



中国移动
China Mobile

6G为什么需要内生AI？

中国移动未来移动通信技术研究院

2021.08

- 内生AI的业务需求分析
- 内生AI的技术挑战分析
- 内生AI的基本特征分析

- “5G AI for network” 促进了移动通信网络的智能化发展；“打补丁”，“外挂”
 - 无法实现高水平网络自治，无法实现预验证、在线评估和全自动闭环优化
 - 模型训练需要大量的数据，集中采集数据困难，传输开销大
 - AI模型的应用效果缺乏有效的验证和保障手段，模型的训练和推理解耦，验证事后进行，对现网影响大
 - 模型迭代周期较长，模型训练开销较大、收敛慢、模型泛化性差
 - 对于AI用例研发采用打补丁，烟囱等方式进行，缺乏统一的框架

- “5G AI for network” 主要用于网络优化，更多的领域需要拓展：
 - 行业用户：助力千行百业的数智化转型
 - 需求越来越多样化：性能差异大（实时性），智能化服务质量差异性大
 - 数据隐私要求高：数据不出园区
 - 终端用户
 - 能力越来越强，如NPU...
 - 数据量大，数据隐私要求高
- “6G Network for AI” 新商业模式：新场景，新能力
 - 行业用户 + 终端用户：化学反应？

■差异化的智能化服务质量需求

- 无线网络波动大：信道时变，用户和业务需求动态性高
- 缺少对AI服务的系统性质量评估系统

■有限的通信和计算资源

- 边缘设备上的算力不足，存储能力弱，无线接入节点与边缘设备之间的带宽有限
- 缺少对AI服务的系统性资源调配系统：算力，数据，通信

■拟真的AI训练与验证环境

- 实现拟真的环境对同步数据的实时性和数据量要求大
- 拟真环境的及时性要求高，如何加速场景遍历，寻找最优模型？

■基于QoAIS的AI全生命周期服务编排

- 构建AI/ML服务质量评估和保障体系
- 全生命周期编排AI/ML 算法模型，算力，数据，通信

问题：如何解决模型收敛慢，泛化性差，训练样本少的问题？

- 迁移学习，元学习：模型迁移，重用或者部分重用模型，小样本训练
- 引入在线/离线学习结合，自动更新和迭代模型

■内生AI计算与通信的深度融合

- 数据、算力、带宽和时延受限，内生AI需要考虑计算和通信资源的编排
- 重构网络架构、协议和功能，适应空口传输，优化内生AI/ML性能

问题：引入联邦学习/多智能体/分布式学习，空口模型训练频繁交互大量中间计算结果，如何降低无线资源开销？

问题：引入联邦学习/多智能体/分布式学习，如何加快模型收敛？

问题：空口分布式模型推理需保证推理时延性能，如何进行模型分割和算力调度？

■ 内生AI与数字孪生的融合

- 数字孪生网络对AI/ML workflow或AI模型的效果进行预验证和优化
- 内生AI提供数字孪生网络所需的AI/ML模型

问题：如何避免AI功能的启用对真实物理网络带来的负面影响和模型收敛及时性要求高？
孪生体中引入分布式强化学习，与虚拟环境交互，加速遍历和寻优模型，实现AI模型预验证

问题：物理网络与数字孪生体需要大量同步数据，如何降低数据同步开销？
利用数据自生成技术（自编码，对抗生成）学习数据分布模型，提供孪生体数据生成

- 明确6G 内生AI的需求
- 明确6G 内生AI的特征和关键问题
- 明确6G 内生AI的架构和功能需求



中国移动
China Mobile

Thanks