人工智能提升网络空间安全防御能力研究

王继荣
云南金质信息技术服务有限公司，云南 昆明 650000

DOI: 10.61369/TACS.2025060010

摘　　要： 随着我国数字化进程的持续推进，网络空间面临的安全威胁也在不断升级，并且展现出复杂化、规模化与隐蔽化特征。传统网络安全防御体系建设主要依靠基于规则和特征匹配的防御策略，但目前难以应对高级持续性威胁（APT）、零日漏洞等新型攻击模式。对此，本文从人工智能视角切入研究，通过阐述人工智能在威胁检测、态势感知、自动化响应等方面的价值与优势，进而提出基于人工智能提升网络空间安全防御能力的策略与方法，以此构建数据处理能力突出、模式识别能力强大与预测能力精准的网络空间安全防御范式，从而强化我国网络防御的主动性、高效性与精确性，打造健康且可持续的人机协同防御生态。

关 键 词： 人工智能；网络空间；安全防御能力；策略

Research on Enhancing Cyberspace Security Defense Capability with Artificial Intelligence

Wang Jirong
Yunnan Hi-Q Information Technology Service Co., Ltd, Kunming, Yunnan 650000

Abstract: With the continuous advancement of China's digitalization process, the security threats faced by cyberspace are constantly escalating, and have shown characteristics of complexity, scale and concealment. The construction of traditional network security defense systems mainly relies on defense strategies based on rules and feature matching, but currently it is difficult to deal with new attack modes such as Advanced Persistent Threats (APT) and zero-day vulnerabilities. In response to this, this paper conducts research from the perspective of artificial intelligence (AI). By expounding the value and advantages of AI in threat detection, situation awareness, automated response and other aspects, it further proposes strategies and methods to enhance cyberspace security defense capabilities based on AI. This aims to build a cyberspace security defense paradigm with outstanding data processing capabilities, strong pattern recognition capabilities and accurate prediction capabilities, thereby strengthening the initiative, efficiency and accuracy of China's network defense, and creating a healthy and sustainable human-machine collaborative defense ecosystem.

Keywords: artificial intelligence (AI); cyberspace; security defense capability; strategies

引言

随着5G、人工智能、工业互联网、云计算、大数据等技术普及，现代社会迎来数字化转型浪潮，不仅极大扩展了网络空间边界，而且也带来了新的安全漏洞与攻击方式。面对当前网络空间安全中面临的攻击数量显著提升、攻击手段快速升级、网络防御人才短缺、海量日志数据难以分析等现实问题，人工智能成为提升网络空间安全防御能力的关键手段，尤其在机器学习、自然语言处理、深度学习、知识图谱技术支持下，为现代网络空间安全系统构建提供了新的范式与方法论。

一、人工智能在网络空间安全防御中的应用价值

（一）增强威胁检测与识别的能力：从“事后响应”到“事前预警”和“事中洞察”

第一，强化异常检测。在无监督学习算法等大规模 AI 模型支持下，人工智能可以自动学习当前网络环境、用户习惯以及系统特征的常规行为基线，从而可以实时识别脱离基线的异常数据[1]，可以显著提升未知威胁与内部威胁的检测与识别效果，具备了应对零日攻击的基础能力，从“事后响应”逐步向“事前预警”发展。

第二，恶意软件识别与钓鱼检测。在深度学习支持下，人工智能可以根据分析操作码、API 调用序列、PE 文件结构等静态数

据，同时结合沙箱的动态变化情况，精准识别不同类型的恶意软件[2]。此外，通过计算机视觉技术与NLP技术，人工智能也可以通过分析邮件信息、链接信息、图片信息等要素，识别出高端复杂的钓鱼邮件。

第三，APT攻击检测与识别。人工智能与大数据技术结合，可以对网络空间相关的海量数据进行分析，包括终端日志、网络流量、威胁情报等。依托大数据模型可以绘制外部攻击链的可视图，从而挖掘出高级可持续的安全威胁因子，达到了“事中洞察”的预警效果。

（二）实现安全运维的自动化与智能化：提升响应效率，解放人力

第一，提升自动化水平与编排和响应速率。SOAR平台可以依托人工智能快速处理大量的低级告警，实现自动封禁恶意IP等效果[3]；同时，也可以利用机器学习与预定义方案建立相对复杂的自动化响应工作流，帮助用户自动解决部分安全问题。比如主机感染后自动隔离、证书到期时自动吊销、定期下发扫描任务等，可以将平均响应时间缩短至分钟甚至秒级。

第二，漏洞管理自动化执行。人工智能具备优先、自动执行漏洞修复的能力，一方面可以通过CVSS评分，对漏洞的可利用性、资产重要性、威胁度等进行分析，以此为不同漏洞给出风险评级，进而按照优先级依次进行修复处理。

第三，智能安全运维助手支持。在大语言模型（LLM）支持下，人工智能还可以构建AI助手，以此担任用户的安全分析师，其主要功能包括基于自然语言查询和分析安全事件、自动总结与生成安全事件分析报告、解读与分析复杂攻击技术与流程、提出科学应对的防御建议等[4]。

（三）提升安全态势感知与决策支持水平：从“碎片化信息”到“全局化认知”

第一，突出的宏观态势感知能力。人工智能可以根据网络空间的内外部数据构建全局知识图谱，进而深度分析用户的资产、流量、外部威胁、暗网等相关数据[5]，进而建立可视化的网络安全防御系统，通过图形直观呈现网络安全状态，显示攻击来源、攻击目标与攻击路径，从而为安全防御提供决策支持。

第二，前瞻的预测性防御。在回归预测与时间序列分析等模型支持下，人工智能还能对当前网络空间可能受到的未来攻击趋势与漏洞情况提出预测，从而达到“防患于未然”的效果。

第三，全局辅助决策。当安全威胁或事故发生时，人工智能同样可以根据当前威胁情况或事故原因进行模拟推演，从而提出不同应对措施导致的不同后果，以此为用户处理安全事件提供以数据为驱动的最佳决策选项。

二、人工智能提升网络空间安全防御能力的有效策略

（一）技术融合策略：构建纵深协同的智能防御体系

第一，建立“AI+”安全防御体系，用人工智能赋能现有各类安全产品。具体来说，可以将AI进行模块化设计，并直接嵌入防火墙、EDR、WAF、IDS/IPS等常见安全产品[6]，以此赋予传统安全产品智能检测、自适应检测等功能。

第二，推进多平台整合，建立一体化安全运营平台。比如可以构建SOC 2.0安全运营平台，一方面以AI为核心引擎，另一方面将各类安全防护平台进行集成，以此解决数据孤岛问题，将网络终端、云空间、应用层等各类数据进行集中采集与分析，进一步提升其智能化效率与响应速度。

第三，建立云地协同体系，发挥云计算与本地系统的优势互补作用。在本地可以部署轻量级的AI模型，主要负责处理实时响应问题；同时可以在云端部署算力强大的AI系统与安全威胁数据库，进而通过模型训练与深度分析，构建“云地一体”的智能化网络安全防御体系[7]。

（二）数据与算法策略：夯实智能防御的基石

第一，建立高质量数据治理方案。在人工智能赋能网络空间安全防御中，安全数据质量是影响AI模型功效的核心因素，因此需要通过全方位、系统化的数据管理策略，提升数据采集质量与分析能力。具体来说，一要建立标准化的数据使用流程与方案，规范数据的采集、清洗、标注与存储等相关规范，避免出现虚假、格式不统一等数据问题[8]。二要推进攻击样本库的构建，并不断提升样本库质量与多样性，同时还应建立正常行为样本系统，以此用于AI模型训练，提升大模型的精准度与可行性。

第二，推进算法模型的持续优化与创新升级。算法模型是人工智能实现自动化、智能化与数据化服务的关键，因此在依托人工智能提升网络安全防御能力时，还应注重算法模型层面的优化升级。一方面要推进轻量化变革，解决IoT等计算资源受限的边缘设备的安全防御问题，通过轻量化AI模型提供智能化防御服务。另一方面要深化可解释性，在AI辅助决策中需要强化相关数据的透明度，并通过文本解释与翻译说明数据背后隐藏的问题与安全风险，从而帮助用户理解“为什么这么做”，以此提升用户对AI的信任度，并进一步辅助审计与问责等环节[9]。此外还应强化对抗性训练水平，通过可能受到的供给类型，采用对应的对抗样本进行攻击，以此不断磨炼模型的鲁棒性，提升其防御水平与抗干扰能力。

第三，建立持续性学习机制。由于网络环境与相关技术发展迅速，人工智能在网络空间安全防御中应用时同样需要与时俱进，通过在线学习与增量学习流程，不断强化模型功能与特征，进而不断学习新的攻击数据与防御反馈，并顺应网络环境与威胁因素进行适应性发展，避免模型出现老化与滞后问题。

（三）人机协同策略：优化组织流程与人才培养

第一，重新定位角色分工。在网络空间安全防御体系建设中，人工智能一定程度上取代了安全分析师的部分工作职责，因此需要重新划定职责边界，以此优化人才培养。一方面，AI主要可以负责处理高重复度与高容量的任务，也可以对大数据进行初步分析。安全分析师则需要从战略决策、规则制定、复杂性攻击调查、模型优化等方面入手，提高自身的工作价值与不可替代性。

第二，推进人才培养与核心技能升级。一要升级安全团队的知识结构，培养既具备网络安全知识又了解AI算法的复合型人

才。二要深化内部培训体系，强化 AI 工具等先进软硬件的应用能力，确保传统安全人员也能利用人工智能辅助解决工作。

第三，优化安全流程。在人工智能融入网络安全空间防御系统后，需要根据 AI 的能力特征对安全事件响应流程进行重新设计与优化完善，尤其要突出自动化响应优势，确保自动化环节与人工审核形成无缝衔接，建立高效工作流系统。

（四）法律与伦理策略：构建负责任的智能安全应用框架

第一，强化隐私保护与合规性。在人工智能应用视域下，网络安全防御必须关注用户隐私与数据跨境等相关问题，尤其要建立隐私保护原则，并采用数据脱敏、邦联学习等基础进行完善，从而符合《数据安全法》《网络安全法》《个人信息保护法》等相关法律法规的规定和要求 [10]。

第二，建立算法安全与问责制度。针对人工智能大模型构建的智能安全算法，还需建立对应的评估与审计制度，一方面要防止算法出现偏差问题，从而导致危险因素误报或漏报问题。另一方面要避免出现算法歧视现象，特别是针对特定群体的网络空间服务中，需要建立统一的业务流程与应对方案。当算法安全出现问题时，需要将对应的职责对应到个人，以此强化相关工作人员的算法伦理意识。

第三，建立防御性应用伦理机制。应倡导在网络安全领域应用人工智能技术，同时也要警惕并防范 AI 技术在恶意攻击、黑客系统等方面的应用。此外，还可以强化与国际相关组织机构的对话交流，并协同构建约束恶意引用 AI 的国际标准，形成国际统一的管理规范。

三、结语

综上所述，人工智能在网络空间安全防御领域有着智能威胁检测、态势高效感知、自动化响应等优势，因此可以全面重构网络空间防御模式，显著提升其防御主动性、精确性与高效性。对此，应从技术融合、数据分析、人机协同以及法律伦理等层面构建完善的网络空间安全防御体系，全面提升其防御能力。在未来，网络空间防御系统必然建立在“人机协同，智慧赋能”的生态系统之上，这是构建更安全、更可信数字世界的追求，同时也是政府、产业、高校通力合作应对挑战的必然之路。

参考文献

[1] 常禾雨，魏祥坡，司念文，屈丹 . 面向网络空间防御的人工智能安全对抗框架 [J]. 信息工程大学学报 ,2024,25(05):574-579.

[2] 华程，卜俊兰，张婷婷 . 人工智能技术在网络空间安全防御中的应用 [J]. 数字技术与应用 ,2024,42(10):60-62.

[3] 刘小莉 . 人工智能技术在网络空间安全防御中的应用 [J]. 信息记录材料 ,2023,24(09):189-191+195.

[4] 刘邦桂 . 基于人工智能的网络空间安全防御策略研究 [J]. 软件工程 ,2023,26(04):52-56.

[5] 郑汉军 . 人工智能技术在网络空间安全防御中的应用 [J]. 网络安全技术与应用 ,2022,(01):100-101.

[6] 朱晨安 . 人工智能技术在网络空间安全防御中的应用分析 [J]. 中国高新科技 ,2021,(21):149-150.

[7] 方志伟 . 基于人工智能技术的网络空间安全防御研究 [J]. 电子技术与软件工程 ,2021,(14):240-241.

[8] 贾焰，方滨兴，李爱平，顾钊铨 . 基于人工智能的网络空间安全防御战略研究 [J]. 中国工程科学 ,2021,23(03):98-105.

[9] 牛文 . 人工智能技术在网络空间安全防御中的实践探究 [J]. 无线互联科技 ,2021,18(08):72-73.

[10] 郁陶 . 人工智能技术在网络空间安全防御中的应用 [J]. 电子世界 ,2021,(06):208-209.

