：AI模型在故障发生时智能分析现象和错误数据，快速推断和匹配问题，迅速确定根本原因，显著缩短停机时间，并提高运维效率。
- **故障预测与早期预警**：利用人工智能算法建立故障预测模型，自动分析现场运行数据，包括硬件和软件性能指标、环境因素如温度和湿度以及历史故障记录，可以准确地预测故障发生的概率并发出及时预警，有效预防因现场故障造成的服务中断。
- **运维可视化**：利用数字孪生技术直观展示复杂的运维数据、设备状态、故障信息和系统运行条件。可视化工具将抽象数据转化为视觉图形，有助于运维人员快速了解系统状态、识别问题并做出决策。

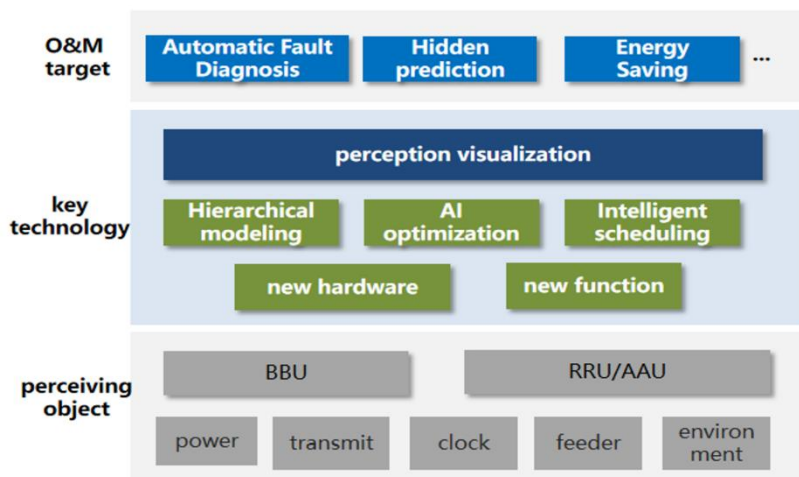


图 4-12 智能网络运维框架

应用程序已在以下场景中实施：数字和智能前传光纤链路。告警和前传光纤链路检测是现有网络的痛点，其特点为高故障率。此外，链路主要由缺乏监控能力的被动设备组成，依靠人工现场故障排除。迫切需要智能和自动检测方法以提高运维效率。通过引入配备反射环回自测功能的智能光模块，结合智能基站故障识别算法和分层数据处理架构，可以自动准确地识别和定位故障。这实现了对前传光路径的智能检测，显著提升了运维效率。

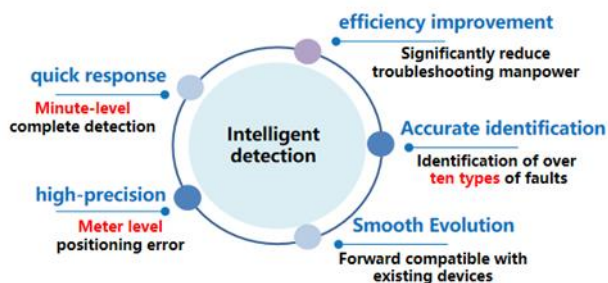


图4-13 数字化与智能化传输

2) 数据中心智能保护

当网络中出现数据中心（DC）级别的灾难时，原本位于已失效DC中的用户将在短时间内尝试重新连接到备用DC。这很可能会导致备用DC的网络负载突然激增，从而可能导致网络过载，进一步扩大服务中断的影响。尽管当前网络具有处理信号激增的预测机制，但这复杂了情况，并对实际切换过程构成了重大挑战。这是因为大规模灾难恢复切换涉及多种类型服务的同时切换，而不是单一服务的直接过渡。

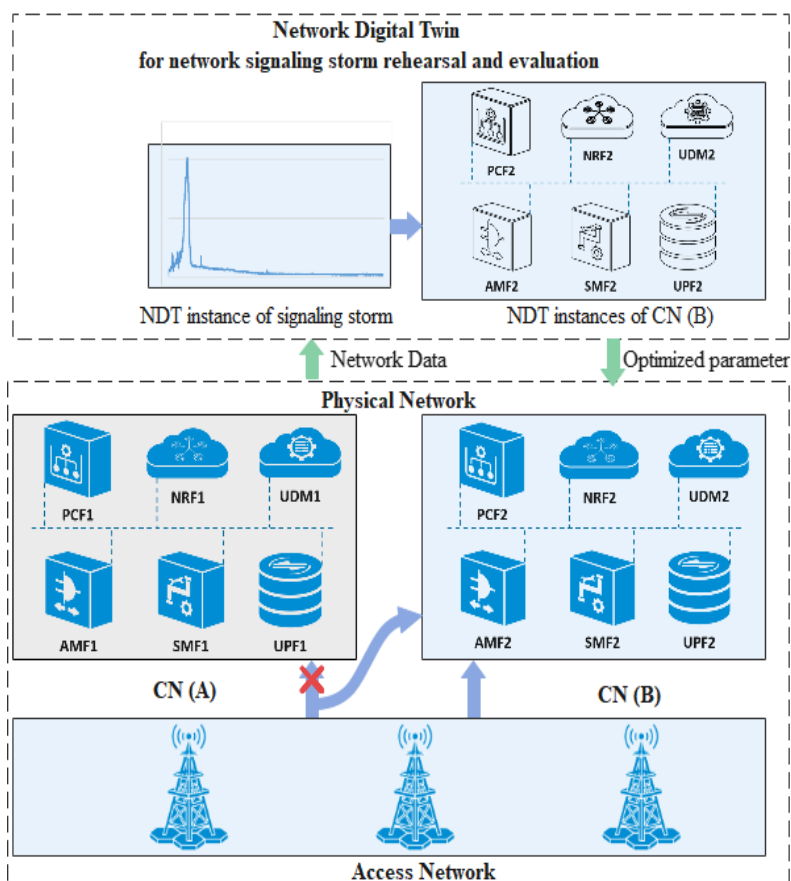


图4-14 数字孪生网络架构

为了更好地预防和控制信令风暴，中国移动与中兴通讯合作，于2024年9月在实际网络环境中对灾难恢复场景下的信令风暴预防和控制解决方案进行了验证。此方案采用人工智能和NDT技术来预测网络中潜在的风险和瓶颈，并启用主动的修复和预防措施，以降低与数据中心灾难恢复切换相关的风险。此外，验证还测试了在极端场景下的云网灾备能力，包括超过340万在线用户的切换以及每秒处理24万条信令消息。

在本解决方案验证中，通过研究资源池的性能组件、资源池内虚拟机/物理机/资源池/存储池之间的性能迭代关系，以及整合云网络功能（NF）的信号影响模型，建立了一个资源负载影响模型。基于NF信号影响模型和资源负载影响模型，构建了一个用于数据中心灾难恢复的网络数字孪生，以实现多云网络性能影响的定量评估和分析。该数字孪生系统已通过能力暴露和网络接口在实时网络系统中进行商业部署。

- a) 部署了一个用于直流灾难恢复的数字孪生系统，实现了对云网络元素信号影响和网络云负载影响的整个评估过程的全面自动化。评估时间已缩短至10分钟以内。
- b) 建立了一个针对网络功能信号影响和网络云的精确评估模型。

负载影响被确定，评估准确性超过95%，从而提高了灾难恢复的可靠性。c) 在切换过程中实时收集了网络功能和服务资源池设备的关键绩效指标（KPIs）。通过比较实际指标与模拟结果，整个切换过程变得完全可见和可控，提高了应急指挥效率和通信安全，同时消除了切换过程中可能出现的潜在二次风险。

关键采用的技术如下列所示：

- **模拟和评估信令风暴：** 数字孪生系统被用于模拟和评估灾难恢复信号风暴的预防和控制效果。它能够自动识别网络性能瓶颈和异常的流量控制参数，从而实现主动风险预防。
- **协同仿真以实现现实影响评估：** 采用一种协作方法，结合核心网络信令风暴影响模拟与资源层流量模拟。这确保影响模拟的波形更接近实际条件，有效提升了灾难恢复的可靠性。
- **人工智能与无损检测：** 对核心网络信令影响的评估采用网络数字孪生模型与人工智能技术的结合。该方法纠正了由终端行为差异和区域服务变化引起的计算偏差，实现模型的自适应。
- **实时监控与决策的可视化：** 利用网络数字孪生仪表板的实时监控能力，切换过程变得可视化和可控。这支持直流切换的实时状态感知和快速决策。

4.6 多模态个人助理

近年来，人工智能代理技术的应用逐渐渗透到日常生活的各个方面。人工智能代理是一种能够感知其环境、做出决策并执行行动的智能实体。人工智能助手服务是人工智能代理技术的一个典型应用。例如，在订购咖啡时，用户只需下达指令，系统便可以自动完成如选择咖啡类型、处理支付和安排配送等任务。这种便捷体验的背后，是人工智能代理技术在自然语言处理、机器学习和大数据分析等领域的快速发展。人工智能代理技术不仅能够理解用户意图，还能根据历史数据和实时环境进行个性化推荐。

2024年10月，智谱公司与清华大学合作推出了AutoGLM工具。通过语音命令，用户可以利用它完成电子商务购物、食品外卖订购、火车票预订、社交媒体互动和发送微信消息等任务。2025年1月，OpenAI发布了AI Agent应用程序Operator，该应用程序可以自动执行各种复杂操作，包括编写代码、预订旅行和自动电子商务购物。在移动设备之外，在智能家居中，人工智能助手服务可以控制家用电器；在车辆系统中，它们可以实现语音导航和音乐播放等功能。随着技术的不断成熟，人工智能助手服务

正在快速融入人们生活的各个方面，市场规模持续扩大，用户数量稳步增加，展现出良好的发展趋势。



图4-15 不同的AI助手服务

AI助手服务的实施主要涉及三种主流解决方案：边缘AI、云AI和混合AI。1) 边缘AI在本地设备上运行，数据处理和模型推理在终端本身完成。这种方法提供了快速响应时间和高隐私性，但受限于设备性能、存储容量和功耗，使其难以处理复杂任务。2) 云AI将所有计算和数据存储在云中，利用强大的云计算资源处理复杂任务。然而，它依赖于网络连接，这可能导致延迟，并引入隐私和安全风险。3) 混合AI结合了两方法的优点。简单任务在边缘执行，而复杂任务则卸载到云端，平衡性能、隐私和网络依赖。总体而言，混合AI有效地弥补了边缘AI和云AI的不足，成为当前AI助手服务的占主导地位的实施方案。

基于混合人工智能解决方案，人工智能助手服务首先在边缘侧分析数据，然后再将其上传至云端进行进一步处理。商业模式转向主要上传数据，对上传速度提出新的要求。目前，上传至云端的边缘侧数据主要是由分析后的关键特征字段或单张图像组成。随着人工智能助手服务的普及，对来自边缘侧的持续图像上传需求将日益增长，需要高上传速度。例如，传输高清图像可能需要高达20 Mbps的上传速度。5G-A网络的高速、低延迟和大连接特性确保边缘侧数据可以快速上传至云端进行推理，同时还能实现对用户云端处理结果的及时反馈。低延迟确保实时交互，大连接能力满足众多智能设备同时访问的需求。5G-A的这些特性为人工智能助手服务的广泛应用和发展提供了强有力的支持。

数字人是人工智能代理技术的另一典型应用，实时和基于流的多元模态人工智能交互是其功能的关键。借助多元模态人工智能模型，人机交互系统可以为用户提供近似人类交互体验，跨越各个领域，如生活方式信息、儿童教育、工作协助以及老年人的医疗保健。在复杂的场景，如工业生产和

智能城市管理，多模态人工智能将突破现有的交互范式，提供更智能、更全面的AI体验。例如，商汤科技推出了中国首个“所见即所得”的原生多模态大型模型——Riri Xin 50+ Ruying 数字人。该产品具备类似人类的实时视觉能力，能够与人们进行流畅的视频互动。它不仅能听、说、看，而且还能做到无延迟。该系统利用边缘和云计算资源快速生成高度逼真的数字人像，并支持与用户进行完全自然、实时的互动对话。

多模态人工智能结合来自各种感官模态的输入，如声音、视觉、手势，甚至触觉反馈，使人机交互更自然、更接近人类。在过去，由于网络速度和稳定性的限制，对这些大规模数据集进行实时处理是一项挑战。5G-A网络的出现使得多模态人工智能能够实时获取、传输和分析大量数据，快速生成响应。在未来，实时移动交互将需要边缘侧模型和通信网络的共同支持。在模型方面，推理延迟应达到人类水平或接近人类水平的响应时间，而系统需要能够瞬间有效地集成和传输来自不同感官通道的信息。从用户终端上传的多模态数据必须快速通过无线信道传输到云端服务器，对上游网络传输提出了更高的要求。

4.7 内在的人工智能

机器人的技术，自20世纪中叶发展以来，已经从简单的自动化设备演变为高度智能和自主的实体智能机器人。实体人工智能指的是机器人物理实体与其环境互动的系统，该系统能够进行环境感应、信息认知、自主决策和动作执行。此系统还能够根据经验反馈不断变得更加智能，并调整其动作。传统机器人根据规则和固定程序操作，适用于具有单一功能和长周期开发的特定场景，但缺乏适应性。相比之下，实体人工智能利用大型人工智能模型，在任务和环境方面提供多面性，具备积极理解和影响世界的能力，同时支持持续的学习和迭代。

人工智能实体化的应用场景涵盖了从日常生活到工业生产的多个领域，展示了巨大的市场潜力和技术优势。在家庭环境中，实体智能机器人能够处理家务劳动，陪伴老人和儿童，极大地提高了便利性。在医疗保健领域，这些机器人协助医生进行手术、护理患者和康复训练。在制造业中，它们执行复杂的组装、检验和物流任务，显著提高了生产线上的自动化水平和产品质量。在农业领域，实体智能机器人自主监测作物、喷洒农药和收获果实。在物流领域，仓库机器人高效地分类和运输货物。在安全领域，巡逻机器人实时监控环境并发出潜在风险的警告。例如，乐聚机器人已与多家汽车工厂合作，在生产线上部署多功能的实体智能机器人。四足机器人作为实体人工智能的一种形式，在恶劣环境中承担了众多任务。例如，优必选的机器人狗在泰山上用于垃圾收集。



图 4-16 库存机器人分类货物的示例

实体人工智能的广泛应用对通信技术提出了更高的要求，尤其是突出了5G-A通信的必要性。利用5G-A网络的大带宽和低延迟特性，结合AI能力，可以通过云协作增强实体智能系统的精确定位、低延迟、实时控制、数据收集和增量训练。这降低了实体智能系统本身的计算能力和能耗。例如，在需要高实时性能和精确控制的场合，如机器人跑酷，5G-A网络的低延迟特性确保机器人能够快速响应指令，实现动作执行流畅。5G-A网络的大带宽特性确保了机器人的大量传感器数据能够实时传输和处理，进一步增强了实体智能机器人的感知和决策能力。

具身人工智能已从室内、小规模应用扩展到室外、长距离场景，这些场景需要广泛的信号覆盖和移动性。5G-A通信完全满足这些需求。例如，华为与乐聚机器人合作，将5G模块集成到具身智能机器人中，而中兴与优必选科技合作，也是为了实现这一目标。这些合作不仅增强了机器人的通信能力，还为它们在复杂和动态的室外环境中的应用提供了技术支持。通过5G-A网络，具身智能机器人可以实现远程控制、实时数据传输和多机器人协同作业，显著提高工作效率和可靠性。

4.8 沉浸式体验

沉浸式实时通信集成了各种先进技术，如虚拟现实（VR）、增强现实（AR）、混合现实（MR）和无眼镜3D，突破了传统显示技术的限制，为用户提供了全新的、高度沉浸式的感官体验。作为高清沉浸式服务，XR和无眼镜3D是元宇宙的初始核心载体，代表了5G服务的主要未来发展方向。它们超高分辨率、高互动性和强计算需求的特点，为5G网络处理此类服务带来了挑战。

典型沉浸式服务应用具有超高清和强交互性，需要5G网络从单维到多维提升服务支持。存在三个主要痛点：1) 网络能力利用率不足：网络利用率低，服务流量增长未达到预期。2) 协调机制不匹配：传输机制低效导致服务质量差和传输效率低。3) 网络特性保障不够精细：网络支持用户和服务的能力不够细致，导致多目标优化复杂度高、资源利用效率低。

端到端的沉浸式实时服务解决方案是一种为高度沉浸式、强交互性和多维多模态通信服务而设计的综合方法。它旨在提升用户体验，确保在传输过程中提供实时、稳定和高品质的服务交付。

中国移动推出了一种以“信息跨层共享、多维度连接增强和量化感知评估”为核心的理念。基于这一理念，他们提出了包括毫秒级服务感知保障、帧级多维度连接增强、实时网络信息开放以及用户体验量化评估在内的多项里程碑式技术解决方案。这些创新旨在构建一个针对XR沉浸式体验的原创网络服务协作技术系统。

1) 解决实时网络服务交互问题

一方面，针对基于网络状态的服务感知，提出了一种毫秒级的服务感知保证技术：为解决服务感知需求的不充分性，提出了一种针对服务帧级特征的实际实时细粒度感知方案，实现了从非实时到毫秒级实时服务质量的飞跃。帧级QoS优化：通过感知帧周期和帧重要性等特征，基于帧级QoS需求进行优化调度；帧识别和保护：基站识别进入的包模式以识别帧周期、帧尾、帧大小和其他服务特征，确保帧级服务保护；跨层信息共享：关键服务信息通过RTP扩展头从应用层传递到UPF，然后通过GTP-U头部字段转发到基站，促进了从应用层到网络层的无缝信息传输。

另一方面，对于基于服务的网络状态感知，引入了实时网络信息开放技术：为了解决服务对网络状态感知不足的问题，提出了用户平面网络信息开放技术，使得应用层能够实时共享拥塞、吞吐量等网络信息。这解决了先前网络能力开放方案在全面性和节点多样性方面的不足，从而提升了网络能力和用户体验。

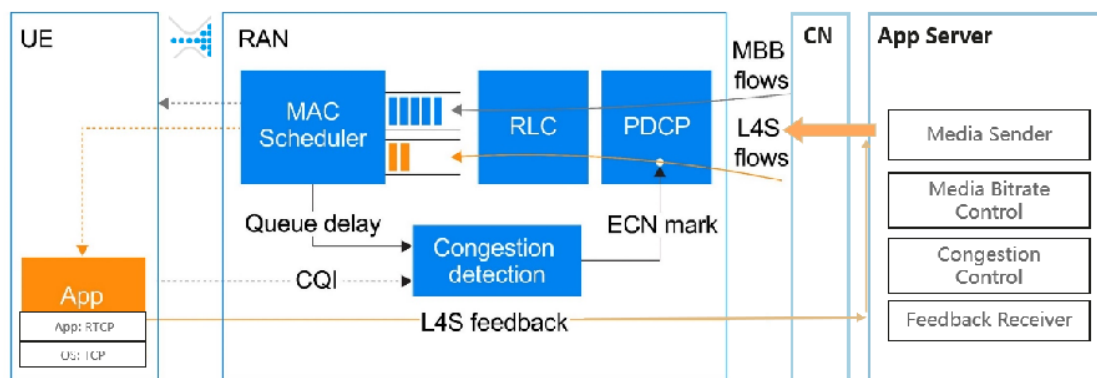


图 4-17 毫秒级服务意识保证技术的示例

2) 最大化连接能力

提议框架级多维度连接增强技术：

- 帧延迟和误码率指导：定义帧延迟和帧误码率等参数，以引导细粒度的帧级保护；
- 优化帧调度和完整性：通过实施帧延迟调度和帧完整性传输，优化帧级QoS指标；
- 多用户交错传输：通过在多个用户之间交错传输，最大化系统容量；
- 交换控制和MCS选择优化：利用基于帧边界的交换控制和调制与编码方案（MCS）选择优化切换延迟，显著提高沉浸式实时通信服务在移动场景中的传输性能和用户体验。

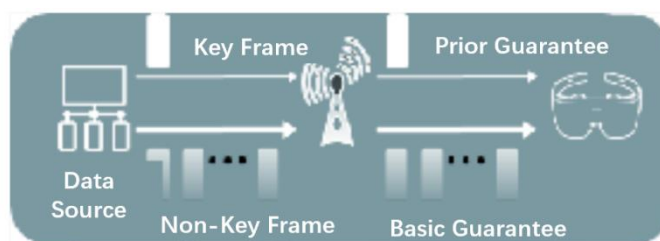


图4-18多流分播示例

3) 标准化经验评估

为了对XR服务体验进行标准化评估，提出了aframe级别的网络指标系统、测试方法和XR服务体验的设备要求。定义了传输延迟、空中接口传输延迟、传输速率和可靠性等frame级别的关键性能指标（KPIs），以定量评估XR服务体验。

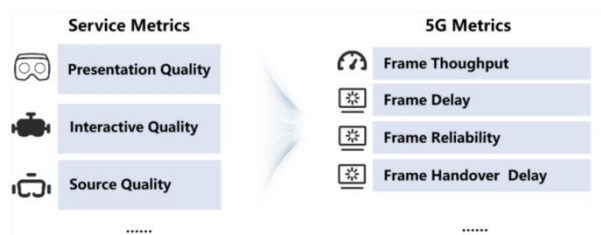


图4-19 5G网络指标系统用于XR沉浸式体验

应用案例 1：北京工人体育场“工体元宇宙”

对于体育行业，中国移动已与中核集团合作，在北京工人体育场部署了基于5G直播网络的全球首个元宇宙应用。这一创新应用集成了网络服务集成、跨层信息共享、基站服务感知和智能数据多流传输技术。它实现了高达100 Mbps的实时体验速度和低于20毫秒的帧级网络传输延迟。这些能力满足了工体元宇宙应用的高流量、强实时交互和高客户移动性需求。现场观众可以享受4K分辨率和90帧每秒的高清XR足球体验。这是在5G直播网络上首次部署元宇宙应用，为加速大规模元宇宙应用提供了示范。它展示了技术创新如何推动商业创新，反之亦然。

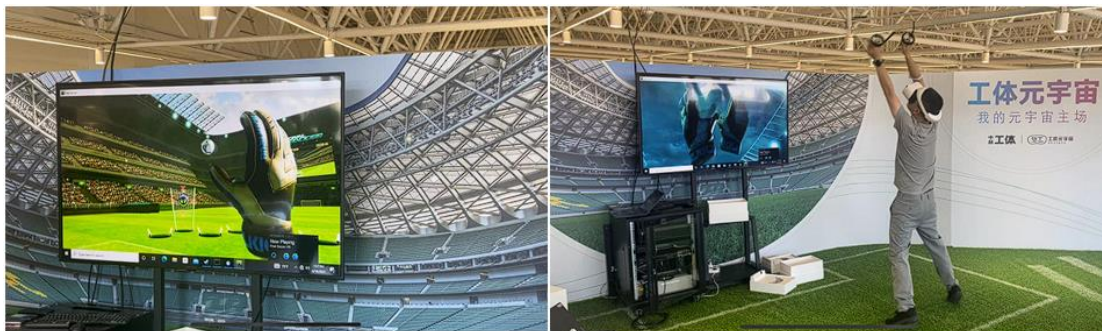


图4-20 北京工人体育场“工体元宇宙”

应用案例 2：首钢一号高炉 SoReal 科幻主题公园

在娱乐行业，中国移动与当红齐天合作，打造了业界首个5G-A室内大空间超密集多用户XR试点应用。对于室内超密集沉浸式XR场景，如重量大的背包设备、散热性能差、电池寿命短以及数十个并发用户的高容量、低延迟和移动性等挑战得到了解决。采用了5G-A智能服务感知、帧级多维连接增强、联合智能波束管理和“无背包”及“无线缆”分布式渲染解决方案等技术。依托5G-A网络提供的大带宽和低延迟，确保了数十个用户实时100 Mbps的体验速率和平均空口传输时延低于10毫秒。服务图像清晰度达到行业标准的4K @90 fps质量。在1000平方米的空间内，用户可以更自由地体验无界限的虚拟世界。



图4-21 首钢一号高炉 SoReal 科幻乐园

应用案例 3：亚洲运动会观看的新视听体验

基于我们公司的咪咕业务平台和5G-A网络的优点，我们创建了新的服务，例如VR电子竞技、VR活动现场直播和亚洲运动会裸眼3D观看。与传统的视频服务相比，XR和裸眼3D技术使用户能够享受逼真的三维视觉效果，自由切换观看角度，并体验高清沉浸式视听内容。这为用户提供了丰富、流畅、逼真的亚洲运动会3D沉浸式观看体验。咪咕移动云VR业务平台为亚洲运动会定制XR观看场景，通过XR技术在一个屏幕上观看多个比赛。这重新定义了观众的“现场”体验，使他们能够感觉到“沉浸式、激动人心、互动性”的亚洲运动会元宇宙。凭借5G-A网络的大带宽和低延迟保证能力，在多用户和多服务并发场景下实现了低于20毫秒的无线传输延迟和125 Mbps帧级保证速度，确保了各种类型高清互动3D云服务的顺畅体验。

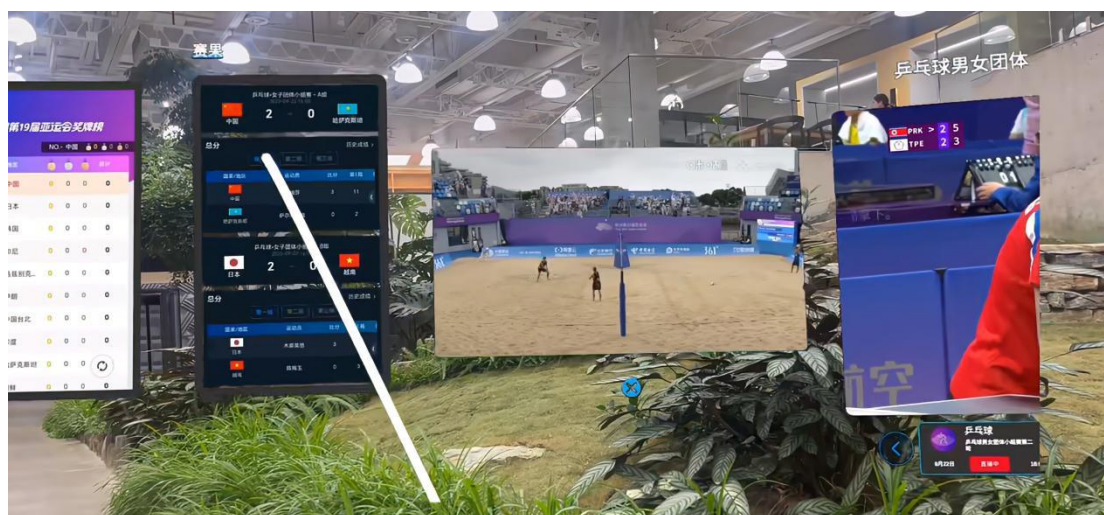


图4-22 亚洲运动会观看的新视听体验

4.9 网络智能驾驶

近年来，车辆变得越来越智能化，将汽车从单纯的交通工具转变为人们出行的“移动第三空间”。5G已成为新车制造商的必备条件，也是高端车辆的优先选择，它提供了更优的车内体验和智能驾驶体验。这使得汽车能够同时为多乘客提供休息区、办公室和游戏室。联网车辆具有三种主要类型的应用：智能驾驶舱、车云协作和车路协作。智能驾驶舱中的大尺寸多屏显示和实时交互增强、车云协作中的云辅助智能驾驶以及在车路协作中的路况感知，都要求5G-A网络具备高速率和低延迟的能力。



图4-23 智能驾驶舱中的多样化车载服务。

随着汽车智能化不断进步，驾驶方式正逐渐从辅助驾驶发展到自动驾驶。以下概述了汽车自动驾驶的五级标准，从L0到L4。根据不同的发展战略路径，自动驾驶的商业路径主要可以分为两大类：从L2（部分自动化）的逐步进化：这种方法涉及从L2开始逐步提升能力，例如特斯拉、理想汽车和蔚来等制造商所追求的；直接进入L4（全自动化）：这种方法涉及直接针对L4全自动化，如Waymo、WeRide、Pony.ai以及百度Apollo等自动驾驶解决方案提供商所采用的。

自动驾驶分类标准

自动驾驶分类	定义
L0：完全人类控制	驾驶员 solely 负责 操作 车辆 在所有道路和环境条件下。

辅助驾驶	驾驶员与车辆共享控制。在特定情况下道路和环境条件下，车辆配备了一个或多个专业自动化控制功能。
L2：部分自动化	车辆主要处于控制状态。在特定情况下...场景下，该车辆可以执行两个或多个自动化操作。驱动功能。然而，驾驶员必须始终保持警觉。环境和做好必要干预的准备。
L3：无需驾驶员监控驾驶过程	车辆在特定的道路条件下完全受控。环境条件，具备自动驾驶能力 的条件自动化。 然而，在紧急情况下，仍需人工干预。情况。
L4：全面自动化	车辆完全处于控制之下，并具备自动驾驶的能力。在各种道路和环境条件下驾驶而不人类干预。

2024年，中国销售的车辆中有超过50%配备了L2+级智能驾驶功能特性。



图4-24 L2+智能驾驶

L4级自动驾驶技术也已成熟，并在无人出租车和无人配送等众多领域实现商业化。截至2024年7月，由萝卜快跑运营的Robotaxi的商业订单已超过600万。基于低延迟和高上行链路容量等关键特性，5G-A网络可以确保上行速度达到20 Mbps，下行延迟为20 ms。这支持了对自动驾驶车辆的实时监控、异常情况下的即时远程控制和异常数据的实时上传到云端。这些功能确保了车辆的持续运行，并赋予了L4级自动驾驶车辆高效运行的能力。



图4-25 机器人出租车应用

智能道路和基于云的大型模型将补充个别车辆的智能，实现“车路云一体化”的最终目标。通过稳定的低延迟5G-A网络，基于云的大型模型的大规模计算能力将增强车辆有限的计算资源。智能道路的实时感知能力也将填补个别车辆超视距场景的盲点，进一步提高旅行安全和效率。例如，车辆可以上传关于复杂交通标志或潮汐车道的实时数据，这些数据无法在本地解析。基于云的AI大型模型将解析和识别这些标志，并立即将结果发送回车辆。这个过程减少了交通事故，允许实时选择最佳路线，并提高智能驾驶的效率。

4.10 低空智能连接

此白皮书描述了利用先进的信息与通信技术、人工智能、大数据以及其他技术手段，在低空场景下构建低空智能网络的解决方案，以应用于通信服务和感知服务。

通讯 在低空飞行环境中，终端设备如无人机面临挑战，如不明确的蜂窝覆盖边界、缺乏主导覆盖小区、众多且无序的邻近小区以及频繁的切换。这些问题可能导致无人机飞行期间频繁切换蜂窝，影响通信的连续性和稳定性，可能导致服务中断、数据丢失以及低空终端不良的通信体验。

为了减少切换次数、提高信号质量和提升用户体验，本文提出了一种针对低空飞行的基于人工智能的持续飞行保障解决方案。通过利用人工智能的预测和决策能力，考虑诸如服务基站和相邻基站的定位、当前RSRP测量值、无人机速度时间戳等因素，实现了沿飞行路径的智能小区切换。根据切换策略预先规划资源，使目标小区能够提前优化资源，确保无人机终端的持续飞行保障。

感知 集成通信和传感技术可以通过利用网络端数据，为低空无人机检测和传感提供服务，满足相关需求。这可以分为两大场景：空域管理和航线保护。空域管理主要用于监控未授权无人机，而航线保护则着重于保护物流或无人机公司的固定航线，以避免碰撞。传统的低速雷达检测技术在城市部署方面存在困难，连续网络构建复杂，且部署成本高昂。集成通信和传感技术通过将传感功能添加到传统基站，实现了通信和传感的集成，提供诸如位置和速度检测、轨迹跟踪和目标识别等能力。它具有连续网络连接和成本效益等优势。

中国移动针对低空和水域的广泛频谱感知需求，开创了中低频集成通信与感知技术和毫米波及Sub-6GHz高低频协调技术解决方案。为解决感知架构和空中接口设计中的问题，如覆盖距离不足等问题，不足的覆盖距离和

高度、低精度、以及复杂的网络和评估，他们提出了包括敏捷部署架构、混合波形新空中接口、超级角度硬件、高精度感知功能、集成通信与感知网络以及新型感知评估模型在内的创新技术和解决方案。作为移动通信网络演进中的新范式，集成通信与感知具备以下技术能力：

- 集成通信与传感空气接口设计：通过共享频谱资源，进行集成空气接口设计，实现通信数据与传感信息的双向整合。关键设计，如“OFDM+LFM混合波”传感波形和集成通信-传感帧结构，使通信和传感性能的资源共享成为可能。
- 集成通信与传感设备能力：通信与传感共享天线、射频模块、一些信号处理设备和其他硬件资源，有效提升设备集成和工作效率。在此基础上，实现了通信和传感功能的软件协调优化。
- 数据与服务开放能力：传感数据不仅可以直接在传感平台上展示，还可以根据客户需求与其他系统集成，实现多源数据AI融合处理。传感结果以不同形式呈现给不同客户，提供定制化服务体验。

应用案例1：浙江低空安全

2024年4月，浙江省杭州市启动了低空安全应用。与浙江移动合作，在一个复杂的城市环境中，围绕奥林匹克体育中心创建了一个低空安全示范区域。这一举措实现了对核心区域（约1平方公里）内无人机的高精度跟踪和识别，协助公共安全部门进行实时监控、报告和对策系统。此外，它还还为近5平方公里的区域提供了无人机感知安全以及及时检测和警报，单个区域可覆盖1.2公里，能在低于300米的高度精确检测到 RCS 为0.01平方米的无人机。

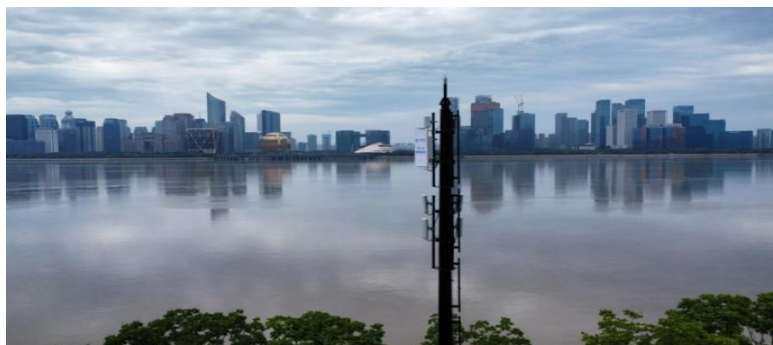


图4-26 低位安全应用案例

应用案例2：福建海洋野生动物保护

2024年，福建省厦门市正在实施海洋野生动物保护的应用。利用集成通信和传感技术，实现了对白暨豚核心保护区的连续超视距监控，同时在指定海域内进行船舶轨迹追踪和预警功能。一个云平台基地，

声学 and 视觉数据中心已为海洋生物建立，并已建立定制的大型模型服务，以创建智能保护区时空基础。



图4-27 海洋野生动物保护应用案例

应用案例 3：江苏长江航道船舶监控

2024年7月，江苏省南京市启动了船舶监控应用。与江苏省海事局合作，在南京长江段部署了三个4.9G集成通信和感知基站，从中山码头到南京长江大桥及八卦洲提供连续覆盖。这实现了对航道内非法船舶、无证船只和渔船的精确定位和全面监控，有效保护了水经济和提升水安全。



图4-28 江苏长江航道船舶监控应用案例

应用案例4：云南机场低空安全问题

在2024年7月，云南省宝山区启动了机场低空安全保障应用。围绕机场建立了一个大规模的低空监控网络和系统平台，实施无人机驱鸟、无人机检查和清晰空域巡逻等功能。实现了对低空无人机、跑道人员与车辆以及park安全的有效监控和预警。通过整合通信与感知技术，机场的低空安全保障能力得到全面升级，打造了智能化服务品牌体验示范区。云南移动与宝山机场签署了首个5G-A商业合作协议，探索了首例商业合作模式，并迅速在其他机场如文山、丽江和大理进行快速复制。



图4-29 机场低空安全应用案例

应用案例5：江苏桥梁微变形监测

2024年，江苏省南京市正在开展桥梁微观变形应用。与苏州科技学院合作，在七里山桥上完成了毫米级微观变形监测的全局首次可行性验证试点。在微风无雨的条件下，变形监测精度达到毫米级，可实现对桥梁结构条件的连续动态监测。试点结果已得到省级交通运输部门的认可。未来的研究将集中在提高恶劣天气条件下的监测能力，为桥梁健康监测提供高效轻量级解决方案，并提升基础设施安全管理。



图4-30 桥梁微观变形监测应用案例

应用案例6：上海低空航线

2024年5月，上海市金山区启动了低空航路申请。与浙江移动合作，上海移动利用多频协调优势，创新性地优化了低空场景的网络。这导致了全球5G-A低空网络覆盖超过100公里的跨海航路，包括沿海、近海和远海区域的5G信号全面覆盖。它为横跨100公里、垂直300米的低空新鲜农产品运输路线提供了强大的网络支持，显著提升了低空网络水平，促进了低空经济发展。



图 4-31 低空航线应用案例

第五次新型产业革命

5G-A和AI的集成可用于个人和行业的各种场景，为用户提供新的能力服务、新的应用服务和新的终端服务。基于这些创新服务，电信运营商可以与5G产业链的上游和下游协同，共同研究和探索创新商业模式，以充分发挥5G-A的新价值。

5.1 关于ToC/BtoC的三个创新商业模式

对于个人用户而言，一方面，除了收取数据流量费用外，电信运营商可以基于新型网络和AI能力，向个人用户提供功能费、服务费、网络差异化套餐费等收费；另一方面，应用服务提供商或终端制造商可以与电信运营商合作，为其高端应用/终端提供网络保障服务，从而获得额外收入。

5.1.1 能力服务

5G 新通话可以为个人客户提供如实时翻译和趣味通话等智能功能。在互联网接入方面，电信运营商可以提供差异化的上行和下行速率、接入优先级和QoS保证。在通话方面，当个人客户使用5G新通话的AI功能时，电信运营商可以收取功能费、服务等。在互联网接入方面，电信运营商根据不同套餐级别提供不同级别的网络能力，从而实现网络能力的货币化。

示例：针对个人客户的差异化网络能力套餐。

某电信运营商推出了两层级的5G套餐。低级别套餐拥有50GB的数据流量，最大下行速率为500Mbps，最大上行速率为50Mbps；高级别套餐拥有150GB的数据流量，最大下行速率为3Gbps，最大上行速率为200Mbps，且高级别套餐还可以为游戏、视频等应用提供优先访问、服务质量保证等服务。

即时通讯。

5.1.2 应用集成服务

5G-A网络可以根据业务类型、位置、时间等因素提供不同级别的网络保证服务。视频、直播、云游戏以及增强现实/虚拟现实等应用服务提供商与电信运营商合作，设计包含5G-A网络权限的应用/会员内容。电信运营商为订阅这些应用/会员的客户提供网络保证服务。应用服务提供商可以根据差异化的网络服务获得更高的收入，而电信运营商也可以根据应用/会员的订阅数量和网络使用量向应用服务提供商收取应用收入分成。

示例：云游戏服务提供商的会员收费

某云游戏服务提供商推出了两个付费会员等级，即VIP会员和超级VIP会员。对于超级VIP会员，云游戏服务提供商与电信运营商合作，为他们提供超高清图像和高级5G-A网络保证服务。超级VIP会员的费用比VIP会员每月高出20元。对于每个新发展的超级VIP会员，电信运营商向云游戏服务提供商每月收取10元作为网络服务费。

5.1.3 终端集成服务

5G-A网络可以根据终端类型提供不同级别的网络保证服务。无人机、智能汽车和人工智能计算机等终端制造商与电信运营商合作，为高端或指定终端提供5G-A网络保证服务。终端制造商可以通过网络服务向单个用户收取更多增值服务费，电信运营商可以根据保证终端的销售量和网络使用情况，从终端制造商那里收取终端销售收益分成。

示例：高端智能无人机的收费

某无人机制造商为直播无人机提供两层产品等级。高端直播无人机产品内置SIM卡，能够提供远程超清晰度和低延迟的实时视频传输服务。当客户购买高端直播无人机时，除了一次性购买成本外，他们还需要每年支付给终端制造商600元作为实时视频传输服务费，同时电信运营商向终端制造商收取每年300元的网络使用费。

5.2 三种创新的企业B2B商业模式

对于工业客户来说，除了提供如5G+DICT解决方案等个性化的定制化服务之外，电信运营商未来发展的趋势是提供轻量级的，

基于网络能力和人工智能应用的标准化和紧密协调的产品服务。这些服务包括以能力平台为主要载体的能力调用服务，基于云服务的任务型服务，以及协调网络能力与客户业务的业务协作服务。

5.2.1 能力调用服务

5G-A网络可以通过API调用来为用户提供切片和边缘计算等能力调用服务。解决方案提供商、系统集成商、行业客户等可基于运营商的网络能力开放平台独立构建个性化网络解决方案，以满足客户对带宽、延迟、可靠性等方面的网络需求。此外，网络能力在需要时可以即时调用，具有短部署周期和快速商业启动的特性。电信运营商可按调用能力的类型、调用次数等因素计费。

示例：Ooredoo 提出的 5G 网络API服务

卡塔尔电信运营商Ooredoo基于5G网络提出了四种API调用服务，包括SIM卡互换、设备状态、设备定位和按需质量。例如，当银行应用希望检查当前用户是否位于特定区域时，它可以使用该API获取该手机号码的定位信息。

5.2.2 基于任务的业务服务

电信运营商整合计算能力、网络、人工智能、大数据等能力和资源，基于云服务向用户提供相对标准化的基于任务的服务（如边缘云渲染、数字人直播、人脸识别等）。当客户使用基于任务的服务时，他们无需关注复杂的底层业务逻辑，可以构建自己的数字智能应用，有效降低应用门槛。任务式服务按照任务调用次数、持续时间或数据量计费，适合不需要太多个性化开发的应用场景。

示例：移动云计算平台提供基于任务的5G服务

中国移动的移动云平台提供了一系列云产品，如弹性计算、边缘云、视频服务和人工智能，包括基于任务的各项服务，例如视频直播、云渲染、语音识别和图像识别。例如，对于语音识别服务，移动云提供标准API接口和常用类型的SDK，并且提供接入指导，企业可以根据自身业务需求灵活定制和方便调用。

5.2.3 商务协作服务

基于5G网络能力与客户业务的紧密耦合，电信运营商和客户通过各自的优势资源相互补充。电信运营商主要专注于5G-A网络的建设，共同建设

智能应用与客户相结合。他们甚至可能在商业运营时进行合作，以确保客户商业运营的有效性，同时分享收益。电信运营商收取网络建设部分的成本，同时，他们根据客户商业运营的实际效果获得收益。

示例：体育场VR服务中的商业合作

一座特定的体育场通过VR服务提升了会员的价值，并提高了体育赛事的票价。电信运营商提供具有大带宽和低延迟的高质量5G-A网络，他们共同开发一个VR应用程序以提升用户的VR体验。电信运营商首先收取部分网络建设费用，然后根据VR服务的收入情况获得一部分收入。

全球行业合作提案

5G-A和AI的整合正在逐步展现其变革潜力。展望未来，5G-AxAI的持续演化为创新提供了激动人心的机会。在5G-A网络基础设施中开发AI代理和大型模型预计将使AI应用更加敏捷和高效。这反过来又将提升5G-A网络的表现力和可靠性。高速连接、实时数据处理和智能决策的结合正在多个行业中解锁新的潜力。为了加速这一进程，GTI提出了三项合作倡议。

联合研究高级技术。 GTI建议建立开放实验室，以完成从早期技术创新到商业验证的优化周期。此倡议旨在通过关注尖端研究来增强技术储备和促进行业发展。

资源共享。 GTI提出构建一个开放且协作的创新社区。这个社区将优化创新投资，最大化资源利用效率。通过资源共享，我们可以降低成本、提高效率，并加速新技术和新应用的开发，从而赋予行业升级和在快速发展的数字景观中保持竞争力的能力。

协同创建应用生态系统。 GTI 鼓励探索具有重大影响的应用场景，并完善可扩展的商业模型以推动工业升级。通过关注实际应用案例并将技术创新与市场需求相结合，该举措旨在确定5G-A和人工智能技术的最具发展潜力的应用领域。

本协作框架旨在利用5G-A与人工智能之间的协同效应，确保技术进步在各行各业转化为实际效益，并为一个更加互联互通和智能的未来贡献力量。

7 术语表

5G-A

5G-Advanced

5QI	5G QoS Identifier
ADC (注: ADC 为英文缩写词, 即Antibody Drug Conjugate, 通常翻译为“抗体药物偶联物”, 但由于其为专有名词, 故在中文文献中通常直接使用ADC)	应用数据通道
人工智能	人工智能
人工智能生成内容 (AI-Generated Content)	人工智能生成内容
人工智能 (Artificial General Intelligence)	通用人工智能
自动导引车 (Automated Guided Vehicle)	自动导引车
AR	增强现实
ASR	自动语音识别
BBU	基带单元
BDC	引导数据通道
CC	组件载体
CNN	卷积神经网络
简历	计算机视觉
DBSCN	基于密度的应用空间聚类
直流电	噪音
DCSF	数据中心
DL	数据通道信令功能
深度神经网络 (Deep Neural Network)	深度学习
双音多频 (Dual Tone Multi-Frequency)	深度神经网络
DTN	双音频多频率
DPD	数字孪生网络
ELAA	数字预失真
增强移动宽带 (Enhanced Mobile Broadband)	极大规模天线阵列
XR	增强型移动宽带
FNN	扩展现实
FTTR	前馈神经网络
FSA	光纤到房间
HOF	灵活频谱接入
GAN (生成对抗网络)	移交失败
GBR	生成对抗网络
IMS	保证比特率
IMS DC	IP多媒体子系统
物联网	IP多媒体子系统数据通道
IOV	物联网
大型语言模型	物联网
MAML	大型语言模型
MF	元动作标记语言
MIMO	媒体功能
MLP	多输入多输出
MOS	多层感知器 (Multilayer Perceptron)
MR	平均观点得分
磁共振成像 (MRI)	测量报告
	磁共振成像

无损检测	网络数字孪生
NF	网络功能
网络功能虚拟化 (Network Function Virtualization)	网络功能虚拟化
自然语言处理	自然语言处理
西北地区国防安全联盟 (Northwest Defense and Security Alliance)	网络数据分析方法
运营与维护 (Operations and Maintenance)	运营与维护
PA	功率放大器
PCF	政策控制功能
PEFT	参数高效微调
可编程逻辑控制器 (Programmable Logic Controller)	可编程逻辑控制器
PM	绩效评估
服务质量 (Quality of Experience)	用户体验质量
服务质量 (Quality of Service)	服务质量
无线接入网络	无线接入网
RLF	无线电链路故障
递归神经网络	循环神经网络
风险回报矩阵	无线电资源管理
SA	服务意识
SDN	软件定义网络 (Software-Defined Networking)
服务级别协议 (Service Level Agreement)	服务水平协议
SUL	补充上行链路
TTS	文本到语音
UCBC	上行链路为中心的宽带通信
UE	用户设备
URLLC	超可靠低延迟通信
UPF	用户面功能
V2X	车与万物
XR	扩展现实

8 参考文献

[1] GTI. "5G-A和人工智能：新时代，新机遇，新价值"。
[2] GSMA。"2024 5G发展状况：GSMA Intelligence 5G连接指数介绍"
[3] 电信。"最大化5G影响：2024年调查报告"
[4] OMDIA."5G-A行业智能成熟度指数"
[5] ZTE公司。5G + AI一体化通信计算。ZTE新闻。[6] R2-2408326讨论和评估关于RM
测量预测，日本NTT DOCOMO， Incorporated。
[7] R2-2405004与其他方面关于RLF/HOF预测的讨论，日本NTT DOCOMO， Incorporated。
[8] R2-2408327关于测量事件预测的讨论，日本NTT DOCOMO， Incorporated。
[9] M.Li等，《对数字孪生网络的自动数据生成与优化》，出自IEEE Transactions on Ser
vice Computing，doi：10.1109/TSC.2024.3522504。[10] Q.Li等，《为第六代网络的本地
网络安全孪生架构：设计至实践》，载于IEEE。

64

《通信杂志》, doi: 10.1109/MCOM.001.2400216.