计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
2024,33(11):38−47 [doi: 10.15888/j.cnki.csa.009702] [CSTR: 32024.14.csa.009702]

E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

边缘计算的安全挑战与解决方法综述①

温木奇[1], 温武少[2]

[1](广东白云学院 电气与信息工程学院, 广州 510450)
[2](中山大学 计算机学院, 广州 510006)
通信作者: 温木奇, E-mail: 718176428@qq.com

摘　要: 相比集中式的云计算框架, 边缘计算在云中心和现场智能设备之间部署了额外的“边缘服务器”, 支持现场智能设备快速、高效地完成运算任务和事件处理. 边缘计算系统中, 现场智能设备数量庞大、边缘计算服务器繁杂, 它们存储的数据敏感和私密性要求高. 边缘计算系统的这些特点, 给网络安全防护带来困难. 解决边缘计算系统的信息和网络安全是边缘计算技术大规模产业化的关键. 而由于边缘服务器设备和现场智能设备的计算能力、网络能力和存储能力的局限, 传统的计算机网络安全技术不能完全满足要求. 分析适合边缘计算系统的联邦学习、轻量加密、混淆与虚拟位置信息、匿名身份认证等有效的敏感数据保护技术, 以及探讨人工智能和区块链等新技术在边缘计算防范恶意攻击的应用, 助力边缘计算的产业化发展.

关键词: 边缘计算; 数据保护; 联邦学习; 区块链; 内生安全架构

引用格式: 温木奇,温武少.边缘计算的安全挑战与解决方法综述.计算机系统应用,2024,33(11):38–47. http://www.c-s-a.org.cn/1003-3254/9702.html

Review on Security Challenges and Solutions to Edge Computing

WEN Mu-Qi[1], WEN Wu-Shao[2]

[1](School of Electrical and Information Engineering, Guangdong Baiyun University, Guangzhou 510450, China)
[2](School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China)

Abstract: Compared with centralized cloud computing frameworks, edge computing deploys additional “edge servers” between a cloud center and on-site intelligent devices to support those devices to quickly and efficiently complete computing tasks and event processing. In an edge computing system, there are a large number of on-site intelligent devices and heterogeneous edge computing servers. Also, stored data is sensitive and requires high privacy. These characteristics of edge computing systems make it difficult to ensure network security. Solving information and network security of edge computing systems is the key to the large-scale industrialization of edge computing technology. However, due to the limitations of computing capacity, network capacity, and storage capacity of edge server devices and on-site intelligent devices, traditional computer network security technology may not fully meet the requirements. Analyzing effective sensitive data protection technologies suitable for edge computing systems, such as federated learning, lightweight encryption, confused and virtual location information, and anonymous identity authentication, and exploring new technologies such as artificial intelligence and blockchain to prevent malicious attacks in edge computing will greatly promote the industrial development of edge computing.

Key words: edge computing; data protection; federated learning; blockchain; endogenous security architecture

① 基金项目: 广东白云学院校级科研重点项目 (2023BYKYZ05)
收稿时间: 2024-05-15; 修改时间: 2024-06-12; 采用时间: 2024-07-04; csa 在线出版时间: 2024-09-27
CNKI 网络首发时间: 2024-09-30

在“万物互联”的背景下, 部署的移动设备和物联网 (Internet of Things, IoT) 设备数量急剧增长, 包括智能家居、车联网、智能制造、健康监控等相关产业的现场智能设备. 这些设备不能完成复杂的运算任务, 由云计算数据中心完成复杂的运算任务. 现场智能设备的数据和需要处理的事件传送到云计算数据中心. 然而, 云计算数据中心与现场智能设备之间存在较大的数据处理和数据传输的延迟, 无法满足实时性要求高的诸多智能业务需求, 例如, 智能驾驶. 为了解决此问题, 在云计算数据中心和现场智能设备之间加上了“边缘服务器”, 实现低时延、高效率地完成现场智能设备的运算任务和事件处理. 边缘计算产业联盟 (ECC) 与工业互联网产业联盟 (AII) 联合发布了边缘计算参考架构 3.0, 3 层结构: 现场智能设备层、边缘计算层、云计算层.

根据国际数据公司 IDC 报告[1], 全球边缘计算支出在 2022 年达 1 760 亿美元, 年增长达到 14.8%. 2024 年, 中国边缘计算市场整体规模预计达到 1 803.7 亿元. 由于边缘计算架构采用分布式部署的特点, 边缘服务器和边缘智能终端分散各处, 更加容易受到各类安全威胁. 随着边缘计算设备数量的激增, 现场智能的动态性和资源约束性为其带来了新的安全隐患[2]. 病毒攻击、分布式拒绝服务攻击 (distributed denial of service, DDoS)、中间人攻击 (man in the middle attack, MITM)、侧信道攻击 (side channel attack, SCA) 和人工智能 (artificial intelligence, AI) 新技术攻击等手段对边缘计算系统的带来巨大的安全挑战. 保障边缘计算系统的安全运行需要解决保障信息的保密性、信息完整性、服务的可用性、系统可控性和系统的可审查性等关键核心问题. 本文重点探讨边缘计算层和现场智能设备层的敏感数据保护、恶意攻击防范的防范手段, 以及探索人工智能、区块链等新技术在解决边缘计算安全中的应用.

1 边缘计算的敏感数据保护

边缘计算层的各个边缘节点相对脆弱而易受攻击, 任何边缘节点中的单个漏洞都可能导致整体隐私信息泄露. 为了解决边缘计算的隐私信息保护的难题, 应该采用更轻量级和更具可扩展性的边缘计算敏感数据保护方法.

1.1 敏感数据信息泄露的安全挑战

存放在边缘服务器和现场智能设备的大量敏感数据, 是系统运行的重要数据组成部分. 保障敏感数据信息的保密性和完整性, 是系统安全运行和用户权益保障的体现. 敏感数据信息包括个人隐私信息、企业或社会机构不适合公布的信息和位置信息等[3].

用户个人信息包括身份信息、账户信息、私密信息、网络行为信息和社会关系信息等, 会存放在边缘服务器或现场智能设备. 在边缘计算系统中, 用户个人信息的泄露主要经由恶意程序、钓鱼欺诈、设备系统漏洞等几个路径, 以及边缘侧与云中心通信过程中的用户信息传送造成个人信息泄露. 用户个人信息的泄露可能给用户带来源源不断的垃圾电话、短信、邮件的骚扰, 也可能给有用户带来祸从天降的办卡透支、被诈骗等经济损失.

在边缘计算环境中, 用户的位置信息通常通过全球定位系统、WiFi、蓝牙、移动基站等手段被边缘计算系统所获取. 用户的位置信息是实现个性化服务和优化资源分配的关键数据, 也是实现智能感知边缘计算系统所需的关键信息之一. 然而, 用户的位置信息被这些设备捕获和记录, 也有泄露自身行动轨迹以及用户关键位置 (即工作场所和家庭住址) 的风险.

企业或者社会机构有其独有的经营信息和技术信息, 是企业重要的商业秘密. 企业或社会机构隐私信息的泄露, 可能造成经济效益不良影响, 不法分子可用此信息进行盗窃或欺诈; 可能造成品牌不良影响, 客户对企业的信任和忠诚度降低; 可能造成企业信息系统安全的不良影响, 不法分子会利用泄露的敏感信息进行网络攻击, 导致企业的信息化系统被攻破.

1.2 敏感数据信息保护手段

1) 联邦学习降低数据泄露风险

边缘计算系统中, 敏感数据的泄露的可能在敏感数据传输过程城中造成, 包括端设备与边缘服务器之间、边缘服务器与边缘服务器之间、边缘服务器与云中心之间的数据传输. 模型训练过程涉及大量的云、边、端之间的数据传送, 存在敏感数据泄露的风险. 联邦学习通过将边缘层和现场终端层直接地纳入机器学习操作, 实现机器学习操作在整个边缘计算中的分布. 联邦学习减少敏感数据的在边缘计算系统中的交互, 实现相关的数据处理安全和隐私保护[4].

谷歌公司首先提出联邦学习[5]. 一种分布式机器学习方法, 在不共享原始数据的情况下, 在本地设备上进行模型训练. 经过特定的时间, 将模型参数进行聚合,

实现全局模型的更新. 联邦学习不必要进行实时敏感数据的集中采集, 是减少通信成本和降低数据隐私泄露的有效方法. 联邦学习的每一次本地更新都会消耗边缘节点的计算资源, 每一次全局聚合都会消耗网络的通信资源. 联邦学习的全局聚合频率、模型训练精度和资源消耗之间存在复杂关系[6]. 资源分配和联邦学习效能之间的合理权衡是联邦学习关注重点之一.

自适应联邦学习 (adaptive federated learning, AdaFL)[7]从理论角度分析了分布式梯度下降的收敛边界, 并在此基础上提出了一种控制算法, 它能在给定的资源预算下, 确定局部更新和全局参数聚合之间的最佳权衡, 以最小化损失函数. 在网络原型系统和更大规模的模拟环境中使用真实数据集进行大量实验证明了 AdaFL 的有效性.

将联邦学习应用至边缘计算中, 避免了用户数据的泄露, 但受到恶意攻击的边缘节点容易发起拜占庭攻击, 对整体任务的效果造成影响. 拜占庭攻击是一般指的是在边缘计算系统中, 存在一些受到恶意攻击的现场智能设备, 上传有害的异常本地模型, 降低整体模型的收敛效率和收敛性. 为了减少拜占庭攻击的可能性, 采用人工智能技术判断边缘节点训练模型的有效性, 检测和摒弃受到恶意攻击的边缘节点的有害的异常本地模型.

联邦学习可支持隐私限制下实现不同组织间机器学习模型协同训练. 相关安全解决办法可从数据分布与共享, 机器学习模型保护, 隐私机制, 安全通信架构, 联邦学习规模与动机这些方面着手[8]. 联邦学习边缘计算系统和云数据中心为训练数据和模型提供高度的隐私保护, 确保数据信息的安全, 鼓励用户加入数据训练. 首先, 通过使用先进扰动技术, 如差分隐私等构造具有复杂数学解的组合定理来保护训练数据集免受数据泄露. 其次, 采用有效的威胁管理和攻击防御解决方案解智能终端设备和边缘服务器端安全问题. 通过在聚合服务器上部署攻击检测机制, 评估每个智能终端的贡献权重, 过滤出恶意客户端并在每个通信回合中识别出模型更新毒害攻击, 数据下毒攻击和逃逸攻击等类型. 在参数交换和更新广播期间, 采用数据加密, 通信认证以及区块链和智能合约的安全分类帐配置以确保无线通信通道的安全和隐私[9].

2) 轻量级加密技术

保障边缘计算的数据安全, 一是完整性, 确保数据在从源点到目的地的传输过程中不被未授权方访问或篡改; 二是保密性, 在边缘节点上进行数据处理时, 数据内容因加密难以被解读. 加密技术是保障数据安全的重要手段. 传统云计算所使用的加密技术虽然保证了通信的安全性, 但计算开销大, 不适用于边缘计算的计算能力和存储能力较低的场景. 轻量级加密技术成为保障边缘计算的敏感数据安全重要技术手段. 引入轻量级的区块链加密技术[10], 可解决物联网和边缘计算的场景, 实现满足业务需求的安全性和隐私性的同时, 只需要引入微不足道的流量, 处理时间和能耗方面开销. Wang 等人[11]提出的在网关上采用轻量级计算的云协助的安全用户认证方案, 以最小的网关计算和存储成本实现相关安全要求, 可有效支持远程用户与物联网设备及边缘服务系统之间的安全访问, 实现前向保密和多因素安全.

在智能驾驶的边缘计算网络中, Basudan 等人[12]设计了一种基于边缘学习的车辆众感框架的无证书路面状态监测聚合签名加密协议, 实现了相互认证、数据保密性和完整性. 在计算开销方面, 车载设备以及边缘服务器仅用执行线性级别的计算, 满足边缘计算系统中的低开销要求. 其优点是在任何情况下都无法获得移动传感器的完整密钥, 从而保证了签名的不可伪造性. 这是一种计算和存储开销小的轻量级加密技术, 保障了联邦学习技术应用过程的模型参数聚合和全局模型更新的敏感数据传输的安全.

3) 基于混淆与虚拟的位置信息保护技术

基于混淆与虚拟的位置信息保护方案, 是通过设计好的算法来隐藏用户的真实信息, 攻击者只能获得被混淆过的位置信息, 从而达到隐私保护的目的. 利用差分隐私来扰动位置, 使用虚拟位置进行位置欺骗或者匿名化, 是一种有效的位置信息保护方法. Yu 等人[13]提出的“一个两阶段动态差分位置隐私框架”适用于边缘计算系统. 在第 1 阶段考虑用户自定义的推断错误阈值和关于用户位置的先验知识, 确定一个位置子集作为保护位置集, 通过增加对手的预期位置推断错误来保护实际位置, 第 2 阶段在保护位置集上实现差分隐私的方式生成伪位置. 该混淆和虚拟技术在位置扰动和计算效率方面均取得了很好的效果.

为了增强位置信息在传输过程的安全性, 对位置信息进行加密. 传统的加密方法, 采用定期发布证密钥、使用消息标识符进行过滤或利用公钥 (完全) 同态

加密 (fully homomorphic encryption, FHE). 传统方法带来繁重认证开销, 不适用于资源受限边缘计算设备. 轻量级的基于同态加密的位置隐私保护技术是边缘计算系统安全保障的一种选择. 轻量级隐私保护身份验证 (lightweight privacy-preserving authentication, LPPA)[14] 设计了一种高效、安全的比较协议, 认证密钥和消息标识符无需服务器和用户之间的交互. 在验证前消除重复和无用的位置信息的加密消息, 实现轻量级的高效加密要求. 还可以采用基于 Paillier 密码系统的利用位置差异的接近检测协议. 在检测过程中, 参数以 Paillier 加密形式在设备之间传输, 以防止外部恶意攻击. 在定位过程中提出了一种实用的对称客户端-服务器协议, 可以保护用户的位置隐私不被泄露给任何一方. 而且, 该方案通过决策树理论推导出检测结果来减少计算量和通信成本.

4) 匿名身份认证技术

基于用户匿名的身份认证是一种用于保护用户隐私信息的安全技术. 其核心目的是在不泄露用户个人信息的情况下, 验证用户的身份. 匿名身份认证需要密码学算法和协议密码算法作为技术基础. 常用的密码学算法包括非对称加密算法、对称加密算法和哈希函数算法. 密码传输常用的关键协议包括传输层安全 (transport layer security, TLS) 协议、第三方认证服务协议 OAuth (open authorization)、OpenID Cennect 协议等. 用户可能拥有一个数字证书或一个加密的数字钥匙, 这些可以在不暴露其真实身份的情况下证明其身份的合法性.

为了节省资源开销, 采用非对称算法中的椭圆曲线密码算法, 它利用椭圆线上的点来执行加密操作或者数字加密认证, 其密钥尺寸更小和安全性更高. 配合传输层安全 (TLS) 协议, 形成一个具有保护数据完整性和保密性的高效的匿名身份认证技术方案. 在车联网的边缘计算系统中, Lo 等人[15]针对车辆传感器网络 (vehicle sensor network, VSN) 提出了一种基于条件隐私的身份认证方案. 该方案采用椭圆曲线加密机制和基于身份的签名机制, 支持匿名认证、数据完整性、可追溯性和批量签名验证. 同时, 签名过程中不需要任何双线性对的操作, 能够大幅度节约时间消耗和计算成本.

1.3 敏感数据信息的监控审计

通过追踪和记录数据处理活动, 确保数据完整性、安全性和合规性. 在边缘计算系统中, 借助第三方审计平台 (third-party audits, TPA) 执行审计是常见做法, 但 TPA 的可信性未知, 可能导致数据隐私泄露和篡改. TPA 需要满足以下两个功能: 一是 TPA 在有效的审计过程中无法获知存储数据的内容, 实现隐私保护; 二是由于数据存储服务器 (边缘节点) 和数据所有者 (现场智能设备) 存在计算能力、存储能力、网络通信能力等方面的限制, 数据的监控审计方案必然不宜过于复杂.

Yang 等人[16]针对数据完整性审计过程中的隐私泄露和批量审计问题, 提出了一种保护隐私的分布式数据审计系统. 系统采用第三方审计平台进行审计. 边缘侧设备将数据存储到云服务器上, 即可删除本地原始数据. 该方案利用同态认证器和随机掩码技术, 保证第三方审计平台在有效的审计过程中无法获知存储数据的内容, 从而达到保护隐私的目的. TPA 具有批量审计特性, 利用双线性聚合签名方法, 将审计协议扩展到多用户设置中, 从而实现分布式多任务审计.

2 边缘计算的恶意攻击与防范

边缘设备的多样性、分布式特性和计算资源局限性, 成为网络安全潜在的薄弱点. 安全薄弱点急剧增加, 边缘计算系统受恶意攻击的概率增加. 网络恶意攻击的形式很多, 包括 DDoS 攻击、绝服务攻击、网络钓鱼攻击、网络监听、物理访问攻击、侧信道攻击、中间人攻击、恶意软件攻击、人工智能技术攻击等. 本文主要讨论对边缘计算系统威胁较大和新型技术的恶意攻击及防范: DDoS 攻击、侧信道攻击、恶意软件攻击和人工智能网络攻击.

2.1 DDoS 攻击及防范

分布式拒绝服务 (DDoS) 攻击是一种来自不同地点的攻击者同时对特定目标发起大量请求, 从而瘫痪合法用户功能的网络攻击方式[17]. 由于边缘计算节点的计算资源有限, 相比于中心化的云计算数据中心, 它们更容易受到资源耗尽的影响. 边缘计算系统具有数量庞大的终端设备, 增加了边缘计算系统的脆弱性. 其中, 泛洪攻击是一种基于大量恶意网络数据包, 以关闭服务器正常服务为目标的 DDoS 攻击, 根据利用的网络协议, 可以分为 HTTP 泛洪、UDP 泛洪、ICMP 泛洪等.

基于历史统计和当前网络数据来建立模型是有效

的 DDoS 攻击检测方案. 在深度学习方法取得重大进展之前, 基于熵的方法是一种典型的检测方案, 在网络流量中, 熵可以用来量化特定特征的分布均匀程度. 如果观察到流量中特定特征的熵突然激烈变化, 这可能表明有大量相似的数据包被发送到网络中, 这是 DDoS 攻击的典型特征[18].

随着深度学习技术的发展, 使用深度学习的方法在边缘计算场景中进行 DDoS 检测. Li 等人[19] 提出了一种名为 IoTEnsemble 的异构网络异常检测框架, 基于树的活动聚类方法来聚合属于同一活动的网络流, 同一簇内的网络流具有相同的流量模式, 如图 1 所示. 由于基于聚类模型, 通过多个子模型集成来检测不同种 DDoS 造成的流量异常, 每个子模型只需要描述一种特定的活动, 这显著降低了对单个模型泛化能力的要求. 此外, 方法中的预处理模块、检测模块可以根据边缘设备或服务器的计算能力分散部署, 这种方案充分考虑到了边缘计算的资源受限场景, 适合部署于计算能力有限的边缘服务器和边缘设备中.

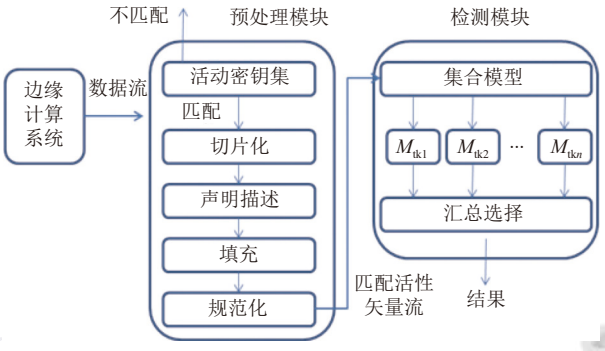

图 1 IoTEnsemble 网络异常检测框架

Li 等人[20]提出了一种名为 ADRIoT 的异常检测框架, 基于深度学习为每种类型的设备构建检测器, 当 ADRIoT 检测到设备已连接时, 从云端下载检测器并在本地执行异常检测, 如图 2 所示. 通过比较捕获的流量与合法的流量模式, 计算出流量偏差值, 进而判断现场智能设备产生的是正常流量还是 DDoS 攻击的异常流量. 针对边缘计算场景中现场智能设备资源有限的问题, 设计一种可变的协作机制来利用本地网络中其他现场智能设备的空闲资源. 协作机制通过资源评估和任务卸载来完成, 资源评估确定现场智能设备上的硬件资源是否已经过载, 如果过载, 将启动任务卸载过程来将检测任务分配给其他具有检测功能的边缘节点.

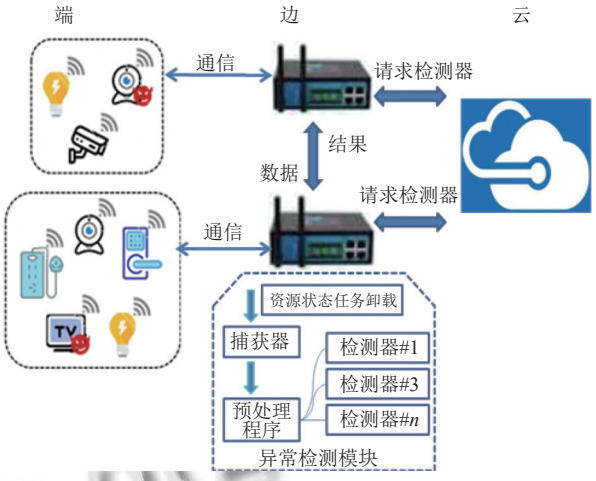

图 2 ADRIoT 架构

边缘计算的 DDoS 防护是持续的研究热点. Li 等人[21]提出了针对移动边缘计算服务应用级 DDoS 攻击的协作防御框架. He 等人[22]采用了一种基于博弈论方法去减轻边缘 DDoS 攻击. Zhou 等人[23]提出了基于在线和预测的云端协同清理的 DDoS 缓解技术实现对边缘计算系统的防护.

2.2 侧信道攻击及防范

侧信道攻击是指利用可公开访问的信息间接危害用户的安全与隐私的一种攻击方式. 在边缘计算中, 它利用了现场智能设备在执行计算任务时产生的非意图信息泄露. 侧信道攻击的一般形式为: 攻击者通过传感器等方式不断从目标边缘服务器或设备中获取某些侧信道信息, 然后将这些可公共访问的信息输入特定的算法或模型中, 让模型尝试学习到公共信息与敏感信息的映射关系. 根据攻击利用的具体渠道, 可以将侧信道攻击分成: 通信过程中的侧信道攻击和针对设备功耗的侧信道攻击. 故障和电源攻击的组合典型的设备功耗侧信道攻击方法, 例如故障模板攻击 (fault template attack, FTA)[24].

针对侧信道攻击的防御方法可以从两个角度来考虑: 对敏感数据施加额外的保护措施和限制对侧信道信息的访问. 在边缘计算系统中, 为了实现对敏感数据施加额外的保护, 采用差分隐私的方法. 差分隐私是在数据中引入一定程度的随机性, 从而影响利用侧信道信息推断隐私信息的过程. 可以添加拉普拉斯分布生成的噪声来实现差分隐私. Roy 等人[25]开发了 Airavat 系统, 底层是增强型安全 Linux 系统 (security-enhanced Linux, SELinux), 增加了强制访问控制的策略, 阻止进

程对隐私数据文件的访问. Airavat 系统是适用于边缘服务器的隐私保护平台. 针对计算结果的隐私泄露风险, Airavat 系统在深度优先搜索 (depth-first search, DFS) 中增加了安全标签, 在映射归约 (MapReduce) 的编程模型中的计算输出需要过滤安全标签, 会导致降低了数据传输和计算的效用. 然后, Airavat 系统集成了访问控制和差分隐私, 在实现差分隐私时去掉安全标签, 提高数据的效用, 可为敏感数据的分布式计算提供强大的安全性和隐私保证. Airavat 系统在 MapReduce 模型中实现了强制访问控制和差分隐私保护, 结合 SELinux、DFS、MapReduce 等保证边缘计算系统的计算过程和结果的隐私安全性.

2.3 恶意软件攻击及防范

恶意软件注入攻击是指攻击者将恶意代码如病毒、木马等植入到现场智能设备或服务器中, 恶意软件对设备构成严重的威胁. 在边缘计算中, 没有强大的计算能力和防御能力, 这种攻击尤其危险.

针对边缘服务器端的注入攻击. 跨站脚本攻击 (cross-site scripting, XSS) 是一种针对边缘服务器发起的攻击, 攻击者将恶意代码注入服务器可以自动执行的数据内容中, 这种攻击是由于边缘服务器不从数据内容中过滤代码引起的. 跨站请求伪造 (cross-site request forgery, CSRF) 和服务器端请求伪造 (server-side request forger, SSRF) 也是两种能够针对边缘计算系统的攻击方法[26]. 在 CSRF 攻击中, 攻击者诱使边缘服务器通过 Web 应用程序执行未经授权的操作, 而 SSRF 攻击则利用边缘服务器来读取和修改内部资源. 这两种攻击本质上都是利用粗粒度的验证机制, 以此来冒充合法的边缘服务器向其他边缘服务器发送命令.

针对现场智能设备端的注入攻击. 由于边缘计算系统中设备的高度异构性, 针对现场智能设备的恶意软件注入攻击具有多种形式. 例如 IoTReaper 病毒, 该病毒利用多种异构的现场智能设备存在的远程代码执行漏洞 (remote code execution, RCE), 通过互联网协议感染了百万级规模的设备[27]. Li 等人[28]发现了 Android WebView 中一个严重的设计漏洞, 攻击者可以远程将恶意软件注入正常 Android 设备中. 这种攻击方式可以实现资源窃取和 UI 劫持等多种攻击效果, 而无需目标主动安装任何恶意程序, 仅通过让目标访问一个恶意网站即可实.

为了应对恶意软件的威胁, 采用深度学习更准确地检测恶意软件. 深度学习技术可以自动学习复杂和高维的特征, 有效识别恶意软件. 在边缘计算系统中, 由于资源的限制, 轻量化部署是边缘计算系统使用深度学习技术的关键. Nguyen 等人[29]提出了一种基于轻量化深度神经网络 (deep neural network, DNN) 识别恶意软件的方法, 可以直接在资源受限的边缘服务器中运行. 但这种方法需要标记数据集来训练恶意软件检测模型, 对于新出现的恶意软件的检测也比较困难. 针对这个问题, Vasan 等人[30]提出在恶意软件检测中使用卷积神经网络 (convolutional neural network, CNN) 的方案, 将二进制恶意软件的一维数组转换为二维的图像数组, 然后使用 CNN 深度学习技术来处理. 这种方法可以清楚地可视化每个恶意软件的特定模式, 采用彩色图像比灰度图像具有更清晰的特征, 将二进制恶意软件转换为彩色图像可以实现更好地检测成功率.

2.4 人工智能技术的网络攻击防范

人工智能技术是防御网络攻击的一项技术, 也为攻击者提供了新的技术手段. 智能攻击能够识别网络的弱点, 选择最有效的攻击路径, 甚至模仿正常的网络流量以规避传统检测. 机器学习模型本身也可能成为攻击目标, 攻击者可能通过输入精心设计的数据来操纵或污染这些模型. 智能攻击的多变性和复杂性使得传统的防御系统难以应对, 新型的边缘计算系统安全需要采用更为先进和动态的防御策略[31]. 基于人工智能新技术的攻击主要体现在几个方面, 包括智能化泛洪攻击、智能化漏洞挖掘和攻击、智能化恶意代码攻击、智能化社工学攻击、智能化网络资产探索攻击, 如图 3 所示.

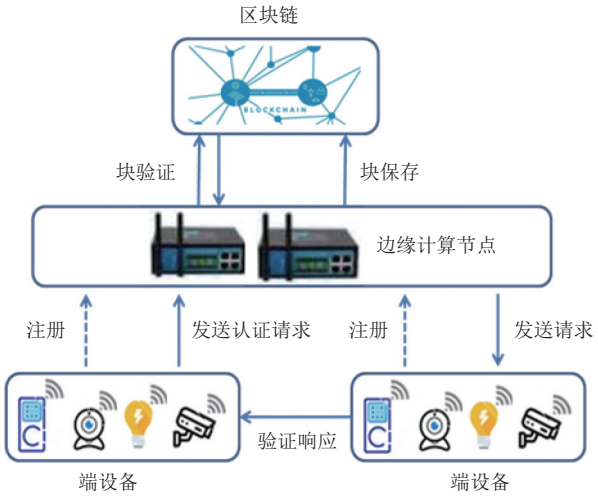

图 3 基于人工智能技术的攻击

解决人工智能技术的网络攻击, 应用人工智能技术防御是一个有效的选择. 分布式的人工智能和机器学习技术能对边缘计算系统的大规模数据实施快速控制和分析, 为检测网络异常显现提供手段, 并且提供预警机制, 从而提高边缘计算系统的安全性能. 使用分布式联合人工智能[32], 边缘设备无需交换数据, 进一步增强了网络安全性. 人工智能和机器学习不仅仅局限于威胁检测, 还被用于智能边缘安全决策. 通过对边缘设备和节点的行为进行学习, 这些技术能够实现智能的访问控制和认证管理, 确保只有合法的设备和用户能够访问网络资源[33]. 利用人工智能和机器学习技术, 对抗训练可以通过向训练数据中注入类似于攻击的扰动样本来增强模型的健壮性[34].

3 边缘计算安全防范的新技术

学术界与产业界目前在边缘计算的研究和开发非常活跃. 研究热点主要围绕着采用 AI 技术支持 5G/6G 网络下的边缘计算微服务[35]及其数据与隐私保护[36], 区块链安全技术应用, 边缘计算内生安全技术框架以及物理层的安全新技术的应用.

3.1 基于区块链的安全防范技术

3.1.1 基于区块链的密钥管理

区块链的不可篡改性使得其成为储存密钥的良好容器, 由于其去中心化的特性, 区块链网络相对传统网络面对恶意攻击有着更强的鲁棒性, 适用于边缘计算密钥加密管理的方法. 区块链技术已经在支持联邦学习的边缘计算机系统中得到应用, Nguyen 等人[29]讨论了该场景下的链设计、通信开销、资源分配、激励机制和安全及隐私保护.

Li 等人[37]设计了一种基于区块链的移动边缘计算密钥管理方案, 边缘计算的现场智能设备默认为一个移动终端设备. 当一个移动终端新加入子网时, 首先生成一个轻量级密钥对, 并将公钥上传到区块链, 打包成一个区块, 并发送到子网中的其他用户. 因此, 所有子网内的移动终端都能使用区块链中的公钥数据进行相互通讯. 当移动设备到另一个子网时, 该子网中的其他设备可以通过更高层次中子网的区块链记录来验证其身份, 由此来解决终端的高移动性问题. 区块链可以安全地储存终端公钥, 使得恶意节点无法获取公钥, 保障网络通信安全. 另外, Gowda 等人[38]不使用区块链来存储所有的公钥, 而是利用单向哈希链生成私钥和公钥对, 并使用椭圆曲线加密技术进行安全共享. 在两个边缘节点建立会话时, 由区块链对这两个节点进行身份验证, 并在验证成功后将两个设备生成的密钥对储存在区块链中. 在这种方案中, 即使恶意节点持有了其他边缘节点的公钥也无法与其进行通讯, 因为这种通讯无法被区块链所验证. 如图 4 所示.

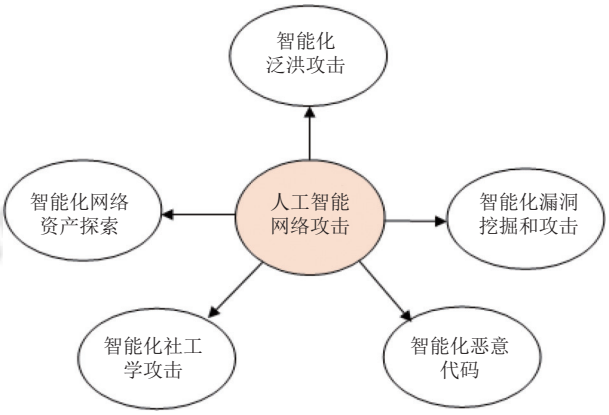

图 4 基于区块链的密钥管理架构

3.1.2 基于区块链的设备管理

边缘节点设备计算资源较弱, 容易受到攻击, 使用区块链技术来管理边缘计算设备. Huang 等人[39]提出了一种基于动态信任的边缘计算资源分配机制, 以解决资源分配过程中的可靠性问题, 并通过信任机制进行动态信用授予, 提高边缘计算的安全性. 此外, Liao 等人[40]提出了一种声誉和投票基础的区块链共识机制, 为了增强边缘计算启用的物联网系统中的安全性. 通过采用简单的声誉评估算法来选择区块提议者, 以减少共识过程的时间消耗, 并设计了过滤算法来检测和过滤具有恶意行为的节点, 从而提高共识的安全性.

在电网设备的具体管理过程中, Gai 等人[41]提出了一种基于区块链的电网设备管理方式. 该方法强调了在电网设备的生命周期管理中的重要性, 包括设备的安装、维护、更新和退役. 所提出的区块链架构不仅涉及了设备的状态跟踪和记录, 还包括了设备性能的实时监 测和维护安排. 该方案特别强调了区块链在确保设备安全和运行效率方面的作用, 它利用区块链的可追溯性来确保电网设备的操作历史被透明且安全地记录, 从而预防针对电网边缘设备 (发电站等) 的恶意攻击.

3.2 探索新型网络内生安全架构的新技术

在 5G 网络的应用业务场景的推动下, 在 6G 网络

计算的新技术研究的支撑下, 新型网络的内生安全架构技术针对边缘计算设备繁杂和算力与带宽资源有限等特性, 为边缘计算的网络安全防范方案提供的新思路和先方法.

新一代的移动通信网络通过构建多模信任的身份管理等基础设施, 从技术上消除当前基础设施存在的安全问题, 实现网络系统内的主体之间都能实现身份的相互识别、数据交互和价值交易等方面的可信交互. 支持多模态网络, 实现集中式、去中心化和第三方信任模式并存. 以数字身份和区块链等去中心化技术, 构建信息基础设施, 形成不同的信任等级和权限. 利用区块链多方信任、透明、不可篡改、可追溯、自动执行等特点, 支持各种情况下的安全、隐私和韧性能力.

新型的网络内生安全架构中包括可信引擎和可信使能单元两类安全能力集, 二者架构上构成了内生安全架构的框架, 内部的承载了可信引擎和可信使能单元, 形成一套移动通信网络安全的自运转体系.

3.3 物理层安全新技术

物理层安全技术 (physical layer security, PLS) 利用无线通信信道固有的随机性和硬件制造过程的难以克隆性, 在要求苛刻的场景中获取熵, 并且提供可验证性、保密性、完整性和隐私性的信息安全保障. 物理层安全技术是对加密技术的补充, 无需复杂的加密, 是系统设计者利用物理模型和环境的优势来防御攻击. 物理层安全新技术包括太赫兹技术 (Tera Hertz, THz)、可见光通信技术 (visible light communication, VLC)、可重构智能表面 (reconfigurable intelligent surface, RIS)[42]等.

未来移动网络将向太赫兹范围 (0.1–10 THz) 的高载波频率发展, 以提高未来无线网络的频谱效率和容量, 并提供无处不在的高速互联网接入. 太赫兹的通信传输具有很强的方向性, 可降低小区间的影响. 太赫兹可在大量子信道上实现跳频, 能增加攻击者检测和干扰信号的难度, 导致窃听和干扰的成本高昂. 可重构智能表面由无源反射元件组成的平面阵列构成, 能够动态调整反射系数, 从而控制反射信号的振幅和相移[43]. 将抗干扰、通信和安全统一在一体化架构下. 利用 RIS 提供的精细化感知和实时重构无线信道的能力, 通过人工塑造无线环境, 主动定制改造无线内生安全属性, 进一步丰富、放大、加速电磁环境的随机性、异构性和动态性.

4 结论与展望

分析边缘计算系统的应用形态, 从敏感数据保护、恶意攻击防范和与安全防范新技术 3 个层面入手, 讨论边缘计算安全面临的挑战和边缘计算安全解决方案. 分析边缘计算的数据信息、位置信息和用户身份信息所受到的威胁和挑战, 研究数据在生成、传输、处理和存储的各环节中的安全性和防泄露. 分析联邦学习机制、轻量级加密技术、混淆和虚拟位置信息保护方法、匿名身份认证技术、敏感数据监控审计等方法, 从而系统性地保障边缘计算的信息安全. 边缘计算系统中设备高异构性和低计算能力的特点, 更加容易受到网络攻击. 本文分析 DDoS 攻击、侧信道攻击、人工智能技术攻击和恶意软件注入攻击等网络攻击手段, 采用有效的防御策略和技术.

新型的网络安全防范技术也为边缘计算系统安全提供了新的手段和方法. 区块链的不可篡改性为数据的安全存储提供了坚实保障, 使得其成为储存密钥的安全容器. 6G 网络的内生安全架构技术, 为边缘计算的网络安全防范方案提供的新思路和先方法. 太赫兹技术和可重构智能表面等物理层安全新技术是对加密技术的补充, 无需复杂的加密, 系统设计者利用物理模型和环境的优势来防御攻击.

参考文献

1 IDC. New IDC spending guide forecasts double-digit growth for investments in edge computing. https://www.businesswire.com/news/home/20220113005169/en/New-IDC-Spending-Guide-Forecasts-Double-Digit-Growth-for-Investments-in-Edge-Computing. (2022-01-13)[2023-04-09].

2 Alsubhi K. A secured intrusion detection system for mobile edge computing. Applied Sciences, 2024, 14(4): 1432. [doi: 10.3390/app14041432]

3 Kumar KP, Prathap BR, Thiruthuvanathan MM, *et al*. Secure approach to sharing digitized medical data in a cloud environment. Data Science and Management, 2024, 7(2): 108–118. [doi: 10.1016/j.dsm.2023.12.001]

4 Arzovs A, Judvaitis J, Nesenbergs K, *et al*. Distributed learning in the IoT-edge-cloud continuum. Machine Learning and Knowledge Extraction, 2024, 6(1): 283–315. [doi: 10.3390/make6010015]

5 McMahan B, Moore E, Ramage D, *et al*. Communication-efficient learning of deep networks from decentralized data. Proceedings of the 20th International Conference on

Artificial Intelligence and Statistics. Fort Lauderdale: PMLR, 2017. 1273–1282.

6 Khan LU, Pandey SR, Tran NH, *et al*. Federated learning for edge networks: Resource optimization and incentive mechanism. IEEE Communications Magazine, 2020, 58(10): 88–93. [doi: 10.1109/MCOM.001.1900649]

7 Wang SQ, Tuor T, Salonidis T, *et al*. Adaptive federated learning in resource constrained edge computing systems. IEEE Journal on Selected Areas in Communications, 2019, 37(6): 1205–1221. [doi: 10.1109/JSAC.2019.2904348]

8 Li QB, Wen ZY, Wu ZM, *et al*. A survey on federated learning systems: Vision, hype and reality for data privacy and protection. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(4): 3347–3366. [doi: 10.1109/TKDE.2021.3124599]

9 Nguyen DC, Ding M, Pham QV, *et al*. Federated learning meets blockchain in edge computing: Opportunities and challenges. IEEE Internet of Things Journal, 2021, 8(16): 12806–12825. [doi: 10.1109/JIOT.2021.3072611]

10 Dorri A, Kanhere SS, Jurdak R, *et al*. Blockchain for IoT security and privacy: The case study of a smart home. Proceedings of the 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). Kona: IEEE, 2017. 618–623.

11 Wang CY, Wang D, Duan YH, *et al*. Secure and lightweight user authentication scheme for cloud-assisted Internet of Things. IEEE Transactions on Information Forensics and Security, 2023, 18: 2961–2976. [doi: 10.1109/TIFS.2023.3272772]

12 Basudan S, Lin XD, Sankaranarayanan K. A privacy-preserving vehicular crowdsensing-based road surface condition monitoring system using fog computing. IEEE Internet of Things Journal, 2017, 4(3): 772–782. [doi: 10.1109/JIOT.2017.2666783]

13 Yu L, Liu L, Pu C. Dynamic differential location privacy with personalized error bounds. Proceedings of the 24th Annual Network and Distributed System Security Symposium. San Diego: The Internet Society, 2017.

14 Zhou J, Cao ZF, Qin Z, *et al*. LPPA: Lightweight privacy-preserving authentication from efficient multi-key secure outsourced computation for location-based services in VANETs. IEEE Transactions on Information Forensics and Security, 2020, 15: 420–434. [doi: 10.1109/TIFS.2019.2923156]

15 Lo NW, Tsai JL. An efficient conditional privacy-preserving authentication scheme for vehicular sensor networks without pairings. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(5): 1319–1328. [doi: 10.1109/TITS.2015.2502322]

16 Yang K, Jia XH. An efficient and secure dynamic auditing protocol for data storage in cloud computing. IEEE Transactions on Parallel and Distributed Systems, 2013, 24(9): 1717–1726. [doi: 10.1109/TPDS.2012.278]

17 Setitra MA, Fan MY. Detection of DDoS attacks in SDN-based VANET using optimized TabNet. Computer Standards & Interfaces, 2024, 90: 103845.

18 Yu S, Zhou WL, Doss R, *et al*. Traceback of DDoS attacks using entropy variations. IEEE Transactions on Parallel and Distributed Systems, 2011, 22(3): 412–425. [doi: 10.1109/TPDS.2010.97]

19 Li RY, Li Q, Huang YC, *et al*. IoTEnsemble: Detection of botnet attacks on Internet of Things. Proceedings of the 27th European Symposium on Research in Computer Security. Copenhagen: Springer, 2022. 569–588.

20 Li RY, Li Q, Zhou JE, *et al*. ADRIoT: An edge-assisted anomaly detection framework against IoT-based network attacks. IEEE Internet of Things Journal, 2022, 9(13): 10576–10587. [doi: 10.1109/JIOT.2021.3122148]

21 Li HJ, Yang C, Wang LM, *et al*. A cooperative defense framework against application-level DDoS attacks on mobile edge computing services. IEEE Transactions on Mobile Computing, 2023, 22(1): 1–18. [doi: 10.1109/TMC.2021.3086219]

22 He Q, Wang C, Cui GM, *et al*. A game-theoretical approach for mitigating edge DDoS attack. IEEE Transactions on Dependable and Secure Computing, 2022, 19(4): 2333–2348. [doi: 10.1109/TDSC.2021.3055559]

23 Zhou RT, Zeng YF, Jiao L, *et al*. Online and predictive coordinated cloud-edge scrubbing for DDoS mitigation. IEEE Transactions on Mobile Computing, 2024. [doi: 10.1109/TMC.2024.3360077]

24 Vafaei N, Soleimany H, Bagheri N. Exploiting statistical effective fault attack in a blind setting. IET Information Security, 2023, 17(4): 639–646. [doi: 10.1049/ise2.12121]

25 Roy I, Setty STV, Kilzer A, *et al*. Airavat: Security and privacy for MapReduce. Proceedings of the 7th USENIX Conference on Networked Systems Design and Implementation. San Jose: USENIX Association, 2010. 297–312.

26 Costin A. IoT/embedded vs. security: Learn from the past, apply to the present, prepare for the future. Proceedings of the 22nd Conference of Open Innovations Association

FRUCT. Jyväskylä: IEEE, 2018. 412–414.
27 Vervier PA, Shen Y. Before toasters rise up: A view into the emerging IoT threat landscape. Proceedings of the 21st International Symposium on Research in Attacks, Intrusions, and Defenses. Heraklion: Springer, 2018. 556–576.
28 Li TX, Wang XQ, Zha MM, *et al*. Unleashing the walking dead: Understanding cross-APP remote infections on mobile webviews. Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. Dallas: ACM, 2017. 829–844.
29 Nguyen XH, Nguyen XD, Huynh HH, *et al*. Realguard: A lightweight network intrusion detection system for IoT gateways. Sensors, 2022, 22(2): 432. [doi: 10.3390/s22020432]
30 Vasan D, Alazab M, Wassan S, *et al*. IMCFN: Image-based malware classification using fine-tuned convolutional neural network architecture. Computer Networks, 2020, 171: 107138. [doi: 10.1016/j.comnet.2020.107138]
31 Mahadevappa P, Murugesan RK, Al-Amri R, *et al*. A secure edge computing model using machine learning and IDS to detect and isolate intruders. MethodsX, 2024, 12: 102597. [doi: 10.1016/j.mex.2024.102597]
32 Tomkos I, Klonidis D, Pikasis E, *et al*. Toward the 6G network era: Opportunities and challenges. IT Professional, 2020, 22(1): 34–38. [doi: 10.1109/MITP.2019.2963491]
33 Sattiraju R, Weinand A, Schotten HD. AI-assisted PHY technologies for 6G and beyond wireless networks. arXiv: 1908.09523, 2019.
34 Tramèr F, Kurakin A, Papernot N, *et al*. Ensemble adversarial training: Attacks and defenses. Proceedings of the 6th International Conference on Learning Representations. Vancouver: OpenReview.net, 2018.
35 Al-Doghman F, Moustafa N, Khalil I, *et al*. AI-enabled secure microservices in edge computing: Opportunities and challenges. IEEE Transactions on Services Computing, 2023, 16(2): 1485–1504. [doi: 10.1109/TSC.2022.3155447]
36 Letaief KB, Shi YM, Lu JM, *et al*. Edge artificial intelligence for 6G: Vision, enabling technologies, and applications. IEEE Journal on Selected Areas in Communications, 2022, 40(1): 5–36. [doi: 10.1109/JSAC.2021.3126076]
37 Li JX, Wu JG, Chen L, *et al*. Blockchain-based secure key management for mobile edge computing. IEEE Transactions on Mobile Computing, 2023, 22(1): 100–114. [doi: 10.1109/TMC.2021.3068717]
38 Gowda NC, Manvi SS, Malakreddy AB, *et al*. BSKM-FC: Blockchain-based secured key management in a fog computing environment. Future Generation Computer Systems, 2023, 142: 276–291. [doi: 10.1016/j.future.2022.12.042]
39 Huang YD, Zhang JR, Duan J, *et al*. Resource allocation and consensus on edge blockchain in pervasive edge computing environments. Proceedings of the 39th IEEE International Conference on Distributed Computing Systems (ICDCS). Dallas: IEEE, 2019. 1476–1486.
40 Liao ZF, Cheng SW. RVC: A reputation and voting based blockchain consensus mechanism for edge computing-enabled IoT systems. Journal of Network and Computer Applications, 2023, 209: 103510. [doi: 10.1016/j.jnca.2022.103510]
41 Gai KK, Wu YL, Zhu LH, *et al*. Permissioned blockchain and edge computing empowered privacy-preserving smart grid networks. IEEE Internet of Things Journal, 2019, 6(5): 7992–8004. [doi: 10.1109/JIOT.2019.2904303]
42 Wu QQ, Zhang R. Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming. IEEE Transactions on Wireless Communications, 2019, 18(11): 5394–5409. [doi: 10.1109/TWC.2019.2936025]
43 Xu JP, Xu AS, Chen LY, *et al*. Deep reinforcement learning for RIS-aided secure mobile edge computing in industrial Internet of Things. IEEE Transactions on Industrial Informatics, 2024, 20(2): 2455–2464. [doi: 10.1109/TII.2023.3292968]

(校对责编: 张重毅)