

AI 原生驱动的 6G 核心网 攻击面多维分析与安全需求建模

Multi-Dimensional Attack Surface Analysis and Security Requirement Modeling for AI-Native 6G Core Networks

陈博文¹, 孔令飞², 郭煜峰³ (1. 中国联通辽宁分公司, 辽宁 沈阳 110000; 2. 中国联通河北分公司, 河北 石家庄 050000; 3. 东北大学, 辽宁 沈阳 110000)

Chen Bowen¹, Kong Lingfei², Guo Yufeng³ (1. China Unicom Liaoning Branch, Shenyang 110000, China; 2. China Unicom Hebei Branch, Shijiazhuang 050000, China; 3. Northeastern University, Shenyang 110000, China)

摘要:

针对 AI 原生 6G 核心网面临的攻击面“质变型扩张”挑战, 提出一种端到端方法, 为 AI 原生 6G 核心网的安全架构设计提供结构化的需求分析与问题定义基础。通过攻击面分类, 系统化分析 AI 模型、开放接口及异构组件共模所面对的攻击面, 明确零信任与内生安全互补的边界判定准则, 构建安全需求矩阵, 揭示出核心网 10 类功能模块中有 8 类需二者协同的量化结论。将安全需求映射至 uRLLC、eMBB 跨域漫游及 AI 智能体决策等典型场景, 验证了协同策略在不同 SLA 约束下的可行性。

Abstract:

To address the challenge of "qualitative expansion" of the attack surface facing the AI-native 6G core network, an end-to-end approach is proposed, which provides a structured requirements analysis and problem definition foundation for the security architecture design of the AI-native 6G core network. Through attack surface classification, the attack surfaces faced by the common-mode coupling of AI models, open interfaces, and heterogeneous components are systematically analyzed. The boundary judgment criteria for the complementarity between zero trust and endogenous security are defined, a security requirements matrix is constructed, and the quantitative conclusion that 8 out of the 10 functional modules of the core network require the coordination of both is revealed. The security requirements are mapped to typical scenarios such as uRLLC, eMBB cross-domain roaming, and AI agent decision-making, thereby validating the feasibility of the collaborative strategy under different SLA constraints.

Keywords:

6G core network; AI-native; Attack surface analysis; Zero trust; Endogenous security

引用格式: 陈博文, 孔令飞, 郭煜峰. AI 原生驱动的 6G 核心网攻击面多维分析与安全需求建模[J]. 邮电设计技术, 2026(9): 89-92.

0 引言

开放架构的引入为 AI 原生 6G 核心网带来全新安全风险^[1]。模型投毒、对抗样本等动态 AI 攻击可突破传统边界防护, 攻击面由网元 NF 细化至容器 Pod 并发生质变扩张, 威胁具备概率驱动、跨域传播、零日短时

响应特征。智能架构使威胁由静态漏洞转为动态漂移风险, 异构跨域部署则进一步放大全网级联攻击危害。学界提出了融合零信任与内生安全构建防御体系的思路, 但现有研究存在三大短板。一是二者融合缺乏清晰范式边界与系统化协同设计方法; 二是尚未形成多维攻击面分类框架, 安全需求分析无结构化依据; 三是缺失攻击面到安全需求的端到端建模手段, 安全架构设计缺少问题导向支撑^[2-3]。

收稿日期: 2026-07-28

关键词:

6G 核心网; AI 原生; 攻击面分析; 零信任; 内生安全

doi: 10.12045/j.issn.1007-3043.2026.09.017

文章编号: 1007-3043(2026)09-0089-04

中图分类号: TN918

文献标识码: A

开放科学(资源服务)标识码(OSID):



1 AI原生6G核心网攻击面分析

1.1 攻击面演化驱动的AI原生6G

1.1.1 5G核心网的攻击面基线

5G核心网采用SBA(Service-Based Architecture)架构,网络功能(NF)通过基于HTTP/2的服务化接口(SBI)进行通信。上述攻击面在5G阶段已得到较为充分的研究,构成了6G安全分析的基线。5G的攻击面本质上仍是静态对抗,网络功能的行为逻辑具有确定性,攻击者利用的是已知漏洞或配置缺陷。

1.1.2 AI原生引入的攻击面扩张

6G核心网的AI原生化过程从3个维度扩张了攻击面。

a) AI模型攻击面。NWDAF、AI智能体等组件引入的ML模型成为新的攻击目标。与传统的确定性NF不同,ML模型的决策边界是数据驱动的、非透明的,主要威胁包括模型投毒、对抗样本以及模型窃取。

b) 开放接口攻击面。6G的开放化架构扩张了API暴露面。主要涵盖跨境SEPP间通信、数字孪生北向接口以及AI模型市场与联邦学习聚合接口。

c) 异构组件共模攻击面。6G的多供应商异构集成在提升弹性的同时,也引入开源组件共模漏洞、AI框架共模漏洞以及硬件共模漏洞等共模风险。

1.1.3 攻击面扩张的量化表征

本文定义攻击面扩张示意指标(Attack Surface Expansion Indicator, ASEI)来半定量分析上述3类攻击面扩张。ASEI是一个示意性指标,用于反映6G相对5G在特定维度上的攻击面增长幅度,而非严格的精确度量。其计算方法为:

$$ASEI_d = \frac{N_{6G}^{(d)}}{N_{5G}^{(d)} + \epsilon} \quad (1)$$

其中, d 表示攻击面类别, $d \in \{\text{AI模型}, \text{开放接口}, \text{异构共模}\}$, $N_{5G}^{(d)}$ 和 $N_{6G}^{(d)}$ 分别为5G和6G在该类别下可识别的攻击向量类型数, $\epsilon=1$ 为拉普拉斯平滑因子。当 $N_{5G}^{(d)}=0$ 时,ASEI的取值上限为 $N_{6G}^{(d)}$,反映的是“从无到有”的质变幅度,而非一个发散的数学奇点。

各类别的计数规则如下。

a) AI模型攻击面以攻击向量类型为计数单位。5G中ML模型不是5G核心网的必要组件, $N_{5G}=0$;6G中可识别的主要攻击向量类型包括模型投毒、对抗样本、模型窃取、推理时规避等, $N_{6G} \geq 3$,故 $ASEI \geq 3.0$,反映出质变型扩张。

b) 开放接口攻击面以新增接口类型为计数单位。5G SBI核心接口约有15类;6G新增跨境SEPP增强、数字孪生北向、联邦学习聚合等关键接口类型, $N_{6G}-N_{5G} \geq 3$, $ASEI \approx 2.0$,反映量变型扩张。

c) 异构组件共模攻击面以共模风险来源类型为计数单位。5G共模风险主要集中于开源库层面;6G新增AI框架共模与硬件指令集共模,共模风险来源从1类增至3类, $ASEI \approx 2.0$,反映结构性扩张。

1.2 多维度攻击面分类法

AI原生6G核心网的攻击面可基于表1所示方法进行分类,以便进行系统化分析。

表1 四维攻击面分类法

维度	符号	定义	取值空间
攻击向量	V	攻击者利用的具体技术手段	{模型投毒,对抗样本,模型窃取,信令注入,API滥用,供应链后门,侧信道,凭证窃取,0-day漏洞,拟态攻击}
攻击面载体	C	被攻击的架构组件或接口	{AI模型,SBI接口,SEPP,数字孪生,NFV基础设施,联邦学习聚合器,硬件加速器,开源组件,执行体池}
攻击阶段	S	攻击在杀伤链中的位置	{侦察,武器化,投递,利用,驻留,横向移动,目标达成}
影响范围	I	攻击造成的安全后果	{机密性破坏,完整性破坏,可用性破坏,模型完整性破坏,隐私泄露,服务降级,合规失效}

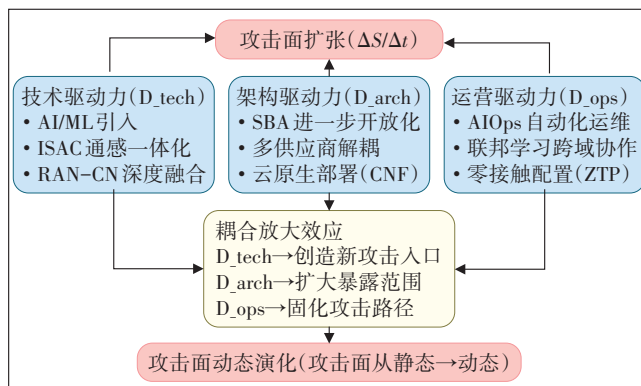
该分类法将任意攻击实例映射为四元组(V, C, S, I)。例如,对NWDAF的模型投毒攻击可刻画为(模型投毒,AI模型,投递,模型完整性破坏)。其攻击面演化过程如图1所示。

1.3 典型攻击向量分析

本节选取3类代表性的AI原生6G攻击向量进行讨论。

1.3.1 AI模型攻击:NWDAF负载预测投毒

攻击向量: V =模型投毒, C =AI模型, S =投递, I =模



1 攻击面演化三维驱动力耦合模型

型完整性破坏。

攻击场景:假设攻击者通过入侵 NWDAF 的训练数据管道,使 NWDAF 的负载预测模型产生偏差。在 6G 中,NWDAF 的预测结果直接驱动 SMF 的会话管理策略和 PCF 的 QoS 策略,因此投毒的影响将从 NWDAF 级联传播至 SMF 和 PCF,最终导致全网资源调度异常。

级联传播模型:设 NWDAF 的预测输出为 $\hat{y} = f_{\theta}(x)$,其中 θ 为模型参数, x 为输入特征。投毒攻击将 θ 篡改为 θ' ,使得 $\hat{y}' = f_{\theta'}(x)$ 偏离真实值。该偏差通过 SBI 接口传播:

$$\hat{y}' \xrightarrow{\text{NnwdaF}} \text{SMF} \xrightarrow{\text{Nsmf}} \text{UPF} \xrightarrow{\text{N4}} \quad (2)$$

1.3.2 开放接口攻击:跨域 SEPP 信令篡改

攻击向量: V =信令注入, C =SEPP, S =利用, I =完整性破坏。

攻击场景:在 5G/6G 跨境漫游中,拜访网络(VPLMN)的 SEPP 通过 N32 接口与归属网络(HPLMN)的 SEPP 交换信令。攻击者可通过以下路径实施攻击。

a) TLS 中间人攻击。利用 N32-f 接口的 TLS 证书验证缺陷,在 VPLMN 和 HPLMN 之间插入恶意代理,篡改漫游 UE 的认证向量或 QoS 参数。

b) JSON Web Token (JWT) 篡改。N32 接口使用 JWT 进行信令保护,攻击者利用 JWT 库的 alg=none 漏洞或密钥泄露,伪造合法的漫游信令。

c) 信令重放。捕获合法漫游信令后,在目标网络中重放以触发 UE 的重复注册或资源分配。

1.3.3 共模攻击:开源组件漏洞的跨 NF 利用

攻击向量: V =零日漏洞, C =开源组件, S =利用, I =可用性破坏。

攻击场景:假设 OpenSSL3.x 中存在一个远程代码执行漏洞,影响所有使用该版本的 NF。在 6G 多供应商环境中,若 3 个不同供应商的 UPF 实例均使用 OpenSSL3.x 进行 TLS 通信,攻击者可利用该漏洞同时攻陷所有 UPF 实例,使异构冗余机制完全失效,因为所有执行体在同一时刻、同一漏洞上共模失败。

2 零信任与内生安全互补需求建模

2.1 零信任与内生安全的内涵对比

2.1.1 概念定义

零信任(Zero Trust, ZT):基于 NIST SP 800-207 的定义,其核心理念是“永不信任、持续验证”——任何

访问请求都必须经过身份验证、授权和加密,且信任状态随时间动态衰减。

内生安全(Endogenous Security, ES):基于动态异构冗余(DHR)架构,其核心理念是架构决定安全。

2.1.2 内涵维度对比

ZT 的高误判风险可被 ES 的物理冗余部分补偿;ES 的共模风险可被 ZT 的细粒度身份校验部分补偿。其典型攻击类型下的联合应用如表 2 所示。

表 2 互补边界判定决策矩阵

攻击类型	可观察性	验证成本	多样性可用性	主导机制
身份假冒(凭证窃取)	高	中	低	ZT 主导
权限滥用(内部威胁)	中	中	中	ZT 主导
已知漏洞利用	高	中	中	ZT+ES 协同
未知漏洞利用(0-day)	低	高	高	ES 主导
供应链后门	低	高	高	ES 主导
AI 模型投毒	中	高	中	ZT+ES 协同
AI 对抗样本	高	中	低	ZT 主导
跨域 APT	中	高	中	ZT+ES 协同
拟态攻击(伪装正常行为)	低	高	高	ES 主导

2.2 顶层设计原则

基于如下 4 项架构设计原则构成完整设计约束。

a) 双层防御原则。AI 原生 6G 核心网的安全架构必须同时具备决策层的零信任控制和执行层的内生安全能力,二者通过明确的协同机制工作。

b) 决策—执行可识别原则。网络功能的“决策逻辑”(策略生成、信任评估、路由选择)与“执行逻辑”(实际数据转发、模型推理)应在架构上可识别——即在设计期明确决策点与执行点的边界。

c) 执行体多样性管理原则。ES 的有效性依赖于执行体在算法实现、运行平台、软件版本、供应商等维度的多样性。

d) SLA 适配原则。ZT 的验证频率与 ES 的冗余度应根据 SLA 约束动态调整。

3 典型场景部署策略

将第 2 章的安全需求映射到 3 个典型 6G 部署场景,验证 ZT/ES 协同策略在不同 SLA 约束下的工程可行性。每个场景标注其主导功能模块,以避免业务场景与功能模块的范畴混淆。具体如表 3 所示。

所有场景均遵循双层防御和决策—执行可识别原则,差异在于 ZT 验证频率与 ES 冗余度。这正是第

表3 uRLLC/eMBB/AI三场景部署策略对比

场景	uRLLC 紧急通信	eMBB 跨域漫游	AI智能体决策
主导模块	UPF 用户面转发	SEPP 跨域信令	NWDAF AI推理
SLA 时延预算	1 ms	10 ms	100 ms
ZT 策略	采样验证 (1/10) <50 μs	全量验证 (每请求) ≤1 ms	基线持续监测 <10 ms
ES 策略	3 异构 (DPDK/ SmartNIC/eBPF) 裕度 10%, ≤10 μs	3 异构 (Open5GS/ Free5GC/OAI) 裕 度 2.9%	3 异构 (Transform- er/GNN/Rule) 裕 度 6.6%, 加权融 合
协同方式	同步指令+异步 报告闭环 ≤200 μs	同步指令+同步 报告闭环 ≤5 ms	同步指令+异步报 告高优优先级告警 ≤ 10 ms
所属象限	象限 I/III	象限 II	象限 IV

2.2 节中原则 b)~d) (SLA 适配) 所预期的: SLA 越严格, ZT 验证频率越低, ES 冗余度要求越高 (闭环时延= $T_{EP} + T_{DP}$; 裕度= $\frac{H_{actual} - H_{min}}{H_{min}} \times 100\%$)。

4 方案对比

ZT×ES 协同策略与现有方案的对比如表 4 所示。

表4 ZT×ES 协同策略与现有方案的定性对比

维度	ZT-only	ES-only (DHR)	朴素叠加	ZT×ES 协同(本文)
防御机制	持续信任评估	异构冗余+裁决	ZT+ES 各自独立	ZT+ES 通过协同接口统一
信任状态	有(随时间演化)	无	各自维护, 互不可见	ES 信息纳入 ZT 信任评估
共模失效防护	无(单执行体)	有(多异构)	有	有(ES 层)+ ZT 可感知多样性退化
状态可追溯	有(CDM 证据链)	无	仅 ZT 侧可追溯	ES 裁决信息纳入证据链
降级路径	无(信任分低→拒绝)	无(DHR 无降级)	各自退化, 无序	有序降级(正常→审计→回退)
接口一致性	N/A	N/A	未定义	明确定义的协同机制
关键弱点	零日漏洞导致单点突破	多样性衰减无感知	ZT 与 ES 判定冲突无仲裁	冗余度固定, 极端场景需扩展

5 结论与展望

AI 原生 6G 核心网的安全挑战源于智能化与开放化带来的攻击面质变型扩张。本文从“攻击面分析→互补需求建模→场景部署验证→方案对比与标准对齐”4 个层次, 系统化地应对了这一挑战。

a) 攻击面维度。建立四维攻击面分类法 (V, C, S, I), 揭示 AI 原生引入的三重扩张, 定义 ASEI 作为攻击面扩张的示意性指标——AI 模型攻击面 ASEI ≥ 3.0, 异构共模 ASEI ≈ 2.0, 开放接口 ASEI ≈ 2.0。归纳了三大演

化趋势及其对应的安全挑战。

b) 需求建模维度。提出 5 条 ZT×ES 互补边界判定准则, 建立“决策可控性×执行可信性”二维安全需求矩阵, 将 6G 核心网 10 类功能模块定位至 4 个象限, 揭示“8/10 模块需要 ZT 与 ES 协同”的量化结论。基于此给出 4 项架构设计原则, 形成从宏观到微观、从静态到动态的完整设计约束体系。

c) 场景验证维度。将安全需求映射到 uRLLC 紧急通信、eMBB 跨域漫游和 AI 智能体决策 3 个典型场景, 验证 ZT/ES 协同策略在不同 SLA 约束下的工程可行性。3 个场景的共性是“双层防御+可识别边界”, 差异是“验证频率与多样性裕度的 SLA 适配”。

d) 标准对齐维度。建立 ZT×ES 协同策略与 3GPP TS 33.501、ETSI TS 28.533、NIST AI RMF 1.0 等标准体系的映射关系, 表明本文的工作是在现有标准自然演进方向上的具体推进。

本文的局限性在于: DC/ET 分值基于设计期估算, 还需运行期数据校准; 互补边界判定准则的定量阈值需运营商根据实际部署环境校准; 3 个场景的验证基于分析性论证, 缺乏原型系统实测。未来工作将围绕如下 3 个方向展开。

a) 原型验证。基于开源 5GC 构建原型系统, 通过实际攻击注入验证协同策略的有效性。

b) 动态 SLA 适配。研究 ZT 验证频率与 ES 冗余度的自适应调整机制。

c) 跨域协同扩展。将本文的单域协同策略扩展到多运营商、多接入技术的跨域协同场景。

参考文献:

[1] 谢泽铨, 张曼君, 徐雷, 等. 6G 网络安全新需求及关键技术研究[J]. 邮电设计技术, 2024(8): 49–52.

[2] Khan R, Kumar P, Jayakody D N K, et al. A Survey on Security and Privacy of 5G Technologies: Potential Solutions, Recent Advancements and Future Directions [J]. IEEE Communications Surveys & Tutorials, 2019, 22(1): 196–248.

[3] Ahmad I, Kumar T, Liyanage M, et al. Overview of 5G Security Challenges and Solutions [J]. IEEE Communications Standards Magazine, 2018, 2(1): 36–43.

作者简介:

陈博文, 助理工程师, 主要从事移动通信核心网分组域关键技术、网络架构演进及网络空间安全保障技术研究工作; 孔令飞, 高级工程师, 主要从事网络空间安全体系架构设计与关键安全技术研究工作; 郭煜峰, 博士, 主要从事智能算法设计、模型构建及人工智能技术在复杂系统中的应用研究工作。