 DOI: https://doi.org/10.12345/bdai.v6i4.30542

Exploration of Data Privacy Protection Strategies in Edge Computing Environments

Ming Deng

China Mobile Communications Group Guangdong Co., Ltd. Guangzhou Branch, Guangzhou, Guangdong, 510220, China

Abstract

With the rapid development of big data and Internet of Things technologies, data privacy protection has become increasingly prominent in the edge computing environment. Edge computing effectively reduces the latency of cloud computing in real-time applications and data processing, and disperses data processing tasks to the network edge. This paper discusses the privacy risks faced by personal sensitive information during transmission, storage and processing in edge computing, as well as the limitations of traditional privacy protection schemes. The research on emerging technologies such as differential privacy, homomorphic encryption and data anonymization applied in the edge computing environment was analyzed, and it was pointed out that these technologies have achieved a balance in maintaining user privacy while enhancing data availability.

Keywords

Edge Computing; Big data; Privacy protection; Differential privacy; Homomorphic encryptio

边缘计算环境下的大数据隐私保护策略探索

邓茗

中国移动通信集团广东有限公司广州分公司，中国·广东 广州 510220

摘　要

随着大数据和物联网技术的迅猛发展，数据隐私保护在边缘计算环境中日益突出。边计算有效减少了云计算在实时应用、数据处理方面的延时，将数据处理任务分散到网络边缘。本文探讨了在边缘计算中，个人敏感信息在传输、存储和处理过程中面临的隐私风险，以及传统隐私保护方案的局限性。在边缘计算环境下应用的差分隐私、同态加密、数据匿名化等新兴技术研究中进行了分析，并指出了这些技术在增强数据可用性的同时，在维护用户隐私方面也得到了平衡。

关键词

边缘计算；大数据；隐私保护；差分隐私；同态加密

1 引言

大数据与物联网技术的蓬勃发展为人类社会带来了前所未有的数字化变革。边缘计算作为一种新兴的计算范式，通过将数据处理任务下沉到网络边缘，有效解决了云计算模式在处理实时应用和大量数据时的延迟与网络拥塞问题 [1]。

2 研究目的与问题

作为一种新兴的数据处理架构，边缘计算带来了新的隐私安全挑战，同时提供了数据处理的效率和实时性。该研究旨在探索数据在传输、存储和处理过程中的数据隐私保护问题，以及如何实现高效的边缘计算，前提是确保数据隐私。

在边缘计算环境中，数据安全与隐私保护面临着独特的挑战。边缘节点通常在本地处理数据，这种分布式的特性虽然降低了数据在网络传输过程中泄露的风险，但也增加了信息泄露和攻击的可能性 [2]。由于边缘设备通常具有有限的计算和存储资源，这限制了其处理大规模数据的能力，同时也给隐私保护技术的实施带来了挑战。

这项研究需要解决的重点问题包括：如何实现数据的高效加密，以及在边缘节点上的安全传输；隐私保护算法适应边缘计算特点如何设计；边缘计算的实时性能如何保证数据安全而不受影响；以及数据获取控制的有效机制如何建立。这些问题的解决对于构建安全可靠的边缘计算环境具有重要意义。

3 边缘计算概述

3.1 边缘计算的定义及特征

边缘计算是在网络边缘节点部署计算资源和存储资源，使数据处理位置与数据源、终端设备更接近的一种去中心化的计算模型。通过在本地设备或边缘服务器上进行数据处

【作者简介】邓茗（1986年12月），男，汉族，广东省吴川市，中级工程师，工程硕士，主要从事大数据技术、数据治理、人工智能和信息安全等研究。

理，边缘计算能够满足实时性要求高的应用场景，同时有效降低网络延迟和带宽消耗。

与云计算的传统相比，边缘计算在技术上的优越性和应用的特点是独一无二的。数据处理方面，边缘计算通过将任务就近传送至边缘服务器，在有效降低云端负载压力的同时，任务传输时延将得到明显降低。同时，这种分布式架构还具有更高的安全性，能够更好地保护用户的数据隐私。在实际应用中，边缘计算主要负责动态数据的聚合和本地决策，并通过分布式推理为智能服务提供支持 [3]。

3.2 边缘计算的应用场景

边缘计算因其分布式特性和实时处理能力，已在多个领域展现出显著优势。在智慧矿山建设中，边缘计算实现了数据流量的动态调整，使得仅必要数据被传输至云端处理，有效降低了数据处理成本和设备能耗。

边缘计算在工业物联的场景下，为实时资料分析提供了一个强有力的依托。通过在边缘节点部署数据处理单元，系统可以对温度传感器、摄像头等设备产生的实时数据进行及时分析和处理。这些数据通过 Wi-Fi、蓝牙等无线通信方式进行传输，在确保数据准确性和完整性的同时，还需要处理异常值和重复值。

同样依靠边缘计算的技术优势来打造智慧城市。边缘节点处理用户侧的数据，只将分析结果和统计数据上传到云端，将涉及隐私的原始数据保存在本地。这种架构不仅能够实现智能交通管理、环境监测等功能，还能有效避免数据在远距离传输过程中可能遭受的泄露或劫持风险。为保护用户隐私，边缘计算节点采用了差分隐私和联邦学习等技术，在不泄露个人敏感信息的前提下完成数据分析和处理任务。

4 大数据隐私保护现状

4.1 当前隐私保护技术分析

在边缘计算场景中，现有的隐私保护技术主要包括基于加密的隐私保护技术、基于差分隐私的技术以及基于 K-匿名的隐私保护技术 (PrivateProtection)。其中，基于加密的技术提供了良好的隐私保护效果和较高的服务质量，但存在密文长度增加和加解密过程带来的计算开销大等问题。

差分隐私技术是建立在严格的数学理论基础上的，它的优点是不受攻击者的背景知识的影响。该技术通过在数据中加入噪音来保护个人隐私，从而使攻击者无法对具体的个人进行信息识别。同时，同态加密可以在加密状态下计算数据，实现数据处理功能的同时保护数据隐私。

边缘计算环境下的隐私保护技术在实际应用中面临挑战，例如数据实时性要求很高，计算资源受到限制。边缘计算中对实时处理的需求，传统的隐私保护方案已经不能完全满足。差分隐私逐渐受到广泛关注，因为其数学框架定义严格，可以对隐私数据进行有效保护。

4.2 法律法规与隐私保护

边缘计算环境下的数据隐私保护不仅需要技术支撑，更需要法律法规的保驾护航。目前，包括身体健康状况、位置信息、日常习惯等敏感信息的收集和使用在内的大量边缘计算场景中的个人数据引发各界广泛关注。这些数据面临隐私外泄的严重危险性在边缘节点的获取、传输及存储过程中都有所面临。

物联网设备的安全设计必须加强，才能应对这些挑战。在设计阶段就应秉承”安全优先”的理念，以安全的硬件和软件组件来确保设备有足够的防护能力。同时，需要在设备中嵌入防火墙、入侵检测系统以及安全的通信协议等有效的安全机制，及时发现和阻止潜在的安全威胁。这些技术措施的实施必须建立在完善的法律法规基础之上，通过制度保障确保各项安全机制得到有效执行，从而构建起全方位的数据隐私保护体系。

5 边缘计算环境下的隐私保护挑战

5.1 数据传输中的隐私风险

5.1.1 安全传输协议的局限性

在边缘计算环境中，像 TLS/SSL 这样的安全传输协议在传输数据过程中扮演着重要角色。通过这些协议，系统能够对数据进行加密保护，防止数据在传输过程中被窃取或篡改。但随着边缘计算应用场景的日益复杂，这些传统的安全传输协议也显现出诸多局限性。

边缘计算环境下的安全传送，面临着别具一格的考验。传统的 TLS/SSL 协议在资源受限的边缘设备上可能会造成较大的计算开销，影响数据处理的实时性能。对于物联网设备而言，这种性能损耗尤为明显，需要在安全性与效率之间寻找平衡点。同时，边缘节点的分散性和异构性也使得统一的安全协议标准难以满足各类设备的需求，增加了协议实施的复杂度。

5.1.2 数据包嗅探与中间人攻击

在资料传送过程中边缘计算设备面临严重安全隐患。数据包嗅探攻击者可能获取用户的敏感信息，通过截取网络进行数据传输。中间人攻击由于大量数据需要频繁地在终端设备和边缘节点之间传输，因此伪装成合法的通信方来窃取或篡改传输中的数据，这在边缘计算环境中是特别危险的。

传统的加密算法在边缘计算场景中可能不再完全适用，因为边缘设备的计算资源有限，复杂的加密过程可能导致性能下降。攻击者可以利用这一特点，通过部署在网络关键节点的嗅探工具，获取用户的个人信息和敏感数据。为了应对这些威胁，在确保数据可用性的同时，还需要使用差分隐私和同态加密等技术来保障数据的私密性。

5.2 计算过程中的隐私威胁

5.2.1 计算节点的安全性问题

边缘计算环境中的计算节点普遍存在资源受限、物理暴露等安全风险。由于这些节点通常部署在接近用户的物理环境中，更容易遭受物理攻击和篡改，对数据隐私保护构成了严峻挑战。边缘计算设备的资源限制特性使其难以部署传统的安全防护措施，现有的集中式安全机制也无法直接应用于边缘计算架构。

对于密码处理器的性能和能耗的权衡，边缘设备的实

时性和资源受限等特征在工业互联网场景中提出了更高的要求。恶意的或被俘获的边缘节点接入工业网络可能造成巨大损失，确保边缘节点的安全接入成为一个亟待解决的问题。边缘节点的物理暴露风险使得攻击者有机会直接访问设备硬件，通过硬件漏洞或侧信道攻击等方式获取敏感数据。

5.2.2 日志数据的隐私泄露风险

边缘计算环境下的日志数据包含了大量用户的敏感信息，如设备位置信息、服务内容和使用频率等个人隐私数据。这些隐私数据由于边缘节点的保护措施不够完善，极易遭到恶意攻击，有被盗取或泄露的危险。一旦泄露包含的用户行为轨迹、使用习惯等信息，作为系统运行的重要记录，日志数据将严重威胁用户隐私。

多个设备在资源共享过程中，通过监视共享资源的流量，在单个边缘服务器共享同一台服务器时，攻击者获取用户的隐私资料的可能性很大。这些数据涉及身份信息、位置信息等敏感内容，需要采取专门的隐私保护措施。尽管边缘节点获取的数据仍面临隐私保护算法失效的问题，但边缘节点通过将计算任务下沉到边缘，一定程度上避免了数据在远距离传输中泄露隐私的风险。

6 边缘计算环境下的隐私保护策略

6.1 数据去标识化技术

6.1.1 伪匿名化方法

通过对原始数据进行智能化处理，在有效降低用户身份识别风险的同时，伪匿名化方法作为边缘计算环境中保护数据隐私的重要技术手段。当前主流的伪匿名化技术包括 k-匿名化和差分隐私两大类，这些技术在数据效用性和隐私保护程度之间寻求平衡。

传统的图像素化、模糊化技术，在实际应用中，虽然识别率可以明显降低，但往往会破坏数据的基本属性，造成数据可用性的明显降低。在边缘计算环境中，伪匿名化方法的选择需要考虑边缘节点的计算能力限制。轻量级的 k- 匿名化技术适合在资源受限的边缘设备上执行，而复杂的智能匿名化算法则更适合在边缘服务器端进行处理。通过合理分配匿名化任务，既能保证数据隐私安全，又能充分利用边缘计算的分布式特性。

当前伪匿名化技术在边缘计算环境中仍面临诸多挑战，如匿名化效率、数据效用性损失等问题亟待解决。未来的研究方向应着重于开发更加高效、智能的匿名化算法，并探索匿名化程度与数据效用之间的最优平衡点。

6.1.2 隐私保护数据湖建设

隐私保护数据湖是一种创新性的数据管理架构，它通过边缘计算技术实现数据的分布式存储和处理，有效降低了数据在传输过程中的泄露风险。数据湖通过将数据处理任务转移到网络边缘，实现了数据的本地化处理，从而减少了敏感数据跨网络传输的频率。隐私保护数据湖的安全机制包含多个层面。在数据传输层面，系统通过安全协议如 TLS/SSL 来加密，以保证传输过程中的数据万无一失。在访问控制层面，通过实施 RBAC 等机制，严格管理用户对数据的访问权限。

6.2 加密与访问控制机制

6.2.1 端到端加密技术

作为边缘计算环境下保护数据隐私的重要手段，端到端加密技术通过确保数据在源与端之间的全程加密传输，有效防止了未授权访问和数据在中间节点上的外泄。在该技术架构下，数据从发送方加密后直至接收方进行解密前始终保持加密状态，任何中间节点都无法获取明文信息。

在传统的数据传输加密方案中，只有在传输过程中才会对数据进行加密，而在源端和端端可能会出现无防护的情况。这种方式容易导致数据在端点遭受攻击时发生泄露。

端到端加密技术的创新之处在于将安全保护范围扩展至整个数据生命周期，显著提升了数据处理各环节的安全性。

加密过程可以用以下数学模型表示：

$$C=E_k(M),M=D_k(C)$$

其中，E_k 表示使用密钥 k 进行加密的函数，表示使用密钥 k 进行解密的函数，M 为明文消息，C 为密文。

在物联网和边缘计算场景中，端到端加密尤其适用于即时通讯、电子商务支付系统等对数据隐私要求较高的应用。通过结合差分隐私 ^DP 等技术，确保数据的可用性的同时，也能对数据隐私进行保护，使隐私保护和数据价值达到平衡。通过在查询结果中加入随机杂讯对个体隐私进行保护，在保持统计特性的准确性的同时，差分隐私是一种数据隐私保护技术。

6.2.2 动态权限管理策略

边缘计算环境中的动态权限管理策略采用基于上下文感知的访问控制机制，根据设备状态、网络环境和用户行为等因素动态调整数据访问权限。通过实时监控和评估系统环境，权限管理系统能够自适应地修改访问规则，在保障数据安全的同时确保系统运行效率。动态权限管理的核心优势在于能够灵活地调整数据流量分配，让只需传输必要的数据到云端处理，就可以根据实时的需求和网络状况进行管理。

7 结论

随着物联网技术的广泛应用，边缘计算节点的数量将持续增长，这使得隐私保护面临着更大的挑战。利用数据加密、身份验证及权限控制等技术，在边缘节点内部及云端传送过程中，为应对这些挑战，建立边缘安全综合治理机制，保证用户数据的安全。

参考文献

[1] 孙剑明，赵梦鑫. 边缘计算下差分隐私的应用研究综述[J]. 计算机科学, 2024.

[2] 徐帅. 数据隐私保护与法律责任：新形势下的挑战与应对[J]. 法制博览, 2024.

[3] 曾文广. 基于边缘计算的互联网系统设计与优化[J]. 电脑知识与技术, 2024.