

基于快慢思考混合模型的核心网运维智能体研究

Research on Core Network Operation and Maintenance Intelligent Agent Based on Mixture of Model

蔡佳翌¹,高 鑫¹,张梦书¹,扶慧湘¹,王惜时¹,魏 宏²,韩 宁²(1. 中国移动通信集团有限公司,北京 100033;2. 华为技术有限公司,广东 深圳 518129)

Cai Jiayi¹, Gao Xin¹, Zhang Mengshu¹, Fu Huixiang¹, Wang Xishi¹, Wei Hong², Han Ning² (1. China Mobile Communications Group Co., Ltd., Beijing 100033, China; 2. Huawei Technologies Co., Ltd., Shenzhen 518129, China)

摘 要:

核心网日常运维活动既有专业高频的状态查询、知识问答、信令分析等任务,也存在网络变更、业务开通等定制性强、流程复杂的任务。通过引入快慢思考混合模型(Mixture of Model)智能体架构,快思考模型处理专业、高频的运维任务,“慢思考”模型则专注于复杂推理任务,从而实现核心网运维全场景的高效自动化处理,并且综合效率和成本最优。未来该方案将进一步向跨域自治和6G内生智能方向演进,助力运营商实现自智网络L5目标。

关键词:

核心网运维;快慢思考混合模型;智能体;自智网络;信令分析

doi: 10.12045/j.issn.1007-3043.2025.08.003

文章编号: 1007-3043(2025)08-0010-04

中图分类号: TN915

文献标识码: A

开放科学(资源服务)标识码(OSID):



Abstract:

The daily operations and maintenance (O&M) of core networks includes professional high-frequency tasks such as status queries, knowledge Q&A and signaling analysis, and customized complex tasks such as network changes, service provisioning. It introduces the Mixture-of-Model (MoM) intelligent agent architecture. Fast-thinking models handle high-frequency O&M tasks with domain expertise and rapid response, and Slow-thinking models focus on complex logical reasoning involving multi-step workflows and configuration dependencies, so as to achieve full-scenario automation in core network O&M, ensuring optimal balance of efficiency and cost. In the future, this solution will further evolve towards cross-domain autonomy and 6G endogenous intelligence, helping operators achieve their L5 goal of self intelligent networks.

Keywords:

O&M of core networks; Mixture of Model; Intelligent agent; Autonomous network; Signaling analysis

引用格式: 蔡佳翌,高鑫,张梦书,等. 基于快慢思考混合模型的核心网运维智能体研究[J]. 邮电设计技术, 2025(8): 10-13.

0 引言

随着5G网络的规模部署和业务快速发展,核心网变得愈发复杂。据统计,近5年中国移动的管理对象持续增长,但运维人员数量长期大致持平,运维工作量和运维人员数量之间的“剪刀差”越来越大。传统依赖专家经验和多工具协同分析的运维方式技能门

槛高,已无法满足日益增长的业务发展诉求。

自智网络(Autonomous Network, AN)作为通信行业的重要发展方向,其目标是通过自动化和智能化手段提高网络的稳定性和可靠性,实现网络的高阶自智。2024年11月, TM Forum联合73家产业伙伴发布了AN L4产业蓝图,明确将智能体技术作为AN L4的核心特征,目标是实现场景的自动化闭环,分为单域自治(2025—2027)和跨域自治(2028—2030)2个阶段^[1]。

收稿日期: 2025-06-30

本文提出一种基于快慢思考混合模型的核心网运维智能体方案,通过将认知科学中的双系统理论与人工智能技术相结合,实现核心网运维场景的高效自动化处理,为运营商实现自智网络 L4 目标提供了新的技术路径。

1 快慢思考混合模型理论基础

1.1 双处理理论概述

快慢思考混合模型的理论基础源于 2002 年诺贝尔经济学奖得主丹尼尔·卡内曼(Daniel Kahneman)的双处理理论(Dual Process Theory)。该理论将人类的决策过程分为 2 个系统。

a) 系统 1(快思考)。快速、直接、直觉的思考方式,通常是无意识的、基于经验和个人习惯的。例如,看到危险躲避、别人叫我们的名字时自动回应等。系统 1 具有高效率、低能耗的特点。

b) 系统 2(慢思考)。缓慢、理性、分析性的思考方式,需要调动注意力来分析,最后作出决定。例如,学习新知识、解决复杂的数学问题等。系统 2 具有强规划、强逻辑的特点,但需要消耗大量认知资源,容易疲劳。

卡内曼的双处理理论指出,大脑在复杂任务处理中,应将简单任务交给“快思考”模型进行快速处理,而将复杂推理任务交给“慢思考”模型进行深度分析。

1.2 AI 领域的快慢思考模型

OpenAI 定义的通用人工智能(AGI)5 级分级标准如表 1 所示,可以看到当前业界模型已经从第 1 阶段

表 1 OpenAI 定义的通用人工智能(AGI)5 级分级标准

级别	聊天机器人 ChatBots	推理者 Reasoners	智能体 Agents	创新者 Innovators	组织者 Organizations
定义	接收输入,理解并生成类人文本的人工智能系统	具备深度理解和推理能力,可处理复杂问题的人工智能系统	根据目标独立执行任务并具备决策能力的人工智能系统	在解决现有问题的基础上,有新想法、创造新解决方案的人工智能系统	管理和执行通常由整个组织执行的广泛任务的人工智能系统
能力	自然语言理解、上下文感知、文本生成、多语言能力和个性化	高级逻辑推理、问题解决能力、理解上下文和细微差别、学习与适应、多领域专业知识、协作式问题解决	自主性、目标导向行为、决策制定、适应性与学习、多领域功能、交互性	创造性思维、研究与开发、自主创新、跨学科整合、对人工智能研究的贡献、与人类协作	集成决策制定、多功能能力、流程自动化、资源管理、持续学习与适应、可扩展性

ChatBots 的概率匹配“快思考”演进到“推理者”阶段,具备了强逻辑推理能力和慢思考能力。

a) 快思考模式。处理日常对话、信息提取等简单任务时,通过领域知识增训微调的 ChatBots 模型具备响应快、结果准确的优势。

b) 慢思考模式。面对数学证明、代码生成、任务规划等复杂逻辑时,Reasoner 模型通过多步骤分析,模型反思输出结果,具备一定的新场景泛化和学习能力。

2 核心网运维现状

随着云原生和全融合的持续演进,核心网复杂度不断提高。近 5 年,中国移动的管理对象持续增长,但运维人员数量长期大致持平。

传统运维方式面临巨大挑战,主要体现在以下几个方面。

a) 效率低下。运维人员每天要处理海量的投诉和告警工单,工作量巨大。例如,投诉处理需要专家工作 4 h 才能完成信令分析,告警处理耗时 1.5 h,难以满足严格的 SLA 要求。

b) 技能门槛高。信令分析等任务依赖资深专家,新员工难以快速上手。运维工作高度依赖专家经验和工具,缺乏对复杂场景的泛化能力。

c) 知识传承困难。专家经验难以有效转化为可复用的知识,新员工培养周期长,技能门槛高。

d) 可靠性风险。网络变更和故障管理依赖人工,易因操作失误或响应延迟导致服务中断。核心网高稳能力是保障业务连续性的关键,但传统方式难以应对。

3 基于 MoM 架构的核心网运维智能体技术方案

3.1 技术架构设计

核心网运维基于分层架构设计如图 1 所示,主要包括以下几个层次^[2]。

a) 网元层。包括核心网的典型网元,是运维智能体的服务对象,可以通过北向接口对上层提供数据采集以及指令执行通道,实现网络的高稳高效运行。

b) 内生运维智能层。厂商域 MDAF,通过收集和各类运维数据,并对数据进行深度分析和推理,对上承接 OSS 层的业务目标,通过 AGI 进行任务的自主规划和自闭环。

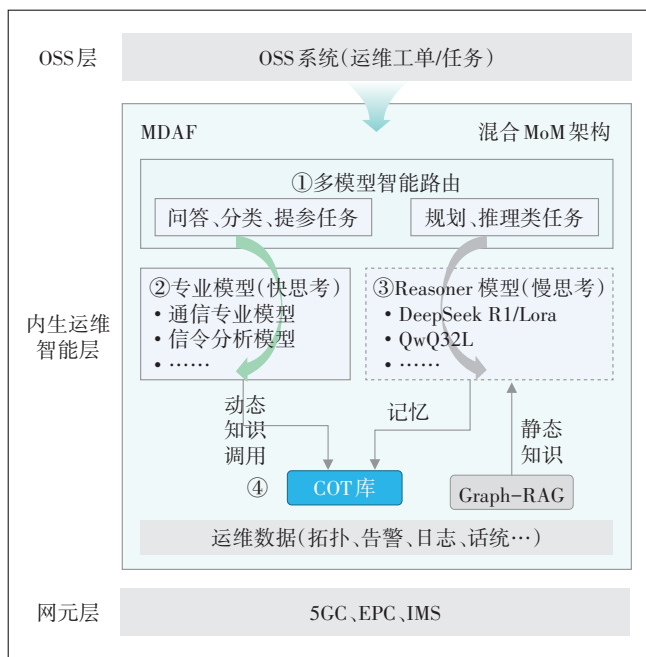


图1 核心网运维智能分层架构

c) OSS层。运维人员入口,通过工单或者基于岗位的智能体与内生运维智能层的Agent协同,实现日常运维流程的E2E编排自动化。

基于快慢思考混合模型的核心网运维智能体方案属于内生运维智能层的实现方案。

3.2 混合模型协同机制

混合模型协同机制是方案的核心,主要包括以下几个方面。

a) 多模型智能路由。根据任务的复杂度,将简单任务分配给“快思考”模型,复杂任务分配给“慢思考”模型。例如,告警分类、信令解析等简单任务由“快思考”模型处理,而根因分析、网络变更规划等复杂任务由“慢思考”模型处理。初期智能体能力有限时,智能路由可以基于任务类型进行规则化的分流,目标状态下智能路由由专业的分类模型来承担。

b) 专业模型(快思考)。核心网运维领域的专业模型,参数量介于1~38 B之间,通过海量的领域专有数据/知识训练而成,典型模型有核心网知识模型、信令模型、KPI模型以及日志模型;其优点是对领域特定任务的准确性高,可快速给出结论。

c) Reasoner模型(慢思考)。通用的LLM模型,具备较强的推理和规划能力,可以基于特定的任务进行自主规划,或者基于已有的线索信息进行推理,综合给出分析结论。Reasoner模型的特点是泛化能力强,

但是因为长思考会导致业务处理流程较长,并且准确率存在幻觉。

d) 记忆转知识。将“慢思考”模型的分析结果经过验证形成思维链,存储到知识库中,供后续“快思考”模型调用,实现由“慢”向“快”的转换。例如,通过多次的根因分析,形成标准故障处理流程,新告警可快速匹配已知模式。

3.3 智能体实现路径

基于MoM架构的核心网运维智能体的实现路径主要有以下几个步骤。

a) 模型选型与训练。

(a) 快思考模型。选择轻量级、高响应速度的模型,如华为盘古大模型,用于处理简单任务。

(b) 慢思考模型。选择参数规模大、推理能力强的模型,如DeepSeek推理模型,用于处理复杂任务。

(c) 数据准备。收集和处理运维数据,包括信令日志、故障案例库、网络拓扑等,构建高质量训练数据集。

b) 智能体框架构建。

(a) 智能体分类。根据运维岗位需求,构建不同类型的智能体,如投诉处理智能体、告警处理智能体、业务开通智能体等。

(b) 智能体交互。设计智能体之间的通信机制,实现信息共享和任务协作。电信领域作为相对封闭的系统,智能体交互建议参考开源A2A协议进行行业适配。

(c) 知识库整合。构建统一的运维知识库,存储模型分析结果和专家经验,支持智能体快速调用。

c) 部署与集成。

(a) 系统集成。与现有运维系统对接,实现工单自动流转、结果自动回填等功能。

(b) 安全加固。采用联邦学习和隐私计算技术,保护运维数据安全,避免敏感信息泄露。

4 未来发展方向与技术展望

4.1 自智网络L4向L5演进

随着自智网络向L5级“无人化”智能运维演进,基于MoM架构的核心网运维智能体方案将进一步发展,其未来发展方向如下。

a) 多智能体协作可编排技术。满足多智能体协同的各种复杂场景需求,实现任务的智能编排和资源的动态分配。

b) 基于多模态大模型的自适应智能体协同技术。使智能体更全面地理解运维场景中的不同环境和任务要求,自动调整策略和行为。

c) 基于 AI 和数字孪生双轮驱动的融合底座。二者结合可实现优势互补,让仿真、验证、预测等能力相互协作、彼此支撑。

4.2 6G 网络内生 AI 技术融合

6G 网络将 AI 技术内生融入网络架构,实现从“外挂式”向“内生式”的根本性转变。内生 AI 已成为 6G 的核心特征之一,将显著提升 6G 的网络性能、用户体验及新业务拓展等多维度的能力。

在 6G 时代,基于 MoM 架构的核心网运维智能体方案的发展机遇和挑战如下。

a) 机遇。

(a) 网络性能提升。AI 技术可优化网络资源调度,提高频谱效率和传输质量。

(b) 用户体验优化。AI 技术可实现用户、业务和网络全方位画像,为运营商提供挖掘连接价值的手段。

(c) 新业务拓展。AI 技术可支持通感算一体调度,催生更多智能化应用场景。

b) 挑战。

(a) 技术复杂度。6G 网络将引入更多新技术,如近场通信、全息通信等,需要更强大的 AI 模型支持。

(b) 算力需求。6G 网络将承载更多 AI 应用,需要更高效的算力分配和调度机制。

(c) 安全风险。AI 技术的广泛应用可能引入新的安全风险,需要更完善的安全防护机制。

为适应 6G 网络的发展需求,基于 MoM 架构的核心网运维智能体方案未来的演进方向如下。

a) 通感算一体调度。将 AI 模型与通信、感知、计算资源深度融合,实现端到端的智能业务编排。

b) 绿色节能机制。通过 AI 预测网络负载,动态调整资源分配,降低能耗,支持“双碳”目标。

5 结论与建议

5.1 研究成果总结

本文提出了一种基于快慢思考混合模型的核心网运维智能体方案,通过将认知科学中的双系统理论与人工智能技术相结合,实现核心网运维场景的高效自动化处理。方案已在 A 省移动和 B 省移动成功部署应用,取得了显著成效,具体如下。

a) 效率提升。投诉处理效率提升 66%,告警处理效率提升 87%,业务开通效率提升 75%。

b) 可靠性增强。网络高稳能力提升,故障恢复时间从小时级缩短至分钟级。

5.2 实施建议

基于方案的研究和实践,提出以下实施建议。

a) 分阶段部署。从简单任务处理开始,逐步扩展到复杂推理任务,降低实施风险和成本。

b) 数据质量优先。构建高质量的运维数据集,包括信令日志、故障案例库、网络拓扑等,为模型训练提供基础。

c) 完善安全机制。采用联邦学习和隐私计算技术,保护运维数据安全,避免敏感信息泄露。

d) 人机协同优化。保留人工干预机制,实现人机协同,避免完全依赖 AI 带来的风险。

e) 知识持续更新。建立知识更新机制,持续将智能体的分析结果和专家经验融入知识库,实现系统自我优化。

总之,基于快慢思考混合模型的核心网运维智能体方案,通过“快思考”和“慢思考”模型的动态协同,实现了核心网运维场景的高效自动化处理,为运营商实现自智网络 L4 目标提供了有力支撑。未来,随着技术的不断发展和应用场景的不断扩展,该方案将进一步向跨域自治和 6G 内生智能方向演进,推动通信网络的数智化转型。

参考文献:

- [1] 华为. TM Forum 和华为联合产业伙伴发布《自智网络等级测评白皮书》[R/OL]. [2025-02-28]. <https://www.huawei.com/cn/news/2024/11/whitepaper-autonomousnetworks/>.
- [2] 华为. 大模型使能核心网运维智能化,助力核心网高稳升级,迈向高阶自智网络[EB/OL]. [2025-02-30]. <https://www.huawei.com/cn/huaweitech/techtalks/has24-highly-stable-core-network>.

作者简介:

蔡佳翌,高级工程师,硕士,主要从事核心网智能运维系统顶层规划、架构设计与系统开发工作;高鑫,硕士,主要从事核心网运维智能体研发工作;张梦书,硕士,主要从事核心网智能运维系统功能架构设计工作;扶慧湘,工程师,硕士,主要从事核心网投诉智能处理方案设计工作;王惜时,初级工程师,硕士,主要从事核心网告警智能处理方案设计工作;魏宏,华为云核心网自动驾驶网络首席架构师,硕士,主要负责核心网自动驾驶网络解决方案的架构设计工作以及通信大模型和运维智能体方向的研究工作;韩宁,华为云核心网网络自动驾驶产品总监,主要负责核心网自动驾驶产品规划开发工作。