人工智能与大数据协同的网络安全防护体系

王树凯

河北省网络安全和信息化技术中心，河北石家庄，050034；

摘要：在数字化转型进程持续深化的时代背景下，网络空间的物理与逻辑边界正不断向多领域拓展延伸。网络安全威胁呈现出显著的复杂化特征，同时伴随规模化扩散趋势与动态化演化规律，传统基于固定规则的防护模式已难以有效应对多维度、多层次的风险挑战。人工智能技术与大数据技术的深度协同与有机融合，为网络安全防护体系的创新性发展提供了核心技术支撑。借助大数据技术的核心优势，能够实现安全相关数据的全面采集、跨域整合与深度挖掘分析；结合人工智能算法所具备的智能分析能力、趋势预测功能与科学决策支撑特性，可构建起一套具备主动感知威胁、精准识别风险、快速响应处置与自适应进化能力的现代化防护架构。文章从技术协同内在逻辑、体系核心架构设计与关键应用实践方向三个核心维度，从理论层面系统阐述人工智能与大数据协同防护的内在作用机理与实践推进路径，为网络安全防护模式的优化升级提供系统性思路。

关键词：人工智能；大数据；网络安全；防护体系；协同技术；主动防御

DOI：10.69979/3041-0673.26.04.078

引言

随着信息技术在经济社会各领域的深度渗透与广泛应用，网络空间已逐步发展成为支撑经济社会正常运行的核心载体。其安全稳定运行状态直接关系到关键基础设施的安全保障、核心数据资源的完整保护以及社会公共利益的有效维护。当前阶段，网络攻击手段正朝着自动化部署、隐蔽化实施、规模化发动的方向加速演进。传统防护模式以预设规则为核心，属于被动防御范畴，因其存在响应流程滞后、威胁识别精度有限、环境适应性不足等固有缺陷，已难以应对日益复杂多变的网络安全威胁。大数据技术的快速发展为海量异构安全数据的高效处理提供了坚实技术基础，能够有效打破不同系统间的数据孤岛现象，实现对网络行为数据、设备运行状态数据、威胁特征数据等多源信息的全面汇聚与系统整合。人工智能技术则凭借其强大的自主学习能力与智能推理分析能力，可从海量繁杂的数据资源中挖掘潜在威胁的演化规律，实现对未知安全风险的提前预判与精准识别。两类技术的协同融合应用，能够有效弥补单一技术在网络安全防护场景中的功能局限性，推动防护体系实现从被动防御向主动防御的转型、从经验驱动向数据驱动的转变、从静态防护向动态自适应防护的跨越，为构建全天候运行、全方位覆盖的网络安全屏障提供关键技术支撑。

1 人工智能与大数据协同的技术逻辑

1.1 技术特性的互补性

大数据技术的核心优势集中体现在对海量异构数据的高效采集、安全存储、精准清洗与系统整合能力方面。该技术能够有效突破传统数据处理模式在处理规模、处理速度与处理类型上的局限，实现对网络环境中多源异构数据的全面捕获。所捕获的数据类型广泛，涵盖网络流量数据、系统日志数据、外部威胁情报数据、设备运行状态数据等多个维度，为网络安全防护工作的开展提供了丰富且全面的数据源支撑。人工智能技术以机器学习算法、深度学习模型、强化学习框架等为核心技术支撑，具备从海量数据资源中自动提取关键特征、发现隐藏关联规律、做出科学智能决策的核心能力。这一能力能够将大数据中蕴含的潜在价值转化为实际可落地的防护效能，为防护行动提供科学指导。两类技术的互补性主要体现在两个方面：一方面，大数据技术为人工智能技术提供了模型训练与推理分析所需的“燃料”，即充足且高质量的数据资源，确保人工智能模型能够基于全面的信息输入开展精准分析；另一方面，人工智能技术为大数据赋予了具备自主思考能力的“智慧大脑”，使海量分散的数据资源从被动存储状态转化为主动支撑防护决策的核心战略资源。

1.2 协同运作的内在机理

人工智能与大数据的协同运作过程遵循“数据采集-智能分析-决策生成-反馈优化”的闭环逻辑体系。首先，

通过分布式部署的大数据采集技术，实现对网络环境中多维度、多类型数据的实时捕获与集中汇聚。采集完成后，经过数据清洗去除无效信息、数据标准化处理统一格式等关键步骤，形成结构化、可分析的安全数据资源库。其次，利用各类人工智能算法对构建完成的数据资源库进行深度挖掘分析，具体包括对正常网络行为模式的建模分析、对威胁特征的提取识别、对攻击路径的溯源追踪分析等多项核心任务。通过这些分析工作，实现对已知威胁的精准识别与对未知威胁的智能预判。再次，基于人工智能算法的智能分析结果，结合网络安全防护的核心需求，生成针对性的防护决策方案。该方案具体包括访问控制策略调整建议、异常流量拦截指令、漏洞修复优先级建议等内容，并通过自动化执行模块将这些决策转化为实际防护动作。最后，将防护动作的执行效果数据、新出现的威胁事件数据等关键信息反馈至数据采集与预处理环节，为人工智能模型的持续优化与迭代升级提供重要依据，最终形成“数据驱动分析过程、分析支撑决策制定、决策优化数据质量”的良性循环机制。

2 人工智能与大数据协同的网络安全防护体系架构

2.1 数据采集与预处理层

数据采集与预处理层作为整个防护体系的基础支撑层，其核心功能定位为实现安全数据的全面采集与标准化处理。为达成这一功能目标，通过在网络关键节点部署分布式数据采集节点，采用日志实时采集、网络流量捕获、API 接口调用、传感器数据接入等多种采集方式，对网络环境中产生的异构数据进行实时、全面采集。采集过程中，严格保障数据覆盖的全面性，确保无关键数据遗漏；同时注重数据获取的时效性，为后续分析工作争取时间优势。采集的数据类型全面覆盖网络层的流量数据、传输层的会话连接数据、应用层的业务交互数据、终端设备的运行状态数据以及外部第三方的威胁情报数据等多个层面。在数据预处理阶段，通过一系列专业处理操作提升数据质量：通过数据清洗剔除冗余信息与噪声数据，通过数据去重消除重复记录，通过数据转换实现格式统一，通过数据归一化处理消除量纲差异。这些操作将异构分散的数据转化为统一格式的标准化数据，同时建立完善的数据质量评估机制，从准确性、完整性、一致性等多个维度对数据质量进行严格把控，确保为后续的智能分析环节提供高质量、高可靠性的数据输入。

2.2 智能分析与决策层

智能分析与决策层是整个防护体系的核心中枢，该层深度融合了人工智能与大数据的核心技术能力，承担着数据分析与决策生成的关键任务。在数据分析环节，首先利用大数据挖掘技术对预处理后的标准化数据进行深度分析，具体采用关联分析挖掘数据间隐藏的关联关系、聚类分析实现数据的分类归整、异常检测识别偏离正常模式的数据等多种分析方法，全面挖掘数据中蕴含的关联关系与异常模式。其次，通过机器学习算法构建高精度的威胁识别模型，基于历史积累的海量数据训练模型，持续提升模型对已知威胁的识别准确率；同时引入深度学习算法，利用其强大的特征提取与模式识别能力，实现对未知威胁的特征自动提取与智能推理分析，显著提升威胁识别的泛化能力。在决策生成环节，将强化学习算法与安全防护规则库有机结合，对数据分析结果进行多维度综合评估，生成最优防护策略方案。该方案具体包括威胁分级处置方案、安全资源动态调度建议、防护策略实时调整指令等核心内容，确保决策结果的科学性与针对性。此外，该层还具备强大的模型自优化能力，能够根据新采集的威胁数据与防护动作的反馈结果，自动调整算法核心参数与模型结构设计，持续提升模型的适配性与准确性，保障防护体系的长期防护效能。

2.3 防护执行与反馈层

防护执行与反馈层是将防护决策转化为实际防护行动的关键执行环节，其核心职责是将智能分析与决策层生成的防护策略精准转化为可操作的实际防护动作。该层通过与网络路由器、防火墙等网络设备，入侵检测系统、防病毒软件等安全设备，以及终端主机系统、服务器系统等进行深度联动，实现防护策略的自动化、精准化执行。具体执行功能包括异常流量实时拦截、恶意代码快速查杀、访问权限动态控制、系统漏洞补丁及时推送等多个方面，确保各类网络威胁能够被快速响应与有效处置。同时，该层具备完善的实时反馈功能，通过部署专门的监测模块，对防护动作的执行效果进行实时监测与全面评估，系统收集威胁处置结果数据、防护系统运行状态数据、新出现的安全事件数据等关键信息，并将这些信息及时反馈至数据采集与预处理层，为防护体系的持续优化提供真实有效的数据支撑，最终形成闭环式的防护运作机制。

3 人工智能与大数据协同的关键应用方向

3.1 威胁智能感知与预警

基于大数据技术所具备的全面数据采集能力，结合人工智能技术强大的智能分析能力，能够实现对网络威胁的全方位、无死角感知与提前预警。通过部署在网络各节点的监测设备，对网络流量数据、设备运行行为数据、用户操作行为数据等多源数据进行持续不间断的监测与收集。人工智能模型基于这些收集到的数据，建立正常网络行为的基线特征模型，该模型能够精准刻画网络在安全状态下的行为模式。当监测到的数据偏离这一基线特征时，模型能够自动识别异常行为，并结合预设的风险评估指标体系对异常行为进行风险等级评估。对于评估后判定为潜在威胁隐患的异常情况，模型能够基于历史积累的威胁数据与威胁演化规律，进行趋势预测与风险预警，及时向安全管理人员发送预警信息，为其提供充足的威胁处置准备时间，将安全威胁遏制在萌芽状态，实现网络安全防护从被动响应向主动预防的根本性转变。

3.2 精准威胁识别与溯源

传统的威胁识别方式主要依赖安全专家人工定义防护规则，这种方式存在明显的局限性，难以应对不断变异、持续更新的新型网络威胁。人工智能与大数据的协同应用，能够有效突破这种规则依赖的局限，实现对各类威胁的精准识别与快速溯源。通过大数据技术的整合能力，将海量的外部威胁情报数据与历史攻击案例数据进行系统整合，构建起全面的威胁数据资源池。利用深度学习算法对这些数据进行深度分析，构建高精度的威胁特征库，该特征库能够实现对已知威胁的精准匹配与快速识别。对于尚未记录在特征库中的未知威胁，通过无监督学习算法对其行为模式与特征参数进行深入分析，实现自动分类与特征标记。同时，结合大数据关联分析技术与攻击路径挖掘技术，能够对网络攻击的源头地址、传播路径、攻击目标、影响范围等关键信息进行全面溯源追踪，为威胁的快速处置与责任认定提供坚实的技术支撑。

3.3 自适应防护策略优化

网络运行环境的动态变化特性与网络威胁的持续演进趋势，对网络安全防护策略提出了自适应调整的核心要求。基于人工智能与大数据的协同技术，网络安全防护体系能够实时感知网络环境的动态变化与威胁演化趋势，实现防护策略的自动优化与动态调整。通过大数据技术对网络拓扑结构变化数据、设备运行状态数据、业务需求调整数据等进行实时分析，精准把握网络环境的当前状态与变化趋势，动态调整访问控制策略的具体规则、安全资源的分配方案等核心防护内容。同时，根据新出现的威胁特征数据与防护效果反馈数据，自动更新防护规则库中的相关条目与人工智能模型的核心参数，确保防护体系能够持续适应不断变化的安全环境，显著提升长期防护工作的有效性与稳定性。

4 结语

人工智能技术与大数据技术的协同融合应用，为网络安全防护体系的转型升级提供了核心技术支撑，推动传统防护模式实现从被动防御向主动防御的跨越、从静态防护向动态自适应防护的转型。通过两类技术在特性上的互补优势与协同运作形成的闭环逻辑，构建起一套涵盖数据采集、智能分析、决策执行与反馈优化等关键环节的完整防护架构。该架构在威胁感知的全面性、威胁识别的精准性、威胁处置的高效性等关键环节实现了防护效能的显著提升。未来，随着人工智能与大数据技术的持续演进与发展，需进一步强化两类技术在网络安全领域的深度融合程度，不断优化技术协同机制，提升人工智能模型的鲁棒性与环境适应性。同时，需高度关注技术应用过程中可能出现的伦理问题与安全风险，通过完善技术规范与管理机制，推动网络安全防护体系向更智能、更可靠、更全面的方向持续发展，为数字化社会的安全稳定运行提供坚实保障。

参考文献

[1]董理君,刘超,张锋,等.大数据与人工智能时代背景下的网络安全课程体系研究[J].软件导刊,2024,23(8):37-42.
[2]崔倩,姜辉,刘雯.人工智能大数据背景下高校网络安全问题与对策[J].微型计算机,2024(12):100-102.
[3]董鹏,孙鹏,李海霞.基于人工智能技术的大数据网络安全防御体系研究[J].张江科技评论,2024(6):65-67.
[4]陆华.基于人工智能与大数据的计算机网络安全防御系统研究[J].信息与电脑，2024，36(3):47-49.
[5]池泽娟.基于人工智能的大数据网络安全解决方案[J].计算机应用文摘，2024，40(2):93-95.