 DOI: https://doi.org/10.12345/bdai.v7i1.35260

Research on software supply chain security and network protection in big data and artificial intelligence environment

Wenqiang Lu

Shanghai Digital Security Technology Co., Ltd., Shanghai, 200072, China

Abstract

With the deepening application of new technologies such as big data and artificial intelligence (AI) across industries, cybersecurity threats have become increasingly diverse, complex, and intelligent. To clarify the comprehensive threats in cybersecurity, identify mainstream threat types, and propose corresponding protective measures, this paper conducts a systematic analysis of network security threats by integrating big data and AI application scenarios, and proposes corresponding protective strategies. The aim is to provide theoretical references and technical guidance from a holistic perspective for building an intelligent, full-chain, and sustainable cybersecurity protection system.

Keywords

big data; artificial intelligence; network security; software supply chain; protective measures

大数据与人工智能环境下的软件供应链安全与网络防护研究

陆文强

上海数字安全科技有限公司，中国·上海 200072

摘　要

随着大数据、人工智能（AI）等新技术在各行业中的深入应用，网络信息安全面临的威胁更加多样化、复杂化及智能化。为了明晰网信安全面临的全貌性威胁、理清主流威胁类型并给出对应防护手段,文结合大数据与AI应用场景，对网络信息安全威胁进行系统分析，并提出相应防护策略”。以期能够从全局角度出发为构建智能化、全链条、可持续性的网络安全防护体系提供一定的理论参考和技术指导思路。

关键词

大数据；人工智能；网络信息安全；软件供应链；防护措施

1 引言

大数据、AI 技术在促进数字化转型过程中衍生出网络信息安全隐患的多面性。数据量巨幅增长、计算能力大幅提升、算法逐步趋于智能化，导致攻击手段愈发多元复杂，针对传统边界防护与静态安全策略不再适用。在此背景下，软件供应链与 AI 模型作为数字化应用的核心基础，逐渐成为新的攻击目标。而就软件供应链、AI 模型等使用对象而言，则可能造成漏洞传播、数据投毒、依赖劫持等安全风险。而为了更好的抵御以上各种风险的发生，迫切需要从全生命周期、全链条的技术及管理体系，实现网络信息安全主动防护以及智能化运维。

【作者简介】陆文强（1989–），中国上海人，本科，战拓主管，从事网络与数据安全、软件供应链相关研究。

2 大数据与 AI 时代网络信息安全威胁全景分析

2.1 软件供应链安全威胁

2.1.1 开源组件与第三方库的风险

企业在系统开发和 AI 平台建设中广泛依赖开源组件及第三方库，这些依赖项存在版本更新滞后、漏洞未修复及被恶意篡改等风险。攻击者可通过供应链环节植入后门或恶意代码，从而渗透生产环境并操控数据。以 2024 年 3 月曝光的 xz/liblzma 供应链后门事件（CVE-2024-3094）为例：攻击者在 xz 5.6.0/5.6.1 中植入恶意逻辑，借助 glibc 的 IFUNC 机制在特定构建与运行条件满足时对 OpenSSH（sshd）的加密流程实施符号劫持与重定向，从而实现未授权访问 / 身份绕过等能力并可能进一步导致远程代码执行的严重后果。所幸该后门并未在主流发行版的稳定渠道大范围落地，但一旦被广泛部署，目标服务器将面临被完全控制的风险，随之而来的关键系统入侵、数据泄露与服务瘫痪等安全隐患将具有高度破坏性和传导性。

2.1.2 AI 模型供应链的可信性问题

AI 模型在训练、评测、打包与部署的各环节均可能遭受数据投毒或模型篡改，导致训练结果偏移或推理决策被操纵。若模型来源不透明、可追溯与验证机制不足（如缺乏数据与权重的来源证明、签名 / 哈希校验与供应链证明），企业将难以及时发现异常，核心算法资产的被破坏、被置换或被误用风险随之上升。

2.1.3 SBOM 缺失与依赖劫持

软件物料清单（SBOM）缺失会使组织对组件版本与依赖关系的整体可视性不足，攻击者可借助依赖劫持或漏洞利用链发起远程入侵。在缺乏持续监控与更新机制的环境中，供应链中的（已知或新披露的）漏洞可能长期滞留并被反复利用，风险呈现隐蔽化与广域扩散特征。

2.2 数据泄露与信息窃取风险

2.2.1 内部人员威胁与权限滥用

在大数据和 AI 驱动的时代，企业内部人员拥有分层权限访问大量敏感信息。若内部人员权限管理不到位或缺乏完善的审计措施，则存在造成客户信息、业务流程及算法参数等被非法查看，即形成高风险泄漏源的可能。不仅如此，频繁的数据调用和系统操作还会导致检测违规访问的难度加大 [1]。

2.2.2 黑客入侵与非法接口调用

网络攻击者可通过漏洞扫描、远程执行、中间人攻击等手段获得对服务器的访问权限，随后利用人工或 AI 辅助的攻击脚本及自动化工具对外部 API 和内部服务接口发起恶意调用，以窃取敏感信息。AI 辅助的攻击与自动化工具提高了攻击的规模与智能化程度，使攻击更易规避基于签名或静态规则的传统防护措施，也使得异常行为更难被实时检测与识别。以实例说明，2020 年 3 月曝出的新浪微博接口被恶意调用事件据报导波及数亿用户数据（包括手机号、账号映射关系等），不法分子利用已泄露的手机号或账号信息批量调用接口进行关联查询，部分数据随后在暗网或论坛上被交易或公开。

2.2.3 数据黑市交易与 AI 自动化窃取

大规模数据的商业价值催生了专业的黑市交易网络，各种黑客组织或者灰产小团体通过暗网来售卖窃取所得数据。在黑市上还可使用 AI 算法实现自动化的数据收集汇总、脱敏和批量分析，以达到更高的数据交易效率以及更高的隐蔽性 [2]。在这一过程中，自动化爬虫、智能代理技术可以大大加速黑市窃取的速度，形成了极大的信息安全风险。

2.3 网络攻击与恶意行为

2.3.1 DDoS 与勒索软件智能化演进

在大数据与 AI 环境中，分布式拒绝服务攻击日益表现出高智能特性，即利用自动化僵尸网络动态调整攻击流量分布、精准消耗目标系统资源、规避传统的监测规则。并且勒索软件利用 AI 技术智能决定加密策略，选择对系统内部重要节点进行加密，针对已经被防护住的网络流选择再次入侵，利用机器学习分析出系统防护薄弱点并将其变成突破口，形成了具有连贯性的攻击链路。

2.3.2 APT 攻击与 AI 驱动的精准渗透

高级持续性威胁 APT 利用人工智能改进攻击途径，在通过了解网络拓扑结构、系统配置以及用户使用习惯后，实现长期、隐蔽的渗透。并且使用智能算法来确定侵入的最佳节点以及攻击的时间点，增强了攻击过程的隐蔽性。

2.3.3 攻击自动化与 AI 恶意使用

随着攻击自动化程度大大提高，使用 AI 生成的恶意脚本和漏洞扫描工具能够快速渗透到大规模的目标当中。并且还能利用 Deepfake 技术伪造语音或视频实施网络钓鱼诈骗，网络攻击效率与隐蔽性极大提高 [3]。2024 年 5 月，美国某知名 AI 平台，攻击者通过自动化脚本与对抗式提示，滥用其公开接口向模型发送特定构造的恶意 prompt，该平台被迫暂停部分接口服务，致使数以万计的用户无法正常使用；并且部分客户因模型返回了私密信息（名字、邮箱等）、代码、SRC 等内容而导致平台损失数百万美元的直接经济收入和平台声誉。

2.4 隐私与合规风险

2.4.1 跨境数据与个人信息保护

随着全球化加速，跨境数据流转呈增长态势。受限于不同法域的数据保护要求差异，若企业未依法完成数据出境安全评估或标准合同备案，亦未建立合法性基础与最小必要的处理机制，同时在技术层面缺乏端到端加密、去标识化 / 匿名化与统一的访问控制和审计，则跨境传输链路将暴露于未授权访问与泄露等高风险之中。据媒体报道 / 公开通报，2025 年 9 月，某企业因未履行出境评估、未取得用户单独同意、且未对出境数据采取必要加密或去标识化措施而将个人信息传输至境外，被依法予以处罚。

2.4.2 AI 模型训练中的数据合规问题

AI 模型训练依赖于海量且种类多样的数据集，在获取、存储和使用过程中必须严格遵守相关的隐私以及数据保护方面的规定。未经匿名化或者脱敏处理的数据，直接用于模型训练的过程中，存在一定的合规风险。

2.4.3 算法黑箱带来的监管挑战

复杂的 AI 算法在决策时展现出了高非线性以及自适应的特点，外部是无法看到其中内部逻辑的，在一定程度上加大了外部监管的难度。而算法本身不透明会导致监管部门一旦发现了问题就很难找到真正的责任所在，企业只能通过建立可解释性的机制和监控来应付这种长期存在的监管压力。

2.5 系统与平台安全漏洞

2.5.1 大数据平台与云计算漏洞

网络攻击者常通过漏洞扫描、远程执行与中间人攻击等手段获取对目标服务器的访问权限。获取访问后，攻击者会利用人工或 AI 辅助的攻击脚本与自动化工具，对外部 API 与内部服务接口发起大规模恶意调用，从而窃取敏感信

息。以一次具有代表性的事件为例，2024 年针对 Snowflake 客户实例的攻击利用被窃凭证和配置弱点，导致数十家客户敏感数据被窃并在暗网或勒索平台上被披露或用于敲诈。为降低此类风险，建议在云与数据平台层面强制实施多因素认证（MFA）、最小权限与会话管理、API 网关限流与白名单策略，并结合行为分析实现基于风险的实时检测与自动化响应。

2.5.2 API 与微服务的安全缺陷

微服务架构中使用的 API 接口存在着输入验证不严格、身份认证缺失以及请求限流不足等问题。攻击者可以通过 API 调用链路或者横向移动等方法实现对关键服务的访问或者数据的抓取。此外，微服务间通信缺乏安全防护措施或采用的加密强度低，可能会被截获和篡改，导致整个系统的攻击面增加。

2.5.3 AI 平台与模型管理系统漏洞

AI 平台与模型管理系统涉及训练 / 部署 / 更新全流程的数据流与集中化参数管理。若平台存在安全漏洞，攻击者可能篡改训练数据、置换模型制品或窃取参数机密，进而破坏模型完整性与来源可信，导致推理结果不可验证、可靠性下降。

3 大数据与 AI 网络信息安全防范措施研究

3.1 软件供应链安全治理措施

3.1.1 软件物料清单（SBOM）与安全成分分析（SCA）

企业可建立 SBOM 并对各个软件组件以及它的软件依赖情况开展完整的记录，以高效跟踪与管理各个组件的版本信息、许可证信息以及安全补丁情况。同时利用 SCA 的安全成分为扫描的软件能够自动高效地识别开源库和第三方组件中的已知漏洞、配置缺陷及潜在恶意代码，以达到在整个组件生命周期内对组件进行持续性地监测和安全评估的目的。

3.1.2 AI 模型供应链安全

针对 AI 模型的开发与部署阶段，以考虑增加模型可信认证的措施来校验训练数据源是否安全、完整和正确实现。可采用链式审计记录以及对数据源进行审核的办法保障模型训练无被投毒或者篡改的情况发生，运用持续监控验证的方法来有效保障模型供应链从生成、传输到部署的全流程安全。

3.1.3 第三方组件与 AI API 的审计机制

对于第三方组件和 AI API 接口来说，要建立起访问日志分析、权限使用审查、接口调用异常等审计工作的周期，并结合一些自动化工具来进行外部依赖的安全检测以及版本的一致性校验，降低由于第三方组件存在漏洞、API 接口被滥用等情况造成的供应链安全问题发生概率。

3.2 数据泄露与信息窃取防范措施

3.2.1 数据分类分级与零信任架构

企业应对数据资产进行分类分级，并根据数据的敏感度确定不同的访问权限、操作控制 [4]。此外，还可应用零信任的安全机制，以持续性认证、最小权限访问、动态安全策略等方式管理企业和外部用户。使用零信任，在多租户、分布式场景下对数据进行分隔，防止数据水平流动，降低人员滥用授权或越权访问的风险。

3.2.2 加密、脱敏与差分隐私

在数据存储、传输、处理过程中使用高安全性加密算法、结合使用字段脱敏、数据屏蔽等手段降低被泄露的风险。此外，在 AI 模型训练、统计分析时使用差分隐私方法，对各字段加入一定的噪声或扰动，以防止个人用户信息能从模型或统计结果中被推算出来，同时不影响数据分析的正确性。

3.2.3 基于 AI 的数据异常检测

可使用机器学习、深度学习等算法进行数据访问和操作的行为分析，建立动态的行为模型，通过对异常访问模式、异常流量以及异常的操作序列来发现可疑的活动。并根据需要将可疑的数据泄露事件触发出来，然后融合日志审计和告警的方式及时做出反应并追查源头，使得数据安全防护更具有主动性、智能化的特点。

3.3 网络攻击与恶意行为防范措施

3.3.1 AI 驱动的自动化检测与修复

在数据层面，基于机器学习的异常模式面对海量日志和交易数据实时扫描，并利用自适应算法将异常流量和正常行为区分开来，对出现的数据访问异常即刻标注和阻断，并对可能存在的泄密风险点启动自动隔离模式。在网络层面，深度强化学习结合流量分析与拓扑映射，可在多节点环境中动态调整防护策略，对端口扫描、恶意请求及僵尸网络行为进行预测性阻拦，同时生成修复建议以优化防火墙和入侵防御系统配置。

3.3.2 自动化入侵检测与溯源

利用 AI 对多源异构数据集的综合研判以及多元分析和模型判定等技术对异常流量、异常访问和潜在攻击行为等进行自动化的检测，并且使用行为分析、异常打分和事件关联的方法完成对入侵行为的准确定位和溯源，一旦被发现有可疑行为立即会触发自动化的处置措施（如：封堵可疑连接，隔离受影响的节点等），避免了人工干预可能造成的延迟，保障了应急响应的及时性和精确度。

3.3.3 红蓝对抗与攻防演练中的 AI 应用

在红蓝对抗演练时，AI 可以模拟出攻击方的策略、生成攻击样本、动态地调整防御策略，防御方运用强化学习、对抗训练来更新自己的防护策略，以提高对于各种类型的真实网络环境下的攻击所做出的防御响应能力。持续的演练使得安全人员不断发掘防御上的漏洞，不断检验防护的效果，并通过持续的磨合而形成攻防双方的一种可循环的防御优化闭环，以此来提升对于未知威胁的抵抗能力 [5]。

3.4 隐私与合规风险防范措施

3.4.1 AI 合规治理框架

企业可以构建 AI 合规治理体系，通过算法透明、算法

可解释性的技术方法，在 AI 模型从训练到推理应用的过程中，全过程记录并可追踪管理模型决策路径、训练数据使用情况以及输出结果。算法可解释性工具能够生成决策因子分析报告，确保在异常行为或监管审查时，模型逻辑和行为路径可验证。

3.4.2 数据跨境合规与 AI 监管协同

针对跨境数据流动，企业可以依据目标国家或地区数据保护法规制定规范的数据访问和审计策略，采用跨境数据管理平台及监管系统对接的方式实现数据跨境流转合法性和异常行为监测。同时使用智能合规算法对跨境数据传输的数据集进行不间断的检查，发现可能存在的合规冲突，并自动生成合规报告用于监管审查。

3.4.3 隐私计算与联邦学习

隐私计算技术与联邦学习机制允许多方在数据不出本地的情况下进行模型训练和分析，降低敏感信息暴露风险。通过同态加密、安全多方计算及差分隐私机制，实现数据使用的安全隔离与联合建模。

3.5 系统与平台安全漏洞防范措施

3.5.1 SDL 与 DevSecOps 中的 AI 测试工具

在软件开发生命周期（SDL）与 DevSecOps 流程中，AI 驱动的安全测试可以自动化完成静态代码分析、依赖库检查和漏洞扫描等操作。随后借助于深度学习来对代码潜在的缺陷和安全隐患点进行检测，在代码发布之前给出相应的修复建议。同时结合 CI/CD 流水线做到从开发到生产的一整条链条的安全防护，大幅缩短了漏洞被曝光的时间窗口期。

3.5.2 云原生安全与容器智能检测

对云原生架构及容器化部署环境，使用 AI 辅助实现自动化、实时监测及策略执行的功能，从容器镜像、运行实例、网络通信等方面综合分析容器数据，并通过行为模型检测出异常进程、非法访问、滥用资源等行为，借助自动化策略执行解决上述威胁。

3.5.3 AI 辅助漏洞扫描与修复建议

首先，采用机器学习与图神经网络等算法对系统的组件与服务依赖关系进行全面扫描，得出漏洞严重性评分及修复优先级表。其次，基于历史攻击情况与修复策略库，提供精确修复策略和建议，并给出系统自身配置优化方案。最后，自动化反馈机制支持多轮扫描与策略迭代，在周期最短时间内发现问题、定位问题、修复问题，提升平台防护弹性。

4 大数据与 AI 环境下的综合治理与发展趋势

4.1 技术与管理的全链路融合

在大数据和 AI 系统的建设和使用过程中，应实现从开发到使用的整个生命周期协同。首先，按照企业风险管理和信息安全管理的要求，建立跨部门的协同工作机制，把安全嵌入开发全生命周期以及运维过程中。其次，从技术上利用自动化监控、智能告警、权限控制等手段来达到数据流、模型训练以及系统访问等各个方面实时受控的目标。

4.2 AI 安全与数据安全的协同治理

首先，根据模型生命周期、数据安全的要求建立统一的监控平台，实现从模型训练到推理，再到数据使用的全流程安全审计。其次，对于模型决策过程、数据使用行为均能进行多源头、多角度的数据融合作业，发现异常行为，并针对异常采取闭环策略加以治理，使得 AI 的决策过程和数据的使用行为均能够在受控环境中运行，达到联防联控的作用，提升治理威胁的能力。

4.3 量子计算、AI 武器化与全球博弈

面对技术前沿和全球安全形势的变化，量子计算可能成为未来的杀手锏，侵入现有的加密手段。基于加密的演进和安全防护的能力，我们需要准备量子抗性算法以及相应的密钥轮换计划。另外，随着人工智能武器化以及跨国网络博弈形成新的博弈场，也给安全政策带来了更多的不确定性，这需要我们尽快制定动态安全政策，并做好相应的国际合作准备。可以利用大数据风控以及 AI 来增强技术方面的能力，综合使用风险评估、态势感知和跨域协同治理的技术，来有效提高对于新型的威胁具有更强的预判性和可控性。

5 结论

基于大数据、AI 深度融合下，网络信息安全威胁呈现出多维度、智能化和动态化的全貌。为此急需构建“全链路、智能化、协同化”的安全防护体系，贯穿数据采集、传输、存储、处理及模型部署全生命周期，通过零信任架构、AI 驱动的异常检测、自动化漏洞修复和供应链审计实现智能化防护，同时将技术手段与管理机制深度融合，形成跨部门、跨系统、跨企业的协同防御能力，本文不仅系统分析了大数据与 AI 背景下的网络安全威胁，还提出了针对软件供应链与 AI 模型供应链的重点防护路径，这是保障未来数字生态安全的关键。

参考文献

[1] 王尚.网络信息安全威胁及防范技术研究[J].计算机应用文摘, 2024, 40(5):113-115.

[2] 卢欣.计算机网络信息安全防护探析[J].信息产业报道, 2024(2): 0057-0059.

[3] 马瓯瑞,吴月文,高和平.5G网络信息安全威胁与防护技术研究[J].工程技术研究, 2023(10).

[4] 李军元,徐磊,王磊,等.分析计算机网络下的信息安全及防护[J].电脑乐园, 2023(3):0028-0030.

[5] 郑玉泽.基于大数据技术的计算机网络信息安全与防护策略[J].信息与电脑, 2024, 36(2):218-220.