多维威胁与融合防护：大数据信息安全技术综述与体系构建

罗金华

云南师范大学信息学院，云南昆明，650500；

摘要：大数据与云计算深度融合的时代，信息安全面临多维威胁。本研究旨在构建一个大数据环境下的动态综合防护体系。首先从技术、管理、人员及环境四个维度系统剖析核心安全威胁。继而综述数据加密、访问控制及零信任架构等热点技术的原理与适用场景。在此基础上构建融合技术、管理、人员与运营的协同防护框架。

关键词：信息安全；加密技术；大数据；防护体系

DOI：10.69979/3041-0673.26.02.013

1 绪论

1.1 研究背景及意义

随着计算机网络技术的深度发展，信息的存储、处理与交互方式发生了根本性变革。无论是个人用户将数据托管于云盘，还是企业机构依托云平台和云计算开展服务，均对安全防护措施提出了迫切需求。在此背景下，其带来的安全与隐私问题已然成为云计算和大数据时代所面临的核心挑战之一，而保护涉及国家机密的敏感信息免受敌对势力窃取，更是信息安全领域的首要任务。数据安全的核心目的是防范数据遭盗窃的风险，保障业务连续性并将损失降至最低。

大数据平台在用户信息搜集与实时跟踪分析中的广泛应用，使得用户隐私保护面临严峻的挑战，现有的防护机制难以全面覆盖数据流转过程中的风险。在早期国家层面安全标准尚未健全、法律监管与技术措施协同不足的背景下，信息隐私泄露侵权事件频发，集中暴露了大数据时代用户隐私安全领域的尖锐问题。2014 年，12306 购票官网发生信息泄露事件，用户购票数据（身份证号码等信息）被非法出售，作为覆盖全国用户的票务公共服务平台，该事件直接凸显了公共领域数据隐私保护的薄弱性；2018 年 9 月，Facebook 遭遇黑客攻击，数百万用户的电话号码与邮箱地址被窃取，导致 4000 万用户账号处于安全风险状态。可见在互联网大数据海量搜集、存储与智能分析挖掘全流程中，用户个人敏感信息随时可能被触及，若任一环节存在安全缺口均可能引发难以挽回的后果。

尽管传统信息安全技术（如防火墙）已发展得相对成熟，但其在面对大数据环境下的新型攻击向量（如高级持续性威胁 APT）时，往往显得力不从心。现有研究缺乏对大数据环境下安全技术体系的系统性梳理，以及针对多层次、协同化防护策略的深入探讨。因此本文旨在系统综述大数据环境下信息安全的热点技术，并在此基础上构建一个整合技术、管理与法规的综合防护框架。

1.2 国内外研究现状

在国内层面，我国信息安全呈现出“法规先行、技术跟进、体系构建”的鲜明路径。自 1994 年法规出台以来，我国已构建起以《网络安全法》为核心，《数据安全法》《个人信息保护法》为支撑的完备法律体系，为信息安全实践提供了强制性框架。早期研究领域集中于网络边界安全与病毒防治，随着大数据时代的到来，研究前沿已拓展至云计算安全、隐私计算等方向。

在国际层面研究起步更早，且更侧重于技术标准引领与前瞻性战略布局。美国长期主导着密码学、可信计算等基础领域的研究，其《电子签名法案》的颁布极大地推动了电子商务的安全发展。国际上对云原生安全以及对抗性机器学习的研究已进入实践深化阶段。

综上所述，尽管国内外在信息安全领域已取得重要成果，但在大数据环境所要求的弹性、自适应与智能化防护方面，仍缺乏一个能够有机融合技术、管理与合规的体系化框架。

2 信息安全的多维威胁模型

信息安全是一个动态对抗过程，其威胁来源日趋复杂与隐蔽。本节从技术、管理、人员及环境四个层面，系统剖析影响信息安全的核心要素。

2.1 技术层面：系统脆弱性与外部攻击

技术层面是安全威胁最为直接的体现，主要包括系统自身的脆弱性及恶意的外部攻击。

（1）软件与系统漏洞：任何复杂的信息系统都不可避免地存在设计缺陷或编码错误所构成的漏洞，无论是操作系统还是物联网设备，这些漏洞为攻击者提供了未经授权的入口。在大数据平台中，组件的多样性与复杂性更显著扩大了攻击面，使得漏洞管理面临严峻挑战。

（2）恶意代码与黑客攻击：大数据时代信息的高速流动与价值密度高的特性，极大地刺激了病毒、木马等恶意代码的传播，以及高级持续性威胁等有组织的黑客攻击。从早期的“熊猫烧香”到如今的勒索软件即服务，攻击手段不断演进，其破坏力、隐蔽性、针对性与持续性均增强，旨在窃取数据和牟取经济利益。

2.2 管理层面：制度缺失与内部控制失灵

再先进的技术也需要依托严谨的管理制度才能发挥作用，管理松懈是许多重大安全事件的根源。

（1）安全策略与制度缺位：许多组织未能建立系统化的信息安全管理制度。具体表现为数据分类分级模糊、安全责任归属不清等。这种制度性缺位导致安全防护工作无章可循，缺乏一致性。

（2）访问控制机制薄弱：访问控制是防止越权访问的核心手段。然而实践中普遍存在权限分配过宽、口令策略松散、特权账号管理混乱等问题。特别是在远程办公常态化的背景下，脆弱的访问控制使得“边界”概念模糊，内部数据更易被非法访问与窃取。

2.3 人为层面：内部威胁与意识不足

人是最关键也是最不确定的安全因素，由内部人员引发的安全风险往往更具有破坏性。

（1）非恶意疏忽：绝大多数员工并非有意破坏，但因安全意识薄弱，容易发生操作失误，如错误配置云存储桶导致数据公开、点击钓鱼邮件导致凭证泄露等。这些无意识的行为为攻击者创造了可乘之机。

（2）恶意内部人员：拥有特定数据访问权限的内部员工可能出于经济利益，故意窃取、篡改或销毁关键数据。由于其对系统架构和防御措施了如指掌，此类威胁检测难度大，造成的损失也尤为严重。

2.4 环境层面：合规挑战与供应链风险

超越组织边界的外部因素，构成了信息安全的宏观背景与新型风险源。

（1）法律法规与合规性压力：随着《网络安全法》等法规的深入实施，数据处理的合规性本身已成为一项核心风险。企业若未能满足监管要求，即便未遭受黑客攻击，也可能面临巨额罚款乃至刑事责任。

（2）第三方供应链风险：现代信息系统深度依赖外部的开源组件、云服务商和软硬件供应商。这些第三方自身的任何安全漏洞或恶意行为，都会沿供应链传导并放大，直接影响本组织的信息安全，如近年来频发的开源软件漏洞引发了全球性安全危机。

3 加强信息安全的措施

在大数据与云计算深度融合的背景下，信息安全问题呈现出多样化、复杂化特征。为此需要从技术防护、管理规范、用户行为与系统运维等多个维度构建协同联动的信息安全防护体系。

3.1 技术防护层构建

3.1.1 强化边界防护与入侵检测

防火墙作为网络安全的第一道防线，其作用不仅体现在访问控制与流量控制，更在于构建内外网之间的安全隔离区（DMZ）。在大数据平台下，建议采用下一代防火墙（NGFW）、集成深度包检测（DPI）、应用识别、入侵防御系统（IPS）等功能，实现对恶意流量的实时识别与阻断。同时应结合入侵检测系统（IDS）与入侵防御系统（IPS）地协同联动构建主动防御机制，提升对 APT 攻击等高级威胁的响应能力。

3.1.2 数据加密与密钥管理

数据加密是保障数据机密性与完整性的核心手段。建议采用分层加密策略：在数据存储阶段采用 AES-256 等对称加密算法对静态数据进行加密；在数据传输阶段采用 TLS 1.3 协议实现端到端的加密；在数据使用阶段引入同态加密或可信执行环境（TEE），实现“可用不可见”的数据处理模式。此外应建立密钥生命周期管理体系（KMS），实现密钥的生成、分发、轮换与销毁全过程可控。

3.1.3 身份认证与访问控制

传统基于用户名和密码的认证方式已难以应对暴力破解与凭证窃取攻击。建议引入多因子认证机制（MFA），结合生物特征（指纹、人脸）、动态令牌（OTP）、设备指纹等多维度信息提升身份认证强度。同时逐步构建零信任安全架构，默认不信任任何用户与设备，所有访问行为均需经过动态身份验证与权限授权。

3.1.4 漏洞管理与数据备份

系统漏洞是攻击者入侵的主要入口。针对大数据平台组件（如 Hadoop、Spark、Kafka 等）更新频繁、依赖复杂的特点，应建立自动化漏洞扫描与补丁管理机制。建议引入 CVSS 评分体系对漏洞进行分级处理，优先修复高危漏洞。数据备份是应对勒索软件、系统故障等突发事件的最后防线。建议采用“3-2-1 备份策略”：至少保留 3 份数据副本，2 种不同存储介质，其中 1 份异地备份。

3.2 管理制度保障

3.2.1 安全制度与数据分级分类

信息安全不仅依赖技术手段，更需要制度支撑。建议依据《网络安全法》《数据安全法》《个人信息保护法》等法规，建立数据分类分级管理制度，明确敏感数据的识别、存储、传输与销毁流程。同时应依据最小权限原则制定访问控制策略，并建立用户权限的动态审计与撤销机制，防止权限滥用与内部泄露。

3.2.2 法律合规与治理体系

随着《个人信息保护法》《数据出境安全评估办法》等法规落地，数据处理的合规性已成为企业安全治理的重要组成部分。建议建立合规性评估机制，定期对数据处理流程进行审计，确保数据采集、存储、共享与跨境传输符合法律要求。

3.3 人员行为层规范

3.3.1 安全意识教育与培训

研究表明超过 60%的信息安全事件与用户操作失误密切相关。为此应建立常态化安全培训机制，通过模拟钓鱼攻击、密码强度测试、安全演练等手段提升用户安全意识，同时建议引入行为分析系统，对异常登录、批量下载、权限越界等行为进行建模识别，及时预警潜在内部威胁。

3.3.2 用户行为分析

为有效识别和预警内部威胁，应引入用户行为分析系统。通过对用户日常操作日志（如登录时间地点、访问频率、数据下载量、权限使用记录等）进行持续监控和建模，建立正常行为基线。一旦检测到与基线显著偏高的异常行为，系统应能自动告警，使安全团队能够及时介入调查，将潜在的内部威胁在早期阶段进行遏制。

3.4 安全运营支撑

3.4.1 安全运营中心与 SIEM

建议构建安全运营中心（SOC），集成日志审计、威胁情报、事件响应等功能，实现对全网安全态势的集中监控与联动响应。通过引入安全信息与事件管理系统（SIEM），对多源日志进行关联分析提升对复杂攻击链的识别能力。

4 数据加密技术体系及其应用演进

4.1 数据加密的基本模型与分类

一个完整的密码系统通常由五个组成部分构成：明文空间（M）、密文空间（C）、密钥空间（K）、加密算法（E）与解密算法（D）。现代加密技术主要分为两大类：

对称加密：加密和解密使用同一密钥（如 DES），其优势在于计算效率高和吞吐量大，适用于海量数据的加密。

非对称加密：使用公钥和私钥这一对密钥如（RSA、ECC），公钥用于加密、私钥用于解密。该方式解决了密钥分发问题，但计算开销大，加密速度远慢于对称加密。

4.2 大数据环境下的加密策略与部署方式

4.2.1 链路加密：保护通信链路

链路加密在 OSI 参考模型的数据链路层或网络层对数据包进行即时加密。数据在离开一个节点时被加密，在进入下一个节点时被解密，随后立即使用下一段链路的密钥重新加密。所有数据均以密文形式在物理链路上传输，这提供了良好的流量保密性，能有效对抗通信分析。

4.2.2 节点加密：对链路加密的增强

节点加密旨在弥补链路加密在节点处的安全缺陷。它通过在节点处部署专用的安全模块，对收到的密文进行解密后，立即在模块内部使用下一链路的密钥进行重新加密。在节点加密中，明文不暴露于节点主机的内存中，从而提升了在不可信网络中的安全性。

4.2.3 端到端加密：保护数据内容

端到端加密在 OSI 参考模型的应用层或传输层实现。数据从发送端始终以密文形式存在，直至到达最终接收端才被解密，通信路径上的任何中间节点均无法获取明文。端到端加密真正实现了数据内容对通信路径的全程保密，而路由信息保持明文以确保网络可达性。

5 结论

当前的安全威胁已形成一个由技术漏洞、管理缺陷、人为风险与供应链环境交织构成的复杂生态系统。任何单一维度的防护都难以应对此种系统性风险，这从理论上论证了构建体系化防御方案的必然性与紧迫性。本文所构建的技术-管理-人员-运营四层综合防护体系表明，先进的技术手段是基石，但必须与严谨的制度化管理、深入人心的安全文化以及智能化的持续运营紧密结合，才能形成动态、弹性的纵深防御能力，数据加密仍然是核心技术但需范式演进升级。面向未来单纯的静态加密已不足够，隐私计算等在保障数据可用性的前提下实现机密性的新技术，代表了大数据与云环境下的重要发展趋势。

参考文献

[1]王魏,赵奕芳.大数据时代计算机网络信息安全及防护策略[J].中阿科技论坛(中英文),2022(01):72-75.

[2]袁素芳.网络信息安全管理下计算机应用策略[J].网络安全技术与应用,2021(12):29-30.

[3]袁一帆.探讨如何实现大数据时代的计算机网络信息安全[J].网络安全技术与应用,2021(12):174-175.

[4]卜水洲.计算机网络信息安全中数据加密技术[J].数字技术与应用,2021,39(11):234-236.

[5]杜娟.大数据时代计算机信息安全防范措施研究[J].无线互联科技,2021,18(05):33-34.

[6]许文杰.计算机信息安全问题及防御[J].电子技术与软件工程,2020(02):238-239.

[7]冯成春.浅谈计算机网络安全漏洞及防范措施[J].计算机光盘软件与应用,2014,17(11):188+190.

[8]许葵元.计算机网络信息安全及防护策略研究[J].赤峰学院学报(自然科学版),2014,30(01):81-83.

[9]黄惠海.探究大数据时代下的企业网络信息安全[J].网络安全技术与应用,2022(02):103-104.

[10]杨虎.基于计算机互联网信息安全的防御技术要点研究[J].信息与电脑(理论版),2021,33(04):214-216.