大数据时代保障信息数据安全的方法分析

王博涵

香港中文大学，中国香港

摘要：在大数据时代，数据的采集、存储与处理规模呈指数级增长，数据应用已渗透至社会各领域，但网络攻击手段升级、数据泄露事件频发，传统安全防护体系难以满足需求，保障信息数据安全成为亟待解决的重要课题。基于此，本文围绕大数据时代信息数据安全保障展开研究，剖析当前数据安全面临的复杂挑战，系统阐述保障信息数据安全的关键方法。从技术防护、制度建设、管理优化等多维度提出应对策略，旨在为提升大数据环境下信息数据安全防护水平提供理论参考与实践指引，助力各领域安全、高效利用数据资源。

关键词：大数据时代；信息数据安全；数据加密；访问控制；安全管理

DOI：10.64216/3080-1486.25.10.040

引言

在大数据技术快速发展和广泛应用的背景下，数据已经成为推动社会经济发展和科技创新的核心资源——数据泄露、恶意攻击、非法访问及其他安全威胁的频繁发生，不仅危害个人隐私和企业利益，而且给国家安全和社会稳定带来了潜在风险。因此，深入探讨大数据时代信息数据安全保障的有效途径，建立全方位多层次数据安全防护体系具有现实意义和迫切性。

1 大数据时代常见的信息数据安全问题

1.1 数据规模和多样性造成管理困境

伴随着物联网和云计算的推广，大数据表现出海量、多源和异构的特点，企业平均每天生成的数据量可达到 PB 级别，涉及文本、图像和视频等不同形式。传统数据管理和安全防护手段很难处理这样一个复杂数据环境，而且数据存储零散且格式不一致，使得安全策略很难得到有效涵盖[1]。与此同时，非结构化数据的比例也在攀升，传统的加密、访问控制技术的保护作用受到了一定的限制，数据分类分级越来越困难，很容易产生安全保护漏洞。

1.2 网络攻击技术升级的外部威胁

黑客的攻击方式在不断地更新和升级，新型的攻击工具如自动化攻击工具和高级持续性威胁（APT）等不断涌现和发展，勒索软件利用加密数据勒索大量赎金；供应链攻击从数据源头潜入，例如在软件更新包中嵌入恶意代码。另外，减少了漏洞利用成本，攻击者可以从暗网中获取攻击脚本，并对大数据平台开源组件和云服务接口等弱点进行精准攻击，数据泄露的风险明显增加。

1.3 内部管理漏洞导致数据泄露等风险

企业内部人员操作不规范、权限滥用等问题成了导致数据泄露的主要原因。有的职工安全意识淡薄，任意利用公共网络进行敏感数据的传递，或者没有及时对系统补丁进行更新，给黑客入侵以可乘之机。另外，对数据访问权限的划分也不够清晰，有的员工权限过大，在离职或者调离岗位时权限没有及时收回，造成非法访问或者恶意窃取数据的行为。

2 大数据时代保障信息数据安全的价值

2.1 保障个人隐私和合法权益

信息数据安全保障，是个人隐私保护的一道关键防线，大数据时代个人信息的采集范围日益扩大，涉及生物特征、消费习惯、健康数据以及其他敏感信息。如果这些资料被泄露出去，就会造成精准诈骗和身份冒用，从而严重侵犯个人的财产安全和生活安宁。通过数据加密和匿名化处理技术手段，可以有效预防个人信息泄露，保障公民合法权益，增强公众对于数字社会的信任。

2.2 维护企业的核心竞争力和商业利益

数据已经成为了企业的核心资产，确保数据安全是一个企业赖以生存和发展的根本所在。对于金融企业来说，客户交易数据的泄漏会造成信任危机并造成客户的损失；对于制造企业而言，研发数据泄露会弱化企业的技术优势。企业可以通过建立完整的数据安全体系来避免数据泄露带来的经济损失、保护商业机密和知识产权、增强企业市场竞争抗风险能力、保障业务的持续平稳经营。

2.3 推动数字经济的健康、持续发展

数据安全是发展数字经济的先决条件，如果数据安全问题经常发生，就会妨碍数据要素自由流动和高效使

用，遏制数字技术创新。一个健全的数据安全保护系统能够为企业创造一个既安全又可靠的数字环境，从而激发企业在数据应用方面的创新活力，并进一步推动大数据、人工智能等新兴行业的发展。与此同时，确保数据安全能够提升国际数据合作互信基础、推动数字贸易的发展、赋予数字经济新动能。

3 大数据时代保障信息数据安全的方法

3.1 对数据进行全生命周期加密

数据全生命周期加密是一种全面覆盖数据从生成、储存、传送到最终销毁各个阶段的保护措施，通过先进的加密技术来确保数据在各个环节都能保持高度的安全性，数据生成阶段，原始数据经过初步加密，例如物联网设备获取数据的过程中，传感器数据经过内置的加密模块实时加密，以避免数据从源头被盗用，在数据存储环节中，通过磁盘加密和数据库加密的方式来保护静态数据，甚至当存储设备失窃或者数据非法拷贝时，未经许可的用户仍然不能阅读数据内容[2]。比如，金融机构在客户的交易记录中使用AES高级加密标准存储加密，以确保敏感信息的安全性，数据传输时，采用 SSL/TLS 协议和其他技术构建安全通道来加密传输的数据，避免了网络传输时数据的侦听和篡改。例如，云服务中的数据传输，采用 SSL 证书进行身份验证和加密传输，保证了客户端和云端的数据安全互动。在数据使用阶段动态地加密授权访问数据，例如利用同态加密技术使用户能够以密文状态进行运算，既保证了数据隐私也不会影响数据处理效率，在数据达到生命周期结束而需破坏时，通过多次覆写和物理粉碎的安全数据擦除技术来保证数据不可修复，彻底杜绝了数据泄露的风险。数据全生命周期加密将数据无缝防护起来，并通过各种加密技术协同运用，建立立体式数据安全防护网络，有效防御各种数据安全威胁。

3.2 零信任的访问管控

零信任访问管控打破传统“内网就是信任”的安全理念，遵循“永远不会被相信，永远被核实”原则，对所有访问请求进行严格身份验证与权限控制，在用户身份验证方面，采用多因素认证机制，除用户名和密码外，还结合生物特征识别（指纹、面部识别）、动态令牌等方式，确保访问者身份真实可信。以企业远程办公系统为例，该系统需要员工登录后不只输入口令，还要进行手机验证码或者指纹验证，以减少账号被窃取的风险，在权限管理方面，零信任架构使用最小权限原则，并依据用户的角色、接入场景进行动态的权限分配[3]。如果员工正在对企业数据进行存取，则系统会依据自身部门职责、任务需求和当前存取设备安全状态等因素动态地赋予相应的权限，并且该权限只对当前存取周期有效。同时，通过持续的访问监控与风险评估，实时监测用户行为，一旦发现异常操作，例如，敏感数据的高频接入在非工作时间，立即终止访问并触发告警。零信任访问管控还对访问设备进行严格管理，只有通过安全检测，例如系统补丁的更新、防病毒软件的状态等的设备才能接入网络，阻止恶意软件被传染的装置成了数据泄露的突破点。这种建立在动态信任评估基础上的访问控制方式有效地阻止了内部人员权限滥用以及外部攻击者横向渗透等行为，构筑了一道数据访问安全屏障。

3.3 AI 智能监测与预警

AI 智能监测预警是在人工智能和机器学习技术的辅助下，对数据安全威胁进行实时监控、精准分析和及时预警，对于数据安全威胁检测，采用机器学习算法分析大量的网络日志、系统行为数据，学习数据的正常访问模式和系统运行规律并构建行为基线模型。当用户或系统出现偏离基线的异常行为时，如在很短的时间内进行敏感数据的海量下载、非授权端口的接入等，AI 系统能够迅速识别并发出告警。比如，银行利用 AI 对客户的交易行为模式进行分析，就可以及时发现异常交易并预防欺诈风险，针对复杂未知威胁，AI 利用神经网络算法等深度学习技术自动提取和识别攻击特征[4]。通过研究海量历史攻击数据，AI 系统可以发现新的攻击手段中可能存在的规则，甚至在遇到未曾发生的攻击方式时可以根据行为模式相似性做出警告。同时，自然语言处理技术能够实时分析安全事件报告和漏洞公告中的文本信息，迅速得到最新的安全情报并及时地更新威胁检测模型。

AI 智能监测预警系统也具有自动化的响应能力，当发现安全威胁时，能够按照预设的策略对被感染的设备进行自动的阻断攻击、隔离、重置用户权限和其他措施，以减少安全事件带来的危害，AI 系统通过不断学习和自我优化，持续提高威胁检测准确率和响应效率，对数据安全进行智能化和主动式保护。

3.4 区块链存证溯源

区块链存证溯源借助区块链技术分布式账本和不可篡改的特点，提供了可信存证和全流程追溯能力。在数据存证方面，将数据哈希值存储在区块链上，形成不可篡改的电子证据，由于区块链是分布式存储的，因此数据哈希值被保存于多个节点上，对任意一个节点进行修改均不能更改整个记录，从而保证了存证数据真实完整[5]。以司法领域为例，以区块链存证的电子合同和版

权作品数据可以成为法律效力证据，在数据溯源中，区块链链式结构全面记录数据运行历史，其中包含数据创建、修改和获取行为。每一项操作都被详细地记录为一个独立的区块，并在其中包括前一个区块的哈希值，从而生成了时间戳的序列。通过对区块链进行查询，可以清楚地回溯数据来源、流转过程和各个环节操作主体。供应链金融下，区块链能够跟踪货物交易全过程数据，保证交易信息的真实可靠，避免数据造假，区块链共识机制确保数据操作合法，只需通过网络内大部分节点所承认的操作才能在区块链中进行记录。结合智能合约技术可以自动和规范化地进行数据操作，例如在数据访问权限改变时智能合约会自动对区块链记录进行更新。区块链存证溯源是数据安全的可信基础支持，提高数据可信度和可审计性并有效防止数据篡改和误用。

3.5 对数据进行分级分类管理

数据分级分类管理是指通过科学分类和合理分级制定差异化安全防护策略，来准确高效地保护数据安全，在数据分类部分，依据其业务属性和内容特征对其进行分类。例如，企业数据可以划分为客户数据、财务数据和研发数据；政府数据可以划分为人口信息、地理信息和政务公开信息。通过确定数据分类标准使得数据管理更条理化，方便后续安全策略的制定，数据分级是根据数据的敏感性和泄露后可能带来的危害程度来进行的，通常会将数据分为公开数据、内部数据、敏感数据和核心数据等不同的安全级别[6]。比如，企业普通宣传资料是公开数据，客户身份证号码和银行卡信息是敏感数据，核心商业机密和技术专利是核心数据等等。不同层次的数据所对应的安全防护要求也不相同，需要对核心数据采取最高层次的加密和访问控制防护措施，同时可以对公开数据适当地减小防护强度。

根据数据分级分类，有针对性地制定了数据安全策略，对访问控制而言，核心数据只允许特定岗位的人员进行访问，需要多级审批；获取敏感数据需要二次身份验证。在数据存储上，对核心数据使用加密存储和异地备份的方式保证数据的安全性和可用性。同时，建立了数据分级分类的动态管理机制，并随业务的开展和数据价值的改变而适时调整数据类别和等级，确保安全策略时刻匹配数据风险。数据分级分类管理使数据安全防护资源得到合理配置，提高数据管理效率和应用价值，同时保证数据的安全性。

3.6 应急响应机制的优化

应急响应机制优化的目标是建立一个快速、高效、协同处理数据安全事件的系统，将数据安全事件带来的危害降至最低，应急响应组织架构需要健全，应急指挥组、技术处置组、后勤保障组及其他队伍职责分工需要明确，以保障安全事件中各个部门之间的协同作战。同时，制定了详尽的应急响应预案，对数据泄露、勒索攻击和系统入侵等各种安全事件进行处理，并对处理流程和技术手段进行界定，之后强化应急演练和训练，通过经常性实战化演练检验并优化应急预案可行性和有效性，增强队伍应急处置能力。训练内容既涉及数据恢复和漏洞修复技术技能，又需要加强人员安全意识和风险防范能力建设，以保证遇到突发安全事件能沉着应对和科学处理，要构建应急响应联动机制并强化与外部安全机构、网络运营商和监管部门之间的交流与合作。当出现重大数据安全事件时，可以迅速获得外部技术支持和资源保障，对攻击进行及时屏蔽、溯源追踪并减小事件的影响范围。通过应急响应机制的不断优化，形成了防范、监控、处置和恢复全过程闭环管理，增强了数据安全风险快速应对和处置的能力，夯实了数据安全保障基础。

4 结束语

综上所述，大数据时代下，确保信息数据的安全是一个系统而又长远的项目，将技术防护、制度约束和管理优化相结合，并与新兴技术趋势相融合，建立一个动态的、智能化数据安全防护系统，可以有效地应对日趋复杂的数据安全问题。今后，我们需要继续重视技术革新和安全需求的变化，继续完善数据安全保障策略，筑牢数字经济良性发展的安全基石。

参考文献

[1]柴梦竹.人工智能技术在大数据网络安全防御中的应用[J].计算机与网络,2021,47(15):44-45.
[2]张晓军.自动化网络安全防御的“是与非”[J].网络安全和信息化，2021(8):111-112.
[3]张荣.一种基于人工智能的多层次网络安全防御模型研究[J].信息与电脑，2021，33(13):180-182.
[4]李跃忠.浅谈大数据时代背景下的数据安全治理[J].中国信息化,2021,(04):76-79+67.
[5]王振中.大数据时代网络信息安全存在的问题及对策[J].软件,2021,42(08):33-38.
[6]郭晓丽.大数据时代企业信息安全保障策略分析[J].中国管理信息化,2021,24(20):115-116.