
网络人工智能应用白皮书

Whitepaper on Network AI Applications

(Rev. 20190626)

中国联通

中兴通讯股份有限公司

2019年6月

目 录

第一章	人工智能技术简介与前沿进展	7
1.1	人工智能发展历史	7
1.2	人工智能前沿进展	8
1.3	网络人工智能 (Network AI)	8
第二章	标准组织与开源组织进展	10
2.1	标准组织	10
2.1.1	3GPP	10
2.1.2	ETSI	10
2.1.3	ITU-T	11
2.1.4	CCSA	11
2.1.5	O-RAN	11
2.2	开源组织	12
2.2.1	Linux 基金会(Linux Foundation) ^[8]	12
2.3	产业联盟	14
2.3.1	中国人工智能产业发展联盟 (AIIA)	14
第三章	人工智能在通信网中的主要应用场景	15
3.1	网络规划	15
3.1.1	智能化容量评估	15
3.1.2	自动站址规划及覆盖效果评估	15
3.2	网络运维	15
3.2.1	智能故障溯源	16
3.2.2	智能健康度预测	17
3.2.3	智能工单管理	18
3.2.4	智能 DevOps	19
3.3	网络优化	19
第四章	人工智能在运营商业务创新中的应用	22
4.1	边缘智能	22
4.1.1	无线感知服务	22
4.1.2	应用使能服务	23
4.2	物联网	24

4.3	创新业务.....	25
4.3.1	自动驾驶.....	25
4.3.2	VR/AR.....	25
4.3.3	智慧家庭.....	26
第五章	面向 5G 的人工智能应用.....	27
5.1	5G 网络编排.....	27
5.2	5G 网络节能.....	27
5.2.1	基于人工智能的网络节能.....	27
5.3	5G 基站调优.....	28
第六章	网络 AI 能力平台.....	29
6.1	中国联通 AI 人工智能平台系统架构.....	29
6.2	基于人工智能网络平台的应用实现流程.....	31
总结	33
缩略语	35
参考文献	36
致 谢	37

图 目

图 1	网络故障溯源方案示意图.....	16
图 2	网络故障预测方案示意图.....	17
图 3	智能工单管理方案示意图.....	19
图 4	无线覆盖智能优化方案示意图.....	20
图 5	智能流量预测及路径调优方案示意图.....	21
图 6	异系统频谱资源共享示意图.....	22
图 7	TCP 跨层优化示意图	24
图 8	网络人工智能平台参考架构.....	29
图 9	智立方平台功能架构（截至 2019 年 6 月 27 日）	30
图 10	人工智能端到端实现流程.....	31

第一章 人工智能技术与前沿进展

1.1 人工智能发展历史

1955年，麦卡锡（John McCarthy）、明斯基（Marvin Minsky）、罗切斯特（Nathaniel Rochester）和香农（Claude Shannon）等科学家在美国达特茅斯学院开会研讨“如何用机器模拟人的智能”，首次提出“人工智能”这一概念，标志着人工智能学科的诞生^[1]。

维基百科将人工智能（AI: Artificial Intelligence）定义为由人制造出来的机器所表现出来的智能，一般情况是指通过普通计算机程序来呈现人类智能的技术^[1]。

人工智能 60 余年的发展历程划分为以下 6 个阶段：

（1）1956 年至 20 世纪 60 年代初期：人工智能概念在 1956 年首次被提出后，涌现出系列研究成果，如机器定理证明、跳棋程序、LISP 表处理语言等，掀起了人工智能发展的第一个高潮。

（2）20 世纪 60 年代至 70 年代初期：科学家们尝试更具挑战性的任务，并提出了一些不切实际的研发目标，从而导致目标难以达成，使人工智能的发展走入了低谷。

（3）20 世纪 70 年代初至 80 年代中期：此阶段出现了专家系统，通过模拟人类专家的知识 and 经验解决特定领域的问题，实现了人工智能从理论研究走向实际应用的重大突破。专家系统在辅助医疗、商业决策等领域取得成功。

（4）20 世纪 80 年代中期至 90 年代中期：随着应用规模不断扩大，专家系统内在的知识获取困难、推理方法单一等问题不断显露，从而限制了基于专家系统的人工智能应用发展。

（5）20 世纪 90 年代中期至 2010 年前后：伴随高性能计算、分布式计算技术的发展，人工智能应用所需的算力水平得到提升，推动人工智能技术进一步走向实用化。1997 年，IBM 公司的深蓝超级计算机战胜了国际象棋世界冠军卡斯帕罗夫，向全世界证明对于特定问题，机器在单一领域的“智能”能够达到或者超过人类的最高水平。

（6）2011 年迄今：经过 60 多年的演进，特别是在移动互联网、大数据、超级计算、传感网、脑科学等新理论、新技术以及经济社会发展强烈需求的共同驱动下，人工智能技术也正在加速发展，呈现出深度学习、跨界融合、人机协同、群智开放、自主操控等新特征。受脑科学研究成果启发的类脑智能蓄势待发，芯片化、硬件化、平台化趋势更加明显，人工智能发展进入新阶段。当前，新一代人工智能相关学科发展、理论建模、技术创新、软硬件升级等整体推进，正在引发链式突破，推动经济社会各领域从数字化、

网络化向智能化加速跃升。

同时，通过开源组织、公有云服务商的努力，人工智能应用的构建变得日趋简单。人工智能技术在信息和通讯、工业制造、安防、金融、自动驾驶、医疗健康等众多领域取得切实应用，人工智能赋能的产品、业态创新带来生态和商业模式的全新变革。

1.2 人工智能前沿进展

全球产业界充分认识到人工智能技术引领新一轮产业变革的重大意义，纷纷调整发展战略。比如，在其 2017 年的年度开发者大会上，谷歌明确提出发展战略从“Mobile First”（移动优先）转向“AI First”（AI 优先）；微软 2017 财年年报首次将人工智能作为公司发展愿景。

人工智能当前前沿的发展方向包含数据平台、开源算法、计算芯片、基础软件、图形处理 GPU 服务器等技术生态系统和智能制造、智能医疗、智能安防、智能零售、智能家居等商业和应用生态系统。

在技术生态方面，人工智能算法、数据、图形处理器（Graphics Processing Unit，简称 GPU）/张量处理器（Tensor Processing Unit，简称 TPU）/神经网络处理器（Neural network Processing Unit，NPU）计算、运行/编译/管理等基础软件已有大量开源资源，例如谷歌的 TensorFlow 第二代人工智能学习系统、脸书的 PyTorch 深度学习框架、微软的 DMTK 分布式学习工具包、IBM 的 SystemML 开源机器学习系统等；此外谷歌、IBM、英伟达、英特尔、苹果、华为、中国科学院的寒武纪等积极布局人工智能领域的计算芯片。在人工智能商业和应用生态布局方面，“智能+X”成为创新范式，例如“智能+制造”、“智能+医疗”、“智能+安防”等，人工智能技术向创新性的消费场景和不同行业快速渗透融合并重塑整个社会发展，这是人工智能作为第四次技术革命关键驱动力的最主要表现方式。

1.3 网络人工智能（Network AI）

电信运营商作为信息与通信产业生态的主体之一，坚持推进人工智能技术与网络运营和业务创新相结合，使得通信产业成为人工智能技术落地应用的主要领域之一。

在信息通信向人类生产、生活方方面面渗透的过程中，电信运营商一直扮演着“水”的角色，在 AI 向各行各业渗透时亦如是，不可或缺。“上善若水，大有可为”。在 2017 年 12 月乌镇世界互联网大会——人工智能分论坛上，中国联通总经理陆益民将中国联通在人工智能中的作用定位为：不仅仅是网络基础设施的提供者、数据管道的提供者，

也是人工智能计算能力的提供者、海量数据的提供者。

2019年6月6日，工业和信息化部向中国联通办了5G牌照；如果说3G、4G开启了移动互联网时代，那么5G将开启人工智能时代。因此，5G网络与人工智能的结合将成为必然命题，运营商应紧紧抓住国家人工智能发展规划带来的历史性机遇，充分利用各方技术、产品、运营实力，促进通信行业向网络智能化、业务个性化、行业应用智慧化和管理智能化转型。通过人工智能技术，提高网络规划、建设、维护等效率，增强网络智能组网、灵活运作、高效支撑业务等能力，降低网络建设维护成本和管理成本，提升用户体验和行业、个人、家庭业务的竞争力，实现网络智能化的转型。

从学科分类来看，人工智能技术是计算机科学的一个分支，致力于了解智能的实质，并生产出一种新的能以人类智能相似的方式做出反应的智能机器，感知环境并采取行动以最大限度地实现目标。

网络人工智能(Network Artificial Intelligence)是指将人工智能技术应用在运营商网络中，通过网络的智能化或智能子系统替代或优化目前依靠人工进行的工作，使运营商能够更加便捷、高效的提供更加优质的网络服务^[2]。

网络人工智能化演进是个分阶段实现的长期过程，基于人工智能技术应用的不断深度，大体可分三个阶段：

第一阶段主要是实现网络资源的统一部署、控制和管理，实现软件定义的业务敏捷部署，释放大量重复、低效的手工劳作和人工交互；

第二阶段引入大数据分析和机器学习等 AI 技术，提供网络预测性感知和前瞻性的保障能力，确保网络可靠、平稳运行并提供优化建议；

第三阶段需要匹配高层次运营意图和策略，网络在自动管控、深度感知基础上，通过对这些意图的持续验证和综合优化，实现运营意图驱动的智能闭环自治，在极大降低 OPEX 的同时最大释放网络潜能和效益。

第二章 标准组织与开源组织进展

2.1 标准组织

2.1.1 3GPP

3GPP SA2 在 2017 年 5 月的杭州会议上成立了关于 5G 网络智能化的研究项目“Study of enablers for Network Automation for 5G (eNA)”^[3]。该项目是 3GPP Rel 16 的项目，SA2 已在 2018 年 1 月的哥德堡会议上正式展开项目讨论。

该项目的立项背景是 3GPP SA2 在 Rel 15 的 5G 核心网中引入了一个 NWDAF (Network Data Analysis Function) 的网络功能。目前该功能主要应用于部署了网络切片的场景，通过对网络切片相关的网络数据的自动化分析，向 PCF 和 NSSF 提供网络切片状态分析结果。另一方面，在 Rel 15 的 5G 网络架构研究中，SA2 引入了一些新的需求，例如按需移动性管理、非标准化 QoS、流量分流和卸载等，在没有网络数据分析的情况下，这些需求难以实际部署和运行。因此，为了让 5G 网络能够更灵活更智能的提供服务，华为在 SA2 牵头成立了 eNA 项目。

eNA 项目的研究目标是，通过 NWDA 对网络数据的收集和分析，生成分析结果，然后利用分析结果进行网络优化，包括定制化的移动性管理，5G QoS 增强，动态流量疏导和分流，UPF 选择，基于 UE 业务用途的流量策略路由，业务分类等。

在 2018 年 6 月的 RAN#80 全会上通过了“RAN-centric Data Collection and Utilization SI”立项，研究面向网络自动化与智能化的无线大数据采集与应用，并且探索在 RAN 侧引入数据分析的潜在影响。

在 2018 年 9 月的 SA WG5#81 会上通过“Intent driven management service for mobile networks”立项，调查意图驱动的移动网络管理场景，研究可用于实现移动驱动目标的、包括 SON 在内的自动化机制，以及描述意图的适当机制。

2.1.2 ETSI

2017 年 2 月欧洲电信标准化协会 ETSI 正式批准成立了新的行业标准组 ENI (全称“Experiential Networked Intelligence”)^[4]。该工作组的工作目标是就网络智能化的场景和需求达成一致并定义基于“观察-判断-决策-行动”这种模拟人脑决策的闭环控制模型，通过对网络环境持续学习和决策结果的持续优化等手段来有效应对复杂的网络管控挑战、提升网络运营管理效率和体验。研究范围包括分析运营商传统网络和 SDN/NFV 网络中运营运维需求，引导构建包括自适应感知、支撑灵活策略定义以及智能化决策和执行的体系架构。该体系架构应能充分支撑运营商灵活的业务策略和自动化、自优化、

自治的智慧网络理念。同时，该架构还将推动创新的遥测遥感、合理的大数据采集和管理、机器学习算法等技术来支撑智能分析和决断，统一的策略模型。

2017年12月欧洲电信标准化协会 ETSI 正式批准成立了新的行业标准组 ZSM（全称“Zero touch network & Service Management”）。ZSM 最初将专注于 5G 端到端网络和服务管理，如网络切片管理，并将扩展未来网络世代的管理。目标是让所有操作流程和任务 - 交付，部署，配置，保证和优化 - 自动执行，理想情况下 100% 自动化。该小组将定义一个新的，面向未来的，横向和纵向的端到端可操作框架，以实现新兴和未来网络和服务的灵活，高效和定性管理和自动化。横向端到端是指跨域，跨技术方面。垂直端到端是指跨资源方面，从面向资源的层到面向客户的层。ZSM 小组还将促进相关标准化机构与开源项目之间的协调与合作。

2.1.3 ITU-T

在 2017 年 11 月召开的 ITU-T SG13 会议期间，来自德国、韩国、中国、突尼斯、非洲等不同国家地区建议成立机器学习-网络焦点组，经 SG13 全会讨论正式批准成立。焦点组正式名称为 Machine Learning for Future Networks including 5G, 简称 FG-ML5G^[5]。

FG-ML5G 是一个对 ITU 成员和非 ITU 成员都开放的工作平台，目标是分析如何在未来网络特别是 5G 网络中应用机器学习来提升网络性能和用户体验，包括分析目前业界标准组织相关工作并与其合作，研究机器学习在未来网络的应用场景、潜在需求、架构，以及具体接口、协议、算法、数据结构和个人信息保护等方面，并分析机器学习对自主网络控制和管理方面的影响。

FG-ML5G 目前作为 SG13 下设的焦点组，研究期为 1 年，如果有必要也可以由 SG13 决定是否延期。FG-ML5G 输出的研究报告和标准草案可以作为后续 SG13 相关研究的输入。

2.1.4 CCSA

已经在多个应用领域展开相关研究工作，在 2017 年 7 月的 TC1-WG1#58 会上讨论通过了“人工智能在电信网络演进中的应用研究”课题立项。2017 年 12 月的 TC5-WG6 #47 会上通过“人工智能和大数据在无线通信网络中的应用研究”课题立项。2017 年 12 月的 TC5-WG12#2 会上通过“智能化 5G 核心网络切片技术研究”课题立项。2018 年 8 月的 TC5-WG6#49 次会议上通过“移动通信网络智能化能力分级研究”课题立项^[6]。

2.1.5 O-RAN

2018 年 2 月，中国移动、美国 AT&T、德国电信、日本 NTT DOCOMO 以及法国

Orange 等 5 家电信运营企业宣布联合成立开放无线接入网络（O-RAN）联盟，目标是使无线网络实现标准化、通用化、开源化、智能化。该联盟计划引入 RAN 智控平台，通过实时数据分析、机器学习及人工智能技术，让网络更具智慧性^[7]。

2.2 开源组织

2.2.1 Linux 基金会(Linux Foundation)

2.2.1.1 Linux 人工智能基金会（Linux Foundation AI）

2018 年 3 月，Linux 基金会下成立了专门研究人工智能技术的深度学习基金会（LFDL: Linux Foundation Deep Learning），以支持人工智能、机器学习和深度学习方面的开源创新，如 Acumos AI、Angel ML 等^[8]。

2019 年 5 月，Linux 基金会正式宣布将 Linux 基金会深度学习基金会更名为 LFAI 基金会(LF AI Foundation)，以支持开源 AI，ML 和 DL，并创建可持续的开源 AI 生态系统，使用开源技术轻松创建 AI 产品和服务，LFAI 的成员包含 AT&T、Orange 等运营商、华为、中兴、爱立信、诺基亚等设备商以及腾讯、百度、滴滴、红帽等互联网企业和 IT 企业。

截至 2019 年 6 月，Acumos AI 是 LFAI 首个和唯一的孵化晋级的开源项目，该项目的初期源代码由 AT&T 和 Tech Mahindra 联手研发并贡献给 Linux 深度学习基金会。该项目旨在针对现代企业引入 AI 技术的共性需求，提供一个共同的架构和平台，以帮助建立和管理人工智能平台。

2018 年 11 月 7 日，Acumos AI 的首个版本 Athena 发布，提供编辑、集成、组合、包装、培训和部署 AI 微服务，基于平台提供的微服务，开发人员能够连接各个应用程序来创建 AI 和机器学习产品。

截至 2019 年月，LFAI 还有 4 个开源项目在孵化中^[9]，分别是：

（1）Angel-ML

该项目旨在提供基于参数服务器（Parameter Server）理念开发的高性能分布式机器学习平台。Angel-ML 由腾讯和北京大学联合开发，由腾讯贡献给 Linux 深度学习基金会，基于腾讯内部的海量数据进行了反复的调优，并具有广泛的适用性和稳定性，模型维度越高，优势越明显。

Angel 的核心设计理念围绕模型。它将高维度的大模型合理切分到多个参数服务器节点，并通过高效的模型更新接口和运算函数，以及灵活的同步协议，轻松实现各种高

效的机器学习算法。

当前，Angel 基于 Java 和 Scala 开发，能在社区的 Yarn 上直接调度运行，并基于 PS Service，支持 Spark on Angel，集成了部分图计算和深度学习算法。

(2) Elastic Deep Learning (EDL)

该项目将基于百度开发的 PaddlePaddle 框架的开发的 AI 模型和应用与 Kubernetes 提供的资源管理、作业调度相结合，通过 Kubernetes controller, PaddlePaddle auto-scaler 组件实现优化的作业调度，从而实现集群硬件资源利用率的最大化。

(3) Horovod

该项目旨在为基于 TensorFlow 的模型训练提供分布式训练的加速方式，该项目的初期源码由 Uber 贡献给 Linux 深度学习基金会。

(4) Pyro

Pyro 是基于 Python 和 PyTorch 实现的深度概率建模工具库，帮助开发人员为 AI 研究创建概率模型。该项目由 Uber 贡献给 Linux 深度学习基金会。

2.2.1.2 LFN (LF Networking Fund)

2018 年 1 月，Linux 基金会为整合产业资源，平衡不同网络开源项目的生态系统，消除不同项目之间的重叠或冗余，创建更高效的流程，加快网络开源发展进程，成立了网络基金组织 LFN (Linux Foundation Networking Fund)，旨在通过统一的董事会管理，协调网络相关的开源项目。

截至 2019 年 6 月，LFN 管理的开源项目报告：ONAP、Open Daylight、OPNFV、FD.io、PNDA、SNAS、Tungsten Fabric 等 7 个网络相关的开源项目。

ONAP 目前是全球最大的 SDN (Software Defined Networks, 软件定义网络) /NFV (Network Function Virtualization, 网络功能虚拟化) 网络协同与编排器开源平台，面向 5G、企业和家庭宽带等场景，打造网络全生命周期管理平台，包括 SDC (Service Design and Creation, 服务设计和创建)、SO (Service Orchestration, 网络服务编排)、VF-C (Virtual Function Control, 虚拟功能控制)、AAI (Active and Available Inventory, 活动和可用目录)、DCAE (Data Collection Analytics and Events, 数据采集分析和事件处理) 等项目。在 ONAP Amsterdam 开源平台虚拟网元部署过程中，通过 SDC 上载各类虚拟网络功能并进行服务编排，然后由 SO 调用 VF-C 进行虚拟网元实例化，其中 Multi-VIM (Multiple Virtualized Infrastructure Management, 多虚拟基础设施管理) 负责电信云平台选择，VIM

(Virtualized Infrastructure Management, 虚拟基础设施管理) 负责具体虚拟计算、虚拟存储、虚拟网络资源池的管理和调度。虚拟网元配置信息, 如 vCER (Virtualized Cloudify Enterprise Router, 虚拟云化企业路由器) 需要的内存、CPU 核数等以 TOSCA (Topology and Orchestration Specification for Cloud Application, 面向云应用的拓扑和编排规范) 或 Heat (开源云计算平台 OpenStack 的一个子项目, 用于进行资源和应用编排) 模板文件上传至 SDC, 而具体的虚拟网元镜像 image 直接上传到 VIM 平台中。ONAP 通过对接厂商自有管理器, 实现对虚拟网络功能的全生命周期管理。

2.3 产业联盟

2.3.1 中国人工智能产业发展联盟 (AIIA)

为加快推动我国人工智能产业发展, 搭建人工智能产业发展公共服务平台, 提升产业发展能力与应用水平, 按照《“互联网+”人工智能三年行动实施方案》有关部署, 在国家发展改革委、科学技术部、工业和信息化部、中央网信办四部委共同指导下, 2017年10月13日, 中国人工智能产业发展联盟 (AIIA) 在北京宣布成立, 中国工程院院士潘云鹤当选为第一届理事长, 中国工程院院士高文当选为专家委员会主任^[10]。

中国人工智能产业发展联盟以国家产业政策为导向, 以市场为驱动, 以企业为主体, 搭建产学研用合作平台, 促进联盟成员的研发、设计、生产、集成、服务等水平, 构建我国人工智能产业生态, 提升我国人工智能产业的竞争力, 强化人工智能与经济社会各领域深度融合, 促进技术进步、提高生产效率, 推动传统行业数字化转型, 支持新技术、新产业、新业态、新模式加快发展。

中国人工智能产业发展联盟的发起单位包括中国信息通信研究院、百度、阿里、腾讯、清华、浙大、科大讯飞、奇虎 360、中科院自动化所、沈阳新松、中兴通讯、上海仪电、中国电力科学研究院、航天科工集团、中国联通、电子四院、电子一所等积极推动人工智能产业发展的代表性企业、高校、科研院所、社会组织。

第三章 人工智能在通信网中的主要应用场景

3.1 网络规划

5G 网络规划的考虑因素相比 4G 有大幅增加，比如有 SA 和 NSA 两种组网方式，需要考虑低、中、高频及非授权频谱多种频段，存在 4G/5G 频谱共享、多阵列天线等多种站型。这一系列新变化给 5G 网络规划带来的复杂性呈指数级增长。

而基于 4G 现网积累的容量、覆盖等方面的历史数据，再结合 5G 新特性，借助 AI 技术进行关联分析、学习训练、智能推理，将有效指导 5G 无线网络规划。

3.1.1 智能化容量评估

对于 5G 高带宽业务场景，可以以 4G 现网的网管、MR 等数据为基础，结合工参、用户业务模型、套餐、网络性能负荷等信息，利用 AI 算法分析热点场景的特性，建立用户需求预测模型，为 5G 网络的容量规划及建设提供指导。

3.1.2 自动站址规划及覆盖效果评估

5G 刚开始建网时，可利用 AI 技术基于 4G 现网 MR 数据进行 5G 站址规划和覆盖评估。综合考虑 5G 覆盖强度及质量要求、室内外覆盖特性、4G/5G 频段与功率差异、现网工参等因素，利用 AI 最优化算法从现有 4G 站址中筛选出推荐列表，优先纳入 5G 站址规划。

基于 4G 现网的 MR 数据、频段、邻区关系、邻区信号、高精电子地图、遮挡物位置及高度等信息，利用 AI 算法对无线覆盖特性进行学习训练和建立模型，然后结合 5G 频段、环境等特性对 5G 网络的覆盖效果做出评估，并可根据实际覆盖效果对所建立的模型做进一步的迭代优化。

3.2 网络运维

随着电信网络云化和 ICT 融合日益加速，加之 5G、IoT、AI、大数据等新兴技术的引入，电信网络的规模和复杂度与日俱增，运营商在网络运维方面面临越来越大的压力：

（一）5G 网络相对于 4G，技术架构革命性变化，导致网络复杂性加大

虚拟化层、硬件层、应用层、管理层等完全解耦，且可能来自不同厂商，原来一个系统内部的问题外化成互联互通问题，增加了故障定位、根因追溯的难度。

云化业务要求支持迁移，软硬件一一对应关系变为软硬件多对多关系，故障线索链存在动态变化的可能。

网络服务化、切片化架构，使得系统日志，告警数量激增，定位根因故障难度加大。

(二) 新业务无历史运维经验：5G 需要支持万物互联，很多业务领域无成熟经验借鉴

而目前以人工为主的传统运维模式已不能适应网络的先进性，传统网络运维困境日益凸显。为解决这些问题，需要借助 AI 能力进行复杂信息处理，从而提升运维效率。依托机器学习/深度学习，可以在海量运维数据中抽取隐含的关联特征和规则，进而追溯事件根因、精准定位故障。同时，通过共性特征的提取，可以对未来事件进行预测，变被动响应式运维为主动式。

3.2.1 智能故障溯源

故障根因分析是网络运维工作的重要方面。现网中故障多种多样，单一故障告警已不能反映出准确的故障根源。这种情况下，运维人员也难以根据经验进行判断，快速准确定位根因。引入 AI 智能分析后，可以从多种多样告警信息中提取共性特征，快速收敛到共性故障点，从而提升处理效率，降低运维难度。

智能故障溯源基于大数据分析和 AI 特征挖掘等技术，依据网络及业务流程上下游关系，结合多个维度的历史数据进行分析，挖掘出依靠传统人工经验很难发现的潜在特征和规则，输出故障和特征匹配的规则库。实际网络中，系统可以根据故障特征自动进行规则匹配和诊断，并给出处理建议。另外，智能故障溯源还可以跟 BOSS 系统配合，精准定位故障并触发工单。如图 1 所示。

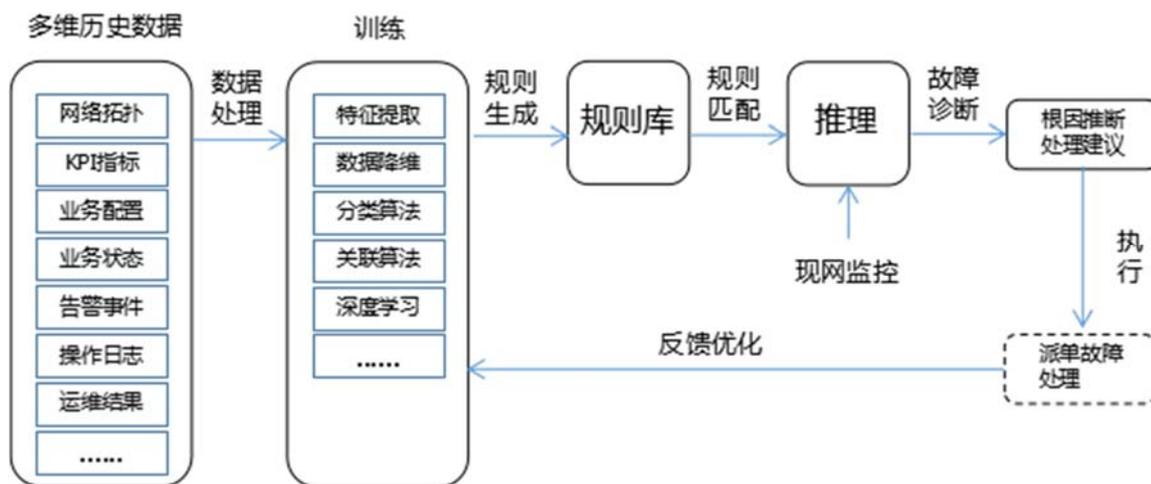


图 1 网络故障溯源方案示意图

智能故障溯源诊断，包括两大部分：

基于机器/深度学习生成诊断规则库

通过获取多维度历史数据，如 KPI、告警、性能、配置、日志、故障记录、工单等，基于这些数据，借助 AI 算法实现特征及规则挖掘。

诊断规则的运行

第一步，通过现网监控获取实时故障及告警；第二步，匹配规则库进行根因分析，智能诊断故障，关联工单系统触发派单；第三步，故障维护完成后，进行维护有效性的反向修正，强化现有规则库，实现自优化闭环。

智能故障溯源极大提升了告警处理及故障诊断效率，典型应用场景有：瞬断告警、告警频发、网元内故障关联分析、上下层业务故障关联分析、跨子网故障分析等等。

3.2.2 智能健康度预测

传统运维关注发现异常及根源，借助运维工具进行异常处理，恢复网络正常运行，保障网络服务体验。在实际网络中，异常发生后通常网络质量和用户体验会受到不同程度影响，网络恢复如果需要调配设备，则时间耗费更长。另外，为保证故障及时恢复，设备资源冗余、人员技能提升等方面也是不小的成本。

网络运维如果从事后被动处理转为事前主动预测/预防，运维成本和用户体验将大幅提升。

智能健康度预测基于 AI 的机器学习/深度学习技术，分析海量运维数据，对故障发生前某一段时间内的网络特征等进行抽取，从而有效进行网络监控、预测/预防网络故障发生。

如图 2 所示，预测系统通过线下历史数据训练，得到预测模型；线上部署系统后，定期采集数据，检测健康度，并用训练好的预测模型进行故障预测。有未预测到的故障发生，则以该故障时间点前的某一阶段数据作为输入，重新训练预测模型，从而实现不断迭代优化。

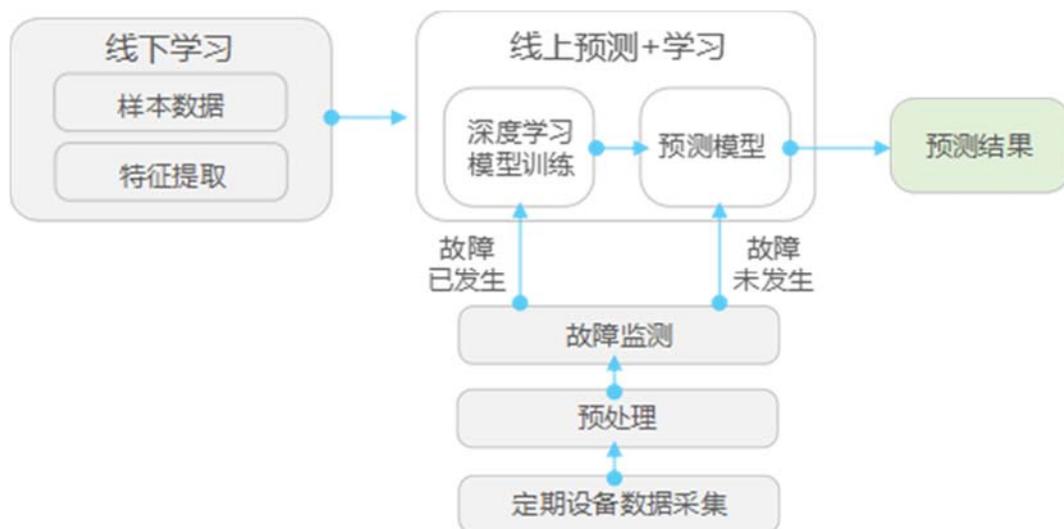


图 2 网络故障预测方案示意图

智能健康度预测有几类典型应用场景：指标趋势及指标异常预测，设备关键模块故障预测等。详细介绍如下：

指标趋势及指标异常预测

线下通过历史数据分析，挖掘网络健康度的指标趋势与网络数据的深层次关联特征。线上通过对网络健康度及关联指标的实时监测，利用训练好的的特征规则进行匹配，预测网络健康度指标趋势。

根据预测出的中长期网络健康走势，可估测网络未来落入健康阈值以下的可能时间点，提前进行资源预留及相应处理。因此，指标趋势预测可弥补网络规划的不足，指导网络扩容、备品备件储备等。

另外，系统还能够预测中短周期内下一步的指标状态，根据指标预测值是否超过阈值来预判故障会否发生，实现事前故障预防。

关键模块故障预测

业务处理芯片、光模块等关键模块的稳定性对电信设备非常重要，一个关键模块故障就会导致大量业务瘫痪。

关键模块发生故障前，通常会存在一些特征性的异常，而特征性异常可以通过特定的指标来表征。比如：芯片故障通常跟寄存器异常有关，光模块故障通常跟光功率异常有关。

通过机器学习/深度学习，挖掘关键模块故障与特征性异常之间的深度关联关系，从而实现关键模块故障预测，提前进行预防。

3.2.3 智能工单管理

工单管理是运维的核心，工单将运维人员和组织有效串联起来，形成任务驱动的运维体系。而运维人力合理调度就成为工单调度的重点。

智能工单管理系统基于电子化数据、准确的位置定位、丰富的交互以及接口能力，实现运维全过程的可视化、可管可控、可分析，从而做到主动运维。

智能运维调度，主要包括两个部分：

基于 AI 的故障工单预警

从历史工单中提取与故障发生有关的特征向量，利用 AI 预测算法，如多项式拟合、神经网络等，生成故障预警模型。通过该预警模型，智能工单管理系统根据当前特征可预判故障发生率。

实时智能调度

根据工单类型、工单技能要求、人员资质、工单优先级、站点位置、运维车辆等信息，采用遗传算法自动规划车辆行驶最佳线路，高效安排外线运维。

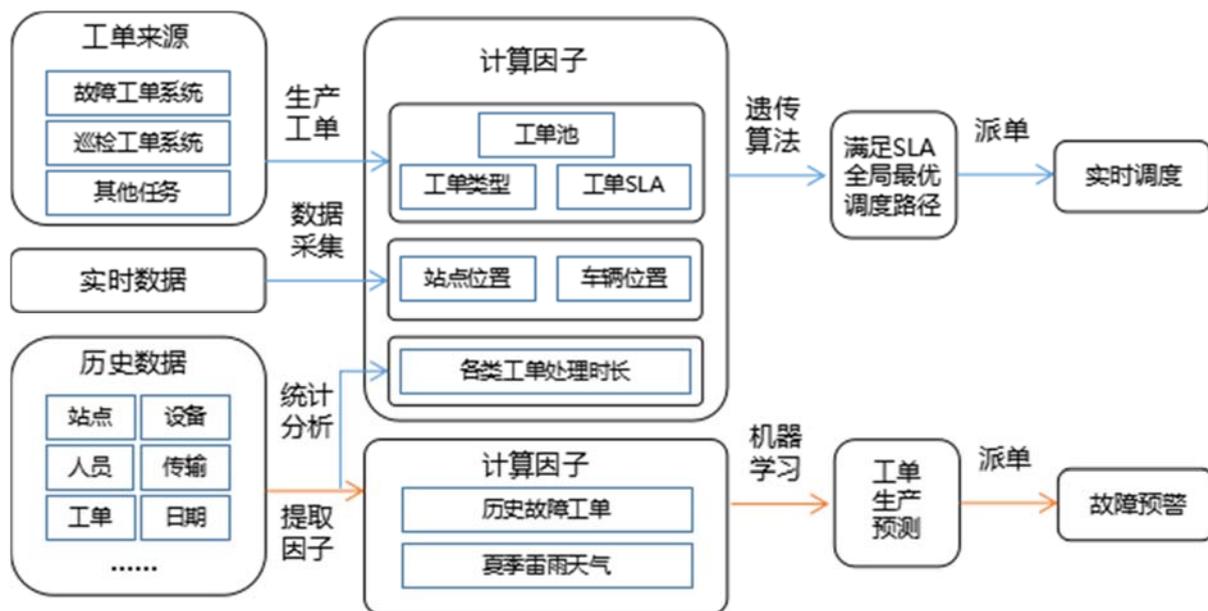


图3 智能工单管理方案示意图

3.2.4 智能 DevOps

DevOps 结合 AI 能力之后，可以实现业务开发端的价值最大化，业务运维端的自动化，业务使用端的体验最优。

业务在线设计阶段，通过分析软件仓库的调用行为数据、网络实例运行数据等，获得应用的价值分布。结合流量趋势，利用机器学习可以预测未来市场热点。基于该预测结果，再叠加 DevOps 能够快速进行业务设计及部署的优势，可以快速开发、上线符合市场热点需求的应用，使得业务开发端价值最大化。

业务上线运营阶段，利用 AI 驱动的 RCA 技术可快速进行根因定位，再结合 DevOps 策略平台及自动弹缩、自愈等技术，可实现基于策略的运维自动化闭环。

通过对业务实例和资源动态数据的学习分析，可预测与业务需求最匹配的资源分配，再结合 DevOps 的自动编排功能，能够做到最优选路、用户面资源适时下沉等等，从而实现用户体验最优。

3.3 网络优化

网络优化是运营商投入大量人力的重点工作，也是提升网络 KPI、保障用户体验的重要一环。网络优化涉及面很广，既包括无线网络优化，也包括网络各层级的 KPI 优化、网络策略优化、传输路由优化等。

3.4.1 无线覆盖智能优化

无线覆盖质量一直是屡遭用户投诉的原因之一，弱覆盖、过覆盖等问题直接影响用户体验。对于覆盖问题，传统上主要依赖路测、呼叫跟踪、投诉处理、人工经验等手段来解决，往往投入成本高、处理周期长、优化效果不理想。

基于 AI 的无线覆盖智能优化，利用 AI 及大数据技术，根据历史上的覆盖数据、关联数据进行学习、训练，生成优化控制模型，可以自动输出参数规划及调优建议，从而实现无线覆盖智能优化。



图 4 无线覆盖智能优化方案示意图

如图 4 所示，整个流程主要包含如下步骤：

- 网络综合数据采集：提供现网 KPI、工程参数、性能/配置管理、MR 报告、呼叫跟踪、DPI 等多个维度的历史数据给大数据平台。
- 数据处理：大数据平台对多个维度的原始数据进行数据分析和预处理。
- AI 模型训练和输出：AI 平台利用大数据系统预处理后的数据进行模型训练，模型训练好后可以进行预测、推理，给出无线覆盖参数优化建议。
- 参数优化实施：根据大数据和 AI 平台提供的优化建议实施网络优化调整。调整方式可以是自动下发执行优化参数，也可以是人工执行的物理调优。

3.4.2 智能流量预测及路径调优

随着网络规模的增长，IoT、网络切片等新兴技术的引入，网络复杂度越来越高。未来网络要满足高带宽、低时延、按需而动的业务需求，在流量调度和路径规划方面使用传统方式已很难实时给出最优方案。基于 AI，可以利用历史数据进行数据分析和模型训练，结合实际网络需求进行流量预测和路径优化推理。

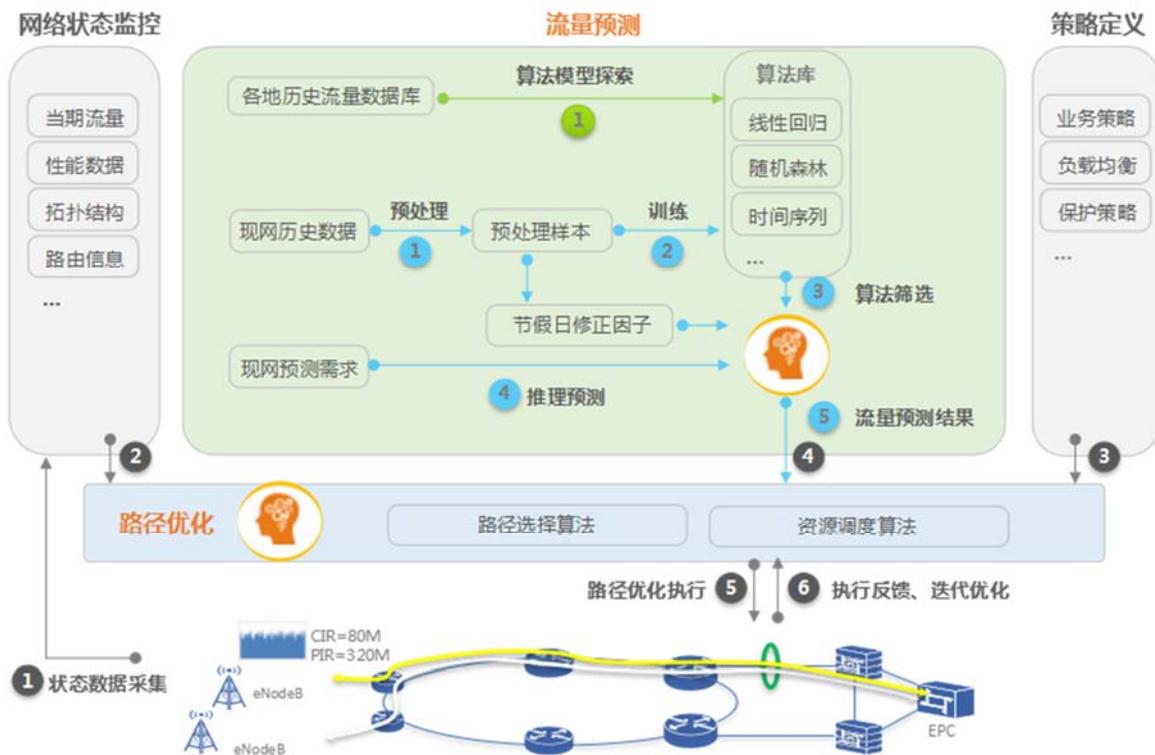


图5 智能流量预测及路径调优方案示意图

智能流量预测及路径调优方案如图5所示。

流量预测：可适用于中长期流量增长预测、短期峰值评估、区域仿真等，进行扩容支撑、规划支撑、主动运维。

流量预测能力建立需要两个阶段工作：

(1) 算法探索阶段：利用各地网络历史数据，对多种算法模型进行训练，训练效果好的算法模型纳入算法库。

(2) 现网应用阶段：将现网历史数据按地域、业务分组等做预处理，获得训练数据；然后对算法库中的多种预测模型作进一步训练，得到现网各地区各业务的最佳预测模型。

最后，根据现网预测需求，选择最佳模型进行推理预测，获得最终的流量预测结果。

路径优化：根据网络状况，动态优化路由策略，提高传输效率，同时动态调度资源，提高网络使用效率。

路径优化能力建立需要对流量、性能、拓扑、路由等网络状态数据进行采集和实时监控，结合业务、负载、冗余等策略规则，以及上文提到的流量预测结果，共同作为路径优化算法模型的输入，将模型推理结果用于现网执行。并且，对执行效果进行反馈，从而不断迭代优化算法模型。

第四章 人工智能在运营商业务创新中的应用

4.1 边缘智能

MEC（多接入边缘计算）作为 5G 关键技术之一，在靠近用户侧的位置上提供 IT 服务环境和云计算能力，能更好地支持 5G 网络低时延、高带宽的业务。随着垂直行业应用对边缘侧网络服务的需求越来越多样化、定制化，MEC 与 AI、大数据等智能技术结合，将在边缘业务场景智能化、无线网络能力开放等方面发挥赋能作用。

4.1.1 无线感知服务

应用一：无线上下文环境感知

MEC 从多个 4G/5G 站点获取无线上下文信息 RNIS，构建无线上下文环境，并统计用户终端特性、业务特性等数据，利用 AI 分析挖掘数据、业务、无线环境之间的内在关联，建立用户特征模型库。MEC 根据实时采集到的测量数据与特征库进行匹配，能够更为准确地预测用户环境与业务特性变化趋势，更好地支撑准实时的用户算法策略和参数配置优化，提供更好的业务体验。

应用二：频谱感知

边缘计算节点根据不同区域无线系统长时间的频谱测量结果，结合对各无线系统在不同区域的无线环境特征、不同时间段内的用户行为特征、以及不同用户的业务特征等的统计数据，利用 AI 和大数据技术进行分析和建模，进行准实时的频谱共享、异系统干扰协同等应用。



图 6 异系统频谱资源共享示意图

系统快速频谱共享

多种制式混合组网时，不同制式小区在同一时刻的忙、闲程度可能存在差异，利用 AI 算法可以在多制式（即异系统）间，分时共享频谱资源。BSC/RNC/eNB/gNB 可以与

部署在 MEC 中的 iCS 智能协调服务器建立连接，通过 MEC 内置的 AI 能力，结合历史统计数据和当前频谱资源使用情况进行分析决策，动态调整相关门限，实现异系统的频谱资源共享。如图 6 所示。

异系统干扰协同

对于无线系统之间的同频干扰和邻道干扰，边缘计算节点可以获取不同系统在不同位置和时间段的射频信号、业务负荷、用户行为等特征信息，并进行统计分析，在此基础上建立网络特征模型。边缘计算节点根据实时的无线频谱测量特征，优化频谱资源规划与调度，降低系统间干扰。

对于无线系统的互调干扰，边缘计算节点为不同系统在不同频段建立频谱特征库，并根据实时收集的无线频谱测量特征来识别造成干扰的无线系统及基站位置，从而解决互调干扰问题。

应用三：业务感知

MEC 结合 AI 及大数据能力，深度挖掘数据、业务、无线上下文环境之间的内在规则，为不同无线环境下的业务感知提供更为精确的业务特征识别，提供更为精准的差异化服务策略。另外，还可以在边缘节点部署高算力的硬件加速器资源（如 FPGA、GPU 等），为业务提供高实时、高性能的 AI 解析能力，更好支撑 5G 时代业务感知。

应用四：用户感知

MEC 通过统计分析用户终端的协议、性能、业务等特性，建立不同用户的特征库，然后根据获取的实时用户测量数据与用户特征库进行比对匹配，从而可以更精确地预测用户业务变化趋势、用户行为对网络负荷的影响，准实时优化用户算法策略及参数配置。

4.1.2 应用使能服务

应用一：定位

无线通信系统中常用的定位方式包括 C-ID, E-CID, OTDOA, UTDOA 和指纹定位。自动化的指纹采集和更新方案是指纹定位实现网络级商用的关键技术。LTE 系统中，借助终端的 MR 上报和 AGPS 测量能力，已经能够完成室外环境指纹库的建立和更新。但是室内场景环境下指纹库的建立依然需要依靠人工。

随着 5G 网络建设的展开和 IoT 终端的规模使用，物联网终端可以作为各类无线通讯系统的射频信号传感器被广泛的部署。通过这些位置已知的物联网终端测量的各无线通信系统信号特征，借助 AI 和大数据对海量终端的信息收集和分析，能够对指纹库的收集和完善起到重要的作用。借助边缘计算节点的实时计算能力，可以对各类终端位置

进行计算和定位。此外，室内的微站之间也可以进行协同，根据 AI 建立指纹模型，也可以针对进行定位。

借助于定位的基础能力增强，上层业务提供商可以提供更多的业务，如室内定位导航，室外室内相结合的定位导航等。

应用二：TCP 优化

目前应用无法感知网络状态，两者信息没有有效对接，导致用户获取到的无线资源没有被充分用于提升视频下载/播放速度。

MEC 基于 RNIS 信息，获取当前基站的资源状况以及视频用户端的带宽信息，同时基于 TCP 透明代理及 DPI 分析，借助 AI 视频识别对业务流进行分析（包括终端、空口、业务类型，请求服务器等端到端信息）。将用户的业务从终端状态、无线资源、应用协议等方面进行大数据分析，通过不停的反复学习训练，从而形成相应的业务行为与无线网络环境匹配的智能化模型。将该模型应用到业务中，辅助应用端进行 TCP 窗口调整及优化。使用快启动，减少用户等待播放的起始时延，并对于重传报文优先发送，减少报文不当丢弃，如图 7 所示。

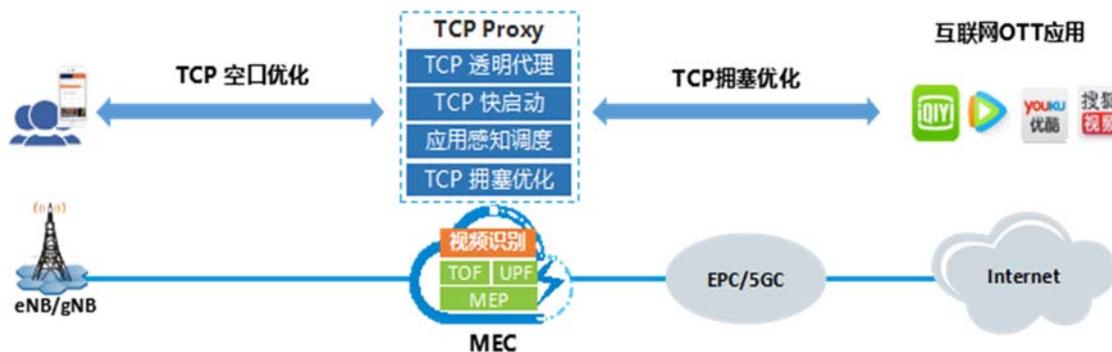


图 7 TCP 跨层优化示意图

应用三：本地缓存

传统 CDN（内容分发网络）目前主要覆盖固网场景，且位置较高。MEC 可进一步使 CDN 更加贴近终端用户。当地市机房中同时部署 MEC 时，可通过 MEC 将内容引入网内，访问 CDN 服务器的用户请求和响应都不必再迁回到省干网络，而可以直接从本地获取数据，大大降低了业务访问时延。为了提高移动边缘缓存命中率，切实减少用户访问时延，可以基于 MEC 使用 AI 对用户的业务流进行预测和分析来确定预先存储的内容。

4.2 物联网

工业自动化的主旨就是工业能够进行自动生产，通过人力的极少参与，实现生产过

程中的各种自动化的过程控制。5G 网络可以为工业自动化提供针对智能设备的宽带数据、针对自动化过程的超高可靠低时延通信以及提供针对大量连网设备的机器类通信，5G 是工业自动化的基础设施。

人工智能将会进一步提升工业自动化的智能水平，比如基于计算机视觉技术的智能化产线物料纠错、基于运动智能的智能 AGV、高实时性的智能机器臂等。

4.3 创新业务

万物互联时代，增强的移动互联网业务和物联网业务将成为未来移动通信发展的主要驱动力。围绕 5G 三大应用方向：移动宽带增强(eMBB)、大规模物联网(mMTC)、超高可靠超低时延通信(uRLLC)，在包括 VR/AR、M2M、智慧城市、智慧家庭、自动驾驶、工业自动化等在内的典型应用场景下，5G 将满足人们在居住、工作、休闲、交通等方方面面多样化的业务需求，提供超高流量密度、超高连接数密度、超低延迟的极致业务体验。

同时，连接类型的泛化、业务的多样性、数据的海量增长，也为人工智能在 5G 垂直方向的应用提供了天然土壤。人工智能和 5G 的垂直应用，在未来的发展是密不可分的。

4.3.1 自动驾驶

未来车联网的普及将彻底改变人类的出行方式和城市交通管理方式。车联网是一张由车辆的各种动态信息，如传感、定位、导航、无线通信等构成的巨大的 V2X 交互网络。一方面，依托 5G 网络的高吞吐、低时延能力，所有自动行驶的车辆随时与云调度平台保持高速互通；另一方面，依托灵活的车辆间数据交换技术，比如 D2D、车与车之间实现实时的互动反应。

自动驾驶本身要求感知、决策与控制三大系统缺一不可，而环境识别能力是感知系统智能水平的关键所在，这也是基于深度学习的人工智能技术在自动驾驶领域的关键应用。决策与控制一方面依赖于 5G 网络的低时延交互技术，如 D2D、边缘计算等，另一方面依托于车辆内嵌的强的人工智能推理加速能力；同时，云平台的统一调度、实时策略等也需要依赖人工智能的跨域分析和推理判决能力。

4.3.2 VR/AR

5G 对于 AR/VR 的优势主要体现在三个方面：更大的容量、更低的延迟以及更好的网络均一性，这些指标的提升能够极大满足大众对于 VR/AR 极致体验的需求，VR/AR

的应用会得到进一步的发展。

VR/AR 整个产业链涉及较多的模块或功能组件，比如：显示屏、摄像头、传感器、体感设备、语音识别、定位、硬件、平台、内容制作等，其中大部分都可以利用人工智能技术提升用户体验和自身价值，比如语音识别、摄像头图像识别、内容推送等。

4.3.3 智慧家庭

人工智能，除了将给网络的运维运营模式带来根本性的变革，在家庭终端方面也会带来全新交互模式，完全颠覆智能设备跟人的互动体验，为日常生活带来极大的便利性。比如：智能终端会比以往更智能，数字虚拟助手会更懂用户，基于人工智能的生物特征识别手段让用户识别更便利更安全。

以用户体验出发，以家庭智能终端设备为切入点，结合人工智能技术，为用户打造全新的智慧家庭体验，包括智控机顶盒、智能摄像头等。

智控机顶盒

机顶盒已经成了当下家庭数字化娱乐的基本配置，丰富的节目内容带来了多样的娱乐体验。人机交互方式正由传统的按键式，向按键结合语音交互转变。基于人工智能的语音识别和语义理解，使得交互更自然、更有效。

通过手机中的机顶盒助手，其中包含了语音识别和语义分析的人工智能模块，通过机顶盒助手上的语音按键完成语音采集。内置的人工智能识别模块对语音进行识别后，将识别后的语音进行语义分析，并将分析结果发送到机顶盒端，机顶盒展示语音和语义识别结果，并执行相应的命令。可实现多种语音交互功能，包括天气查询、电视选台、电视剧选集等。

智能摄像头

联网摄像头如今已越来越多的走进日常生活，成为家庭安防的重要组成部分。尤其是结合了人工智能技术的智能摄像头，应用人脸识别、物体识别、语音识别等技术，让摄像头更加智能，使用更加方便。结合手机 APP 的远程访问，更可随时随地，实时查阅、监控。

第五章 面向 5G 的人工智能应用

5.1 5G 网络编排

运营及编排层主要包括：业务和资源的设计、调度及管理，如全局业务编排（编排、激活、生命周期管理等），全局资源编排（NFV 资源编排、SDN 资源编排等），以及运营支撑相关的组件，如外部门户、工单系统、客户服务系统等，同时也包括专用的数据集中管理平台，如大数据系统，负责数据统一采集、统一存储、数据智能挖掘分析等。

对于该层面的人工智能使能策略，如果运营商网络先期已经部署了成熟商用的大数据平台，可优先在大数据系统引入 AI 引擎，对 OSS 侧和 BSS 侧数据做更深度、智能化挖掘，从而指导运维和运营；如果还未部署大数据系统，可以优先考虑部署融合的大数据和人工智能平台产品，实现设备层面、网络层面、业务层面、用户终端层面、运营层面及异厂家跨制式的全方位数据感知和分析，从而提升运营智能化。

同时，随着 5G 及虚拟化网络的建设，编排系统可逐步叠加 AI 能力，提升产品编排、业务编排、端到端资源编排的自动化、智能化，通过动态的、智能的策略来实现网络的随需而动，同时对业务量的变化做前瞻性的智能预测，并动态规划、管理配套资源。

另外，对于部分运营支撑系统，可按需引入 AI 基础技术和通用算法，提升运营效率，如智能工单及调度系统、智能语音客服系统等。

5.2 5G 网络节能

5.2.1 基于人工智能的网络节能

网络能耗是运营商的重要 OPEX 支出之一，提高能源效率亦是未来网络运营的关键要求。通过网络流量数据、资源利用率数据统计关联分析，网络忙闲状态与人们的工作和休息节奏具有高度的相关性。尤其是学校、CDB、商业区、大型居民区等热点区域具有显著的潮汐效应。网络容量目前都是按照网络业务的峰值要求建设的，包括无线基站以及集中部署了未来 NFV 网络基础设施的数据中心，如何基于网络业务量对基站和服务器进行休眠和唤醒操作，在满足实时业务需求的同时实现绿色节能的目标，需要进一步研究。

基于 AI 的网络智能节能技术，充分考虑了网络覆盖、用户分布，业务特征，辅以负荷预测，精准预测和选择热点小区，适时选择休眠和唤醒设备时机，充分提供可靠和无缝的用户体验，使得区域网络资源利用率更高，用户体验保障更好，区域内网络能耗有效降低。

5.3 5G 基站调优

人们对高速移动数据的渴求是无止境的。可是在城市环境中可用 RF 频谱已经饱和，为了满足业务需求的急速增长，4G 和 5G 时代增加天线的数目不可避免。Massive MIMO (大规模天线技术)是 5G 时代的核心技术之一。

为了充分发挥 Massive MIMO 的优势，需要对 Massive MIMO 的基站的各项参数进行灵活的适应性调整，应因对业务的变化。面对如此多的参数组合和快速的业务变化，人工调整参数将是噩梦，迫切需要更加智能的可以自我优化的机制。引入 AI，可以自动识别 MM 站点的场景，比如体育馆、CBD、车站，并且能够这些场景对话务进行预测，以及用户分布和业务等，获取当前最优参数值配置，比如通过天线权值调整方向，倾角等，通过 MM 天线权值的调整，实现 Massive MIMO 的自优化配置，改变 MM 小区的覆盖分布，及时适应用户分布和业务类型，有效提升资源利用率和用户体验。

第六章 网络人工智能平台

网络人工智能平台是提供 AI 即服务（AiaaS: AI as a Service）的云化平台，是构建智能化是应用所需要的核心组件。

6.1 网络人工智能平台系统架构

常见的网络人工智能平台参考架构如图 8 所示，包括基础设施层、存储计算层、服务支撑层、应用层、及系统管理等多个层次。



图 8 网络人工智能平台参考架构

(1) 基础设施层

支持 CPU/GPU/FPGA 等异构的基础设施，支持裸机、基于虚拟化的虚拟机、容器化等多种部署方式，可以帮助运营商根据自身需求，灵活选择部署方式。

(2) 存储计算层

支持多种数据采集方式，提供完善的大数据 ETL 解决方案；

支持 HDFS 分布式文件系统，分布式数据库 RDBMS，集群资源管理；

支持批处理、流处理、交互 SQL、OLAP、分析挖掘等；

支持 Hadoop, Spark，支持 Spark ML 和 MLlib 以及 deeplearning4j, Caffe on Spark, TensorFlow on Spark 等深度学习框架；

支持主流的 TensorFlow、Caffe 等深度学习开源框架，并对开源框架做了大量自研的优化增强。

(3) 服务支撑层

支持数据驱动的数据治理、实时数据生产和质量监控、完善的元数据管理、统一高

效管理数据、算法和任务；

支持基于 OpenAPI 的数据开放、多租户管理；

支持一体化可视化建模工具，支持模型探索、开发、部署，集成种类丰富的算法库，支持硬件加速；

丰富的组件能力，提供丰富的基础技术组件，实现多应用方向的支持。

(4) 应用层

基于丰富的技术组件，支持面向各行业及垂直领域的丰富的应用集成。比如，以电信行业为例，支持智能网络运维、智能网络优化、智慧运营、智能客服、智慧家庭等具体应用。

(5) 系统管理与非功能属性

提供面向人工智能平台的系统安全管理和维护管理。

6.2 中国联通网络人工智能平台智立方（CUBE-AI）

中国联通于 2019 年 6 月 27 日在上海 MWC 展会期间发布了网络人工智能平台智立方（CUBE-AI），这是中国联通参考 Linux AI 基金会（LFAI）开源项目 Acumos 的设计理念，自主研发并开放源代码的网络人工智能平台，提供 AI 模型开发、模型共享和能力开放等网络 AI 即服务（NAIaaS: Network AI as a Service）的云化平台。

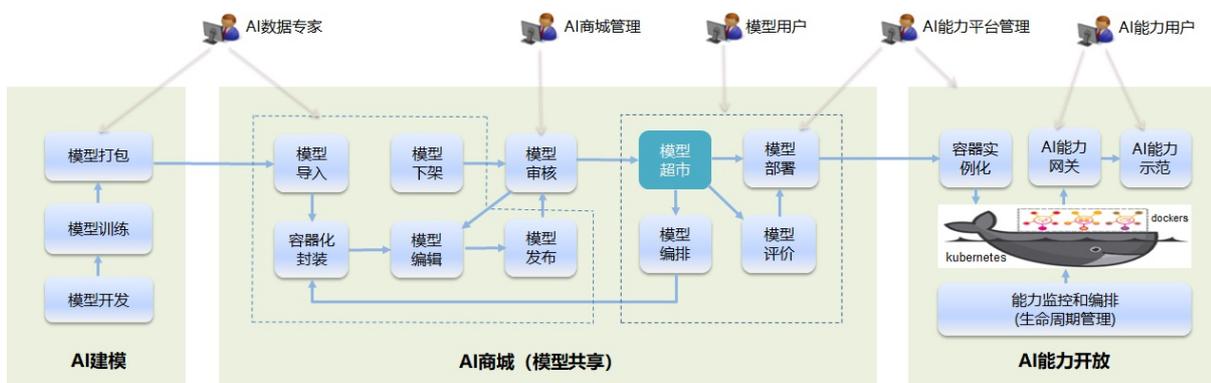


图 9 智立方平台功能架构（截至 2019 年 6 月 27 日）

如图 9 所示，截至 2019 年 6 月，智立方平台提供 AI 模型打包、模型导入、容器化封装、模型编排、模型发布、模型搜索、模型部署、AI 能力开放、能力编排、能力演示等功能，支持 AI 模型的 docker 容器化封装和微服务化部署。基于智立方平台，能够在网络 AI 模型开发者和模型的实际使用者之间架设一条互通的桥梁。开发者无需关心具体的部署环境，最终用户无需了解 AI 算法的具体实现细节，使得开发者和使用者能够专注于各自最擅长的领域进行创新，从而加速 AI 创新和应用进程，促进 AI 算法从设

计、开发直到部署、应用整个生命周期的快速迭代和演进。

2019年6月27日，智立方平台的全部源代码已在 github 上开源共享。

运营商网络智能化演进离不开产业生态的共同努力，基于智立方平台，中国联通发起了网络 AI 合作伙伴计划，首批共有中兴、华为、爱立信、诺基亚等 22 家合作伙伴签约加盟，共同基于智立方平台建设网络智能化应用，切实推进人工智能在运营商网络落地应用、赋能业务创新。

6.3 基于人工智能网络平台的应用实现流程

人工智能端到端工作流程，总体分为三大部分：数据处理阶段、AI 训练阶段和应用推理阶段，如图 10 所示。中国联通人工智能方案，基于大数据和人工智能融合平台，支持人工智能的端到端工作流程。

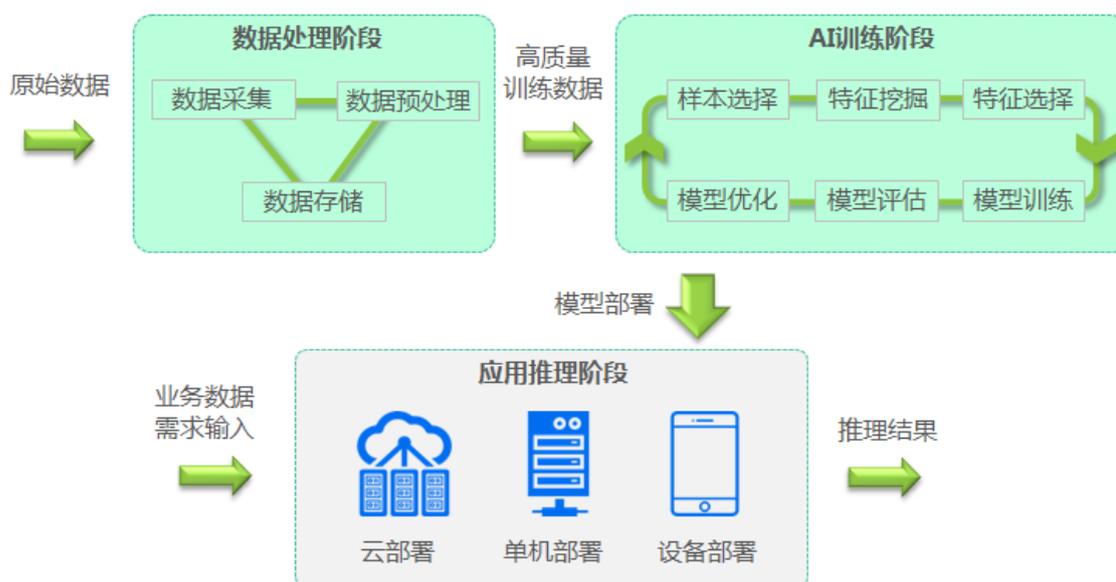


图 10 人工智能端到端实现流程

(1) 数据处理阶段

通过数据采集工具进行数据采集，并对采集到的原始数据如文本、图像或者应用数据进行清洗和预处理，清理异常数据，对数据进行标准化，最终生成高质量的数据供后续模型训练使用。

(2) AI 模型训练阶段

使用分布式训练引擎，可以高效完成超大规模的 AI 模型训练。引擎根据模型计算图特征、异构计算设备网络拓扑结构、参考历史性能数据，自动规划最优分布式计算方案，实现高效的自动化分布式训练。

(3) 应用推理阶段

支持在云侧、单机和端侧内嵌等多种部署方案，满足客户多场景需求。比如：针对中心 DC，提供 GPU 加速方案，提供强大的处理能力；针对边缘 DC，提供 FPGA 加速方案，降低能耗，降低运营成本；针对端侧，实现模型在嵌入式设备上运行等。

支持通信网络不同应用场景的推理平台部署，比如：支持大数据推理平台，实现智能运维和运营支撑；支持网管单机推理平台，实现准实时推理；支持网元内嵌的实时推理等。

总结

中国联通作为网络强国和数字中国建设的主体之一，积极落实国家重大战略部署，积极推进人工智能、大数据、云计算等前沿技术与网络生产运营相结合，通过智慧运营实现网络价值增值的同时，赋能 5G、IoT、MEC 与 AI 融合业务和商业模式创新。

伴随着 5G、IoT 开启万物智联的时代，网络人工智能的应用程度将在某种程度上决定一个运营商的网络运营能力，因此，推动网络人工智能与网络运营的深度融合，是在 5G 时代构建差异化优势的主要技术路径之一。

2018 年以来，中国联通以 CUBE-Net2.0+ 为总体目标，设计了面向互联网化运营转型支撑的开放网络服务体系（ONS：Open Network Service），实现智能化、自动化的网络运营支撑系统，推动网络运营效率提升的同时，也助力业务质量、客户感知与市场效益的提升。

同时，我们也要指出，由于通信网络的复杂性和长期以来形成的封闭性，网络人工智能的普遍应用还存在诸多问题，譬如：

一、网络 AI 的应用显效是个较为长期的过程。正如机器学习需要经历学习的过程，AI 融入电信网络进行赋能，也需要一个渐进、迭代的过程。目前 AI 在电信行业的战略地位尚不明确，需要加强 AI 顶层设计，在业界形成网络 AI 化的共识，才能充分发挥 AI 使能作用。

二、网络 AI 人才缺乏，市场缺乏通用产品。目前缺乏相应人才帮助运营商将 AI 技术引入电信网络。单纯的 AI 技术公司不能很好理解运营商的网络需求；而运营商自身对 AI 技术又缺乏足够了解，所以如何构建合理人才机制非常关键。

三、网络 AI 应用场景众多，算法迁移相对困难。网络是否是一个适合 AI 解决的商业问题还有待细分领域的继续探索实践。目前网络 AI 应用重点还是在效率提升和成本降低方面，如智能运维、智能节能等。

四、高质量网络数据采集投入较大，限制网络 AI 应用场景。数据质量决定着玩过 AI 应用的效果，保障网络数据采集的实时性、完整性、高质量都需要较大的持续投入，网络数据处理对算力要求较高，某种程度上限制了网络 AI 应用的普及。

缩略语

3GPP	3rd Generation Partnership Project	第 3 代合作伙伴计划
5G	5th Generation	第 5 代移动通信系统
AI	Artificial Intelligence	人工智能
AIIA	Artificial Intelligence Industry Alliance	中国人工智能产业发展联盟
AR	Augmented Reality	增强现实
CCSA	China Communications Standards Association	中国通信标准化协会
CPU	Central Processing Unit	中央处理器
ENI	Experiential Networked Intelligence	体验网络人工智能
ETSI	European Telecommunications Standards Institute	欧洲电信标准协会
FG-ML5G:	Focus Group on Machine Learning for Future Networks including 5G	包含 5G 的未来网络中的机器学习焦点组
GPU	Graphics Processing Unit	图形处理器
ITU-T	ITU Telecommunication Standardization Sector	国际电信联盟-电信标准化部门
KPI	Key Performance Indicators	关键性能指标
MIMO	Multi-input Multi-output	多入多出
NFV	Network Functions Virtualization	网络功能虚拟化
NWDA	NetWork Data Analytics	网络数据分析
O-RAN	Open Radio Access Network	开放无线接入网
QoS	Quality of Service	服务质量
RAN	Radio Access Network	无线接入网
SDN	Software-defined networking	软件定义网络
SON	Self Organizing Networks	自组织网络
TCP	Transmission Control Protocol	传输控制协议
UE	User Equipment	用户设备
V2X	Vehicle-to-everything	车辆到 x
VR	Virtual Reality	虚拟现实

参考文献

- [1] Wikipedia. Artificial Intelligencehttps://en.wikipedia.org/wiki/Artificial_intelligence
- [2] 网络人工智能应用白皮书. SDN/NFV产业联盟,2018.
- [3] Yao H J C Q. Developing Networks using Artificial Intelligence . Springer,2019.
- [4] Etsi. 网络人工智能需求标准（ENI: Experiential Networked Intelligence）. 2019.
- [5] Itu-T. FG-ML5G: Focus Group on Machine Learning for Future Networks including 5G
<https://www.itu.int/en/ITU-T/focusgroups/ml5g/Pages/default.aspx>
- [6] 中国通信标准化协会(Zhong Guotongxinbiaozhunhuaxiehui). 关于印发中国通信标准化协会2018年第三批研究课题项目计划的函（通标发[2018]272号）. 2018.
- [7] O-Ran. <https://www.o-ran.org/>
- [8] . Linux Foundation<https://www.linuxfoundation.org/>
- [9] . LF AI Foundation<https://lfai.foundation/>
- [10] 中国人工智能产业发展联盟. 2019.

致 谢

本白皮书在撰写过程中参阅了 GSMA、ITU、ETSI、CCSA 等国内外标准组织、Linux 人工智能基金会、ONF 等开源组织以及北京邮电大学、中国信息通信研究院等科研院所和 AIIA 等行业协会发表的研究成果及行业报告，在此表示感谢。

本白皮书的部分案例源于中兴、华为、诺基亚、爱立信等行业合作伙伴与中国联通联合开展的合作，对上述合作伙伴的支持表示感谢！

本白皮书由中国联通和中兴通讯股份有限公司共同编撰，感谢中兴通讯的汤楠、张嗣宏、刘长鹏、张健等专家对本书编写的贡献。