面向大规模物联网的零信任管理研究综述

邢方圆[1],董 傲[1],孙羽羿[2],童 飞[1],贺诗波[3],程 光[1]
(1. 东南大学网络空间安全学院,江苏南京 211189;2. 杭州师范大学信息科学与技术学院,浙江杭州 311121;
3. 浙江大学控制科学与工程学院,浙江杭州 310058)

摘 要: 随着物联网(Internet of Things,IoT)技术的快速发展和广泛应用,大规模IoT(Large Scale IoT,LS-IoT)的部署已成为实现智能化、高效化社会基础设施的必然趋势. 然而,由于大规模网络具有异构化、高时变性和广分布的特点,导致网络与信息安全问题日益凸显. 传统的基于边界防护(Perimeter Based Security,PBS)的安全模型难以有效应对LS-IoT中复杂且动态的威胁. 零信任架构(Zero Trust Architecture,ZTA)强调"永不信任,始终验证"的安全理念,为保障LS-IoT的安全提供了一种潜在解决方案. 本文首先系统综述了ZTA的三大核心能力,包括软件定义边界(Software-Defined Perimeter,SDP)、身份识别与访问管理(Identity and Access Management,IAM)、微隔离(Micro-SeGmentation,MSG). 其次,结合LS-IoT的特点和需求,深入探讨了实现ZTA核心能力所需的七大关键技术,包括持续身份认证、动态访问控制、轻量加密技术、身份治理与管理(Identity Governance and Administration,IGA)、终端安全、网络隔离以及持续监控. 再次,以ZTA在工业IoT、5G医疗、自动驾驶和远程办公四个典型场景的实际应用为例,探讨了ZTA在提升网络安全性方面的有效性. 最后,文章分析了大语言模型(Large Language Model,LLM)、生成式人工智能(Artificial Intelligence,AI)、可解释性人工智能(eXplainable Artificial Intelligence,XAI)、边缘计算和后量子加密(Post Quantum Cryptography,PQC)等前沿技术与ZTA的融合,并展望了ZTA未来的发展方向. 通过上述工作,旨在为ZTA的实际应用和LS-IoT的安全保障提供参考.

关键词: 大规模物联网(LS-IoT);零信任架构(ZTA);网络安全;智能化
基金项目: 国家自然科学基金(No.62472085)
中图分类号: TP393.0 **文献标识码:** A **文章编号:** 0372-2112(2025)08-2993-33
电子学报URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20250228

A Survey of Zero Trust Management for Large-Scale Internet of Things

XING Fang-yuan[1], DONG Ao[1], SUN Yu-yi[2], TONG Fei[1], HE Shi-bo[3], CHENG Guang[1]
(1. *School of Cyber Science and Engineering*, *Southeast University*, *Nanjing*, *Jiangsu* 211189, *China*;
2. *School of Information Science and Technology*, *Hangzhou Normal University*, *Hangzhou*, *Zhejiang* 311121, *China*;
3. *College of Control Science and Engineering*, *Zhejiang University*, *Hangzhou*, *Zhejiang* 310058, *China*)

Abstract: With the rapid development and extensive application of internet of things (IoT) technologies, the large-scale deployment of IoT (LS-IoT) has become an inevitable trend for building intelligent and efficient social infrastructure. However, the heterogeneous, time-varying, and widely distributed nature of large-scale networks has led to increasingly prominent network and information security issues. Conventional perimeter-based security (PBS) models struggle to address complex and evolving threats in LS-IoT environments. The zero trust architecture (ZTA), which emphasizes the security principle of "never trust, always verify", provides a potential solution for ensuring the security of LS-IoT systems. Initially, this paper systematically reviews the three core capabilities of ZTA, including software-defined perimeter (SDP), identity and access management (IAM), and micro-segmentation (MSG). Subsequently, aligning with the characteristics and requirements of LS-IoT, we delve into seven critical enabling technologies for implementing ZTA, including continuous identity authentication, dynamic access control, lightweight encryption technology, identity governance and management (IGM), terminal security, network isolation, and continuous monitoring. Then, throught practical applications in four representative scenarios, such as industrial IoT, 5G-enabled healthcare, autonomous driving, and remote work, this paper illustrates the effectiveness of ZTA in enhancing network security. Ultimately, this paper explores the integration of emerging technologies, such as large language models (LLM), generative artificial intelligence (AI), explainable machine learning

收稿日期:2025-03-26;录用日期:2025-07-28;责任编辑:覃怀银

(XML), edge computing, and post-quantum encryption (PQC) with ZTA, and discusses the future development directions of ZTA. This work aims to provide valuable insights for advancing ZTA implementation and strengthening security assurance in large-scale IoT.

Key words: large-scale internet of things (LS-IoT); zero trust architecture (ZTA); network security; intelligence
Foundation Item(s): National Natural Science Foundation of China (No.62472085)

1 引言

物联网(Internet of Things,IoT)设备的爆炸性增长,从智能家居到智慧城市、工业控制系统和医疗设备,形成了庞大且复杂的网络系统.这些设备不仅数量庞大,而且设备类型异构,操作系统、通信协议、硬件架构等差异显著,给系统和网络安全管理带来了前所未有的挑战.随着云计算、移动办公和IoT设备的普及,传统的基于边界防护(Perimeter Based Security,PBS)模型,如集中防火墙、入侵检测/防御系统,以及以虚拟专用网络(Virtual Private Network,VPN)为代表的隧道式远程接入,主要依赖明确且静态的"内部可信—外部不可信"安全边界进行防护[1,2].当数以亿计的IoT设备跨域、跨域主体、随时上下线地接入网络时,这种单一边界的假设不再成立.因此大规模IoT(Large Scale IoT,LS-IoT)的独特性决定了其需要一种全新的安全架构,以满足对异构、动态变化的LS-IoT设备的精细化管理.

零信任架构(Zero Trust Architecture,ZTA)的概念最早由Forrester Research的分析师John Kindervag于2010年提出[3],其核心理念是"永不信任,始终验证",强调对每一个用户、设备和数据流进行持续的安全监控和验证.最初,ZTA主要针对企业内部网络和云计算环境的安全需求而提出,旨在解决传统"信任但验证"模式在复杂多变的网络环境中存在的诸多安全隐患.与传统模式不同,零信任完全打破了固定的信任边界,强调无论内部还是外部的任何访问请求都必须经过严格的验证和授权.随着LS-IoT环境的快速发展,ZTA的核心理念被进一步扩展和应用于IoT安全管理.IoT环境中的设备数量庞大、类型多样且动态变化,传统的基于PBS的安全模型难以有效应对这些特性带来的复杂安全挑战.因此,ZTA特别适用于LS-IoT这种动态且复杂的环境.ZTA通过软件定义边界(Software-Defined Perimeter,SDP)、身份识别与访问管理(Identity and Access Management,IAM)、微隔离(Micro-SeGmentation,MSG)三大核心能力实现对设备的精细化管理和数据流的严格控制.例如,在工业IoT环境中,不同生产设备之间频繁的数据交互涉及高实时性要求,而设备通常功耗和计算能力受限,导致传统的安全防护手段在此类场景中表现不足.零信任通过微分段技术和持续身份验证,有效地保护了LS-IoT环境中的每一个节点,确保了设备之间的通信安全.表1从信任假设、控制对象粒度、策略动态性、横向渗透防护、资源消耗以及适用场景六个维度直观对比了ZTA在LS-IoT中的安全优势.

表1 ZTA与安全模型对比

安全模型	传统PBS[4~6]	网络接入控制/VPN[7,8]	基于区块链的分布式信任[9,10]	ZTA
信任假设	内网默认可信,外网不可信	设备接入通过认证后即被信任	链上共识保障数据完整性	默认不信任任何实体,访问前后持续验证
控制对象粒度	子网级/IP级	端口级/会话级	交易级	用户、设备、进程、流四维细粒度
策略动态性	低,依赖人工变更	中,接入时评估一次	中,区块达到共识后固化	高,基于实时上下文与风险评分自动调整
横向渗透防护	弱,侧向流量可自由传播	取决于访问控制列表(Access Control List,ACL)与虚拟局域网(Virtual Local Area Network,VLAN)规则设计	中等,共识机制使数据层不可篡改,但智能合约等逻辑控制层可能存在漏洞	通过MSG严格限制东西向流量
资源消耗	低	中等	计算/存储开销高	可通过轻量级认证与边缘计算优化,降低能量消耗
适用场景	封闭园区网,流量形态简单	远程办公,中小规模园区	价值链溯源,数据不可篡改	海量、异构、跨域、动态的LS-IoT

近年来,学术界和工业界围绕如何将ZTA应用于云计算、LS-IoT等新兴技术领域展开了广泛研究,尤其在身份验证、访问控制以及微分段技术方面取得了显著进展.在应用方面,随着企业对安全需求的不断提升,ZTA正逐步成为网络安全领域的主流趋势.例如,美国国防部和国家安全局等机构正在积极推进零信任

的实施,并发布了相关指导文件《零信任架构》[11]. 此外,市场上已有多家技术提供商推出了基于零信任的解决方案. 例如,Zscaler的云安全平台、Palo Alto Networks的Prisma Access等. 其中Google的BeyondCorp项目是ZTA在实际企业环境中的成功应用范例,展示了去边界化安全模式的可行性和优越性[12].

围绕IoT安全领域,目前已有部分基础研究,具体应用场景包括智能电网[13]、工业IoT[14~17]、低空IoT[18,19],以及边缘计算[20,21]等. 在智能电网方面,文献[13]提出了一种特征注意力蒸馏防御方案,以降低后门攻击成功率. 针对工业IoT场景,文献[14,15]分别研究了跨域协同中的认证与隐私包含方案和入侵检测中的三阶段未知攻击检测架构;文献[16]设计了一种虫洞路由环路检测与虫洞-灰洞复合攻击路由环路检测方案;文献[17]设计了一种基于弱监督的时序异常检测方案. 在低空IoT中,文献[18,19]分别提出了基于OODA(Observe-Orient-Decide-Act)循环决策模型与协同干扰框架的混合攻击抑制方案,以及细粒度隐式表征与粗粒度体素表征结合的无人机三维场景重建策略. 针对边缘计算场景,文献[20]提出了一种基于移动目标防御的分布式拒绝服务(Distributed Denial-of-Service,DDoS)攻击缓解策略;文献[21]设计了一种数据、结构双复合的DDoS攻击检测方案.

通过归纳总结现有基础研究发现,当网络节点扩展至更大规模时,将出现身份爆炸、策略规则指数级膨胀、横向渗透扩大等问题. 针对LS-IoT场景,身份认证与管理、权限分配、内网横向防御等关键技术亟待解决. 因此,对ZTA在LS-IoT中的适配性与前沿进展进行系统且全面的调研分析十分必要.

然而,目前部分ZTA综述在整体全面性和前瞻性方面仍存在不足. 尽管这些综述奠定了理论基础,但大多聚焦于企业网络或中小规模IoT,尚未充分关注LS-IoT场景下新的安全需求. 表2从文章视角与关键覆盖范围出发,总结对比了2020—2025年4篇代表性零信任综述的局限性. 与此相比,本文首次从更大范围、更高异构度和持续动态性的LS-IoT视角,统一梳理零信任的三大核心能力与七大关键技术,并重点分析了这七大关键技术应用于LS-IoT场景所需的持续性、动态性、细粒度、轻量级等特征. 此外,结合大语言模型(Large Language Model,LLM)、生成式人工智能(Artificial Intelligence,AI)、可解释性人工智能(eXplainable Artificial Intelligence,XAI)、边缘计算以及后量子加密(Post Quantum Cryptography,PQC)等最新趋势,为未来LS-IoT场景面临的关键痛点挑战提供了一体化研究框架和实践路线.

表2　零信任综述对比

文献	切入视角	关键覆盖范围	局限	与本文差异
《工业物联网零信任安全研究综述》[22]	工业IoT垂直领域	核心技术、典型工业场景(电力,车联网等)	局限于工业垂直领域; 缺少云-边-端协同与大规模异构问题研究	本文扩展到异构LS-IoT、海量终端和高动态场景; 同时关注多行业应用
《零信任研究综述》[23]	初始概念与技术梳理	概念级框架、核心组成要素	早期奠基性梳理,聚焦概念定义,不含具体场景实践; 缺少对LS-IoT应用的细化研究	本文在其术语基础上,进一步细化面向异构LS-IoT的关键技术栈; 验证ZTA在典型LS-IoT场景的实用性
《零信任安全架构研究综述》[24]	企业能力栈	企业数字化、云迁移安全	缺少前沿技术的未来展望; 时效性不足,更多关注已有方案的总结	本文在三大核心能力基础上,引入生成式AI等前沿探索; 结合IoT海量异构性做出新问题剖析
《零信任网络综述》[25]	网络层实践路线	IAM、MSG、SDP; 云计算/5G	偏网络层实现,身份治理、终端安全等跨层分析不足; 对ZTA现有方案局限与挑战的深度讨论相对欠缺	本文强化身份生命周期管理与硬件/终端安全,同时分析ZTA关键痛点

本文首先系统地总结了ZTA的三大核心能力:SDP、IAM、MSG,以及实现ZTA的所需要的七大关键技术:身份认证、访问控制、加密技术、身份治理与管理(Identity Governance and Administration,IGA)、终端安全、网络隔离以及持续监控. 其次,分析了ZTA在工业IoT、5G医疗、自动驾驶以及远程办公四大经典场景中的应用价值,并探讨了零信任与AI、边缘技术、后PQC等前沿技术的融合. 最后,本文梳理了ZTA在实际部署过程中的技术难题,并对未来的发展方向进行了展望,旨在推动零信任在LS-IoT的实际应用.

为清晰表明本文的逻辑主线,图1展示了本文的章节结构与主要内容. 第1节引言介绍了研究背景;第2节阐述ZTA三大核心能力(SDP、IAM、MSG)及其内在联系;第3节探讨LS-IoT的特点及其与ZTA的匹配度,分析ZTA在海量异构、资源受限、高动态等场景下的安全需求;第4节详细剖析实现ZTA所需的七大关键技术(身

份认证、访问控制、加密技术、IGA、终端安全、网络隔离、持续监控);第5节列举工业IoT、5G医疗、自动驾驶、远程办公4个典型应用场景,结合实证案例说明ZTA落地成效;第6节分析前沿技术(LLM、生成式AI、XAI、边缘计算、后PQC)与ZTA的融合;第7节讨论分析当前ZTA面临的关键挑战,并展望未来趋势;第8节总结全文.

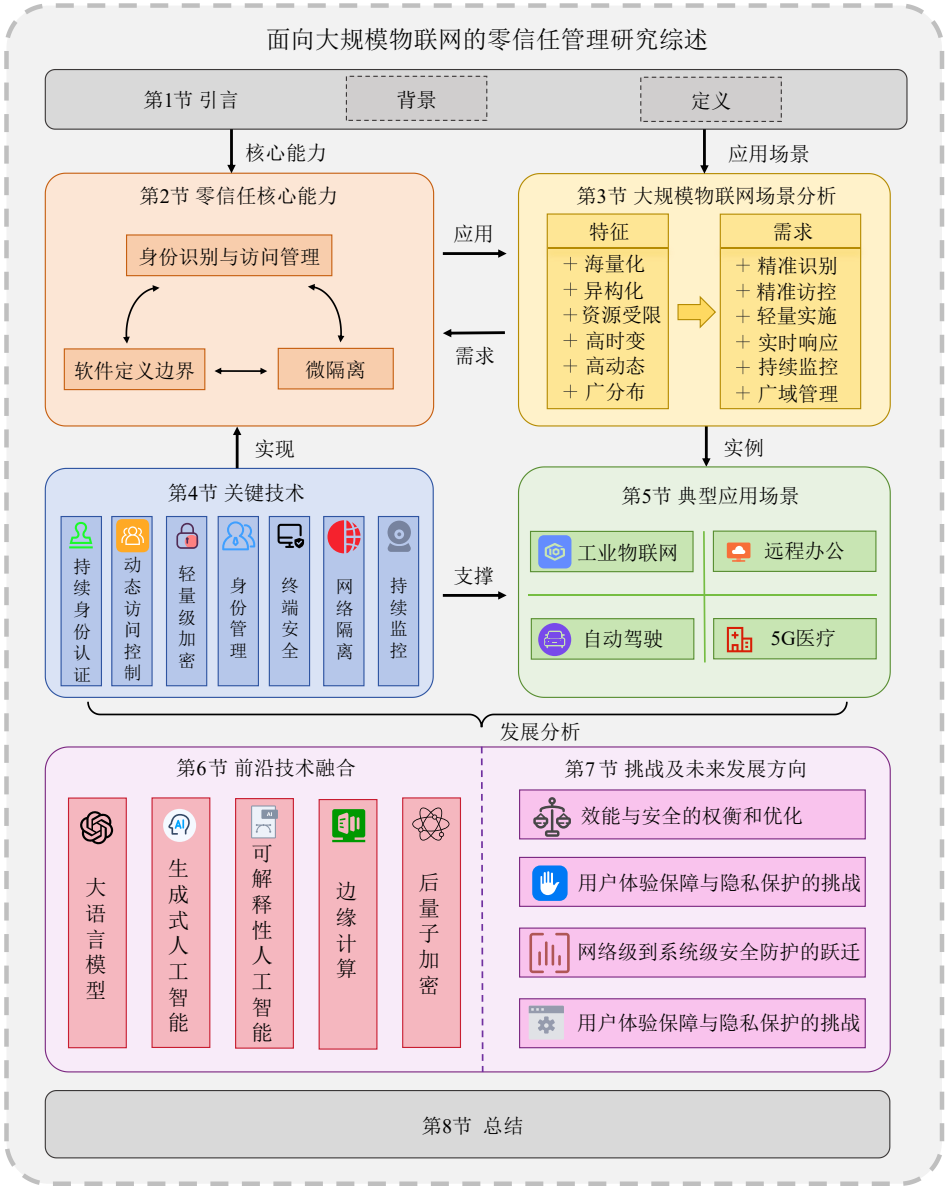

图1 文章结构与主要内容

2 零信任核心能力

如图2所示,ZTA包含三大核心能力:IAM、MSG以及SDP. IAM作为ZTA的基础和核心,为SDP和MSG提供身份验证与访问控制支持,保障资源间安全访问授权;MSG聚焦于东西向流量的控制,确保服务器间的通信得到严格保护;而SDP专注于数据南北向流量的安全性,保护用户与服务器之间的交互.

(1)IAM

IAM是ZTA中的核心模块,负责对用户身份、权限和资源访问的全面管理. IAM的核心任务是在确保用户经过严格认证的基础上,赋予其适当的访问权限,并根据其身份和环境动态调整访问控制策略,确保每次访问都符合最小权限原则.

IAM工作过程主要包括认证管理、IGA、审计管理、授权管理四个模块,这些模块协同工作以实现不同身份实体对资源的访问控制. 首先,用户需要经过认证管理模块,该模块不仅对"人"进行认证,还对设备以及应用进行认证,并通过持续认证方案最终确定用户的身

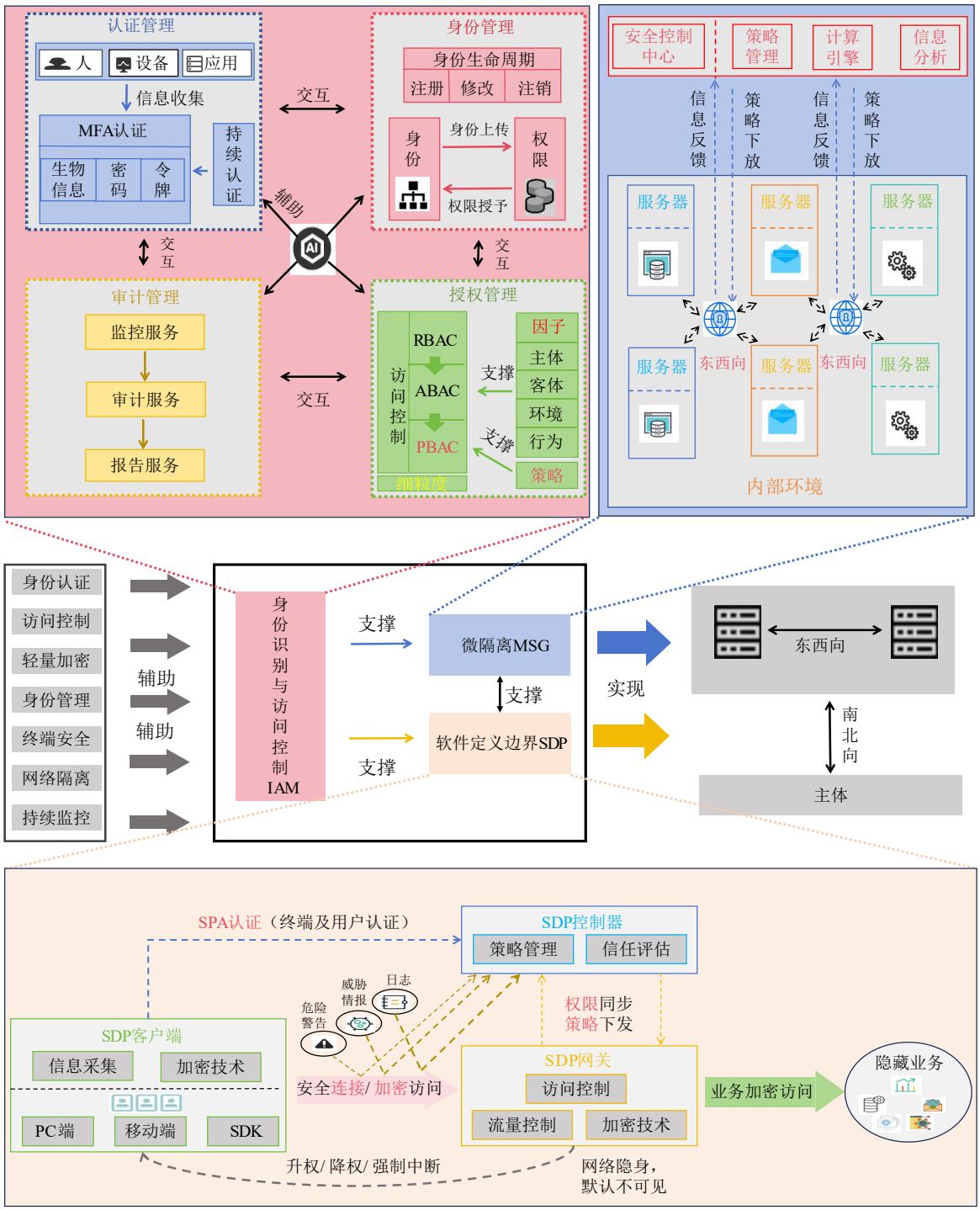

图2　ZTA图

份. 其次,IGA模块根据认证结果分配相应的权限,允许用户访问受保护的资源. 权限的具体确定由授权管理模块负责,通过多种访问控制方式为不同身份分配相应的权限. 最后,为了实现增强的IAM功能,ZTA还包括审计管理模块. 该模块收集来自多种渠道的反馈信息,实现对主体访问资源的动态授权. 与传统的IAM相比,AI技术的引入使得IAM能够更加灵活地应对复杂的访问场景. 在AI的辅助下,各模块能够更高效地分析和处理大量数据,实现细粒度的权限控制和动态策略调整,从而提升整体的安全性和管理效率.

(2)MSG

MSG通过将数据中心业务单元进行分组以限制横向移动,降低内部威胁风险.传统数据中心是基于VLAN进行子网划分,粒度比较粗,而MSG采用最低权限原则,为数据中心的内部流量提供精细化控制,避免单一系统被攻破后波及其他系统.

MSG控制东西向流量,部署在数据中心内部.MSG通过网络隔离将数据中心划分为多个微小计算单元,形成多个节点.节点间的资源访问需要经过MSG网关的认证,安全控制中心会收集流量信息并进行分析,基于分析结果下发动态策略.可信的MSG网关会根据这些策略进行访问控制,未通过认证或不具备权限的访问请求会被拦截.这种精细化控制有效防止单一系统被攻破后影响其他系统,减少内部威胁.

(3)SDP

SDP是一种创新的网络安全模型,通过构建隐藏且动态的网络边界,将受保护资源完全隐藏在内部网络中,避免其暴露于外部威胁之下,确保只有经过严格验证的用户和设备才能访问内部资源,实现"隐形化"保护[26,27].

SDP与MSG的功能相似,都用于实现流量控制,但各自的侧重点不同.SDP主要控制南北向流量,部署在框架边缘.SDP架构由SDP客户端、SDP控制器和SDP网关三部分组成.SDP控制器负责策略的制定和信任评估,是SDP架构的核心.SDP客户端用于收集访问主体的信息以进行身份认证,而SDP网关充当连接访问主体和受保护资源的桥梁.通过这三个组件的协同工作,SDP可以在每一次访问请求时动态评估并调整网络连接,确保资源的最小暴露和最大安全性.尽管SDP框架与VPN较为相似,但SDP安全性远高于VPN.VPN使所有连接的用户都能够访问整个网络,但SDP不会共享网络连接,且SDP还会同时验证设备和用户,从而使攻击者更难仅凭盗取的凭证入侵系统.

3 LS-IoT场景分析

在LS-IoT环境中,IoT技术的迅速发展带来了设备连接数量的爆炸式增长,形成了庞大而异构化的网络系统[28].IoT的应用广泛,覆盖工业、医疗、智能家居、智慧城市、海洋勘探监测等领域[29].然而,IoT快速发展的同时也带来了严峻的安全挑战,本节将分析LS-IoT的特点及安全需求.如图3所示,相较于传统网络,LS-IoT有着显著的区别与特点,这些特点不仅决定了其应用场景的广泛性,还直接影响了其安全需求和应对策略.

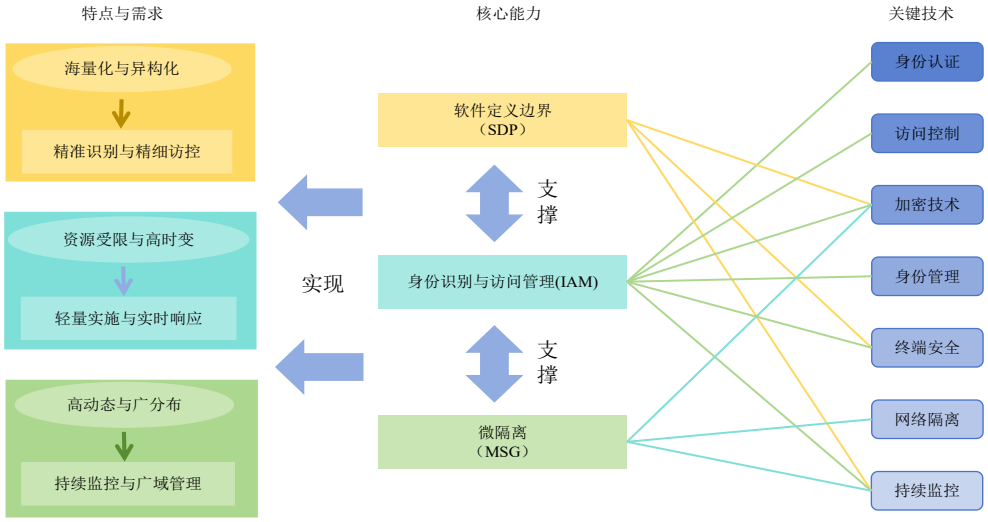

图3 核心能力、关键技术以及大规模物联网特点需求对应图

(1)海量化与异构化

LS-IoT环境的首要特征是设备数量的爆炸式增长与类型的高度异构化[30].成千上万的传感器、嵌入式系统和控制设备各自采用不同的硬件平台、操作系统和通信协议[31,32].这使得传统的统一标准安全策略难以高效适用.面对海量化与异构化的LS-IoT环境,首要的安全需求是实现对用户和设备的精准识别、对资源访问进行精细访控.为实现精准识别与精细访控,需要持续身份认证、细粒度访问控制等关键技术的协助.

(2)资源受限与高时变

IoT设备在多数应用场景中受限于计算与存储资源,且能耗敏感.在某些关键领域,如工业控制、远程医疗或自动驾驶,实时性要求甚至达到毫秒级响应.这就要求安全机制必须足够轻量、快速,以适应资源受限条件并实现近实时的安全响应[33~35].轻量实施与低延迟处理成为应对受限性和高实时性需求的重要准则,因

此需要轻量高效的技术方案[36,37].

(3)高动态与广分布

LS-IoT环境还呈现出高度动态与地理分布广泛的特征,设备频繁接入或离线,网络拓扑与通信模式频繁变化,且设备可能分布于跨地域、多域乃至跨组织的广域场景中. 在如此复杂的环境下,传统的固定安全管理策略必然滞后[38,39]. 相应地,安全需求转向了持续监控与广域管理,即通过对网络状态和设备行为的实时观测和策略动态调整,保障安全措施能够与环境变化同步演进. 动态性IGA相关等技术是满足LS-IoT持续监控与广域管理安全需求的必要条件.

(4)ZTA关键技术

在LS-IoT环境中,有效地实现ZTA的核心能力需要综合多种关键技术. 这些技术不仅要支持ZTA的基本原则,还必须适应IoT特有的广分布、异构性和动态性需求. 身份认证是IoT设备面临的主要挑战之一. 身份认证技术通过简化认证过程和优化算法,使其适合异构化的环境. 例如,分布式身份认证机制采用区块链等分布式技术,建立去中心化的身份认证体系,避免单点故障,提高系统的可靠性和安全性[40]. 访问控制同样需要灵活且精细. ZTA中细粒度的访问控制,根据用户或者设备的属性、行为以及环境等因素动态决定访问权限,确保最小化权限原则的实施. 结合机器学习和AI技术,自动化的策略管理能够动态调整权限,提高管理效率和安全性. 加密技术的优化也是关键研究方向. 轻量级加密算法在保证安全性的同时,降低了计算和能耗开销,适合IoT设备的性能需求. 为防止攻击在网络内部的横向扩散,网络隔离技术同样至关重要. 例如,网络功能虚拟化通过将网络功能虚拟化并部署在通用硬件设备上,实现了灵活的网络安全功能部署和管理. 利用网络功能虚拟化,可以实施细粒度的MSG策略. 通过大数据分析和机器学习算法,对设备行为和网络流量进行实时监控和分析,自动识别异常行为和潜在威胁. 一旦检测到异常,安全事件响应机制能够迅速采取措施,如隔离设备、调整安全策略等,防止威胁扩散.

考虑到LS-IoT海量异构、资源受限、高时变、高动态以及广分布的特征及其安全需求分析,本研究进一步凝练出八大痛点,包括海量身份认证精度低、权限分配粒度宽泛、策略可解释性差、跨域身份映射弱、边端资源开销高、南北向链路易被窃听、东西向横向渗透快、规则爆炸与冲突. ZTA依托八大关键能力:持续身份认证、动态访问控制、轻量级加密、IGA、终端安全、网络隔离、终端安全防护以及持续监控,并叠加LLM、生成式AI、XAI、边缘计算、后PQC等前沿增益技术,能够有效解决LS-IoT存在的痛点问题. 表3直观展示了“痛点—基准关键技术—前沿增益技术”的对应关系,各痛点与其在基准关键技术和前沿增益技术的具体解决思路,将在第4节和第6节分别详细阐述.

ZTA凭借其“永不信任,始终验证”的理念,能够有效满足LS-IoT的多方面安全需求. 传统的PBS难以有效管理所有设备的访问权限. SDP通过构建动态、隐匿的网络边界,确保只有经过严格验证的设备和用户才能访问特定资源,为多样化设备提供了有效的边界控制,防止未授权访问. IAM通过多因素认证机制,确保每个设备和用户的身份真实性,并支持基于策略和细粒度访问控制,实现对IoT设备和用户的精细化权限管理. IoT设备之间的通信频繁且复杂,MSG通过将网络划分为多个独立的微段,每个微段之间的通信均需经过严格的安全策略验证,这种机制特别适用于保护IoT网络内部的设备交互,确保一旦发生安全事件不会影响全局系统. 因此,新兴的ZTA及其核心能力非常契合LS-IoT的特点,能够满足其安全需求.

4　关键技术

在零信任安全架构中,关键技术的实施直接关系到核心能力的实现. 尤其是在LS-IoT环境下,设备数量庞大、类型多样且资源受限,使得单一安全技术难以满足其需求. 如图4所示,本节将详细探讨ZTA在LS-IoT中的七大关键技术支撑,包括持续身份认证技术、动态访问控制技术、轻量级加密技术、IGA技术、终端安全技术、网络隔离技术以及持续监控技术. 其中身份认证、访问控制、IGA技术等与IAM模块密切相关. 终端安全、网络隔离等在SDP和MSG的具体实施中不可或缺,加密技术和持续监控技术贯穿整个架构,它们共同支撑图2所示ZTA的三大核心能力运转,并为LS-IoT提供持续、动态、细粒度的安全防护.

4.1　持续身份认证技术

在LS-IoT场景中,节点类型多达百万级,且上线/离线节奏呈秒级波动. 传统的一次性凭证无法覆盖“海量异构+高动态”的双重挑战. 同时,大量低功耗终端又承受不起频繁重认证的计算负担. 持续身份认证通过把合适的指标转化为随时间衰减的动态信任分数,能够在会话全周期内实时进行身份风险验证. 在图2所示的ZTA中,持续身份认证技术主要对应IAM模块,并与SDP控制器所需的动态策略评估部分相互配合. ZTA的“永不信任,始终验证”理念在LS-IoT中需要通过持续身份认证来落地. 在实践中,持续身份验证往往依赖两大关键思路:其一是通过量化信任分数,对身份可信度进行动态、细粒度的管理,从而实现灵活的权限决策;其二是结合多种认证策略(如基于时间、生物特征、行为特征、风险评估等)持续监测并验证用户和设备的身

表3 LS-IoT主要痛点与技术对策映射表

核心能力	主要痛点	基准关键技术	前沿增益技术
IAM	海量异构设备 身份认证精度低	持续身份认证:通过多模态设备指纹与动态信任评分,保证会话期间身份始终可验证	生成式AI:根据原始数据特征合成数据或补全缺失特征 LLM:利用用户/设备的数据特征与合成特征进行推理融合
	权限分配粒度宽泛	动态渐进式访问控制:RBAC初映射→ABAC细化→风险-阈值调整,将权限细粒度下沉到会话级	生成式AI:自动生成策略片段 LLM:推理校验策略片段,自动调整差异补丁,确保权限随风险动态变化
	策略可解释性不足	反馈式访问控制:结合XACML规则库,把决策阈值与上下文绑定	XAI:每次授权输出特征重要度
	跨域身份映射准确性低	分布式IGA:在各域保留身份映射表	LLM:根据跨域特征进行推理 生成式AI:利用推理结果,输出/更新“身份-能力-域策略”映射表
SDP	南北向通信链路易被窃听	轻量级加密:对长链路通信数据进行分阶段轻量级加密,保证数据机密性	后PQC:为长期数据生命周期提供抗量子保护
	边-端资源受限	轻量级模型与加密算法:减少端侧计算/能耗	边缘计算:将高开销决策与策略缓存下沉到边缘设备
MSG	东西向横向渗透	网络隔离:信任分数驱动的微分段按节点实时划域并更新访问流表	LLM:对流量模式做异常报告,检测未知模式并输出策略,自动生成流表补丁,实时下发
	分钟级拓扑变化 规则爆炸与冲突	多级分层访问控制:先基于角色粗粒度过滤,再基于属性细粒度过滤,减少策略冗余与冲突	LLM:语义压缩相似策略聚合成“元策略” 生成式AI:根据“元策略”生成具体的策略集合

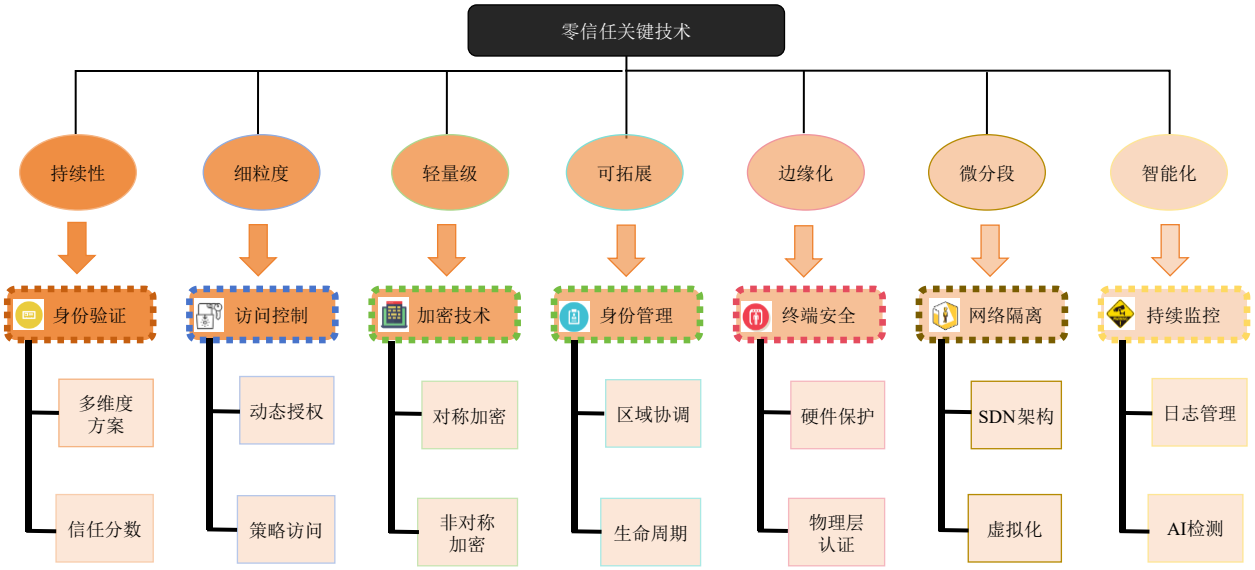

图4 ZTA关键技术

份合法性,从而在各种场景和条件下有效降低安全风险. 以下将分别介绍动态信任分数评估与认证机制,以及多元持续认证策略在LS-IoT中的应用.

4.1.1 动态信任分数评估与认证技术

在LS-IoT环境下,单纯依靠预定义规则的标准认证机制往往过于僵化,难以适应动态多变的网络生态. 此时,基于信任分数的认证方法应运而生. 其核心思想是在认证过程中对多个维度的信任指标(如设备属性、历史行为、环境上下文、网络状态等)进行量化,并将这些分值综合为一个动态变化的信任分数. 如图5所示,基于分数的认证流程分为三步,信任指标的选取与量化、综合信任分数的计算、阈值与信任值的优化与判定. 值得注意的是,阈值的设定也需要考虑IoT设备的特性和安全需求.

信任指标的选取是认证的第一步,选择不当可能导致信任评估不准确,使认证无效. 常见的信任指标维

度包括主体、客体、环境、行为以及物理实体. 通常选取指标时需遵循5大原则,分别是目的性原则、完备性原则、可操作性原则、独立性原则以及显著性原则[41]. 对于IoT设备,信任指标的选取需要考虑设备的资源受限性和多样性. 常用的信任指标包括设备的硬件特征[如媒体访问控制地址(Media Access Control, MAC)、设备指纹等]、通信行为特征(如数据包大小、通信频率等)、环境特征(如地理位置、网络环境等),以及历史行为(如过去的访问记录、异常行为等). 在IoT环境中,物理实体的特征也尤为重要,许多设备可能具备独特的物理属性,如传感器读数、硬件时钟偏移等. 信任指标的量化方式视指标特性而定,主要包括二元或多元方式. 量化后,将各维度属性的量化结果进行存储,方便信任分数的计算与优化.

综合信任分数的计算是认证的核心步骤,不同计算方案不仅决定着身份验证的正确率,还决定认证的效率. 一般首先进行初始分数的计算,例如,通过线性加权法,模糊逻辑或者贝叶斯推理方法得到初步的信任值. 其次,对各维度属性进行权重分配,不同维度信任值的权重影响综合信任分数. 权重确定方法主要分为主观确权和客观确权. 主观方法(如层次分析法)灵活性高,适合根据设备特性和场景进行调节;客观方法(如熵权法)基于样本数据更具科学性. 最后,通过确权算法后生成综合信任值.

阈值与信任值的优化与判定是认证的最后一步. 上一步得到的综合信任值可通过时间窗口衰减机制、异常行为惩罚机制与上下文感知检测机制等方法优化或修正分数,从而得到最终的信任分数. 在一般情况下,阈值是根据经验设定的,是固定的. 但由于IoT设备可能频繁加入或离开网络,且安全威胁多样,静态的阈值可能无法适应实际情况. 动态阈值与信任值的生成类似,针对不同的应用场景的安全需求,通过基于环境感知[42]等方法不断优化初始阈值,生成此阶段最合适的阈值. 最后通过与信任分数的比较可以决定是否允许访问.

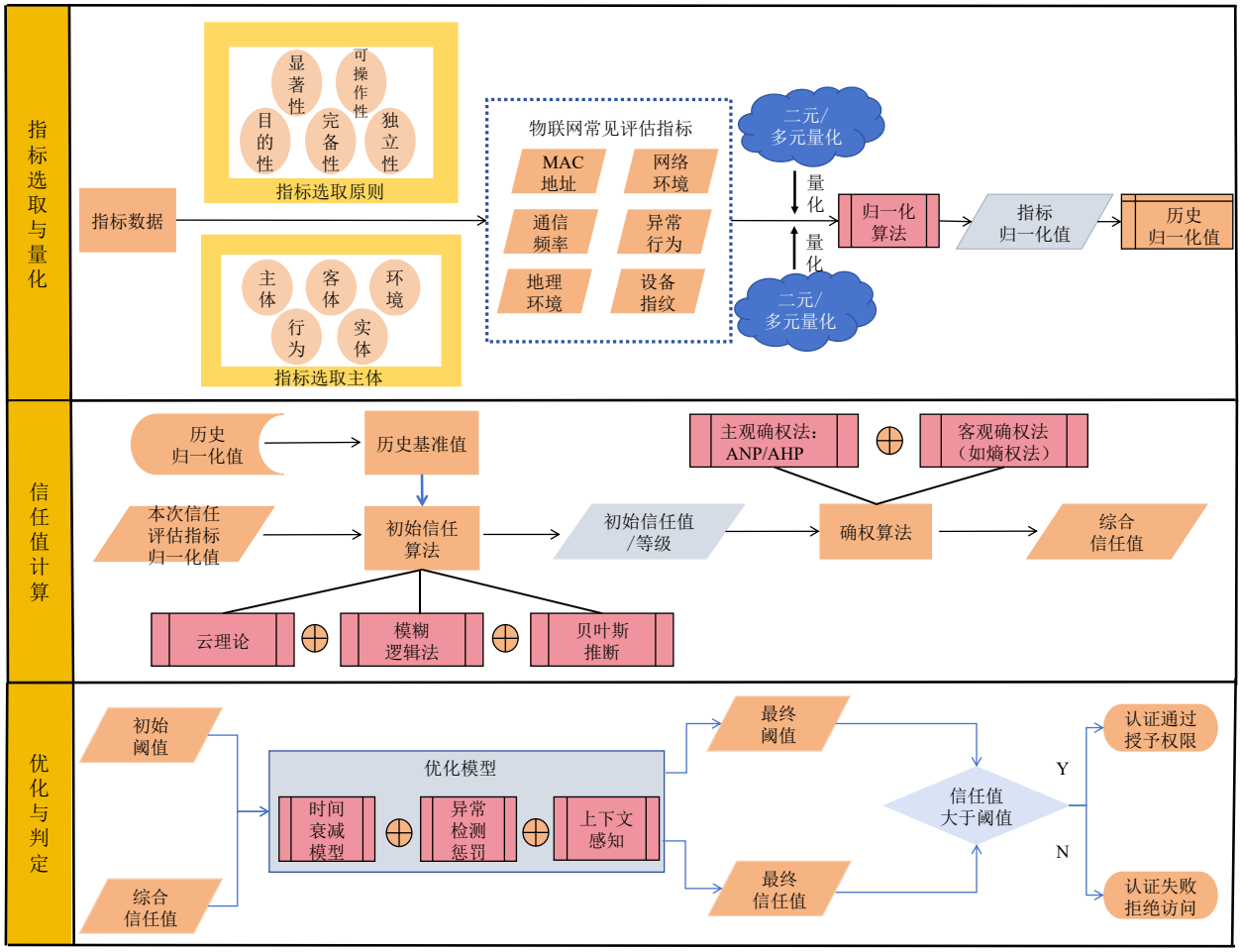

图5　基于分数的认证流程图

针对IoT环境,已有部分学者提出具体有效的基于分数的认证方案. 例如,文献[43]提出了一种面向IoT环境的基于GAme理论的ZEro Trust轻量级认证框架,该框架将用户与设备的信息作为输入,然后设计基于

动态博弈理论的信任策略对访问者进行评分，根据评分结果决定是否允许访问. 通过博弈论模型，对设备的信任分数进行动态调整，平衡安全性和资源消耗. 此外，最常用的是基于生物信息的信任分评估方案. 例如，文献[44]提出了一种基于血压、脉搏等生理信息的评分方案以及文献[45]介绍的一种基于手势、步态的生物行为的评分方案. 文献[46]提出了一种基于风险机制的评分方案，首先根据主体用户的身份、操作环境、资源位置等信息评估风险值，然后根据风险值决定身份是否合法. 文献[47]提出了一种基于可验证随机函数和声誉投票的随机声誉投票机制和区块链方案，同时，引入节点信用评分机制，动态评估节点信用. 上述基于信任分数的认证方案在LS-IoT中具有重要意义. 它们能够在保证安全性的同时，兼顾设备的资源限制和网络的可扩展性.

为便于工程落地，算法1按“指标量化→加权求初分→时间衰减与异常惩罚修正→动态阈值优化→决策判定”五步动态更新信任值，该算法可直接嵌入边缘节点或认证网关.

算法1 动态信任分数计算与认证算法

输入：k个特征$I=(i_1,i_2,\cdots,i_k)$，指标权重$w=(w_1,w_2,\cdots,w_k)$，历史信任分数T_{prior}，基础信任阈值θ_{base}，时间衰减因子$\alpha\in(0,1)$，异常惩罚因子$\beta\in(0,1)$，二次认证阈值系数$r\in(0,1)$

输出：(T_{new},decision):更新后的信任分数与认证决策

步骤1：指标量化

for n =1 to k do

针对特征i_n，利用模糊逻辑或贝叶斯概率等量化方法计算多级信任分数q_n;

end for

步骤2：计算初始综合信任分数

for n =1 to k do

计算特征i_n的信任分数$Q_n=q_n\times w_n$;

计算总信任分数$Q_{n,\text{sum}}=Q_{n,\text{sum}}+Q_n$;

end for

计算瞬时信任分数$Q_{\text{instant}}=Q_{n,\text{sum}}/W_{\text{sum}}$;

步骤3：动态修正信任分数

根据时间衰减因子α和历史信任分数T_{prior}更新时间修正信任分数T_{cor};

if 存在异常 then

根据异常惩罚因子θ和瞬时信任分数Q_{instant}更新异常修正信任分数Q_{cor};

end if

融合T_{cor}与Q_{cor}更新最新信任分数T_{new};

步骤4：动态阈值优化

根据当前环境调节基础阈值θ_{base}得出动态阈值θ_{dyn};

步骤5：决策判定

If $T_{\text{new}}\geqslant\theta_{\text{dyn}}$ then

decision=认证通过;

else if $T_{\text{new}}\geqslant r\theta_{\text{dyn}}$ then

decision=二次认证;

else

decision=拒绝访问;

end if

4.1.2 多元持续认证技术

除了依赖信任分数进行动态调节，ZTA中的持续身份验证还需要从多种策略入手，以适配不同应用场景和设备特征. 在LS-IoT中，持续认证不再局限于最初登录时的身份校验，而是要求在整个访问期间甚至设备运行全生命周期中不断验证. 本节将详细介绍符合上述特点的持续认证方案，包括基于时间、基于生物特征、基于风险和基于活动分析的认证.

基于时间的认证是持续认证中最早应用的方案之一. 最典型的例子是时间同步的一次性密码. 在该方案中，系统和用户设备之间保持时间同步，通过短时有效的一次性密码来进行认证. 这种方式简单有效，但其适用性局限于安全要求较低或用户行为较为稳定的场景. 除了时间同步的一次性密码，学术界也提出了更先进的基于时间序列的轻量级认证方案. 例如，文献[48]提出了一种轻量级的持续认证方法，利用预定的伪随机二进制序列进行认证. 当用户的访问时间序列与系统中的唯一序列完全匹配时，才允许其访问. 文献[49]也提出了一种基于时间访问序列的轻量级持续认证，然而，时间序列所需要的种子的获取方式有所不同. 文献[48]基于支持向量机(Support Vector Machine，SVM)将多种不同信道特征转化为时间序列种子，而文献[49]从近期序列拓展池中生成种子，然后从码本矩阵中获得时间序列，这种方式避免了信道的间歇性的缺点，具有更高的鲁棒性.

基于生物特征的认证技术是持续认证中的核心领域之一. 在LS-IoT中，除了设备的存在，“人”也是需要纳入认证的一个重要方向，认证方案利用用户独特的生理或行为特征进行身份验证[50]. 该技术分为两类：基于生理特征和基于生物行为特征. 基于生理特征是指通过血压、脉搏等生理数据进行持续监测，例如，基于脉搏的持续认证方法. 该类方案准确性较高，但对设备的传感器要求较高，从而增加部署成本[44]. 基于生物行为特征是指通过用户的行为(如步态、触摸手势等)进行身份验证. 基于步态和手势的持续认证方案，能够持续跟踪用户行为特征，实现身份认证[45]. 相比生理特征，行为特征认证更加适合大规模应用，并且在用户体验上具有一定的优势. 然而，生物信息认证的主要挑战在于数据隐私保护. 如果生物数据被窃取或篡改，那么可能会导致不可逆的安全问题. 因此，在ZTA下，生物

特征认证需要与隐私保护机制相结合,确保数据的安全性.

基于风险评估的认证技术是近年来应用较为广泛的持续认证方式. 该技术通过实时评估用户的操作环境、行为特征等,生成风险分数,并根据该分数动态调整认证策略. 文献[46]提出的基于风险机制的认证方案,通过风险分数触发不同级别的认证方式. 当风险分数中低风险时,系统可以进行二次验证,而当为高风险时,直接判定非法. 此类方案在安全性和灵活性方面表现优异,尤其适合复杂的动态IoT环境.

基于活动分析的认证是ZTA中最前沿的持续认证技术之一. 在IoT环境中,设备的正常活动行为具有一定的规律性. 通过对设备与用户行为、网络流量、系统活动等进行实时监控和分析,系统能够识别潜在的威胁并进行动态调整. 机器学习和大数据技术可用于建立用户行为的基线模型. 一旦检测到异常活动,即可降低信任分数,系统会自动触发进一步的认证措施. 活动分析认证的优势在于无需频繁打断用户操作,能够提供较为平滑的用户体验.

这些多元持续认证策略既可以独立部署,又可以有选择地组合应用. 例如,对于资源受限的传感器节点,可采用轻量级的时间序列匹配或风险驱动策略;对于医疗设备或关键工业设备,则可以结合生物特征与行为分析以增强安全性与可靠性. 基于此,本文设计了一种基于时间、行为、风险评估和活动分析的多元持续认证算法,在会话循环中定期收集上下文数据,分别调用时间、生物、风险、行为等策略并行评估. 若存在一个策略判定认证失败,则中断会话;只有全部策略判定认证成功,会话才能被认定为正常. 此外,为减少系统将合法用户判定为非法用户的概率,本算法设定:若多数策略判定可疑,则触发“额外认证”,重新给予用户二次认证,从而以降低误报率. 具体实过程见算法2.

算法2　多元持续认证

输入:当前会话标识sessionID,监测时间间隔或事件触发周期Δt,
多元策略集合S={基于时间,基于生物,基于风险,基于活动}
输出:decision ∈{正常,额外认证,异常中断}
while is Active (sessionID)　do
步骤1:收集上下文数据
　　收集会话上下文数据ctxData;
步骤2:多策略评估
for strategy∈S do
　　逐个策略评估ctxData,其结果f记录在列表authFlags中;
end for
步骤3:综合判断
if 存在f=fail then
　　decision =异常中断,会话结束;
else if 所有f=success then
　　decision=正常;
else if authFlags中f=suspicious数量过半 then
　　decision=额外认证;
end if
步骤4:等待下一个检测周期Δt
end while

4.2　动态访问控制技术

ZTA的核心理念之一是“最小权限原则”,访问控制通过某种途径显示准许或限制主体对客体访问的能力范围,防止非法用户的入侵和合法用户误操作,从而保证系统资源的安全受控. 访问控制技术是IAM模块细粒度授权落地的核心实现手段之一,也可与MSG和SDP结合,对外部南北向与内部东西向流量加以严格限制. 由图2可知,该技术贯穿ZTA各个核心领域,为ZTA在LS-IoT环境中的“最小权限原则”提供了技术支撑. LS-IoT业务链条长、层级深,一旦越权访问发生,极易触发级联事故. 同时,节点属性与运行环境随任务快速切换,使静态角色或白名单难以及时反映当前风险. 融合多种经典访问控制的多级自适应授权的方式,能够实现“粗粒度角色先筛、细粒度属性再限”的细粒度控制. 将实时信任分数及环境上下文共同作为授权因素,实现风险阈值动态调整的渐进授权策略,能够进一步确保最小权限. 当前常用的访问控制技术主要有三种,分别为基于角色的访问控制(Role-Based Access Control,RBAC)模型、基于属性的访问控制模型(Attribute-Based Access Control,ABAC)模型以及基于策略的访问控制(Policy-Based Access Control,PBAC)模型. 然而,简单应用这些模型难以满足LS-IoT环境下“最小权限”的需求,需要在原有技术的基础上进行优化,实现更细粒度和动态的访问控制.

4.2.1　经典访问控制技术

为了构建符合ZTA的安全体系,有必要对现有的访问控制模型进行分析和评估. 以下将探讨RBAC、ABAC以及PBAC模型.

(1)RBAC模型

RBAC模型是一种较为传统的访问控制模型,它通过预定义的角色来管理用户对系统资源的访问权限. RBAC模型的核心概念是信息技术(Information Technology,IT)权限分配给角色而不是直接分配给用户,间接层级可以提供更简单的安全管理,简化了权限管理的复杂性[51]. RBAC模型广泛应用于企业信息系统、医院、简单web环境等场景,特别适合权限需求较为固定的环境[25,52]. 然而,在LS-IoT环境中,RBAC模型主要依赖预先定义的静态角色和权限分配,这在面对设备状态频繁变化、需求多样化以及安全策略实时调整等复杂场景时,在灵活性和响应速度上存在一定的局限性.

IoT设备数量庞大且类型多样,手动为每个设备分配角色和权限是不可行的.此外,IoT设备的移动性和环境的动态性也使得基于固定角色的访问控制难以适应.

(2)ABAC模型

ABAC模型通过基于主体的属性来控制对资源的访问,其可以提供更细粒度的访问控制,并支持动态的访问控制决策[53].这种模型使得访问控制更具灵活性和可扩展性,可以更好地满足复杂的访问控制需求.然而,ABAC模型的实施也面临挑战.首先,ABAC模型的计算和管理成本较高,特别是在需要实时处理大量访问请求的LS-IoT中.此外,IoT设备资源受限,可能无法支持复杂的属性评估和策略执行.为了解决这些问题,可以采用轻量级的ABAC模型简化属性集合,或者将策略评估任务转移到边缘计算节点.

(3)PBAC模型

当前最适合ZTA的访问控制模型是PBAC模型,PBAC模型与ABAC模型紧密相连.ABAC模型可以看作是PBAC模型的一种实现形式,策略可以包含对属性的检查.因此,PBAC模型的使用规则中也包含主体、资源、环境等属性,但PBAC模型更关注策略的编写和管理[54].这些策略通常是以可读形式编写的规则,描述了访问用户、可访问资源和访问条件,能表达更复杂的业务逻辑.PBAC模型除了ABAC模型这种实现形式,还包括基于风险的访问控制,其特点是能够实时评估访问请求的风险等级,根据风险动态调整访问权限.基于风险的PBAC模型在检测到高风险活动时,可以自动触发额外的安全措施或限制访问.显而易见,在LS-IoT环境中,PBAC模型的优势尤为突出,其可以通过灵活的策略管理,实现对设备的动态访问控制.其优越性决定了可以应用于多个领域,尤其是在需要基于复杂规则和多种因素管理访问权限的场景.例如,文献[55]提出了将PBAC模型应用到复杂的机器人系统中,一定程度上利用PBAC模型解决了机器人系统存在的安全问题.综上所述,PBAC模型的优点:一是更加灵活,能够实现细粒度与粗粒度结合,并且可以利用多种因素,如用户属性、环境条件、时间限制等;二是可用自然语言设置策略,基于任务或事件等其他不同的场景灵活配置管理[56].表4具体描述了三种访问控制模型的适用性和局限性.

表4 经典访问控制模型对比

访问控制模型	优点	缺点	适用场景
基于角色访问控制[51]	简单易用,易于管理 权限分配集中,便于进行大规模管理 适合静态环境,权限不常变化	灵活性较低,角色需要提前定义; 无法满足细粒度的访问需求; 难以适应动态多变的环境	企业内部层级结构清晰的组织 权限需求较稳定的环境 需要快速设置权限的中小企业
基于属性访问控制[53]	灵活性强,可基于用户属性、环境、资源动态调整权限 支持更细粒度的访问控制 适合复杂多变的环境	实现复杂,管理成本较高; 配置策略可能较复杂,不易理解; 存在性能开销,尤其大规模场景	金融、医疗等对安全要求高的行业 动态用户和数据需求较多的环境 需满足多种复杂访问条件的系统
基于策略访问控制[54]	自然语言设置策略,策略驱动,灵活性较高 能够应对复杂的权限配置和动态场景 更易于实现细粒度访问控制	策略管理较为复杂,需有完善的策略编写与管理体系; 需要更多资源进行策略配置和维护	对权限和数据安全要求严格 应对复杂访问需求的大型系统

4.2.2 细粒度访问控制技术

在“最小化权限”原则的要求下,仅依靠当前常见的粗粒度、静态的访问控制技术无法满足LS-IoT的需求,需要不断优化当前的访问控制技术,才能进一步提高访问控制的适用性.经调研,目前实现更加细粒度的访问控制模型的方法主要有三种,包括基于多级授权的、基于反馈式动态优化的和基于渐进式的访问控制策略.

基于多级授权的访问控制策略通过结合多种访问控制模型(如RBAC模型与ABAC模型)来实现更精细的权限管理.具体而言,首先,定义基础角色;然后,在每个角色的基础上,通过属性进一步细化权限.例如,可以定义一个“传感器管理员”角色,并基于设备的属性(如设备类型、位置、状态)来限定其访问范围.此外,设计多级授权层级,每个层级对应不同的权限级别,初级层级允许基本访问权限,随着用户行为和信任度的提升,逐步授予更高层级的权限.为了减少手动操作,提高管理效率,可以利用自动化工具或脚本,根据设备和用户的属性动态分配角色和权限.在技术实现方面,使用如可扩展访问控制标记语言(eXtensible Access Control Markup Language,XACML)等策略定义语言来定义基于角色和属性的访问策略,并部署一个高性能的策略引擎[如开放策略代理(Open Policy Agent,OPA)]用于解析和执行多级授权策略.同时,引入信任管理系统,实时评估用户和设备的信任度,根据信任度动态调整其权限.这种方法在实际应用中,例如,文献[57]提出的基于信任的动态RBAC模型,通过实时监控设备行为,动态调整其所属角色,从而实现灵活的权限管理,展示了

其在IoT场景中的适用性. 然而,多级授权策略可能变得复杂,难以维护,可以使用模块化策略设计,将策略分解为可重用的模块,并利用可视化工具辅助策略管理. 此外,为了应对大规模环境下的性能瓶颈,可以通过分布式策略引擎部署、缓存常用策略结果以及优化策略查询算法来提升性能.

算法3通过结合RBAC模型与ABAC模型,先完成粗粒度的"角色初映射",再通过属性匹配(设备、用户、环境)细化权限,最后对请求者实时评分,并以双阈值(θ_1,θ_2)实现最小化权限管控.

算法3　多级RBAC-ABAC访问控制算法

输入:访问请求(userID,resID,action,ctx),角色库为RoleDB,属性库为AttrDB,策略集合为PolicySet_XACML,信任评分函数为TrustFunc,动态阈值函数为Threshold
输出:decision∈{允许授权,额外判定,拒绝授权}
步骤1:角色初映射(粗粒度)
　根据userID查询角色库RoleDB获取基础角色baseRole;
if baseRole==空 then
　decision=拒绝,结束;
end if
步骤2:属性细分(细粒度)
　根据资源resID查询属性库AttrDB获取资源属性resAttr;
　根据userID查询属性库AttrDB获取用户属性userAttr;
　读取上下文环境获取环境属性envAttr;
　合并resAttr、userAttr、envAttr得出属性集attrs;
　用baseRole和attrs在策略集合PolicySet_XACML中获取匹配策略strategy;
if strategy==空 then
　decision=拒绝,结束;
end if
步骤3:多级信任门控
　利用TrustFunc(userID,resID,strategy)获得信任分数T;
　利用Threshold(strategy)得到双阈值θ_1(低阈值)和θ_2(高阈值);
if $T\geqslant\theta_2$ then
　decision=允许授权;
else if $\theta_1<T<\theta_2$ then
　decision=额外判定;
else if $T\leqslant\theta_1$then
　decision=拒绝授权;
end if

基于反馈式动态优化的访问控制策略通过实时监控访问请求和环境变化,利用反馈机制对访问控制策略进行动态调整. 具体实施步骤包括部署监控系统,实时收集访问请求、用户行为、设备状态和环境信息等数据. 进而采用基于博弈论[43]或机器学习等方法设计风险评估模型,实时评估每个访问请求的风险等级,并根据风险评估结果动态调整访问权限. 例如,一种方法是基于博弈论的模型将用户和系统视为博弈参与者,定义其策略和收益函数,通过计算纳什均衡点确定最优的访问控制策略,并根据均衡策略的收益期望调整访问控制授权. 另一种方法是利用机器学习模型,通过特征提取、模型训练和实时预测,对访问请求进行风险评估并据此调整权限. 为了确保系统的实时性,可利用边缘计算分布式部署风险评估模块,同时采用数据匿名化、加密传输和访问日志审计等技术,确保数据隐私与安全.

算法4展示了反馈式访问控制的具体流程. 在边缘端,先进行特征提取,随后调用风险评估模型计算风险值R,系统根据本地策略阈值将会话划分为"正常、限制、中断"三级,并根据警告信息,对误报或漏报实时调节阈值并增量训练模型,实现"监控→评估→决策→校正"的闭环优化.

算法4　反馈式动态优化访问控制算法

输入:访问请求(userID,resID,action,ctx),策略风险阈值PolicyRisk(resID,op,Rmin,Rmax),风险评估模型RiskModel,误报/漏报反馈信息Alert
输出:state∈{正常,限制,中断}
步骤1:特征提取
　提取当前事件特征featrue;
步骤2:风险评估
　调用风险模型RiskModel获取风险分数R;
步骤3:策略匹配
　在PolicyRisk中根据resID、op查找对应策略的风险下限R_{min}与上限R_{max};
if $R<R_{min}$ then
　state=正常;
else if $R_{min}<R<R_{max}$ then
　state=限制;
else
　state=中断;
end if
步骤4:策略自优化
for each alert ∈ Alert do
　if alert.type==误报 then
　　将R_{min}与R_{max}同比上调;
　else if alert.type==漏报 then
　　将R_{min}与R_{max}同比上调;
　end if
　更新风险模型RiskModel;
end for

基于渐进式的访问控制策略通过分阶段授予权限,基于用户或设备的行为和信任水平动态调整其权限等级. 具体实施包括根据用户或设备的初始身份和基本属性授予最低权限等级,持续监控其行为,收集相关数据,并基于这些数据评估其信任水平. 信任评分系

统通过定义关键指标,如设备的稳定性、历史行为记录、响应时间和异常检测等,采用加权评分、贝叶斯网络或基于规则的评分算法,对各指标进行综合评估,计算总信任分数,并根据信任评估结果动态提升或降低权限等级. 例如,当设备的信任水平达到特定阈值时,自动授予更高权限;当信任水平下降时,减少或撤销权限[58]. 为了提高信任评分的准确性,可以通过多源数据融合、引入专家知识和持续优化评分算法来实现. 同时,采用实时数据处理技术,如流处理框架 Apache Kafka 和 Apache Flink,减少权限调整的延迟,并结合多因素认证、行为基线和异常检测技术,提高系统的抗攻击能力,防止恶意用户操纵行为数据以提升信任评分.

表5总结了三种细粒度优化策略的关键特点、技术措施和优点. 通过完善和优化上述优化策略,细粒度访问控制在LS-IoT环境中能够更有效地实现“最小权限原则”. 具体的实现细节和技术措施确保了策略的灵活性、动态性和高效性,同时通过性能优化和安全性增强措施,提升了系统整体的可靠性和安全性. 综合应用这些优化策略,可构建一个适应复杂、多变IoT环境的高效、安全访问控制体系,为ZTA提供坚实的技术支持.

表5 细粒度访问控制技术对比

优化策略	关键特点	实施步骤/技术措施	优点
基于多级授权的访问控制策略	结合多种主流访问控制模型,实现更精细的权限管理	定义基础角色并基于属性细化权限,设计多级授权层级,使用自动化工具动态分配权限,采用XACML等策略语言和高性能策略引擎,引入信任管理系统动态调整权限	灵活细致的权限管理;提高管理效率
基于反馈式动态优化的访问控制策略	实时监控与反馈机制动态调整策略	部署监控系统收集数据,利用博弈论或机器学习进行风险评估,动态调整权限,优化算法性能并采用边缘计算,确保数据隐私与安全	动态调整权限;提高安全性和抗攻击能力
基于渐进式的访问控制策略	分阶段授予权限,基于行为和信任动态调整	初始授予最低权限,持续监控行为并评估信任水平,动态提升或降低权限,采用多源数据融合和实时数据处理技术,结合多因素认证和异常检测增强安全性	动态调整权限;提高安全性和抗攻击能力

4.3 轻量级加密技术

在ZTA中,加密技术主要与保护通信安全的需求相呼应. 对数据的加密与解密贯穿南北向、东西向流量的整个过程,是实现“永不信任,始终验证”的关键组成之一. IoT终端普遍资源受限[59],却又面临侧信道窃密的双重压力. 轻量级加密算法成为ZTA中必不可少的技术之一. 本节将探讨适用于LS-IoT环境的ZTA加密技术,重点介绍轻量级加密算法及其在IoT环境中的应用,传统加密算法与轻量级加密算法的对比如表6所示.

轻量级加密算法专为资源受限设备设计,能够在有限的硬件条件下提供有效的数据加密和解密. 加密算法主要分为对称加密与非对称加密. 前者适用于数据量大的传输场景,因为其加密和解密速度较快;后者适用于需要数字签名或密钥交换的场景,尽管其加密和解密速度相对较慢,但在处理简短数据时,其安全性更高. 在IoT环境中,设备可能彼此陌生,无法事先共享密钥. 因此通常先采用非对称加密算法交换对称密钥,再通过对称加密算法实现高速的数据加解密. 传统的非对称加密算法[例如,RSA(Rivest-Shamir-Adleman)[60]]与对称加密算法[例如,AES(Advanced Encryption Standard)[61]];DES(Data Encryption Standard)[62]计算复杂度高,不太适合资源受限的IoT设备. 采用轻量级的非对称加密算法,如椭圆曲线加密(Elliptic Curve Cryptography,ECC)[63],其计算效率更高,适合嵌入式设备和传感器节点,在提供强大安全性的同时,显著降低了计算和存储需求. 在对称加密算法方面,基于可编程门阵列(Field-Programmable Gate Array,FPGA)和65 nm技术可以实现轻量级的AES对称加密算法,通过优化硬件架构和资源利用,提升了加密速度和效率,并且通过并行处理和流水线设计,该方案大幅减少了逻辑资源的占用和功耗,同时保持了AES算法的安全性[64]. 文献[65]介绍了两种轻量级对称加密算法:一是普雷森特(PRESENT)加密算法[66],它是一种专门为IoT设备设计的加密技术采用简洁的置换-代换网络结构和较小的密钥及块大小,通过优化的S-盒设计和低门电路复杂度,PRESENT在硬件实现中占用资源少,进一步降低了功耗和成本;二是扩展可变长度分组密码算法(eXtended Tiny Encryption Algorithm,XTEA)加密算法[67],通过优化的代换-置换结构和硬件友好的设计,实现了在低资源环境中的快速和安全加密. 其低延迟设计和并行处理能力使其能够在实时应用中高效运行,同时灵活的密钥管理提升了资源利用效率,适用于需要高效处理的场景. 它们在设计时兼顾了安全性和效率,尤其适合IoT环境中的大规模数据传输.

轻量级加密算法在面向LS-IoT的ZTA中起着至关重要的作用[68]. 针对大量资源受限的IoT设备,这些算法(如PRESENT和XETA等)能够在有限的硬件条件下

提供高效且安全的加密与解密功能[31]. 它们不仅满足了低功耗和高效率的需求,还确保了数据传输的安全性和实时性,从而支持IoT环境下的ZTA,保障系统的整体安全和可靠运行.

表6　轻量级加密与传统加密对比

加密算法类型	算法名称	计算效率	密钥生成时间	优点	缺点	适用场景
传统加密算法	AES[61]	较高	中等	安全性高,广泛应用,标准化支持	计算复杂,功耗较高,适合资源充足的环境	大规模数据加密,金融、银行、政府应用
	DES[62]	较低	快速	简单,计算速度快	密钥长度短,安全性差	已不推荐用于高安全性场景
	RSA[60]	低	慢	安全性高,支持数字签名和密钥交换	计算复杂度高,密钥长度长,处理速度慢	数据传输安全、数字签名、证书颁发
轻量级加密算法	PRESENT[66]	非常高	快速	高效,适用于资源受限设备	安全性不如传统加密算法	IoT设备、嵌入式系统
	XETA[67]	高	快速	能耗低,适合低功耗设备	尚未大规模应用,安全性研究有限	轻量级嵌入式应用、传感器网络
	ECC[63]	较高	较快	密钥长度短,安全性高,适合资源受限设备	实现复杂度高,需要精确的数学运算支持	移动通信、IoT、安全交易

4.4　IGA技术

在ZTA中,IGA与IAM相辅相成,共同实现对海量用户、设备、应用的全生命周期管理. IAM的认证与访问控制技术更侧重“认证与授权逻辑”,而本节的IGA则聚焦“身份生命周期及跨域管理”. 在ZTA中,IGA技术需要具备高效性、可扩展性和轻量级的特性,是实现“最小权限原则”的关键[69]. 身份生命周期管理包括身份创建、维护、监控和撤销,确保只有经过严格认证的身份信息才能够在受控条件下访问资源. 在身份创建方面,首次创建的用户、设备或应用身份,应确保唯一性. 在身份维护方面,随着身份信息在系统中持续存在,其权限可能随着角色或任务的变化而调整. IGA系统需根据实时情况动态调整权限,确保不会授予超出需要的权限. 在身份监控方面,身份的活动和访问行为需要持续监控,任何异常行为都应立即触发安全警报或强制进行额外验证. 在身份撤销方面,当身份不再需要访问资源时,必须及时撤销其权限,防止遗留权限导致的安全风险. 传统的集中式IGA系统在LS-IoT中存在性能瓶颈和单点故障问题,分布式IGA可以将IGA任务分散到网络的边缘节点进行管理[70]. 如利用区块链的去中心化,实现设备身份的分布式管理和可信认证,通过智能合约,自动执行身份注册、认证和撤销等操作,确保设备身份的安全性和一致性.

在LS-IoT环境中,设备可能跨越不同的网络域或组织,联合IGA允许不同域之间共享身份信息,支持设备在多个域中进行认证和授权[71]. 跨平台IGA也带来了用户隐私泄露的风险,目前保护隐私的前沿技术主要有两种. 一种是基于区块链的去中心化技术,允许用户掌控自己的身份信息,而不是由中心化机构管理. 这与零信任的“永不信任、始终验证”理念相契合,用户可以随时验证自己的身份而不依赖第三方[72]. 另一种是联邦学习技术,允许多个实体在不共享原始数据的情况下,协同处理身份认证和授权任务. 这种技术确保了用户隐私的保护,同时可以在多方协作环境中执行安全验证[73]. 例如,文献[74]提出了利用集中式联邦学习解决了传统联邦IGA系统中的隐私暴露的问题. 需在多个不同平台之间协调IGA,以确保用户、设备和应用在任何环境下都能得到一致的安全保护. 为了确保各个系统的安全性,频繁的身份验证可能导致用户体验降低. 其中,单点登录是提高用户体验和操作效率的关键技术之一[75]. 它允许用户使用一组凭据访问多个系统和应用程序,简化了身份验证的过程.

表7总结了IGA技术部署的相关功能,通过采用轻量级身份认证协议、分布式IGA架构、单点登陆等技术,结合区块链等前沿技术,可以有效地实现对IoT设备的IGA,保障网络的安全性和可靠性.

4.5　终端安全技术

随着远程工作和移动办公的普及,终端设备成为用户访问组织资源的主要入口,也是网络攻击的重点目标. 在LS-IoT场景中,大量现场节点暴露在恶劣环境和长寿命周期下,一旦被物理篡改或植入恶意固件,侧向渗透将突破网络层防线. 通过在硬件层嵌入可信固件,并结合实时监测系统,可将“硬防护-软监测-快响应”闭环前移到终端侧,从而大幅度提高终端安全性. 在图2所示的ZTA中,终端安全既与SDP控制器的“访问主体安全性”评估相关,也与MSG中对东西向流量的设备端保护相关,二者协同保障终端侧不会成为网络攻击的跳板. 表8总结了不同终端安全技术的特点与适用性,包括基于物理层的无密钥认证、基于可信平台模块(Trusted Platform Module,TPM)和TrustZone、终端监控等终端安全方案. 这些技术的结合为LS-IoT中的终端提供了更高效的安全保障,同时也更好地满足了ZTA的需求.

表7 IGA技术总结

技术/概念	描述	关键步骤/特点	挑战
身份全生命周期管理	确保最小权限原则,通过对身份的创建、维护、监控和撤销,实现对身份的全面管理	身份创建;身份维护;身份监控;身份撤销	设计轻量级机制,适用于资源受限的设备;处理大量身份数据的复杂性
分布式IGA[70]	将IGA任务分散到网络的边缘节点或通过去中心化方式进行管理.利用区块链的去中心化,实现设备身份的分布式管理	通过智能合约,自动执行身份注册、认证和撤销等操作确保设备身份的安全性和一致性	性能开销大,影响系统效率;实现复杂度高,技术门槛高
联合IGA[71]	允许不同网络域或组织之间共享身份信息,支持设备在多个域中进行认证和授权	支持跨域的IGA和认证,解决设备跨越不同网络域或组织的认证需求	用户隐私泄露的风险;需采用隐私保护技术,如区块链去中心化、联邦学习等
隐私保护技术[73]	在IGA和认证过程中保护用户隐私,防止敏感信息泄露	基于区块链的去中心化;联邦学习技术;匿名认证与零知识证明	技术实现复杂,需专业知识支持;可能增加系统开销,影响性能
单点登录[75]	允许用户使用一组凭据访问多个系统和应用程序,简化身份验证的过程,提高用户体验和操作效率	简化身份验证流程,减少用户频繁身份验证;结合多因素认证,保持高安全性同时提升用户体验	凭据集中管理存在安全风险;需要平衡安全性和便捷性

4.5.1 基于终端物理层特征的安全技术

目前主流的终端认证方案分为基于密钥的和无密钥的物理层认证两类. 基于密钥的认证方案的核心思想是通过共享一个秘密密钥,在密钥的帮助下产生鉴别信号,并将其加到消息信号上. 尽管理论上安全性较高,但在实际应用中,基于密钥的认证可能带来较高的计算开销和延迟. 大部分基于密钥的认证方案需要一个可信的第三方来管理密钥的分发和更新,但是零信任的核心原则之一是“从不信任”,这与零信任网络有一定的冲突性,现已不适合零信任网络.

越来越多的研究开始关注无密钥的物理层认证方法. 这些方法通过利用通信链路或设备的物理属性来进行认证,具备较低的计算开销和网络负担. 例如,文献[76]提出了利用设备硬件的时钟偏移特征实现设备认证;文献[77]提出了利用移动设备常见的扬声器-麦克风系统中声学非线性失真特性实现认证;文献[78]使用令牌和设备的上下文信息来实现对设备的持续认证. 当前新认证方法通过捕捉设备在物理层级的独特性,实现对设备的持续监控与认证. 相较于传统方案,基于物理层特征的认证更加轻量,适合资源受限的IoT设备. 但是在复杂多变的动态通信环境中,单一物理层属性可能难以全面反映设备特性,从而影响认证的准确性和系统的鲁棒性. 因此,越来越多的研究开始探索将多个物理层属性相结合的联合认证方法. 例如,文献[58]利用共形预测器对发射方多种物理层属性进行联合估计和分类,从而实现动态调整设备信任等级. 总体来看,采用无密钥的物理层认证技术为提高IoT终端的安全性、保障网络整体的安全提供了有效途径.

表8 终端安全技术

类别	技术/算法		核心思想	优点	适用场景
基于终端物理层特征的安全技术	基于密钥的认证方案		通过共享密钥进行认证,确保只有知晓密钥的用户才能通信	理论安全性高,广泛应用	传统网络环境,不适合零信任网络
	无密钥物理层认证		利用设备的物理属性进行认证,无需共享密钥	计算开销低,网络负担轻,适合资源受限设备	LS-IoT、动态通信环境
基于终端硬件级别的安全技术	TPM[79]		提供硬件级别的安全功能,增强系统安全性	提供多层次安全保障,保护敏感数据	各类计算设备,保护数据和身份
	TrustZone[80~82]		提供安全世界和非安全世界隔离,保护敏感数据	硬件级隔离,提高整体安全性资源受限的微控制器和IoT设备	资源受限的微控制器和IoT设备
	监控技术	行为分析入侵检测系统[83]	通过数据挖掘监控并分析终端行为,判断入侵行为	能及时发现异常,减少潜在威胁	IoT终端的安全监控
		基于Intel SGX的监控方案[84]	利用硬件安全特性完成终端认证,确保设备身份真实性	提高认证的真实性与完整性,增强系统安全性	需要高安全性认证的IoT环境

4.5.2 基于终端硬件级别的安全技术

硬件安全技术保护设备免受恶意软件和未经授权的访问. 这些技术包括设备管理、漏洞管理、终端防护软件等,有助于确保终端设备的安全性. TPM与TrustZone是常用的硬件安全技术. TPM是一种专门的硬件组件,用于增强计算机系统的安全性. TPM芯片通常嵌

入在计算机主板上,它提供了一组安全相关的功能,如加密密钥生成、存储和管理,确保数据的完整性和机密性[79]. 为了增强终端的安全性,TPM具有安全启动、加密密钥管理、数据保护、身份验证等功能,TPM作为一个硬件安全模块,为终端设备提供了多层次的安全保障,有效防范各种类型的攻击,保护系统和数据的安全性,特别是在防止硬件级别的攻击和保护敏感数据方面.

TrustZone通过在硬件级别提供隔离和安全功能,确保敏感数据和操作免受未经授权的访问和恶意软件的侵害,其最大的特征是TrustZone将处理器分为安全世界和非安全世界,确保安全世界中的敏感操作和数据不会被非安全世界访问或篡改[80]. 这种隔离在硬件级别实现,相较于纯软件解决方案更安全. 利用隔离硬件级执行环境的方式能够有效地提升终端设备的整体安全性. 通过直接虚拟化TrustZone来实现安全操作系统的私有化,防止攻破安全操作系统将导致所有可信应用程序被攻破,避免安全操作系统成为整个TrustZone安全的单点故障[81]. 针对轻量级的IoT设备,特别是微控制器上运行的设备软件的攻击,若IoT设备使用资源受限的微控制器,则很难进行防御. 文献[82]介绍了一种基于TrustZone-M启用微控制器的IoT设备综合安全框架,能够有效地保护资源受限的IoT设备,契合LS-IoT轻量级的要求. 加强对终端的监控同样是保障终端安全的重要手段之一,文献[83]介绍了一套IoT终端安全监控系统,用于持续监控终端的安全情况. 该系统的核心是基于行为分析的入侵检测技术,利用数据挖掘算法,对终端异常数据进行分析,从大量不规则、无序的数据信息中找出数据之间的规律;利用改进后的算法提取各种终端行为的特征模式,从而判断入侵行为,实现终端入侵检测. 文献[84]提出了一种基于Intel SGX的远程终端监控与IoT设备认证方案,作为对传统终端安全方案的改进. 这一方案通过利用Intel SGX的硬件安全特性,能够在可信环境中完成终端认证,并及时发现异常,减少潜在威胁对终端安全的影响,确保设备身份的真实性与完整性.

4.6　网络隔离技术

海量异构设备的横向渗透风险与分钟级拓扑变化,使传统VLAN/ACL难以在足够细粒度和足够速度上完成分段. 基于动态信任分的微分段策略,可按节点实时评分将其划入逻辑安全域,并在控制平面一次性下发流表到数据平面;若监控发现异常,仅需重算受影响段的策略而非整网调整,从而实现最小攻击面与快速阻断. 网络隔离是实现MSG核心能力的直接方式,同时为SDP提供南北向入口的“最小暴露面”. 在LS-IoT环境中,往往面临大量设备的接入和复杂的网络结构,网络隔离技术显得尤为重要.

软件定义网络(Software Defined Network,SDN)是一种灵活的网络架构,通过解耦数据传输与控制,实现数据层与控制层的隔离,为细粒度网络隔离的实现提供强有力的支持. 通过SDN,网络可以被划分为多个独立的微分段,每个段内的通信受到严格控制和监控. 这种细粒度的隔离策略能够有效减少攻击面,防止攻击者在不同网络段之间自由移动. SDN集成了先进的流量分析技术,能够对复杂的网络流量进行精确的分类和预测[85,86]. 这使得网络隔离策略不仅能够细粒度地实施,还能根据实际流量情况进行优化,在确保安全性的同时提升网络效率. SDN允许网络管理员通过集中控制器实时调整网络拓扑和流量路径,根据安全策略和实时威胁情报动态实施网络隔离. 例如,当检测到异常流量时,SDN控制器可以立即重新配置网络,阻止可疑流量进入关键资源区域[87]. 然而,SDN的集中管理也可能带来系统效率和单点故障的问题. 为了解决这些问题,优化SDN网络资源分配模型是一种有效的方法,例如,文献[88]设计了一个基于价格的SDN网络资源联合分配模型,该模型能够保障安全性的同时提高系统的效率.

算法5展示了信任驱动的网络隔离算法. 按节点信任分数T_i及相似阈值τ自动划分安全分段,构造访问矩阵$\boldsymbol{A}$区分允许/拒绝,实现对跨段通信的精确管控与实时调整.

算法5　信任驱动的网络隔离算法

输入:节点$i\in V$的信任分数T_i,段$j\in S$的信任分数T_j,信任相似阈值τ,跨段白名单W

输出:访问矩阵$\boldsymbol{A}=\{a\}$

步骤1:基于信任的动态分段

for each $i\in V$ do
　for each $j\in S$ do
　　if $|T_i-T_j|\leq\tau$ then
　　将i纳入段j;
　　end if
　end for
　若所有段都不匹配,则新建段k,将i纳入段k;
end for

步骤2:构建访问矩阵$\boldsymbol{A}=\{a\}$

for each 访问对(s, d) do
　if s和d在同一段中或s和d在跨段白名单W中 then
　　a_{sd}=允许;
　else
　　a_{sd}=拒绝;
　end if
end for

虚拟化和容器化技术提供了更深层次的隔离. 这些技术不仅可以提供基础设施级别的隔离,还能实现应用级别的分离,确保不同应用或工作负载之间的通信安全. 虚拟机隔离允许在同一物理主机上运行多个独立的虚拟环境,每个虚拟机可以有独立的操作系统、应用和网络策略[89]. 虚拟化技术确保即使某个虚拟机被攻破,其他虚拟机也不会受到影响. 而容器化技术进一步细化了虚拟化的概念,允许多个容器在同一操作系统实例上运行[90]. 容器化适用于需要密度和灵活性更高的环境,能够更有效地管理和隔离应用之间的通信.

技术在LS-IoT环境中,网络隔离是实现ZTA的关键手段. 通过微分段策略,结合先进的网络技术,能够有效降低网络攻击风险,保障设备和数据的安全. 未来,随着技术的发展,网络隔离技术将进一步演化,为IoT安全提供更强有力的保障.

4.7 持续监控技术

持续监控技术横跨整个ZTA,向三大核心能力的实现提供可量化证据、向网络层提供策略触发条件,是"永不信任,始终验证"实时落地的基石. 对于LS-IoT而言,海量异构设备与高速高动态数据流的特性进一步凸显了持续监控的重要性. 与传统周期性或被动式的安全策略不同,持续监控强调实时性和主动性,通过不断收集并分析网络、系统以及设备的运行数据,及时发现与处置安全威胁,从而有效保障网络和数据的安全. 本节将介绍几种适用于LS-IoT的持续监控技术.

4.7.1 流式数据处理与实时检测

为了应对LS-IoT中存在高吞吐量的实时数据流,流式数据处理技术能够对来自各类传感器、边缘设备和网络节点的实时数据进行持续监控和分析[91]. 相比批处理模式,流式处理可以在毫秒级或秒级的时间对数据进行筛选、聚合和分析,从而在更短的窗口内发现异常事件或潜在威胁. 例如,在智能交通系统中,利用流式数据处理,对车辆和道路传感器的数据进行实时分析,可以快速定位交通异常、预测路段拥堵,并识别潜在的安全风险. 这不仅提升了告警的实时性,还便于与上层安全策略(如微分段策略)相互配合,做到即时响应.

4.7.2 AI驱动的威胁检测

由于LS-IoT环境的复杂性和攻击手段的多样化,传统基于规则或特征匹配的监控系统已经不足以应对动态和隐蔽的安全威胁. 基于AI与机器学习的威胁检测模型能够从大量异构数据中学习"正常行为"模式,并持续更新和优化侦测规则[92]. 在异常检测方面,无监督算法(如聚类、孤立森林等)可以建立系统或设备的正常行为基线,一旦出现偏离正常基线的行为即可触发告警. 在入侵检测方面,利用深度学习网络,对历史攻击数据进行训练,识别常见攻击模式;并且在新威胁出现时,模型可通过在线学习或增量学习机制自适应地更新检测能力.

在零信任环境下,这类智能化检测不仅可以帮助LS-IoT系统实时发现异常设备或恶意流量,还能将检测结果与动态信任评估及访问控制策略联动,自动触发更高等级的防护措施,实现全方位的防御.

4.7.3 日志管理

日志管理是持续监控的核心技术之一. 部署专门的安全信息和事件管理(Security Information and Event Management,SIEM)系统,收集和分析来自IoT设备、网络设备和安全设备的日志和事件信息,提供全局的安全态势感知.

在集中化管理方面,所有设备及应用的安全日志会被持续归集在SIEM平台,以避免各组件相互独立、信息碎片化的问题. 另外借助大数据分析技术,系统对实时日志进行聚合,将不同来源的事件进行关联,发现隐藏的攻击链条,及时给出全网安全态势,并辅助管理者在短时间内定位安全问题根源. 例如,通过关联多个系统中的日志,可以揭示复杂的攻击路径,从而提高对网络入侵和数据泄露的响应速度. 诸多行业都有严格的合规审计要求,通过SIEM的日志归档与报表功能还可以简化审计流程,并提升监管透明度.

在LS-IoT环境中,海量日志会带来更严峻的存储和运算挑战,对于日记数据的容量高效率压缩也是必不可少的重要技术. 文献[93]提出了一种LogBlock方法,通过预处理日志头和重新排列日志内容来降低日志的重复性,从而提高日志文件的压缩比,相较于传统的数据压缩方式,LogBlock显著提高了日记数据处理速度与降低了数据容量,为海量日志的后续实时分析提供保障.

5 经典应用场景

本章节探讨ZTA在LS-IoT环境中的经典应用场景,包括工业IoT、5G医疗、自动驾驶以及远程办公,展示其如何为这些行业提供强有力的安全保障,并引用思科两篇关于零信任安全的调研报告《Cisco's Guide to Zero Trust Maturity》[94]与《Security Outcome for Zero Trust: Adoption, Access and Automation Trends》[95](以下简称《思科ZTA调研报告》)支持ZTA的实用性与可行性.

图6展示了ZTA在LS-IoT中的层级设计,包含终端感知层、传输层和控制层,结合了云端的数据存储、分析与决策支持. 边缘感知层位于图的底部,负责数据采集和设备交互,成为IoT环境中的数据来源和关键入口. 传输层位于中间,负责将终端层的数据可靠传输到上层. 该层通过网关等网络节点实现多层次、多协议的

数据传输,以保障数据的实时性和准确性. 控制层是系统的核心,包括策略管理、认证与授权. 控制层依托零信任策略,通过动态的授权机制管理对资源的访问. 云端数据分析与决策支持通过AI、威胁评估、大数据分析等模块,对收集到的数据进行深度分析,并基于分析结果进行策略调整.

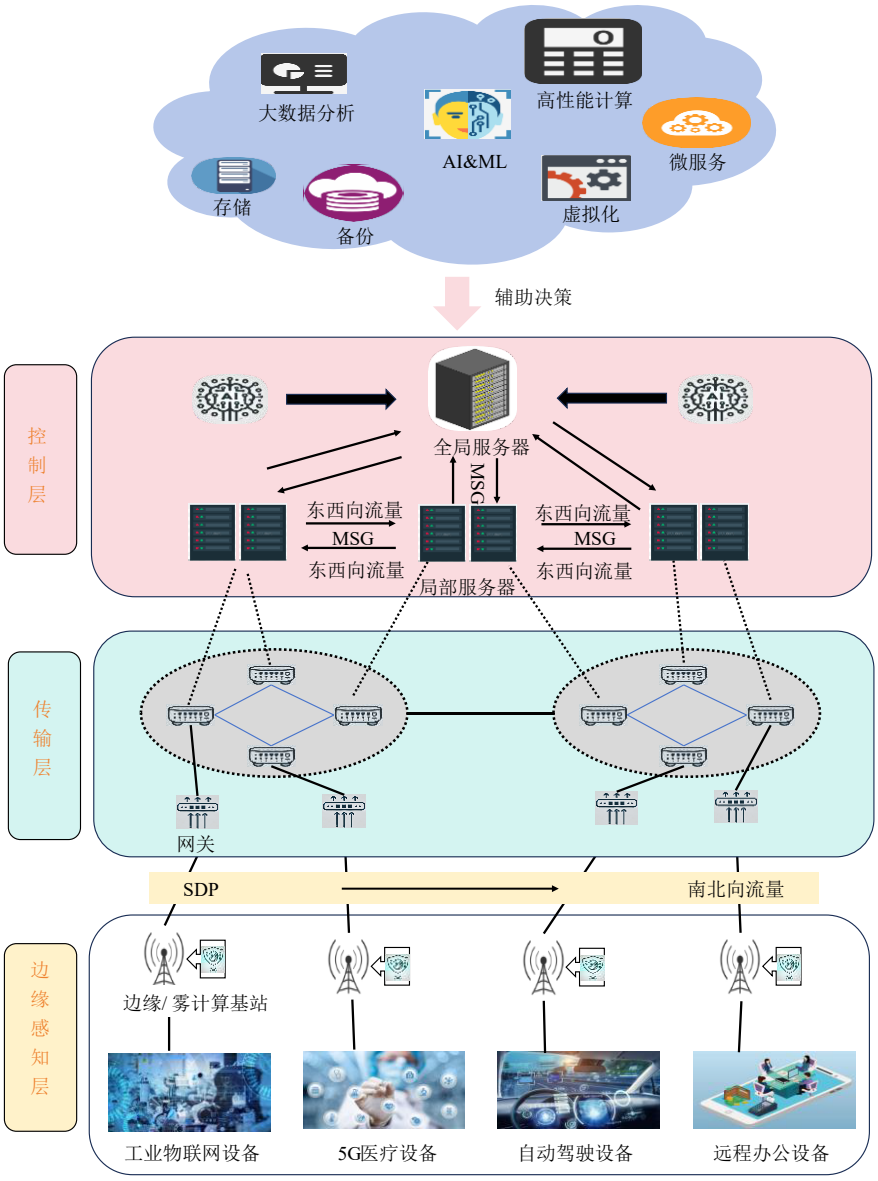

图6　ZTA在典型应用场景中的层级设计

5.1　案例场景:工业IoT

工业IoT将传统的工业控制系统与互联网连接,使设备、传感器和系统能够实时通信,极大地提高了工业自动化和效率. 然而,设备的互联互通也带来了显著的安全挑战,特别是设备认证、数据传输保护和网络攻击的防御方面[96~98]. 由于工业设备的多样性和通信环境的复杂性,传统的边界防御策略难以有效防护.《思科ZTA调研报告》显示,实施零信任"网络和工作负载"支柱后,恶意内部人员攻击的可能性减少了9%,同时,通过自动化与编排,企业适应外部变化事件的能力增加了14%[94,95].

在工业IoT场景中,ZTA可以通过多层访问控制引擎和MSG技术来提高网络安全性. 每个设备和用户都需要经过多因素身份验证,访问权限则根据设备的信任评分动态调整. 采用ABAC模型,系统可以对设备的访问权限进行细粒度管理,仅允许授权设备访问特定资源. 文献[99]通过多层访问控制和基于物理模型的策略优化,在工业IoT企业中部署更安全的ZTA,大幅提升了网络安全性. 西门子在其工业自动化过程中,结合了ZTA,通过细粒度访问控制技术,通过这些措施,

西门子将IT领域的零信任安全措施直接应用于操作技术(Operational Technology,OT)环境,提升了对生产环境的控制和保护能力,确保“最小权限访问”原则[100,101]. 此外,西门子对工厂设备和网络通信进行细粒度的分段,每个工业设备和控制系统都需要经过多层身份验证,确保设备之间的通信是安全的. 通过这种方法,即便某个设备被攻击,也不会影响其他设备的安全运行. 类似地,通用电气Predix平台[102]采用了基于大数据分析的解决方案,实现了对生产过程中持续的数据监控和保护,确保工业环境的持续稳定性和数据的传输安全性.

这些实际案例展示了ZTA在工业IoT中的成功应用,特别是确保关键设备和数据免受网络攻击方面. 《思科ZTA调研报告》表明实施零信任后,数据泄露的成本降低近50%,投资回报率达到191%,安全运营中心的效率提高了90%,成熟的零信任企业比初步实施的企业更有可能实现业务连续性,业务韧性提升可达63.6%[94,95].

5.2 案例场景:5G医疗

5G技术的普及为医疗行业带来了巨大的变革,特别是在远程诊疗、医疗设备互联和大规模数据处理方面. 然而,医疗行业对数据隐私和安全有着极高的要求,患者信息和实时诊疗数据的安全传输是核心问题. 同时,实时数据处理和设备互联带来的安全复杂性也是主要挑战之一. 医疗设备需要在毫秒级别的时间内响应数据,而传统的加密算法可能会增加通信延迟,影响诊断结果的准确性. 因此,需要采用高效的加密算法以确保数据传输在毫秒级内安全传输.

随着网络攻击的大幅增加,医疗行业需要确保系统和设备的安全. 医疗机构通常拥有数百甚至数千台医疗设备,包括植入设备到服务器系统等,都面临多种漏洞. 为确保这些设备的安全,实施零信任安全架构是有效技术之一.《思科ZTA调研报告》显示,ZTA采用轻量级多因素身份验证和基于信任的访问控制模型,确保每个医疗设备和用户的合法性,降低了勒索软件攻击的概率8%[94,95]. 在身份安全方面,世界范围内,很多企业提出了相应的安全方案,如IBM提供了基于零信任理念设计的身份安全解决方案IBM Security Verify[103]. 所有医生的操作请求都必须通过多因子认证和基于行为的动态访问控制模型,这确保了敏感的患者数据在传输和存储时的安全性,贯彻了零信任中的最小权限和持续身份验证原则. 云安全联盟(CSA)发布的《基于ZTA的医疗设备安全》报告[104]旨在指导医疗服务机构如何在医疗设备中实施零信任安全架构,以应对日益增加的网络攻击和设备漏洞. 报告基于零信任成熟度模型的五个支柱:身份、设备、网络、应用程序和数据,详细分析了每个支柱在医疗设备安全中的应用. CSA报告认为,通过识别所有设备、实施访问控制、MSG和持续监控等手段,ZTA可以大大增强医疗设备的安全性,识别漏洞并应用补救措施,是目前最佳的安全方案. 在学术方面,5G医疗也是热门领域,例如,文献[105]探讨了在5G智能医疗环境下,如何通过ZTA提升系统的安全性和隐私保护,详细分析了5G智能医疗所面临的多种具体应用场景,包括但不限于远程会诊、远程手术、远程教学、远程急救以及远程监护,并证明了基于ZTA的安全系统在功能性和性能上均符合预期,能够有效应对5G智能医疗环境下的各种安全挑战. 最后《思科ZTA调研报告》中指出,持续用户验证在减少医疗事故发生概率方面表现突出,特别是在数据安全方面,降低了5.3%的网络数据泄露风险[94,95].

5.3 案例场景:自动驾驶

自动驾驶技术的快速发展使车辆的智能化和互联化成为现实. 然而,由于车辆与外界的通信高度依赖网络,网络攻击带来的安全风险也随之增加. 自动驾驶系统面临的主要威胁包括数据篡改、设备劫持和通信信号攻击. 自动驾驶的核心挑战在于实时数据处理和低延迟通信的安全保障. ZTA需要在不影响车辆响应速度的前提下,确保安全性.

ZTA通过多层身份验证和加密通信确保每个通信节点的合法性. 车辆的每个访问请求都需要经过轻量级多因素身份验证,保障只有经过授权的设备和用户可以访问控制系统[106]. 在自动驾驶过程中,数据的实时性和准确性至关重要,ZTA利用基于行为的动态验证技术来保证车辆通信链路的安全. 通过实时监控车辆的行为和通信模式,系统能够识别并响应潜在的网络攻击,如对车辆通信信号的干扰或篡改. 在科研领域,文献[107]设计了一种适合自动驾驶的ZTA,其利用相关零信任组件实现了身份认证和行为识别,能够准确地识别车辆的各种行动,比如超速、非法驾驶等并通过合理的方式进行提示或阻止,以此来保证更加安全的自动驾驶. 在具体实践方面,亚信安全[108]在自动驾驶和车联网领域积极推进ZTA的应用,旨在提升智能网联汽车的整体安全性. 在车云连接的场景中,亚信安全强调以身份为基石、以场景为框架、以数据为驱动、以访问规制为核心的原则. 通过车端代理、身份中心、安全代理网关、安全运营中心,确保车云之间形成可靠连接. 在车内,亚信安全通过构建车载诊断(On-Board Diagnostics,OBD)、蓝牙钥匙、车机、网关、遥测盒(Telematics Box,TBOX)、域控制器等的访问和刷写零信任环境感知能力,外联车联网零信任管理平台及亚信安全的安全运营中心,构建真正的车内安全防护体系.

5.4 案例场景:远程办公

随着企业业务的扩展,远程办公已成为常态. 然而,员工设备离开企业网络安全边界,传统的VPN技术无法完全保障数据安全,导致信息泄露的风险大幅提升. ZTA为远程办公环境提供了有效的安全保障. 通过SDP技术,每个设备的访问请求都需要通过动态身份验证和访问控制策略. 员工可以通过SDP与公司资源建立安全连接,而每个设备的访问权限会根据实时的信任评估结果进行调整. 同时,ZTA采用终端安全技术实时监控设备状态,防止恶意软件和未经授权的访问. 《思科ZTA调研报告》指出,企业通过实施ZTA,能够将未计划工作的频率减少43%,成本效率提高了47%,显著提升员工的工作效率[94,95]. 在企业应用方面,Google的BeyondCorp专门针对远程办公场景,应用ZTA技术,摒弃了传统VPN,确保员工无论使用何种设备,均能够通过强身份验证和基于风险的访问控制安全地访问公司资源. 如图7所示,展示了BeyondCorp核心组件和具体的工作流程.

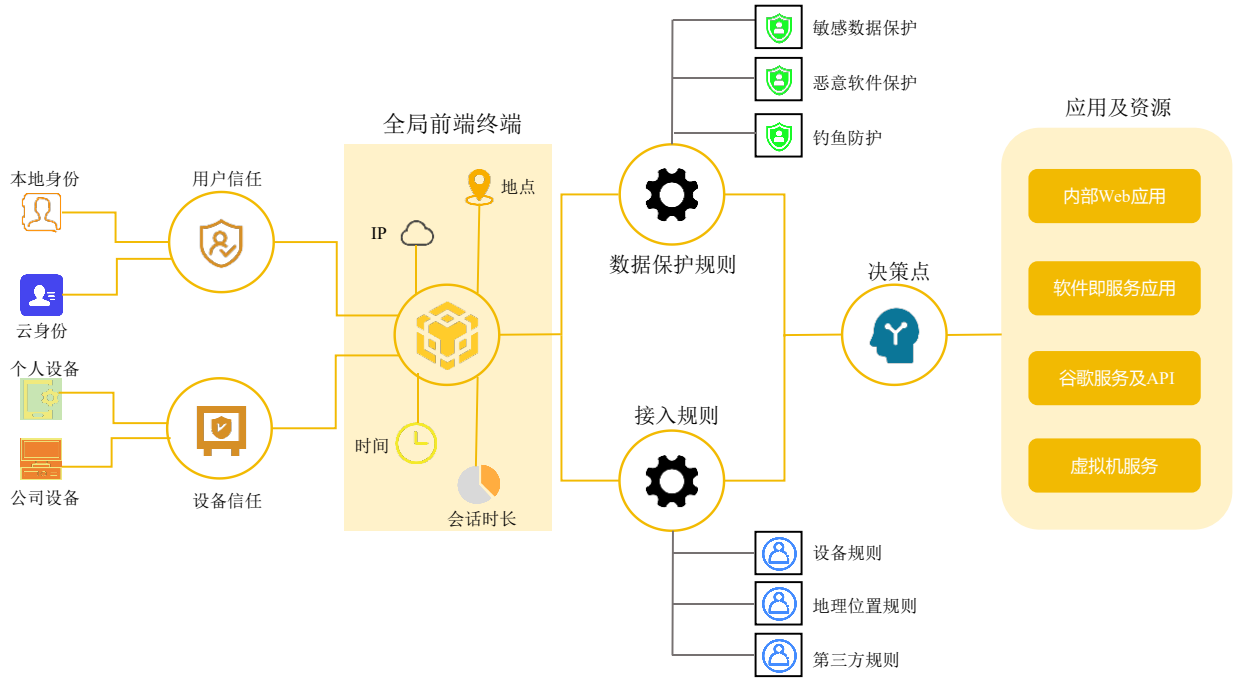

图7　BeyondCorp核心组件及工作原理图[109]

此外,为了减少ZTA部署对正常业务带来的影响,BeyondCorp采用自动化分段式迁移功能模块,逐步将核心模块其他功能迁移到ZTA中[110~113]. 研究表明:通过这种部署方式,技术问题的发生率从0.8%降低到0.3%,与谷歌类似的大规模内部IT变革相比,BeyondCorp导致的支持问题减少了30%[110,111]. 谷歌的BeyondCorp项目为ZTA的部署提供了具体的实践方案,展示了零信任在远程办公场景中的应用效果,确保了灵活办公环境下的网络安全.

6 前沿技术融合

随着网络安全威胁的持续演变,ZTA已经成为确保LS-IoT中数据安全的核心战略. 然而,面对日益复杂的IoT环境,整合最新的前沿技术能够进一步有效解决ZTA的三大核心能力面临的痛点问题. 本节将探讨5项关键的前沿技术,包括LLM、生成式AI、XAI、边缘计算以及后PQC,并分析这些技术如何与ZTA相融合,为未来的网络安全提供更强有力的保障. 图8展示了ZTA应用于LS-IoT的典型网络结构. 在数据采集层中,大量IoT设备实时采集数据并进行初步处理,需利用后PQC技术保证底层数据的传输安全. 在智能分析层中,LLM与生成式AI协同对海量异构数据进行异常检测和策略生成,并借助XAI来实现对模型与策略的可解释性与可验证性. 在策略执行层中,边缘节点作为ZTA的本地执行者,落实访问控制、流量隔离等安全策略,并与云端或其他节点之间的通信采用PQC进行加固. 整合LLM、生成式AI、XAI、边缘计算以及后PQC等前沿技术,可为LS-IoT环境提供更高水准的安全保障与可扩展性.

6.1 LLM

LLM近年来在自然语言处理和智能化决策中表现出了强大的能力,其在ZTA中的应用可大大提升系统在身份验证和动态访问控制中的响应能力. 如图8所示,处于“智能分析层”的LLM,能够在多源异构的IoT数据中快速进行聚合与语义理解,从而识别潜在的安

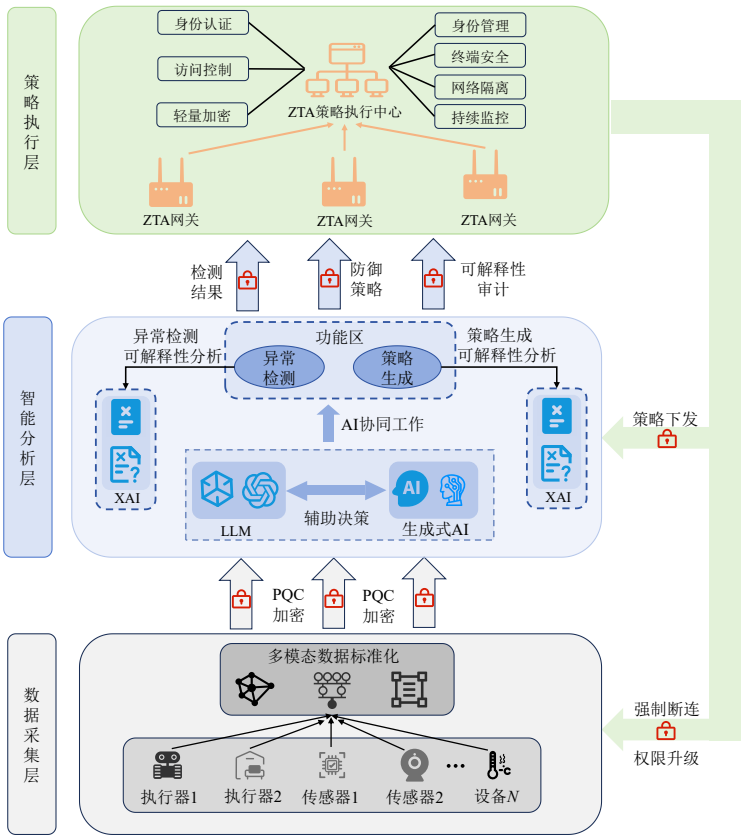

图 8 ZTA 应用于 LS-IoT 的典型网络结构图

全风险，并为后续策略制定提供深度语义支撑.

由于 IoT 设备数据分布广泛且类型多样，使得 IoT 数据特征难以捕获，进而难以准确评估模型安全性和性能[114,115]. LLM 具备强大的数据分析能力，能够从大量的非结构化数据中提取有用信息，可以用于分析海量的 IoT 数据. 此外，LLM 还具备强大的泛化能力，能够在不同场景和环境下应用其训练所得的知识. 这一特性使得 LLM 在应对新型和未知的安全威胁时，能够迅速适应并提供有效的解决方案. 例如，在工业 IoT 中，工业控制设备和传感器实时监测生产环境与设备状态. LLM 可以将这些设备的状态数据整合到一个统一的模型中，并通过历史数据进行设备的故障预测. ZTA 则利用这些模型结果，动态地调整设备的访问权限，确保只允许健康状态下的设备继续访问敏感数据或进行关键操作. 泛化能力还使得 LLM 能够在不同类型的 IoT 环境中保持高效的性能，无需针对每一种新环境进行大量的重新训练，从而显著提升了 ZTA 在多样化 IoT 系统中的适应性和灵活性.

将 LLM 集成到面向 IoT 的 ZTA 中，首先需要考虑 IoT 设备面临资源受限的问题. 因此，需要轻量化模型适配. 目前，一种有效可行的方式为设计基于知识蒸馏的轻量级 LLM（如 TinyBERT）. 通过提取预训练模型的语义特征，该方法能够实现设备行为语义的实时解析. 具体而言，模型蒸馏方法是将一个大型预训练模型（如 GPT-4、BERT）作为教师模型，TinyBERT 作为学生模型，通过训练蒸馏过程，提取出语义特征并针对不同的 IoT 场景优化其适应性，最终生成适合边缘设备的轻量级 LLM 模型. 结合 LLM 后的系统架构如图 9 所示. 在此架构中，LLM 被嵌入 ZTA 的 IAM 模块中，构建动态风险评估引擎. 设备或用户的认证请求首先经过传统 MFA. 其次，由 LLM 分析其历史行为日志，包括访问时间、操作频率、网络流量模式等，生成动态信任评分（0~1）. 最后根据信任评分进行风险决策与授权管理，如触发二次认证或权限降级.

LLM 在“智能分析层”中为 ZTA 提供了高效的异常检测和知识推断能力，而在后续的“生成式 AI”节中，这种语言理解与推断能力可以进一步与自动化防御策略生成结合起来，实现真正的智能安全决策.

6.2 生成式 AI

与 LLM 相同，生成式 AI 同样可在威胁场景识别与风险评估中发挥作用. 例如，通过生成对抗网络（Generative Adversarial Network，GAN）等模型，可模拟多样化的攻击路径，帮助零信任系统对抗潜在威胁. 与此同时，生成式 AI 也能借助大规模参数学习，生成针对特定攻击的安全防御策略. 生成式 AI 具有强大的学习和生成能力，在 ZTA 中有着广泛的应用潜力，尤其是在模拟

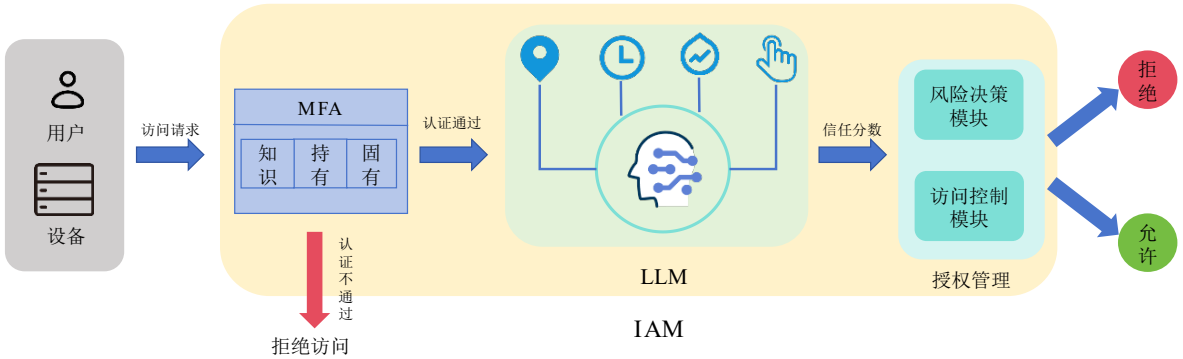

图9　结合LLM的ZTA

攻击、自动生成安全策略以及隐私保护方面,生成式AI展现了其独特的优势.

生成式AI在LS-IoT中的主要作用可以体现在定制化安全策略生成、自动化设备配置与智能模拟攻击与防御上. IoT系统庞大且复杂,传统的手动配置和管理方式已难以应对IoT设备数量的增长和安全需求的复杂性,而生成式AI可以根据不同设备的属性和行为,动态生成定制化的安全策略,帮助ZTA实现高效的自动化管理. 例如,在智能电网系统中,电表和电力控制设备通过IoT技术与中央控制系统连接. 生成式AI可以自动为每台电表生成安全策略,确保每个设备都只访问其所需的资源. 若检测到某个电表的行为异常,则生成式AI能够根据实时数据自动生成一套调整后的安全策略,并与零信任系统协同工作,自动更新设备的访问权限.

利用生成式AI技术可以实现自动化攻防推演与策略优化,具体的技术路线为通过GAN模拟攻击链,构建基于GAN的攻击代理,生成APT攻击路径(如钓鱼邮件→横向移动→数据窃取),输出攻击概率图. 其中,模型的输入为网络拓扑、设备漏洞库、历史攻击数据;输出为攻击路径集合$P=\{p_1,p_2,\cdots,p_n\}$,每条路径包含攻击步骤及成功概率. 最后根据红队模型提出的攻击方式利用强化学习优化ZTA策略,状态空间可以为网络拓扑、设备信任评分、实时威胁情报;将调整微分段策略、升降权限、隔离设备作为动作空间,接着奖励函数可以设置为与安全效益和性能损耗等相关的函数,最后使用PPO(Proximal Policy Optimization)算法在模拟环境中迭代优化策略.

然而,自动生成的安全策略仍可能存在不可预期的误差或不合理之处. 为进一步保障策略的可靠性,下一步需引入XAI,以对生成过程进行审计.

6.3　XAI

随着AI技术在ZTA中的广泛应用,系统的透明度和可解释性成为新的挑战. 设备的安全行为和网络流量非常复杂,传统的AI模型往往是“黑盒”式的,难以解释决策过程. 在“智能分析层”,LLM与生成式AI分别为ZTA提供了异常检测与策略生成功能,但策略的透明度和可解释性却长期是安全领域的一大痛点. 本节探讨的XAI能够使研发者与使用者追溯并理解AI决策背后的逻辑.

XAI可以帮助解释ZTA中的复杂决策过程,尤其是在医疗IoT环境中. 成千上万的医疗设备通过网络传输敏感的患者数据,XAI技术可以对这些设备的行为进行分析,并解释为何将某个设备被判定为存在安全风险. XAI技术可通过可视化的方式解释设备的行为变化,并帮助系统管理员理解该设备的异常操作是如何影响整个网络的安全性的. 在医疗IoT中,数据的高度敏感性和设备的复杂性使得可解释性尤为重要,帮助系统管理员快速识别和响应异常行为,确保患者数据的安全和医疗系统的稳定运行. 在工业IoT中,复杂的生产流程和多样化的设备类型也需要XAI来及时发现和解决潜在的安全问题,以提升整体系统的可靠性和安全性. 因此,XAI不仅提升了ZTA系统的透明度和可信度,还在关键场景中发挥了至关重要的作用,帮助构建更加可靠和安全的IoT环境.

目前,利用XAI透明化ZTA安全决策最常用的方式主要有两种. 一种是基于LIME(Local Interpretable Model-agnostic Explanations)的访问控制解释,根据输入,即设备访问请求,如属性、行为、环境,该模型生成决策解释,例如,“拒绝设备A访问:因其地理位置(北京→纽约)突变,且过去1 h通信频率异常(+300%)”,最后输出可视化决策报告,辅助管理员快速响应. 另一种是SHAP(SHapley Additive exPlanations)值驱动的策略优化,SHAP值可以量化设备属性(如IP信誉、固件版本)对信任评分的影响,若某特征(如固件版本过旧)SHAP值过高,自动触发固件升级策略.

通过XAI,管理员可以更容易地定位异常行为及产生的根本原因,并评估防御策略的合理性,不仅有助于在“策略执行层”落地防御措施,也使得后续的边缘计算节点能够更安心地执行这些策略,形成云-边协同的

完整安全闭环.

6.4 边缘计算

在LS-IoT环境下,边缘设备数量巨大,且这些设备通常具备资源受限的特性.前述的LLM、生成式AI和XAI主要在“智能分析层”发挥关键作用,而IoT应用中的大部分实时决策与策略执行往往需在靠近终端的边缘侧完成.本节将讨论如何结合边缘计算为ZTA提供低延迟和高可扩展性的支撑.利用边缘计算实现分布式可行技术路线可以为构建三级信任锚点架构.在该架构中,端侧利用物理层指纹等完成设备本地轻量级认证;边缘侧设置信任中心(如Edge-ZTA),执行动态信任评估与策略缓存;云端则作为全局策略库与威胁情报中心.具体的协同认证协议如下:步骤一,设备向Edge-ZTA提交认证请求(含设备指纹、行为摘要);步骤二,Edge-ZTA调用本地模型快速验证,若置信度不足则请求云端仲裁;步骤三,云端返回全局策略,Edge-ZTA更新本地缓存.

在具体应用方面,例如在智慧城市场景中,路灯、智能交通信号灯、环境传感器等大量设备分布在不同位置,每个设备的数据量小但通信频繁,且设备异构性强.传统的集中管理方式会造成过高的网络负荷.通过在各个区域的边缘节点部署ZTA,可以本地化处理这些设备的身份认证和数据访问请求.例如,当一个传感器发送数据到控制中心时,边缘节点对该设备的各种属性进行收集,形成信任证据,然后进行动态认证,验证其是否有权限上传数据.如果该设备未通过认证,边缘节点将拒绝其访问请求,从而保护整个网络免受潜在的攻击.在学术方面,边缘计算发展迅猛,可以利用现有的成果完善ZTA.例如,文献[116]定义了一种称为边缘智能的范式,最大限度地利用边缘设备内的可用资源,提高了节点运算效率以及降低了能量消耗.文献[117]介绍了边缘节点进行合理优化也是降低节点能耗的途径之一.例如,将传感器和加速器与非标准接口连接起来,对从外设流出的数据进行实时预处理,以及加速近传感器分析、加密和机器学习任务.借助边缘计算,将大部分计算和决策下沉到本地,既能满足LS-IoT的实时需求,也能为后续“后PQC”提供更灵活的密钥管理策略.

6.5 后PQC

随着量子计算技术的不断演进,传统的RSA和ECC等公钥算法在量子计算机的攻击下将变得脆弱.对于节点数量庞大且数据生命周期更长的LS-IoT而言,这种威胁尤其严重.因此,为ZTA引入具有抗量子攻击能力的PQC是保护未来LS-IoT通信的关键.为了在ZTA中实现对量子攻击的有效防护,表9展示了能应用于ZTA的几类典型的PQC.

表9 典型后PQC算法

类别	代表算法	优点	适用场景
基于格的加密算法	CRYSTALS-Kyber[118]	密钥体积较小,计算效率高,易于在受限设备或嵌入式环境中部署	适用于资源受限的IoT环境,如IoT终端—边缘节点加密通信、密钥协商
	SABER[119]	避免了随机噪声采样,兼具安全性与较优的性能,在带宽有限的场景中表现良好	对功耗和通信容量均敏感的IoT设备,能够在保证安全性的同时降低软件/硬件实现开销
	NTRU[120]	最早的基于格的公钥体制之一,数学架构成熟,抗量子分析研究较完善	适合对速度要求高的IoT应用,如大规模数据采集和实时密钥交换场景
基于编码的加密算法	Classic McEliece[121]	安全边际极高(至今无实用量子攻击);封装速度快,解封速度特别快	用于对安全性要求极高且存储充足的场景
基于格的签名	Falcon[122]	Falcon基于NTRU格的签名,生成极小签名,适合资源受限场景	适用于LS-IoT环境中设备固件签名、数据完整性验证等
基于哈希的签名	SPHINCS+[123]	仅依赖哈希函数,理论最稳健,抗量子性强	适用于低频率签名但长期抗量子安全的场景

此外,将多种加密技术进行组合以进一步提升安全性,已成为当前的研究热点.文献[124]对多种后量子数字签名与密钥交换机制的组合进行了性能测试,指出Falcon签名算法与CRYSTALS-Kyber密钥交换的结合在未来无线传感器网络部署中或可提供最优安全性.为应对量子计算威胁,美国国家标准与技术研究院(National Institute of Standards and Technology,NIST)发布了关于过渡到后量子密码学标准的ZTA加密演进路线.其中,短期过渡方案采用混合加密模式(如CRYSTALS-Kyber与ECDH的结合)兼容现有设备;中长期方案是全面部署PQC算法,例如,使用Falcon进行数字签名并采用NTRU进行密钥交换,以全面取代传统加密算法.

图10展示了在LS-IoT中PQC增强ZTA安全的部署图.IoT设备节点采用轻量级PQC完成与边缘节点之间的初始密钥交换;若边缘节点在本地无法快速完成风险评估,则将设备的认证需求上传至云端中的零信任策略库,由全局威胁情报中心判定设备是否可信;一旦通过认证,设备、边缘节点和云端之间的后PQC通信通道将保持实时安全通信,避免量子攻击导致信息泄露.

在设备出现安全风险、固件更新或长期使用后量子签名时,这些操作也可在边缘端完成签名验证或密钥刷新,以减轻云端计算负担. PQC在未来具有广泛的实用性. 例如,在智能交通系统场景中,车路协同技术需要高安全和强可靠的数据通信. 自动驾驶车辆与路侧基础设施之间的通信通过后PQC保护,即使在未来量子计算成熟的情况下,经后PQC保护的通信数据仍难以被攻击者破解,从而确保智能交通系统的安全运行. 除自动驾驶外,PQC还适用于工业控制、智慧城市和医疗IoT等需要长期数据保护的场景. 后PQC将为ZTA中的通信加密提供长期保障,从而提升LS-IoT系统在量子时代的安全性和稳健性.

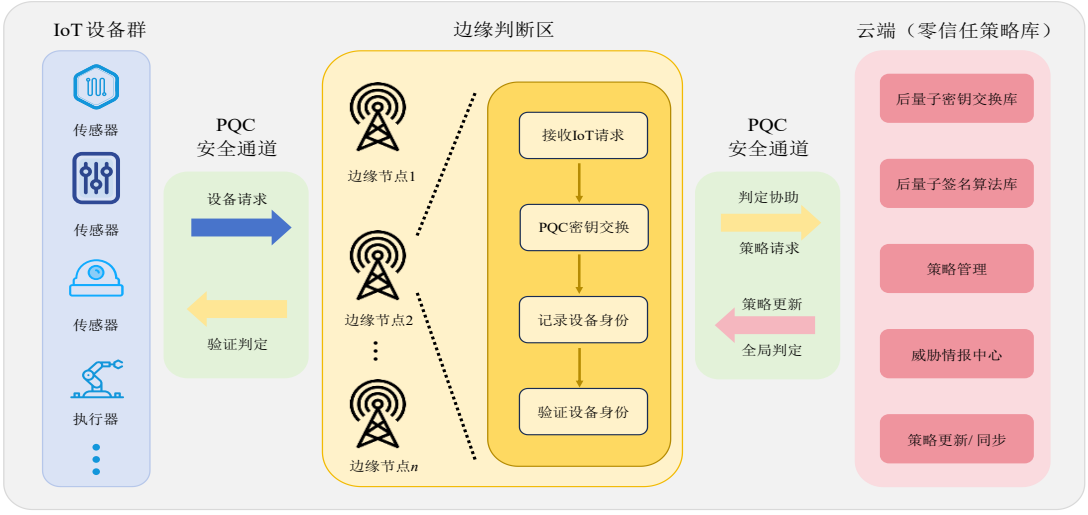

图10　后PQC在LS-IoT中的典型部署图

7　挑战及未来发展方向

7.1　效能与安全的权衡和优化

ZTA的核心原则"永不信任,始终验证"决定了系统为了确保信息安全,需要进行持续监控以及身份验证. 因此,系统需要消耗大量计算机资源处理大量数据,这可能导致系统性能下降. ZTA在LS-IoT环境中的应用加剧了效能与安全之间的矛盾. IoT设备数量庞大、异构性强,且设备资源受限. 持续的监控和身份验证增加了系统的开销,这在LS-IoT环境中尤为突出,尤其是在边缘节点或低功耗设备上,传统ZTA的高复杂性和高延迟可能会影响系统整体性能.

针对效能与安全的权衡问题,已有在LS-IoT中进行了初步探索. 例如,在工业IoT跨域认证场景中,文献[14]先将初始化、应用、认证与确认四大核心函数做成智能合约,实现一体化的授权-认证流程,有效排除伪合法设备,保障了安全性;再通过批量签名与并行操作的方式保持了高吞吐的优势. 在低空IoT场景中,文献[18]通过风险-功耗联合优化与三级分层控制方案,实现了自适应功率分配来平衡安全与能耗. 实践表明:效能-安全的矛盾很难通过静态方案设计直接解决,而应该在运行时引入性能监测进行动态调优,实现随负载和环境自适应的实时平衡.

此外,还可以从多方面加以应对. 例如,应用轻量级加密与认证算法,尽可能在不牺牲安全性的情况下降低计算负担. 当前在轻量级加密算方面有很大的进展,例如,PRESENT算法[66]通过采用简化的S盒和低轮数的置换网络,显著降低了计算复杂度,同时保持了较高的加密效率. 此外,融合边缘计算与云协作技术,利用边缘计算处理简单的验证任务,将复杂的数据处理和行为分析移交给云端或中央服务器,减少系统延迟,并降低轻量设备的负担. 同时,这类分布式架构也有助于将数据传输和安全控制分散化,避免单点故障. 为了进一步提高系统性能,通过采用批量处理设备身份验证或使用异步认证机制,可以减少实时处理的负担,有效提升整体系统性能.

然而,这些优化策略在实际应用中面临诸多挑战. 例如,设计既能保证安全性又具备高效性的轻量级算法仍需进一步研究;边缘计算资源的合理管理和分配也是一大难题;此外,批量处理和异步认证可能导致认证延迟,影响用户体验和系统的实时响应能力. 效能与安全的平衡是ZTA在LS-IoT中成功部署的关键,优化轻量级算法和边缘计算资源管理是未来的重要发展方向之一.

7.2　用户体验保障与隐私保护的挑战

在LS-IoT环境中,部署零信任安全框架的过程中,迫切需要考虑用户体验,包括企业用户、工业设备操作员、智能家居用户和各类移动设备用户. 频繁的多因素身份验证将严重影响用户体验,而对IoT设备与用户行

为的持续监控也带来了隐私保护的难题,尤其是在家庭和医疗等敏感场景中,如何在不牺牲安全性的前提下保障用户体验并维护用户隐私是关键问题.

现有部分研究围绕隐私保护方面陆续展开. 例如,在移动群智感知收集数据场景中,文献[125]提出了一种结合区块链与可信执行环境的“双层智能合约”方案;此外,还可以利用差分隐私的方式,通过设计在线激励决策机制与Stackelberg博弈模型引导自适应注噪,在保障数据质量的同时,保护用户隐私[126,127].

但现有研究仅考虑了用户隐私保护,未考虑用户体验保障. 用户体验是LS-IoT场景能否长久运行、规模扩张的先决条件. 系统操作被频繁打断或业务时延超标,一方面会降低用户体验,另一方面用户为保障系统连续性,可能会绕开安全机制,导致设备风险增加,最终削弱整个系统防御强度. 用户体验与隐私保护之间还存在一定的对抗性,提升安全通常依赖更细粒度的身份验证和持续监控,但这些手段会收集并处理更多个人或设备侧敏感数据;若缺乏隐私友好的处理与透明可解释的告知机制,用户对系统的信任将迅速下降,进一步恶化体验.

用户体验保障与隐私保护在实际应用中也面临诸多难题. 持续多因素认证频次高、交互频繁,容易打断系统进程,影响用户操作;同时,强加密与差分隐私注噪会提高握手时延与功耗,影响实时业务. 若使用户信任系统,需告知数据用途与决策逻辑,但过度披露又可能泄露安全策略细节,增加系统风险. 提升用户体验与保护隐私的平衡,是未来ZTA在IoT中应用的重要发展方向,需要在技术创新和用户体验设计上持续努力.

7.3 网络级到系统级安全防护的跃迁

LS-IoT系统不仅需要在网络层面进行防护,还需深入到系统内部,尤其是在工业自动化、远程办公等场景中,不仅网络通信需要严格保护,操作系统层级的安全也应受到重视. 当前主流的ZTA本身也存在着“不完善”的问题. 例如,传统的ZTA主要关注网络层的安全,忽视了系统层面所需要的细粒度保护. 因此,ZTA的网络PBS不足以应对IoT环境中的系统级威胁,如设备内部的恶意行为、系统级别的漏洞利用等.

为了实现IoT设备能够在系统级安全得到保障,文献[128]引入了一个全新的系统流抽象概念,设计了一种SysFlow的框架,用于捕捉系统内部的所有活动,并通过流量建模实现精确的权限控制,增强了系统的安全性. 同时,在工业IoT中,结合SIEM工具,对设备进行动态监控和更新,确保漏洞能够及时修补,减少内部威胁的攻击面,也是实现从网络级到系统级的全面安全防护的重要方法. 在系统级安全防护方面,文献[129]提出了一种SpecLFB硬件防御方案. 该方案在处理器的缓存区中集成轻量级安全检查,延迟并标记不安全的投机加载指令,从而阻断投机缓存侧信道攻击的建立. 还可与SysFlow等技术互补,一旦底层微架构通过SpecLFB验证,上层可降低动态度量频率,从而减轻性能负担,同时提升从系统层到应用层的纵深防护强度.

然而,现有研究仍存在诸多问题. 例如,如何准确识别和检测系统内部的复杂威胁,如恶意软件和内部攻击;其次,类似SysFlow框架依赖于准确的流量建模,如何在动态和异构的IoT环境中保持模型的准确性和适应性;再次,系统级监控和流量建模可能增加系统负担,影响设备的实时性能,如何在不显著降低系统性能的前提下,实现全面监控,并确保及时响应与处理安全事件,这一挑战仍需进一步解决;最后,不同设备和系统之间的安全机制和数据格式差异较大,如何实现无缝的系统级安全集成和协同工作,均是亟待解决的关键问题.

从网络级到系统级的安全防护是ZTA在IoT中实现全面安全的重要跃迁,类似SysFlow框架和全系统监控技术是实现这一目标的关键手段. 然而,系统级威胁检测和流量建模的准确性仍是主要挑战,未来需要进一步深入研究和创新.

7.4 技术集成和标准化

零信任是安全的,同时也是复杂的,仅单靠一种网络安全技术无法完成. 其部署需要多种技术的相互协助,涉及身份认证、访问控制、加密通信、数据分析等多个技术领域的协同,如何实现这些技术的无缝融合并确保其高效、稳定运行是一大挑战. 此外,LS-IoT系统异构性强,设备的操作系统、通信协议、硬件架构千差万别,其规模和复杂性使得部署ZTA的时间和资源成本显著增加. 在全球范围内,没有统一的行业标准限制了ZTA的在LS-IoT中推广.

当前技术集成与标准化存在多种潜在解决方案. 首先,推动IoT安全标准的统一,参考电气和电子工程师协会(Institute of Electrical and Electronics Engineers, IEEE)和NIST等组织的标准化工作,建立通用的认证与授权框架;其次,采用分阶段部署策略,优先保护关键设备和核心网络,根据设备优先级和安全需求,逐步扩展ZTA的应用范围;再次,开发兼容多种设备和系统的平台,确保不同设备和系统之间的安全机制能够无缝集成和协同工作;最后,推广标准化的工具和方法,简化ZTA的部署和管理,降低企业的实施成本和复杂性.

然而,在实现上述潜在解决方案的过程中仍然面临诸多挑战. 第一,IoT设备种类繁多,操作系统、通信协议和硬件架构各异,制定统一的安全标准需要跨行业和跨领域的广泛合作;第二,不同安全技术和系统之

间的兼容性问题可能导致集成困难,影响ZTA的整体效能和可靠性;第三,统一标准和技术集成需要大量的时间和资源投入,尤其对于中小企业来说,实施成本可能较高;第四,IoT环境具有高度的动态性,标准和技术需要具备良好的适应性,以应对设备的频繁加入和移除以及环境的不断变化.

8 结论

LS-IoT通过海量设备的互联互通,实现数据的高效采集、传输与分析,从而推动智能化决策、资源优化配置以及各行业生产效率的全面提升,成为数字化转型和智慧社会构建的核心驱动力.ZTA因其"永不信任,始终验证"的安全理念,逐渐成为应对复杂网络威胁的有效方案.本文对ZTA的核心能力、关键技术及其在不同领域的应用进行了全面的分析和总结.深入探讨了SDP、IAM、MSG等核心能力,以及身份认证、访问控制、加密技术、IGA、终端安全、网络隔离以及持续监控七大关键技术对ZTA的重要性.通过工业IoT、5G医疗、自动驾驶、远程办公场景4个典型应用场景,展示了零信任在实际应用中的可行性和挑战.为进一步提高网络与系统的安全性和智能化水平,探讨了ZTA与LLM、生成式AI、边缘计算、PQC等前沿技术融合的重要性与难点,并总结了ZTA未来发展方向.本研究通过全面总结面向LS-IoT场景的ZTA的核心能力、关键技术、应用场景等,旨在推动ZTA在IoT场景的实际应用,实现更全面的网络与系统安全保护.

参考文献

[1] LOWDERMILK J, SETHUMADHAVAN S. Towards zero trust: An experience report[C]//2021 IEEE Secure Development Conference. Piscataway: IEEE, 2021: 79-85.

[2] YAN W, ZHANG N, NJILLA L L, et al. PCBChain: Lightweight reconfigurable blockchain primitives for secure IoT applications[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2020, 28(10): 2196-2209.

[3] KINDERVAG J. Build security into your network's DNA: The zero trust network architecture[EB/OL]. (2010-11-05)[2025-02-23]. https://www.virtualstarmedia.com/downloads/Forrester_zero_trust_DNA.pdf.

[4] 蹇诗婕, 卢志刚, 杜丹, 等. 网络入侵检测技术综述[J]. 信息安全学报, 2020, 5(4): 96-122.
JIAN S J, LU Z G, DU D, et al. Overview of network intrusion detection technology[J]. Journal of Cyber Security, 2020, 5(4): 96-122. (in Chinese)

[5] 卿斯汉, 蒋建春, 马恒太, 等. 入侵检测技术研究综述[J]. 通信学报, 2004, 25(7): 19-29.
QING S H, JIANG J C, MA H T, et al. Research on intrusion detection techniques: A survey[J]. Journal of China Institute of Communications, 2004, 25(7): 19-29. (in Chinese)

[6] 钱伟中, 王蔚然, 袁宏春. 分布式防火墙环境的边界防御系统[J]. 电子科技大学学报, 2005, 34(4): 513-516.
QIAN W Z, WANG W R, YUAN H C. Boundary defense system based on DFW[J]. Journal of University of Electronic Science and Technology of China, 2005, 34(4): 513-516. (in Chinese)

[7] ALSHALAN A, PISHARODY S, HUANG D J. A survey of mobile VPN technologies[J]. IEEE Communications Surveys & Tutorials, 2016, 18(2): 1177-1196.

[8] SEXTON C, KAMINSKI N J, MARQUEZ-BARJA J M, et al. 5G: Adaptable networks enabled by versatile radio access technologies[J]. IEEE Communications Surveys & Tutorials, 2017, 19(2): 688-720.

[9] 任彦冰, 李兴华, 刘海, 等. 基于区块链的分布式物联网信任管理方法研究[J]. 计算机研究与发展, 2018, 55(7): 1462-1478.
REN Y B, LI X H, LIU H, et al. Blockchain-based trust management framework for distributed Internet of Things[J]. Journal of Computer Research and Development, 2018, 55(7): 1462-1478. (in Chinese)

[10] SAAD M, SPAULDING J, NJILLA L, et al. Exploring the attack surface of blockchain: A comprehensive survey[J]. IEEE Communications Surveys & Tutorials, 2020, 22(3): 1977-2008.

[11] ROSE S, BORCHERT O, MITCHELL S, et al. Zero trust architecture[EB/OL]. (2020-08-01)[2025-02-23]. https://nvlpubs. nist. gov/nistpubs/SpecialPublications/NIST. SP. 800-207.pdf.

[12] WARD R, BEYER B. BeyondCorp: A new approach to enterprise security[J]. The Magazine of USENIX & SAGE, 2014, 39(6): 6-11.

[13] LI S Z, MENG W C, LIU C, et al. Feature attention distillation defense for backdoor attack in artificial-neural-network-based electricity theft detection[J]. IEEE Internet of Things Journal, 2024, 11(24): 39880-39889.

[14] TONG F, CHEN X, HUANG C, et al. Blockchain-assisted secure intra/inter-domain authorization and authentication for internet of things[J]. IEEE Internet of Things Journal, 2023, 10(9): 7761-7773.

[15] YANG X J, TONG F, JIANG F, et al. A lightweight and dynamic open-set intrusion detection for industrial Internet of Things[J]. IEEE Transactions on Information Fo-

rensics and Security, 2025, 20: 2930-2943.

[16] TONG F, CHEN C, PAN J P. A novel detection and localization scheme for wormhole attack in internet of things[J]. IEEE Internet of Things Journal, 2024, 11(4): 7141-7152.

[17] LIU C, HE S B, LI S Z, et al. Time-series multi-instance learning for weakly supervised industrial fault detection[J]. IEEE Transactions on Industrial Informatics, 2025, 21(4): 3326-3335.

[18] CHEN Y C, LIU G B, ZHANG Z, et al. Improving physical layer security for multi-UAV systems against hybrid wireless attacks[J]. IEEE Transactions on Vehicular Technology, 2024, 73(5): 7034-7048.

[19] ZENG J, LI Y X, RAN Y L, et al. Efficient view path planning for autonomous implicit reconstruction[C]//2023 IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 2023: 4063-4069.

[20] ZHOU Y Y, CHENG G, OUYANG Z, et al. Resource-efficient low-rate DDoS mitigation with moving target defense in edge clouds[J]. IEEE Transactions on Network and Service Management, 2025, 22(1): 168-186.

[21] SHAO Z L, CHEN T Z, CHENG G, et al. AF-FDS: An accurate, fast, and fine-grained detection scheme for DDoS attacks in high-speed networks with asymmetric routing[J]. IEEE Transactions on Network and Service Management, 2023, 20(4): 4964-4981.

[22] 王航宇, 吕飞, 程裕亮, 等. 工业物联网零信任安全研究综述[J/OL]. 计算机研究与发展. (2025-04-08)[2025-07-28]. https://kns.cnki.net/KCMS/detail/detail.aspx?filename=JFYZ2025040300A&dbname=CJFD&dbcode=CJFQ.
WANG H Y, LYU F, CHENG Y L, et al. Review on zero trust security of industrial internet of things[J/OL]. Journal of Computer Research and Development. (2025-04-08)[2025-07-28]. https://kns.cnki.net/KCMS/detail/detail.aspx?filename=JFYZ2025040300A&dbname=CJFD&dbcode=CJFQ. (in Chinese)

[23] 张宇, 张妍. 零信任研究综述[J]. 信息安全研究, 2020, 6(7): 608-614.
ZHANG Y, ZHANG Y. A survey of zero trust research[J]. Journal of Information Security Research, 2020, 6(7): 608-614. (in Chinese)

[24] 张泽洲, 王鹏. 零信任安全架构研究综述[J]. 保密科学技术, 2021(8): 8-16.
ZHANG Z Z, WANG P. A survey of zero trust security architecture[J]. Secrecy Science and Technology, 2021(8): 8-16. (in Chinese)

[25] 诸葛程晨, 王群, 刘家银, 等. 零信任网络综述[J]. 计算机工程与应用, 2022, 58(22): 12-29.
ZHUGE C C, WANG Q, LIU J Y, et al. Survey of zero trust network[J]. Computer Engineering and Applications, 2022, 58(22): 12-29. (in Chinese)

[26] MOUBAYED A, REFAEY A, SHAMI A. Software-defined perimeter (SDP): State of the art secure solution for modern networks[J]. IEEE Network, 2019, 33(5): 226-233.

[27] LUCION E L R, NUNES R C. Software defined perimeter: Improvements in the security of single packet authorization and user authentication[C]//2018 XLIV Latin American Computer Conference. Piscataway: IEEE, 2019: 708-717.

[28] 杨冬, 程宗荣, 田伟康, 等. 广义确定性标识网络[J]. 电子学报, 2024, 52(1): 1-18.
YANG D, CHENG Z R, TIAN W K, et al. Generalized deterministic identification networks[J]. Acta Electronica Sinica, 2024, 52(1): 1-18. (in Chinese)

[29] QIU T, ZHAO Z, ZHANG T, et al. Underwater Internet of Things in smart ocean: System architecture and open issues[J]. IEEE Transactions on Industrial Informatics, 2020, 16(7): 4297-4307.

[30] 胡向东, 张琴. 基于特征组合优化的工业互联网恶意行为实时检测方法[J]. 电子学报, 2024, 52(9): 3075-3085.
HU X D, ZHANG Q. Real-time detection method of malicious behaviors in industrial internet based on feature combination optimization[J]. Acta Electronica Sinica, 2024, 52(9): 3075-3085. (in Chinese)

[31] CAO B, ZHANG Y T, ZHAO J W, et al. Recommendation based on large-scale many-objective optimization for the intelligent internet of things system[J]. IEEE Internet of Things Journal, 2022, 9(16): 15030-15038.

[32] ZHANG Y Y, HUANG Y, HUANG C, et al. Joint optimization of deployment and flight planning of multi-UAVs for long-distance data collection from large-scale IoT devices[J]. IEEE Internet of Things Journal, 2024, 11(1): 791-804.

[33] YANG Z, ZHANG J, JIANG Y L, et al. An energy-efficient convolution-based partitioned collaborative perception algorithm for large-scale IoT services[J]. IEEE Transactions on Industrial Informatics, 2024, 20(5): 7404-7413.

[34] EJAZ W, NAEEM M, ZEADALLY S. On-demand sensing and wireless power transfer for self-sustainable industrial internet of things networks[J]. IEEE Transactions on

Industrial Informatics, 2021, 17(10): 7075-7084.

[35] 童率, 王继良 . 低功耗广域网 LoRa 技术进展与研究挑战[J]. 电子学报, 2024, 52(10): 3623-3642.
TONG S, WANG J L. Progress and challenges of LoRa low power wide area networks[J]. Acta Electronica Sinica, 2024, 52(10): 3623-3642. (in Chinese)

[36] YU Y, LIU S M, YEOH P L, et al. LayerChain: A hierarchical edge-cloud blockchain for large-scale low-delay industrial internet of things applications[J]. IEEE Transactions on Industrial Informatics, 2021, 17(7): 5077-5086.

[37] XIE X, WANG H, LIU X J. Scheduling for minimizing the age of information in multisensor multiserver industrial internet of things systems[J]. IEEE Transactions on Industrial Informatics, 2024, 20(1): 573-582.

[38] 樊琳娜, 李城龙, 吴毅超, 等 . 物联网设备识别及异常检测研究综述[J]. 软件学报, 2024, 35(1): 288-308.
FAN L N, LI C L, WU Y C, et al. Survey on IoT device identification and anomaly detection[J]. Journal of Software, 2024, 35(1): 288-308. (in Chinese)

[39] PAL S, RABEHAJA T, HITCHENS M, et al. On the design of a flexible delegation model for the internet of things using blockchain[J]. IEEE Transactions on Industrial Informatics, 2020, 16(5): 3521-3530.

[40] 程冠杰, 邓水光, 温盈盈, 等 . 基于区块链的物联网认证机制综述[J]. 软件学报, 2023, 34(3): 1470-1490.
CHENG G J, DENG S G, WEN Y Y, et al. Survey on blockchain-based internet of things authentication mechanisms[J]. Journal of Software, 2023, 34(3): 1470-1490. (in Chinese)

[41] PENG Z, ZHANG A, WANG S, et al. Designing principles and constructing processes of the comprehensive evaluation indicator system[J]. Science Research Management, 2017, 38: 209-215.

[42] LOGENTHIRAN T, SRINIVASAN D, KHAMBADKONE A M. Multi-agent system for energy resource scheduling of integrated microgrids in a distributed system[J]. Electric Power Systems Research, 2011, 81(1): 138-148.

[43] GE Y F, ZHU Q Y. GAZETA: GAme-theoretic zero-trust authentication for defense against lateral movement in 5G IoT networks[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 540-554.

[44] MENG L, HUANG D C, AN J H, et al. A continuous authentication protocol without trust authority for zero trust architecture[J]. China Communications, 2022, 19(8): 198-213.

[45] ABUHAMAD M, ABUSNAINA A, NYANG D, et al. Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A contemporary survey[J]. IEEE Internet of Things Journal, 2021, 8(1): 65-84.

[46] LIU J X, SIMSEK M, KANTARCI B, et al. Risk-aware fine-grained access control in cyber-physical contexts[J]. Digital Threats: Research and Practice, 2022, 3(4): 1-29.

[47] ZHANG P Y, YANG P, KUMAR N, et al. RRV-BC: Random reputation voting mechanism and blockchain assisted access authentication for industrial internet of things[J]. IEEE Transactions on Industrial Informatics, 2024, 20(1): 713-722.

[48] FANG H, WANG X B, AL-DHAHIR N, et al. Joint design of multi-dimensional multiple access and lightweight continuous authentication in zero-trust environments[C]// GLOBECOM 2023 - 2023 IEEE Global Communications Conference. Piscataway: IEEE, 2024: 3366-3371.

[49] KHAN S, THAPA C, DURRANI S, et al. Access-based lightweight physical-layer authentication for the internet of things devices[J]. IEEE Internet of Things Journal, 2024, 11(7): 11312-11326.

[50] HARBACH M, DE LUCA A, EGELMAN S. The anatomy of smartphone unlocking: A field study of Android lock screens[C]//Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. New York: ACM, 2016: 4806-4817.

[51] COYNE E, WEIL T. An RBAC implementation and interoperability standard: The INCITS cyber security 1.1 model[J]. IEEE Security & Privacy, 2008, 6(1): 84-87.

[52] BHATTI R, BERTINO E, GHAFOOR A. A trust-based context-aware access control model for web-services[J]. Distributed and Parallel Databases, 2005, 18(1): 83-105.

[53] BOBBA R, FATEMIEH O, KHAN F, et al. Attribute-based messaging: Access control and confidentiality[J]. ACM Transactions on Information and System Security, 2010, 13(4): 1-35.

[54] TANG C L, FU X L, TANG P. Policy-based network access and behavior control management[C]//2020 IEEE 20th International Conference on Communication Technology. Piscataway: IEEE, 2020: 1102-1106.

[55] ZONG Y, GUO Y, CHEN X Q. Policy-based access control for robotic applications[C]//2019 IEEE International Conference on Service-Oriented System Engineering. Piscataway: IEEE, 2019: 368-3685.

[56] URIARTE M, ASTORGA J, JACOB E, et al. Expressive policy-based access control for resource-constrained de-

vices[J]. IEEE Access, 2018, 6: 15-46.

[57] GHAFOORIAN M, ABBASINEZHAD-MOOD D, SHAKERI H. A thorough trust and reputation based RBAC model for secure data storage in the cloud[J]. IEEE Transactions on Parallel and Distributed Systems, 2019, 30(4): 778-788.

[58] FANG H, WANG X B, HANZO L. Adaptive trust management for soft authentication and progressive authorization relying on physical layer attributes[J]. IEEE Transactions on Communications, 2020, 68(4): 2607-2620.

[59] ZHU H, XUE X S, XU M M, et al. Zero trust consumer IoT with robust federated learning over main-side blockchain[J]. IEEE Transactions on Consumer Electronics, 2025, 71(1): 1180-1189.

[60] RIVEST R L, SHAMIR A, ADLEMAN L. A method for obtaining digital signatures and public-key cryptosystems[J]. Communications of the ACM, 1978, 21(2): 120-126.

[61] REIS D, GENG H R, NIEMIER M, et al. IMCRYPTO: An in-memory computing fabric for AES encryption and decryption[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2022, 30(5): 553-565.

[62] SMID M E, BRANSTAD D K. Data encryption standard: Past and future[J]. Proceedings of the IEEE, 1988, 76(5): 550-559.

[63] MENEZES A, VAN OORSCHOT P, VANSTONE S. Elliptic curve public key cryptosystems[J]. IEEE Transactions on Information Theory, 1993, 39(5): 1719-1724.

[64] SHAHBAZI K, KO S B. Area-efficient nano-AES implementation for Internet-of-things devices[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2021, 29(1): 136-148.

[65] SYED N F, SHAH S W, SHAGHAGHI A, et al. Zero trust architecture (ZTA): A comprehensive survey[J]. IEEE Access, 2022, 10: 57143-57179.

[66] CHATTERJEE R, CHAKRABORTY R. A modified lightweight PRESENT cipher for IoT security[C]//2020 International Conference on Computer Science, Engineering and Applications. Piscataway: IEEE, 2020: 1-6.

[67] SEKAR G, MOUHA N, VELICHKOV V, et al. Meet-in-the-middle attacks on reduced-round XTEA[M]//Topics in Cryptology-CT-RSA 2011. Berlin, Heidelberg: Springer, 2011: 250-267.

[68] HABIB M A, AHMAD M, JABBAR S, et al. Speeding up the Internet of Things: LEAIoT: A lightweight encryption algorithm toward low-latency communication for the internet of things[J]. IEEE Consumer Electronics Magazine, 2018, 7(6): 31-37.

[69] 吴海博, 许瑶恭, 李俊. FPTC: 一种信息中心物联网缓存策略[J]. 软件学报, 2022, 33(12): 4816-4837.
WU H B, XU Y G, LI J. FPTC: An ICN-IoT caching scheme[J]. Journal of Software, 2022, 33(12): 4816-4837. (in Chinese)

[70] KUPERBERG M. Blockchain-based identity management: A survey from the enterprise and ecosystem perspective[J]. IEEE Transactions on Engineering Management, 2020, 67(4): 1008-1027.

[71] 魏欣, 王心妍, 于卓, 等. 基于联盟链的物联网跨域认证[J]. 软件学报, 2021, 32(8): 2613-2628.
WEI X, WANG X Y, YU Z, et al. Cross domain authentication for IoT based on consortium blockchain[J]. Journal of Software, 2021, 32(8): 2613-2628. (in Chinese)

[72] XIE H R, WANG Y J, DING Y, et al. Industrial wireless Internet zero trust model: Zero trust meets dynamic federated learning with blockchain[J]. IEEE Wireless Communications, 2024, 31(2): 22-29.

[73] 葛丽娜, 栗海澳, 王捷. 基于多级代理许可区块链的联邦边缘学习模型[J]. 通信学报, 2024, 45(4): 201-215.
GE L N, LI H A, WANG J. Federated edge learning model based on multi-level proxy permissioned blockchain[J]. Journal on Communications, 2024, 45(4): 201-215. (in Chinese)

[74] HUSSAIN M, PAL S, JADIDI Z, et al. Federated zero trust architecture using artificial intelligence[J]. IEEE Wireless Communications, 2024, 31(2): 30-35.

[75] ZHANG Y, XU C X, LI H W, et al. PROTECT: Efficient password-based threshold single-sign-on authentication for mobile users against perpetual leakage[J]. IEEE Transactions on Mobile Computing, 2021, 20(6): 2297-2312.

[76] SHANG C, CAO J, ZHU T, et al. CADFA: A clock skew-based active device fingerprint authentication scheme for class-1 IoT devices[J]. IEEE Systems Journal, 2024, 18(1): 590-599.

[77] JI X Y, ZHOU X Y, YAN C, et al. A nonlinearity-based secure face-to-face device authentication for mobile devices[J]. IEEE Transactions on Mobile Computing, 2022, 21(4): 1155-1171.

[78] BADHIB A, ALSHEHRI S, CHERIF A. A robust device-to-device continuous authentication protocol for the Internet of Things[J]. IEEE Access, 2021, 9: 124768-124792.

[79] RAO SP, LIMONTA G, LINDQVIST J. Usability and security of trusted platform module (TPM) library APIs[C]// Proceedings of the Eighteenth USENIX Conference on

Usable Privacy and Security (SOUPS). California: USENIX Association, 2022: 213-232.

[80] ZHAO B, XIAO Y, HUANG Y Q, et al. A private user data protection mechanism in trustzone architecture based on identity authentication[J]. Tsinghua Science and Technology, 2017, 22(2): 218-225.

[81] KWON D, SEO J, CHO Y, et al. PrOS: Light-weight privatized se cure OSes in ARM trustzone[J]. IEEE Transactions on Mobile Computing, 2020, 19(6): 1434-1447.

[82] LUO L, ZHANG Y, WHITE C, et al. On security of trustzone-M-based IoT systems[J]. IEEE Internet of Things Journal, 2022, 9(12): 9683-9699.

[83] SHU Z M, LIU Y G, WANG H N, et al. Research on the feasibility technology of Internet of Things terminal security monitoring[C]//2021 6th International Symposium on Computer and Information Processing Technology. Piscataway: IEEE, 2021: 831-836.

[84] WANG J, HONG Z, ZHANG Y H, et al. Enabling security-enhanced attestation with intel SGX for remote terminal and IoT[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2018, 37(1): 88-96.

[85] NIE L S, WANG X J, WANG S P, et al. Network traffic prediction in industrial internet of things backbone networks: A multitask learning mechanism[J]. IEEE Transactions on Industrial Informatics, 2021, 17(10): 7123-7132.

[86] ZHAO R J, HUANG Y T, DENG X W, et al. A novel traffic classifier with attention mechanism for industrial internet of things[J]. IEEE Transactions on Industrial Informatics, 2023, 19(11): 10799-10810.

[87] 肖警续, 郭渊博, 常朝稳, 等. 基于SDN的物联网边缘节点间数据流零信任管理[J]. 通信学报, 2024, 45(7): 101-116.
XIAO J X, GUO Y B, CHANG C W, et al. Zero trust management of data flow between IoT edge nodes based on SDN[J]. Journal on Communications, 2024, 45(7): 101-116. (in Chinese)

[88] FENG T, BI J, WANG K. Allocation and scheduling of network resource for multiple control applications in SDN[J]. China Communications, 2015, 12(6): 85-95.

[89] BARI M F, BOUTABA R, ESTEVES R, et al. Data center network virtualization: A survey[J]. IEEE Communications Surveys & Tutorials, 2013, 15(2): 909-928.

[90] GOETHALS T, DE TURCK F, VOLCKAERT B. Extending Kubernetes clusters to low-resource edge devices using virtual kubelets[J]. IEEE Transactions on Cloud Computing, 2022, 10(4): 2623-2636.

[91] GU L, ZENG D Z, GUO S, et al. A general communication cost optimization framework for big data stream processing in geo-distributed data centers[J]. IEEE Transactions on Computers, 2016, 65(1): 19-29.

[92] SHEN M, YE K, LIU X T, et al. Machine learning-powered encrypted network traffic analysis: A comprehensive survey[J]. IEEE Communications Surveys & Tutorials, 2023, 25(1): 791-824.

[93] YAO K D, SAYAGH M, SHANG W Y, et al. Improving state-of-the-art compression techniques for log management tools[J]. IEEE Transactions on Software Engineering, 2022, 48(8): 2748-2760.

[94] CISCO. Cisco's guide to zero trust maturity[EB/OL]. (2025-01-01)[2025-02-23]. https://www.cisco.com/c/en/us/products/security/zero-trust-maturity-guide.html.

[95] CISCO. Cisco security outcomes for zero trust: Adoption, access, and automation[EB/OL]. (2023-11-01) [2025-02-23]. https://www. cisco. com/c/en/us/products/security/zero-trust-outcomes-report.html.

[96] SERROR M, HACK S, HENZE M, et al. Challenges and opportunities in securing the industrial internet of things[J]. IEEE Transactions on Industrial Informatics, 2021, 17(5): 2985-2996.

[97] LIN K, GAO J, HAN G J, et al. Intelligent blockchain-enabled adaptive collaborative resource scheduling in large-scale industrial internet of things[J]. IEEE Transactions on Industrial Informatics, 2022, 18(12): 9196-9205.

[98] CAO Y, JIA F, MANOGARAN G. Efficient traceability systems of steel products using blockchain-based industrial internet of things[J]. IEEE Transactions on Industrial Informatics, 2020, 16(9): 6004-6012.

[99] FENG X M, HU S Y. Cyber-physical zero trust architecture for industrial cyber-physical systems[J]. IEEE Transactions on Industrial Cyber-Physical Systems, 2023, 1: 394-405.

[100] SIEMENS. Siemens news[EB/OL]. (2022-01-01) [2025-02-23]. https://w1.siemens.com.cn/news/news_articles/17509.aspx.

[101] 大咖科技TechChic.零信任架构在安全中的应用[EB/OL]. (2022-03-24)[2025-02-23]. https://baijiahao.baidu.com/s?id=1728132845684495531.

[102] GE DIGITAL. Predix:工业互联网[EB/OL]. (2016-03-01) [2025-02-23]. https://file.caixin.com/file/topic/ge2016new/predix.pdf.

[103] IBM. IBM newsroom[EB/OL]. (2023-06-06) [2025-02-

23]. https://china.newsroom.ibm.com/2023-06-06-IBM-%2CIBM.

[104] CSA工作组著. CSA大中华区零信任工作组译. 基于零信任架构的医疗设备安全[EB/OL]. (2023-09-01)[2025-02-23]. https://www.c-csa.cn/mobile/research/results-detail/i-1900.

[105] CHEN B Z, QIAO S Y, ZHAO J, et al. A security awareness and protection system for 5G smart healthcare based on zero-trust architecture[J]. IEEE Internet of Things Journal, 2021, 8(13): 10248-10263.

[106] 邹光南, 尤启迪, 金星虎, 等. 面向车联网车辆的轻量级持续身份认证协议[J]. 电子学报, 2024, 52(6): 1903-1910.
ZOU G N, YOU Q D, JIN X H, et al. Lightweight continuous authentication protocol for vehicles in vehicular networks[J]. Acta Electronica Sinica, 2024, 52(6): 1903-1910. (in Chinese)

[107] ANDERSON J, HUANG Q Q, CHENG L, et al. A zero-trust architecture for connected and autonomous vehicles[J]. IEEE Internet Computing, 2023, 27(5): 7-14.

[108] HANGYAN. Hangyan charts[EB/OL]. (2024-10-16)[2025-02-23]. https://www.hangyan.co/charts/3479776161188284049.

[109] GOOGLE CLOUD. Zero trust and BeyondCorp: Google cloud[EB/OL]. (2022-08-29)[2025-02-23]. https://globalcloudplatforms.com/2022/08/29/zero-trust-and-beyondcorp-google-cloud/.

[110] PECK J, BEYER B, BESKE C, et al. Colin beske and max slatonstall migrating to BeyondCorp maintaining productivity while improving security[EB/OL]. (2017-07-01) [2025-02-23]. https://research. google/pubs/migrating-to-beyondcorp-maintaining-productivity-while-improving-security/.

[111] VICTOR ESCOBEDO, BETSY BEYER, SLATONSTALL MAX. BeyondCrop5: The user experience[EB/OL]. (2017-08-01)[2025-02-23]. https://www.usenix.org/publications/login/fall2017/escobedo.

[112] HUNTER K, MICHAEL J, BETSY B, et al. BeyondCorp: Building a healthy fleet[EB/OL]. (2018-08-01)[2025-02-23]. https://www.usenix.org/system/files/login/articles/login_fall18_05_king.pdf.

[113] Barclay O, Justin M, Betsy B, et al. BeyondCorp: Design deployment at Google[EB/OL]. (2016-01-01)[2025-02-23]. https://research.google.com/pubs/pub44860.html?hl=zh-cn.

[114] MEI M Y, YAO M W, YANG Q H, et al. On the statistical delay performance of large-scale IoT networks[J]. IEEE Transactions on Vehicular Technology, 2022, 71(8): 8967-8979.

[115] NABIL Y, ELSAWY H, AL-DHARRAB S, et al. Data aggregation in regular large-scale IoT networks: Granularity, reliability, and delay tradeoffs[J]. IEEE Internet of Things Journal, 2022, 9(18): 17767-17784.

[116] AJIRLOU A F, KENARANGI F, SHAPIRA E, et al. NoD: A neural network-over-decoder for edge intelligence[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2022, 30(10): 1438-1447.

[117] SCHIAVONE P D, ROSSI D, DI MAURO A, et al. Arnold: An eFPGA-augmented RISC-V SoC for flexible and low-power IoT end nodes[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2021, 29(4): 677-690.

[118] RAVI P, HOWE J, CHATTOPADHYAY A, et al. Lattice-based key-sharing schemes: A survey[J]. ACM Computing Surveys, 2022, 54(1): 1-39.

[119] MAHDI L H, ABDULLAH A A. Fortifying future IoT security: A comprehensive review on lightweight post-quantum cryptography[J]. Engineering, Technology & Applied Science Research, 2025, 15(2): 21812-21821.

[120] RIBEIRO L A D S, SILVA LIMA J P DA, DE QUEIROZ R J G B, et al. SABER post-quantum key encapsulation mechanism (KEM): Evaluating performance in ARM and x64 architectures[J]. Journal of Cryptographic Engineering, 2024, 14(1): 35-41.

[121] LIU F X, ZHENG Z Y, GONG Z X, et al. A survey on lattice-based digital signature[J]. Cybersecurity, 2024, 7(1): 7.

[122] FITZGIBBON G, OTTAVIANI C. Constrained device performance benchmarking with the implementation of post-quantum cryptography[J]. Cryptography, 2024,8(2):21.

[123] BERNSTEIN D J, HOPWOOD D, HÜLSING A, et al. SPHINCS: Practical stateless hash-based signatures[M]// Advances in Cryptology—EUROCRYPT 2015. Berlin, Heidelberg: Springer, 2015: 368-397.

[124] SEÑOR J, PORTILLA J, PORTELA-GARCÍA M. Performance analysis of postquantum cryptographic schemes for securing large-scale wireless sensor networks[J]. IEEE Transactions on Industrial Informatics, 2024, 20(10): 12339-12349.

[125] ZHOU Y H, TONG F, KONG C M, et al. Towards efficient, robust, and privacy-preserving incentives for crowdsensing via blockchain[J]. IEEE Transactions on Mobile Computing, 2025, 24(8): 7136-7151.

[126] ZHANG M Y, YANG L, HE S B, et al. Privacy-preserv-

ing data aggregation for mobile crowdsensing with externality: An auction approach[J]. IEEE/ACM Transactions on Networking, 2021, 29(3): 1046-1059.

[127] YANG G, SHI Z G, HE S B, et al. Socially privacy-preserving data collection for crowdsensing[J]. IEEE Transactions on Vehicular Technology, 2020, 69(1): 851-861.

[128] HONG S, XU L, HUANG J W, et al. SysFlow: Toward a programmable zero trust framework for system security[J]. IEEE Transactions on Information Forensics and Security, 2023, 18: 2794-2809.

[129] CHENG X, TONG F, WANG H. et al. SpecLFB: Eliminating cache side channels in speculative executions[C]// Proceedings of the 33rd USENIX Security Symposium. California: USENIX Association, 2024: 631-646.

作者简介

邢方圆　女,1991年8月出生于辽宁省抚顺市.2020年博士毕业于大连理工大学,现为东南大学网络空间安全学院副研究员.主要研究方向为物联网安全与资源优化等.中国电子学会会员编号:E190072491M.

E-mail: fangyuanxing@seu.edu.cn

董　傲　男,2002年1月出生于山东省菏泽市.现为东南大学网络空间安全学院硕士研究生.主要研究方向为网络安全等.

E-mail: 220245619@seu.edu.cn

孙羽羿　女,1993年9月出生于贵州省贵阳市.2021年博士毕业于浙江大学控制科学与工程学院,现为杭州师范大学信息科学与技术学院讲师.主要研究方向为物联网、无线通信.

E-mail: yuyisun@hznu.edu.cn

童　飞　男,1987年6月出生于安徽省巢湖市.2016年博士毕业于加拿大维多利亚大学,现为东南大学网络空间安全学院副教授.主要研究方向为物联网系统安全.

E-mail: ftong@seu.edu.cn

贺诗波　男,1983年7月出生于湖南省衡阳市.2012年博士毕业于浙江大学,现为浙江大学控制学院教授.主要研究方向为工业大模型、大数据、物联网.

E-mail: s18he@zju.edu.cn

程　光　男,1973年出生于安徽省黄山市.2003年博士毕业于东南大学,现为东南大学网络空间安全学院教授.主要研究方向为网络流量安全分析、网络安全主动防御等.

E-mail: chengguang@seu.edu.cn