大数据环境下的电子信息安全策略分析

孙堂元　陆亚飞
华信咨询设计研究院有限公司　浙江　杭州　310014

摘　要：为了构建电子信息安全防护体系，文章围绕大数据环境下电子信息安全策略展开分析，首先阐述了大数据环境下电子信息安全的特征，论述了数据泄露、网络攻击及安全管理基础设施缺陷等主要威胁，并从数据加密、访问控制、入侵检测、差分隐私等技术角度提出安全防护策略。研究结果表明，在大数据环境下，数据安全面临多重挑战，传统安全手段难以完全应对，需要构建以技术、管理、意识教育相结合的安全体系，本研究为电子信息安全的防护体系构建提供理论支持和实践指导。

关键词：大数据环境；电子信息安全；数据加密；网络攻击

引言

随着大数据技术的迅速发展，数据已成为社会经济、科技创新的重要资产。然而，在数据存储、处理、共享及传输过程中，信息安全问题日益凸显，数据泄露、隐私侵犯、网络攻击等威胁层出不穷，给政府、企业及个人带来巨大风险。传统的信息安全防护体系难以适应大数据的高并发、异构性、分布式存储等特点，亟需新的安全策略来应对挑战。当前，数据加密、访问控制、入侵检测、防御系统、差分隐私等技术已逐步应用于数据安全管理，但仍存在实践难度与适应性问题。因此，本文结合大数据环境的特性，分析电子信息安全面临的主要威胁，并探讨有效的安全策略，以期为相关领域的安全治理提供参考和理论支持。

1 大数据环境下电子信息安全的特征

在大数据环境下，数据不仅仅是传统结构化数据库中的信息，还包括非结构化数据（如文本、音频、视频）以及半结构化数据（如XML文件、JSON格式数据等）[1]。随着云计算、物联网、人工智能等技术的发展，大数据的影响力进一步扩大，成为推动社会经济、科技进步和企业决策的重要驱动力。大数据的特征主要表现为数据体量大（Volume）、数据流转速度快（Velocity）、数据类型多样（Variety）和数据价值密度低（Veracity），即通常所说的“4V”特性。

（1）海量性

随着互联网、社交媒体、智能终端、物联网设备等技术的普及，全球数据总量呈指数级增长。根据国际数据公司（IDC）预测，全球数据总量在未来几年内将达到数百ZB（Zettabyte），远超传统数据存储和管理能力。例如，社交平台每天产生数亿条信息，智能设备不断上传传感数据，企业经营过程中产生的交易数据、用户行为日志等也在不断积累。如此庞大的数据规模要求数据存储和处理技术的突破，如分布式存储、云计算、大数据架构Hadoop等的广泛应用，以应对超大规模数据处理需求。

（2）多样性

在传统数据管理系统中，数据通常以结构化形式存储，如关系型数据库中的表格数据。然而，在大数据环境下，数据来源广泛，包括社交媒体文本、图片、音频、视频、网页日志、传感器数据等非结构化和半结构化数据。这些不同类型的数据需要多样化的存储、转换和分析方法。例如，在医疗领域，患者的电子健康记录（EHR）、医学影像、基因测序数据等均属于不同的数据类型，需要结合结构化数据库和分布式文件系统进行存储和分析。

（3）高速性

在物联网、金融交易、网络安全监测等场景中，数据是以毫秒级甚至纳秒级的速度产生和传输的。例如，高频交易系统需要对市场数据进行实时分析，以便在极短时间内作出交易决策；智能交通系统需要对传感器数据进行实时计算，以优化交通流量调度。传统的批量处理模式已难以满足此类高时效性需求，因此，流数据处理、边缘计算、分布式计算等技术逐步兴起，能够支持高并发、低延迟的数据处理能力。

（4）真实性

大数据并非所有数据都具备高质量，部分数据可能存在冗余、噪声、不一致性等问题，因此，如何确保数

作者简介：孙堂元（1984.05-），男，汉族，籍贯：甘肃省武威市，本科，工程师，研究方向：通信技术、大数据、物联网

据的真实性和可靠性成为大数据应用的关键挑战。例如，社交媒体中的用户评论可能存在虚假信息，网络爬虫收集的数据可能包含大量冗余内容，物联网设备上传的数据可能受传感器误差影响而不够准确。为了提高数据质量，需要采用数据清洗、数据去重、异常检测等方法，结合机器学习和自然语言处理技术进行数据预处理，以提高数据的可信度，为数据分析和决策提供可靠依据[2]。

（5）价值密度低

尽管数据量庞大，但其中真正有用的信息仅占很小比例。例如，在视频监控数据中，大多数时间的画面没有异常，仅少量镜头捕捉到关键事件；在企业运营数据中，大量日志数据只是记录系统状态变化，只有部分数据对业务优化具有直接指导意义。因此，大数据分析的目标不仅是存储和管理数据，更重要的是如何从海量数据中挖掘有价值的信息，提高数据的利用率。

2 大数据环境下电子信息安全面临的主要威胁

2.1 数据泄露与隐私侵犯

数据泄露与隐私侵犯在大数据环境下日益严重，主要体现在数据存储、传输及访问过程中存在的安全隐患。由于大数据系统需要整合多源异构数据，数据在采集、处理和存储过程中可能涉及多个环节，任何一个环节的安全漏洞都可能导致数据泄露。例如，数据存储通常依赖于云平台，而云存储本身面临未授权访问、恶意攻击及数据篡改等威胁，若访问控制策略不当或权限管理存在缺陷，黑客或内部人员可能利用安全漏洞窃取或滥用敏感信息。数据共享机制的开放性使得数据泄露风险进一步增加，尤其是在跨机构数据流转过程中，数据访问权限未严格控制或数据脱敏处理不足，可能导致用户隐私信息被非法利用。大数据分析依赖于海量用户行为数据，而这些数据通常包括个人身份信息、社交关系、消费记录、位置信息等，一旦发生数据泄露，不仅会造成个人隐私暴露，还可能引发身份欺诈、金融诈骗、精准网络攻击等更严重的安全问题[3]。

2.2 网络攻击形式的多样化

在大数据环境下，网络攻击的形式呈现出更高的复杂性、多样性和隐蔽性，使得传统安全防护手段难以有效应对。首先，分布式拒绝服务（DDoS）攻击成为一种常见的攻击方式，由于大数据平台通常依赖于云计算和分布式存储系统，攻击者可以利用僵尸网络生成大量虚假流量，使得系统负载骤增，导致正常用户无法访问关键资源[4]。数据投毒（DataPoisoning）攻击在机器学习和数据分析系统中尤为突出，攻击者通过篡改训练数据，干扰算法的学习过程，进而影响预测结果，可能导致自动化决策系统作出错误判断。另一方面，勒索软件攻击在大数据环境中威胁日益加剧，攻击者通过加密企业或机构的重要数据，要求受害者支付赎金以恢复数据访问权限，由于大数据存储系统中包含大量关键业务数据，一旦遭受攻击可能导致严重业务中断。与此同时，零日漏洞攻击和高级持续性威胁（APT）攻击在大数据环境中也广泛存在，黑客利用软件或系统中尚未修复的安全漏洞，长期潜伏在目标系统内，窃取敏感数据或操纵系统运行。跨站脚本（XSS）和SQL注入等攻击方式依旧对大数据平台构成威胁，攻击者利用数据接口或数据库查询语句的安全漏洞，在系统内植入恶意代码或篡改数据内容，从而控制数据流动并获取未授权的信息访问权限。

2.3 数据安全管理基础设施的缺陷

数据安全管理基础设施的缺陷主要体现在安全架构不完善、权限管理机制不健全、数据加密与存储策略不足、日志监控与审计体系不完善等方面。首先，大数据系统通常采用分布式架构，由多个数据中心和计算节点共同组成，在数据存储、计算和传输过程中，安全防护措施可能存在不一致性，部分节点可能因配置错误或安全策略缺失成为攻击的突破口。此外，权限管理体系的漏洞是数据安全管理基础设施的重要短板，传统的访问控制模型难以适应大数据环境下复杂的用户身份认证与数据权限分级管理，导致数据访问权限过度开放或权限分配不合理，使得未授权用户或恶意内部人员能够访问敏感数据，从而引发数据滥用或泄露风险。与此同时，数据加密与存储机制的不完善也加剧了安全隐患，部分企业或机构在存储数据时未采取严格的加密措施，或者由于计算性能需求而降低了数据加密强度，使得攻击者能够通过数据库漏洞、内存取证等方式直接获取明文数据。另一方面，大数据系统的日志监控与安全审计机制往往不够完善，由于数据流转速度快、涉及系统庞杂，传统的日志分析手段难以实时检测异常行为，攻击者可以利用系统监测盲区长期潜伏，窃取数据或操纵系统而不被发现。

3 大数据环境下电子信息安全策略

3.1 数据加密与访问控制

数据加密作为最基础的安全措施，可以通过对称加密（如AES）和非对称加密（如RSA）等算法保护数据在存储和传输过程中的安全，同时结合哈希函数和数字签名技术，确保数据完整性和来源的可信性[5]。基于属性的加密（ABE）和同态加密等新型加密技术进一步增强了数据的安全防护能力，支持精细化的访问控制策略。

在访问控制方面，传统的基于身份的访问控制（IBAC）和基于角色的访问控制（RBAC）在大数据环境下存在局限，因此，基于属性的访问控制（ABAC）成为更适用的模型，通过定义访问者的多个属性（如身份、角色、设备、时间、地点等）来动态调整权限，提高安全性与灵活性。为防止访问权限滥用，可以结合多因素认证（MFA）、零信任架构（Zero Trust）以及区块链技术，以增强身份认证的可靠性和透明度，确保只有合法授权的主体才能访问敏感数据。同时，访问控制应采取最小权限原则（Principle of Least Privilege, POLP），限制用户对数据的操作权限，防止内部人员或恶意攻击者非法篡改或泄露数据。

3.2 入侵检测与防御系统

在大数据环境下，入侵检测与防御系统（IDS/IPS）能够实时监测数据流和系统活动，识别潜在的攻击行为并进行响应。当前的入侵检测系统分为网络入侵检测（NIDS）和主机入侵检测（HIDS），其中，NIDS主要监测网络流量，识别异常数据包或恶意流量，而HIDS则聚焦于系统日志、文件完整性及行为模式分析，以发现主机层面的安全威胁。在大数据环境下，IDS系统结合机器学习和人工智能技术，可通过深度学习、支持向量机（SVM）和随机森林等算法分析历史攻击数据，建立动态入侵检测模型，提高攻击识别的准确性。入侵防御系统（IPS）作为IDS的补充，可在检测到可疑活动时自动采取响应措施，如阻断恶意流量、限制可疑用户权限等，以降低攻击影响。随着分布式计算架构的普及，基于云计算和边缘计算的入侵检测系统也逐步应用，通过在不同节点部署轻量级检测模块，实现全局监控与快速响应。

3.3 差分隐私技术

差分隐私技术主要通过引入随机噪声来掩盖个体数据的真实值，从而在保证数据可用性的同时防止恶意分析者推断敏感信息，其基本原理是在查询结果或数据发布过程中，依据拉普拉斯机制或高斯机制添加随机噪声，使得外部攻击者无法精准还原原始数据，同时保证统计分析的有效性。与传统数据匿名化技术（如数据脱敏和k-匿名）不同，差分隐私提供了严格的数学安全性保证，即攻击者即使掌握所有其他数据，也无法确定某个特定个体是否存在于数据集中，因此，在保护数据隐私的同时，能够抵御重识别攻击和数据关联分析。

结语

综上，本文主要围绕大数据环境下的电子信息安全问题，系统分析了数据泄露、隐私侵犯、网络攻击及安全管理基础设施缺陷等主要威胁，并提出基于数据加密、访问控制、入侵检测、防御系统、差分隐私等多层次安全策略。研究表明，面对大数据环境下日益复杂的安全风险，单一技术手段难以全面应对，需要建立技术、管理、法规、教育等多维度协同的安全防护体系。未来，随着人工智能、区块链、量子密码学等新兴技术的发展，电子信息安全防护手段将进一步优化，但仍需在安全性、可操作性、资源消耗等方面进行权衡与改进。本研究为电子信息安全防护体系的构建提供了理论支持，也为未来研究方向奠定了基础。

参考文献

[1]潘娟娟,李明.基于大数据技术的电子支付信息安全加密系统[J].现代电子技术,2021,44(13):71-74.

[2]张艳.大数据时代个人电子信息安全保护——以智能手机为例[J].科技资讯,2022,20(8):19-21.

[3]权莹.电子信息工程领域中大数据传输的网络安全与效率优化[J].网络安全和信息化,2024(6):146-148.

[4]翟敏.电子信息工程领域中大数据传输的网络安全分析[J].电脑采购,2023(51):25-27.

[5]马骥.大数据背景下电子信息档案管理创新研究[J].电子元器件与信息技术,2024,8(2):133-136.