XXXX

人工智能赋能网络攻防平台的路径与进展

王群[1,2]，李馥娟[1]，高光亮[1,2]
（1. 江苏警官学院计算机信息与网络安全系，江苏 南京 210031；
2. 江苏省电子数据取证分析工程研究中心，江苏 南京 210031）

摘　要：网络攻防平台是支撑网络安全对抗演练、提升实战能力的关键载体。本文聚焦于人工智能技术对网络攻防平台的赋能路径与体系化构建方法，针对传统平台在动态演化、场景多样性与评估维度等方面的局限，研究融合多智能体系统、强化学习与大语言模型等关键技术，提出一种集成智能攻击模拟、自适应防御决策与自动化评估反馈的分层可配置平台架构，并设计了基于模仿学习与终身学习的层次化智能体训练机制。该平台能够动态生成高逼真攻击链、实现跨域协同防御与多维度量化评估，从而有效提升攻防演练的实战性与训练精准度。通过策略库、环境仿真及评估体系的场景化配置，该架构可灵活适配教育、科研与产业演练等多元应用需求。本文进一步剖析了平台面临的可解释性、仿真真实性、数据隐私与伦理规范等关键挑战，并从可解释人工智能、数字孪生、隐私计算及标准化协同等方面展望了未来研究方向，以期为构建下一代自适应、可持续演进的新型网络防御体系提供理论支撑与实践参考。

关键词：人工智能；网络攻防平台；多智能体系统；自适应防御；人才培养
中图分类号：TP391
文献标志码：A
doi: 10.11959/j.issn.1000−0801.

The Path and Progress of Empowering Network Attack and Defense Platforms with Artificial Intelligence

WANG Qun[1,2], LI Fujuan[1], GAO Guangliang[1,2]
1. Department of Computer Information and Cybersecurity, Jiangsu Police Institute, Nanjing 210031,China
2. Jiangsu Electronic Data Forensics and Analysis Engineering Research Center, Nanjing 210031, China

Abstract: Cyber range platforms serve as critical infrastructure for conducting cybersecurity exercises and enhancing operational readiness. This paper focuses on the enabling pathways and systematic methodologies for empowering

收稿日期：2026−01−27；**修回日期：**2026−02−23
通信作者：王群(1971-)，男，甘肃天水，博士，教授，邮箱:wqun@jspi.edu.cn；李馥娟（1974-），女，硕士，教授，邮箱:lifujuan@jspi.edu.cn；高光亮（1988-），男，山东潍坊，博士，邮箱:gaoguanliang@jspi.edu.cn
基金项目：国家自然科学基金（72401110）；江苏省高校优秀科技创新团队；公安技术、网络空间安全“十四五”江苏省重点学科；教育部人文社会科学研究规划基金（24YJAZH158）。
Foundation Items: The National Natural Science Foundation of China (72401110), the Excellent Scientific and Technological Innovation Team of Jiangsu Universities, Key disciplines of Jiangsu Province in the 14th Five-Year Plan:Public Security Technology and Cyberspace Security, Ministry of Education Humanities and Social Sciences Research Planning Fund(24YJAZH158).

such platforms through artificial intelligence (AI). To address the limitations of conventional platforms in dynamic evolution, scenario diversity, and evaluation dimensions, we integrate key technologies including multi-agent systems, reinforcement learning, and large language models. We propose a hierarchical and configurable platform architecture that incorporates intelligent attack simulation, adaptive defense decision-making, and automated evaluation and feedback. Furthermore, a layered intelligent agent training mechanism based on imitation learning and lifelong learning is designed. The proposed platform can dynamically generate high-fidelity attack chains, achieve cross-domain collaborative defense, and perform multi-dimensional quantitative evaluation, thereby significantly improving the realism and precision of cyber exercise training. Through scenario-specific configuration of strategy libraries, environment simulation, and evaluation systems, the architecture can be flexibly adapted to diverse application needs in education, scientific research, and industrial drills.This paper also examines key challenges faced by the platform, including model interpretability, simulation authenticity, data privacy, and ethical compliance. Future research directions are outlined, encompassing explainable AI, digital twins, privacy-preserving computation, and standardization collaboration. The study aims to provide theoretical support and practical guidance for building a next-generation adaptive and sustainably evolving cyber defense system.

Key words: artificial intelligence, cyber offense and defense platform, multi-agent system, adaptive defense, talent cultivation

1 引言

网络安全本质上是攻防双方在知识、技能与策略上的动态博弈，而网络攻防平台正是复现这一博弈过程、开展对抗演练与效能评估的数字试验场[1]。传统平台多基于虚拟化及仿真技术构建静态环境，依赖预定义脚本与固定规则开展演练。此类平台虽能满足端口扫描、基础漏洞利用等入门训练需求，但其静态特性与网络攻击的智能化演进趋势已明显脱节[2-3]。传统网络攻防平台的局限性主要集中在：一是环境缺乏动态演化能力，无法复现高级持续性威胁（APT）中攻击者根据防御反馈动态调整路径的博弈过程[4]；二是演练场景同质化严重，固定拓扑与预设脚本导致训练者难以应对跨域攻击、供应链攻击等新型威胁，如何将理论有效应用于实践，仍需深入探索[5]；三是评估体系单一固化，仅以攻击成功率、漏洞利用数量等显性指标衡量训练效果，无法量化训练者的威胁分析、应急决策及策略优化等技术应用和分析能力[6]。

人工智能（AI）技术的自主学习、动态博弈与数据挖掘等特性，恰好匹配网络攻防的对抗性、演化性、实战性等关键需求，推动攻防平台从静态的训练载体升级为具备自主进化能力的智能对抗系统[7]。多智能体系统（MAS）通过构建攻击者、防御者、裁判等交互主体，实现复杂攻击链的系统性仿真；强化学习驱动的智能体可通过环境反馈迭代优化方式，有效提升演练的真实度[8]；生成对抗网络（GAN）能够构建高逼真的异构网络场景，破解传统平台场景的固有难题[9]；大语言模型（LLM）则在攻击脚本生成、安全日志解析及防御策略解释等环节发挥作用，降低操作门槛并提升可解释性[10]。在国内研究方面，刘艾杉等人[11]对深度强化学习的对抗攻防研究进行了全面梳理，系统阐述了从状态、奖励到动作的多种攻击路径及相应防御策略。王立夫等人[12]分析了网络拓扑可辨识性从同构到异构情境下的双向演变规律，为复杂多智能体系统的拓扑辨识理论提供了普适性分析框架。张学旺等人[13]提出了一种融合图节点中心性分析与大语言模型的源码漏洞检测数据增强方法DA_GLvul，在多个评估指标上有效提升了检测性能。在国外

研究方面，Maddireddy 等人[14]提出了强化学习作为动态网络防御核心技术的理论依据，阐明其在入侵检测、恶意软件识别等场景中的自适应防御机制。Sarhan 等人[15]将联邦学习与通用数据格式相结合，解决了跨域协同训练中的隐私保护与数据异构难题。Lanka 等人[16]提出了一种融合大语言模型与蜜罐数据分析的新型威胁检测方法，实现了对攻击者 TTP（tactics、techniques、procedures，战术、技术、程序）[17]的快速提取与实时攻击识别。尽管上述研究展示了人工智能在网络攻防特定环节（如漏洞检测、入侵响应、威胁识别）的应用潜力，但聚焦于构建集成化、可演进智能攻防平台的研究仍面临一些共性挑战。首先，在架构层面，多数平台或侧重于攻击模拟（如自动化渗透测试工具），或侧重于防御决策，未能将攻击、防御、推演、测评等核心功能在一个统一、松耦合的架构中进行有机集成。其次，在智能体能力层面，现有系统的智能体多针对单一、特定任务进行训练（如利用某个漏洞），缺乏在复杂、动态、多阶段攻击链中进行自主策略规划和跨智能体协同的能力；同时，智能体的行为安全性常被忽视，存在模拟攻击溢出演练环境的潜在风险。最后，在评估维度层面，评估多集中于技术指标（如检测率、成功率），缺乏对操作过程质量、人员认知能力提升以及平台自身演进效果的综合性、可解释性评估体系。这些局限性使得现有平台难以支撑从基础教学到高端科研，再到大规模产业演练的全系列、可持续演进的需求。本文旨在系统回应这些问题，重点解决如何构建安全可信、具备跨域协同能力且能持续演进的智能攻防平台等主要问题。通过探索多种 AI 技术的融合路径与演进脉络，为下一代自适应网络防御体系的建设提供理论支撑与实践指引。

2　人工智能赋能网络攻防实验平台的关键技术

为应对 AI 模型安全、可解释性及跨域协同等挑战，本节系统阐述智能攻击模拟、自适应防御决策与自动化评估反馈等关键技术，并通过其深度融合构建平台的支撑和服务能力。

2.1　智能攻击模拟技术

智能攻击模拟技术[18]是一种用于评估和提升网络防御体系能力的方法，其核心在于利用人工智能动态生成并执行模拟攻击。智能攻击模拟技术超越了静态扫描和固定脚本模拟，能够实现自主学习和自适应功能，从而复现更复杂、隐蔽的攻击行为。其核心在于构建一个双向的智能系统：对内，系统能自主感知目标网络环境的变化，实时调整攻击策略；对外，能够从每次攻防交互中通过学习获取信息，积累经验，优化后续行动。智能攻击模拟的关键在于引入了多种 AI 技术，使其具备感知、决策、学习和演化的能力。作为现代网络安全防御体系的前沿方向，代表了从传统被动防护到主动、自适应防御的理念转变。

（1）多智能体系统建模与协同攻击。通过构建异构智能体间的交互环境，为模拟复杂网络攻击行为提供高度拟真的仿真框架。该框架通常包含攻击者、防御者、目标系统与环境以及裁判 4 类核心智能体。在此框架下，各智能体通过感知、决策、行动的自主循环进行交互，能够模拟诸如 APT 攻击中常见的横向移动、权限提升及数据渗漏等复杂、多阶段的协同攻击链。具体而言，攻击者智能体通过实时感知目标系统的漏洞态势、服务配置与防御规则，能够动态规划并调整攻击路径。防御者智能体则依据攻击行为动态调整防护策略，而裁判智能体基于预设的安全指标（如攻击影响范围、行为隐蔽性等）对攻防行为进行量化评估。例如，灵御（PandaGuard）[19]平台便采用此类多智能体框架，集成了 19 种攻击

算法与12种防御机制，实现了对49个大语言模型的系统性安全评估。通过智能体间的协同与对抗，能够自主演化出超出预设脚本的复杂攻击场景，从而有效解决传统网络攻防平台因依赖静态脚本而导致的场景固化与对抗性不足的缺陷。研究表明，在多智能体框架中，攻击行为的适应性与隐蔽性可得到有效提升，为评估和增强防御体系的鲁棒性提供了有效途径。

（2）基于强化学习的攻击策略优化。基于强化学习的攻击策略具备自主优化能力，其主要机制在于智能体通过与环境的持续交互，并依据奖励信号优化决策，从而推动攻击策略从静态预设向动态生成转变。在技术实现上，攻击智能体将渗透测试过程建模为一个序列决策问题：其状态空间定义为目标网络的拓扑、服务配置及防御规则等环境信息；动作空间则覆盖漏洞扫描、利用代码执行及权限提升等攻击操作；奖励函数通常设置为攻击成功率、行为隐蔽性与操作效率等多目标综合指标。研究表明，基于近端策略优化（PPO）的攻击智能体不仅能发现并利用防御体系中的未知脆弱点，其行为模式与真实APT攻击的相似度也显著提高。此外，进化策略（ES）可与PPO等算法结合，在保持早期样本效率的同时，实现了更稳定、高效的后期策略优化，为攻击策略的自主进化提供了另一可行路径[20]。这种自我演进能力使攻防训练得以持续逼近并超越现有威胁水平，成为平台实现持续演进的重要基础。

（3）大语言模型赋能的攻击脚本生成与语义理解。LLM通过其强大的代码生成与自然语言推理能力，为网络攻击模拟实现了从语法层到语义层的功能提升[21]。在攻击脚本生成方面，利用海量漏洞报告、渗透测试手册及攻击代码库对训练领域大模型进行微调，能够将高级别的自然语言攻击意图（如利用Log4j漏洞获取服务器权限）直接编译为可执行的具体攻击序列，包括漏洞检测、载荷构造、权限提升与持久化维持等步骤。此过程不仅自动生成了多样化的攻击样本，更能通过语义理解生成对抗性变体，有效规避基于静态特征的检测规则，从而提升攻击模拟的覆盖范围与逼真度。在可解释性方面，LLM的核心贡献在于其内生的因果推理与说明能力。LLM能够将抽象的、二进制的攻击行为，转化为人类可理解的战术逻辑链。例如，在生成利用Log4j漏洞的脚本时，模型可同步解释其选择JNDI注入而非RCE其他载体的决策依据，并评估该操作在特定防御策略（如WAF规则、IDS签名等）下的被检测概率。这种动态的、伴随式的语义注解，将攻击模拟从一个黑盒操作过程，转变为透明的白盒分析过程，极大地强化了安全人员在对抗中对攻击者思维模式与TTP的理解深度。

2.2 自适应防御决策技术

自适应防御决策是实现网络安全防御智能化的关键技术，其目标是让防御体系具有认知与决策能力，从而完成从机械执行到智能响应的转变。通过构建集感知、分析、决策与行动于一体的闭环系统，最终形成对未知威胁的自主应对与跨域协同防御。

（1）多源融合的实时威胁感知。多源融合的实时威胁感知技术是构建自适应防御体系的基石，其核心在于通过异构数据源的协同分析与深度学习模型的时空特征提取，实现对安全态势的全面、精准、动态认知。具体而言，该系统通过并行采集网络流量、终端日志、用户行为数据（UEBA）以及外部威胁情报等多维数据，构建统一的数据表征层。在此基础上，系统采用复合深度学习架构进行特征挖掘：利用卷积神经网络（CNN）从网络流量中提取空间局部特征（如数据包大小分布、协议类型异常等）；通过长短期记忆网络（LSTM）或时序Transformer模型捕捉系统日志与用户行为中的长程时序依赖关系，有效识别如潜伏性横向移动等复杂攻击序列；同

时，借助图神经网络（GNN）对威胁情报中的实体关系进行建模，实现攻击组织的溯源与关联分析。这种多源异构数据的深度融合，不仅突破了传统单点检测导致的数据孤岛困境，更通过特征层面的交叉验证，可有效提升威胁检测的置信度。研究表明，基于多源融合感知的检测模型对未知攻击的识别率可达89.2%，误报率较传统方法降低60%以上[22]。该能力为实现跨域协同防御奠定了至关重要的感知基础。

（2）基于强化学习的动态策略生成与优化。强化学习通过智能体与环境的持续交互，为构建自主决策、动态响应的网络安全防御体系提供了可行的技术路径。其核心框架将网络防御场景建模为一个部分可观测的马尔可夫决策过程。防御智能体通过感知网络安全状态（如网络流量异常、系统日志告警、威胁情报指标等），进而执行防御动作（如调整防火墙规则、隔离受感染主机、更新入侵防御系统IPS的签名等[23]），最终从环境中获得奖励或惩罚信号，并通过逐步学习优化出一套能够在复杂、动态的威胁环境中实现最大化的安全策略[24]。在这一框架下，奖励函数的设计是引导智能体实现系统可用性损失最小化和威胁遏制最大化双重目标的关键[25]。研究表明，采用相对稀疏的奖励信号（例如，重点奖励智能体成功维持网络未被破坏的状态），相较于设计过于复杂的稠密奖励函数，反而能训练出更有效、更稳健的防御智能体[25]。尽管强化学习在网络安全领域展现出巨大潜力，但其在实际应用中仍面临样本效率、策略稳定性、模型可解释性等多重挑战。为此，近年来研究者们提出了一系列优化路径与前沿方法，主要工作包括提升样本效率与训练稳定性[26]、降低计算开销与增强实时性[27]、引入因果推理提升决策可解释性[28]、自动化算法设计与发现[29]等方面。

（3）人机协同的智能决策支持。人机协同智能决策支持系统[30]作为现代网络安全防御体系的核心组成部分，正在推动防御方式从传统人工主导向着人机智能融合转变，实现了人类认知优势与机器计算能力的高效互补，为解决复杂网络环境下的动态威胁提供了创新性解决方案。在技术实现层面，人机协同系统主要提供了3个主要功能：动态策略推演、可解释性交互和自适应任务分配。其中，动态策略推演依托蒙特卡洛树搜索（MCTS）和贝叶斯推理等算法，能够对防御策略的长期效果进行大规模并行模拟和不确定性量化。研究表明，基于MCTS的推演系统可以评估超过数千种可能的攻击路径，为决策者提供具有前瞻性的防御方法建议[31]。在可解释性方面，系统通过分步式决策指导框架，将人工智能的复杂推理过程依次转化为修复优先级建议、具体规则配置与检测工具部署等可操作序列，这种基于自然语言生成的解释机制有效提升了决策过程的透明度。自适应任务分配机制通过实时监测网络态势和专家认知状态，动态调整人机分工边界。当面临常规威胁时，系统可自主执行日志分析、规则配置等任务；而在处理新型复杂攻击时，则自动提升专家的参与度。这种弹性（如资源伸缩、模型切换）分工机制既确保了响应效率，又保障了关键决策的可靠性。目前，对于网络攻防平台中人机协同研究面临的挑战主要包括人机信任校准、决策责任界定以及系统适应性提升等问题。未来发展方向应着重于构建更完善的信任评估模型，建立决策溯源机制，并加强系统在开放环境中的持续学习能力。随着大语言模型等新技术的发展，人机协同系统有望实现更自然的交互方式和更深入的认知协作，最终建立真正意义上的混合增强智能防御体系。

2.3　自动化评估与反馈机制

自动化评估与反馈机制是实现网络攻防平台智能化演进的核心环节，其作用已超越简单的胜负判定，已深入到攻防全过程。该机制通过对攻击、防御及系统状态的多维指标进行实时采集、

量化分析与动态反馈，构建起一个覆盖训练前、中、后的完整评估闭环，从而驱动防御策略与人员能力的协同优化与自主演进。

（1）多维度可量化的评估指标体系。在人工智能赋能的网络攻防平台中，多维度可量化的评估指标体系已超越传统的二元胜负判定，发展成为涵盖技术效能、过程质量与认知能力的综合评估框架[32]。这套体系旨在从多维度客观衡量平台在提升用户实战能力方面的成效，是解决当前AI赋能网络攻防平台评估标准缺失问题的关键。其中，在技术效能方面确保了基本战术执行的有效性，主要关注攻防行动的直接效果与质量。对攻击方，可评估其攻击路径的复杂性（例如，是否绕过多种防御机制）、漏洞利用的效率（如从渗透到提权的平均时间）以及行为的隐蔽性（如是否触发入侵检测警报）。对防御方，则需评估其威胁检测率、平均响应时间及防御策略的优化度（如策略调整后风险降低的百分比）。在过程质量方面，侧重于攻防任务执行的过程流畅性与资源利用效率。例如，可以评估任务链的完成度（预定义攻击或防御步骤是否全部完成）、工具使用的熟练度与合理性，以及资源消耗（如CPU/内存占用与网络带宽开销等）。在认知能力方面，是衡量高阶思维能力的关键，包括应对未知威胁的能力、策略创新的程度（例如，是否采用非标准方法突破防御或化解攻击）以及技能的跨场景迁移能力。例如，AIRTBench基准通过分析模型在多样化挑战中的表现，来评估其自主发现和利用漏洞的认知能力[33]。当前研究趋势表明，评估体系正在与业界标准框架深度整合。其中，与MITRE ATT&CK[32]框架的对齐使得评估能够系统化地覆盖各类攻击技术和防御措施，大大提升了评估的全面性和实用性。同时，新兴的专用评估基准如AIRTBench[33]、OCCULT[34]等，为衡量自主攻防能力提供了标准化方法，这些基准通过设计多样化的挑战场景，能够有效评估系统在复杂环境下的表现。为客观、全面且无偏差地评估攻防演练成效，本文构建了一套涵盖技术效能、过程质量与认知能力的三级量化评估体系（如表1所示）。该体系超越单一的胜负判定，从技术效能、过程质量、认知能力三个维度进行解构，每个维度下设具体、可量化且力求客观的指标。与此同时，为了确保评估的公平性，提出了4重避免评估偏差的机制：□ 动态权重适配。根据演练目标动态调整各维度权重，避免固定偏好；□ 基线归一化。将指标与标准场景基线值对比，消除场景固有难度差异；□ 对抗性压力测试。使用元评估智能体尝试暴露评估体系盲区，驱动其迭代优化；□ 多方参照校准。关键指标与行业基准或同行评议结果进行相关性验证。

（2）基于实时数据的动态反馈与个性化引导。基于实时数据的动态反馈与个性化引导是实现从标准化训练向精准化教学转变的核心技术。该机制通过构建感知、分析、决策、执行的闭环系统，能够对学员在模拟对抗中产生的操作行为、策略选择以及最终效果进行持续监测与即时解析[35]。当系统通过实时分析发现学员多次未能有效防御某类特定攻击（例如基于Log4j漏洞的利用或鱼叉式钓鱼攻击）时，其内置的智能诊断模块会精准定位其知识或技能短板，并随即自动触发干预机制[36]。这种干预并非简单的错误提示，而是一个深度集成的教学过程，系统会动态地向学员推送与该漏洞原理或攻击手法相关的背景知识、技术分析文档以及针对性的专项训练场景。这种边练边学、动态适配的个性化引导机制，极大地优化了传统固定课程的教学流程，使得人才培养的效率和精准度均获得有效提升。同时，生成式AI（AIGC）在此过程中扮演了“私人教练”的角色，它不仅能模拟多样化的攻击路径，还能基于对实时训练数据的分析，动态生成高度定制化的攻击场景和修复建议。一个典型的应用是，在事后分析与复盘阶段，系统可以利用

表 1　多维度量化评估指标体系

一级指标	二级指标	主要量化测量项（示例）	客观性依据与数据来源
技术效能(衡量直接影响)	攻击有效性	攻击链完成度、关键资产渗透率、漏洞利用成功率	基于MITRE ATT&CK框架映射，由裁判Agent对系统状态和攻击日志进行自动判定
	防御有效性	威胁检测率、平均响应时间(MTTR)、误报率	遵循NIST SP 800-94等指南，通过对比注入攻击样本与告警日志自动计算
	系统影响	服务可用性下降百分比、业务中断时长	由部署在仿真环境中的监控探针自动采集性能数据计算得出
过程质量(评估执行过程)	资源效率	CPU/内存占用率、网络带宽峰值、扫描冗余流量比	通过容器编排平台监控接口与流量分析工具进行实时自动化采集
	策略特性	攻击路径长度、使用战术技术数量、防御规则变更频率	通过自动化分析日志并映射至ATT&CK矩阵计算，结果可复现
	操作合规	违反安全策略次数、操作日志完备率	由平台裁判Agent根据预定义规则进行自动审计
认知能力(衡量高阶思维)	应对未知	对未知攻击的首次有效响应时间、无脚本策略生成时间	通过“突袭测试”记录时间戳差，场景与训练集隔离
	策略创新	采用非标准方法的比例、策略跨场景迁移成功率	通过对比操作序列与基础策略库的相似度进行量化
	决策解释	（对AI）决策特征贡献度熵值；（对人）有效因果分析条数	使用SHAP/LIME等公认XAI算法计算，或引入专家盲审评分

大语言模型的内生因果推理与自然语言生成能力，自动为学员生成详细的、可解释的评估报告。报告不仅会指出操作中的具体失误，更能深入分析攻击步骤间的因果关联与潜在风险，从而将一次简单的攻防对抗转化为深度的学习体验。

（3）基于LLM增强的可解释性评估报告生成。基于LLM增强的可解释性评估报告生成技术，是实现从结果统计到深度认知决策支持的关键提升。该技术的核心在于利用LLM强大的自然语言生成与因果推理能力，将攻防对抗中产生的海量、多维度的底层性能数据（如漏洞识别率、攻击路径复杂性、平均响应时间等指标），转化为包含过程复盘、原因分析与个性化改进建议的评估报告[36]。这种超越胜负判定的解释能力，极大提升了评估结果的可信度与行动指导价值，使安全人员能够精准定位能力短板。在技术实现路径上，该功能依赖于专门的解释性工具集与可解释性增强方法。例如，IBM开发的ICX360工具包就提供了多种解释LLM生成文本的技术[36]，包括通过扰动输入来分析关键成分的黑盒方法，以及利用梯度信号的白盒方法，以此揭示模型决策所依据的输入部分。同时，诸如自解释性增强（SEER）等方法，通过在LLM的表示空间中聚合相同概念并对不同概念进行解耦，能够同步生成与模型输出相匹配的可信的解释，从而增强其内在的可解释性[37]。此外，将评估体系与MITRE ATT&CK等行业知识框架对齐[38]，可以确保根因分析（RCA）在专业术语和攻击技战术语境上的准确性。最终，系统能据此构建动态的、个性化的学习路径，例如推荐针对性的训练模块、工具使用教程或模拟特定攻击场景等，从而将每一次攻防对抗都转化为一个监测、验证、优化的完整学习过程，这不仅可以加速安全人才培养进程，也为构建安全、可信、持续进化的智能攻防训练环境奠定了坚实基础。

表2从核心功能、技术优势、典型应用场景及关键技术等方面系统性地对比分析了构建智能攻防平台所需的各项关键技术。

2.4　与现有智能平台的比较分析

为清晰界定本文所提方法体系的创新特征，本节选取近年来典型的、已发表的智能网络攻防演练或自动化渗透测试平台作为参照，从核心架

表 2　AI赋能网络攻防平台关键技术对比分析

技术类别	核心功能	技术优势	关键技术	典型应用场景
智能攻击模拟	动态生成高逼真、自进化的攻击行为，以评估并训练防御体系	实现自主演化，模拟复杂、隐蔽的APT攻击链，提供持续逼近现实的对抗压力	多智能体系统，强化学习，LLM	主动安全评估，防御体系验证，红队演练与渗透测试
多智能体系统建模与协同攻击	通过多个智能体的博弈，模拟复杂的协同攻击行为	攻击路径的动态规划，解决场景固化问题，提升攻击的适应性与隐蔽性	感知、决策、行动循环、智能体协同算法	复现完整APT攻击链，多阶段攻击演练
强化学习驱动的攻击策略进化	使攻击智能体通过与环境交互来自主优化攻击策略	从预设到生成的转变，发现并利用未知脆弱点，行为模式与真实APT相似度高	近端策略优化，进化策略，奖励函数设计	探索新型攻击路径，测试防御体系的鲁棒性
基于LLM的攻击脚本生成	将自然语言攻击意图编译为可执行脚本，并提供语义解释	自动化生成多样化攻击样本，提升攻击模拟的覆盖范围与逼真度，提供白盒化的攻击逻辑解释，增强可解释性	代码生成、因果推理、自然语言处理	自动化渗透测试，安全教学与攻击链理解
自适应防御决策	构建具备感知、分析、决策与行动能力的闭环智能响应系统	从静态规则匹配到动态智能响应，对未知威胁的自主应对，跨域协同防御	深度学习，强化学习，人机协同	入侵检测系统，恶意软件检测，网络安全管理与运营
多源融合的实时威胁感知	融合异构数据源，通过深度学习模型构建全面的安全态势画像	突破数据孤岛，通过特征交叉验证，提升检测置信度，对未知攻击具有高识别率、低误报率	CNN，LSTM/Transformer，图神经网络	全网安全态势感知，高级威胁狩猎 (Threat Hunting)
基于强化学习的动态策略生成	使防御智能体学习能够最大化长期安全收益的策略	实现系统可用性与威胁遏制的双目标平衡，避免过度防御，具备持续演进和免疫记忆能力	部分可观测马尔可夫决策过程，奖励函数设计	动态防火墙策略调整，自动化入侵响应
人机协同的智能决策支持	将AI的推演能力与专家的决策权结合，提供可解释的决策指导	降低防御决策门槛，融合人类认知与机器计算优势，通过弹性分工，兼顾效率与可靠性	蒙特卡洛树搜索(MCTS)，贝叶斯推理，自然语言生成	安全事件应急响应，新安全人员培训与指导
自动化评估与反馈	对攻防全过程进行精细化度量、分析与指导，构建评估闭环	从结果统计深化为过程治理，驱动防御策略与人员能力的协同优化，实现个性化教学、加速人才培养	多维度指标体系，实时数据分析，LLM	攻防演练效果评估，网络安全技能认证与培训，安全能力成熟度评估
多维度可量化的评估指标体系	建立涵盖技术、过程、能力的综合评估框架，替代二元胜负判定	解决评估标准缺失问题，客观衡量平台成效与用户实战能力，与业界标准对齐、提升权威性	MITRE ATT&CK框架，AIRTBench、OCCULT等专用基准	红蓝对抗考核，安全产品能力评测
基于实时数据的动态反馈与引导	在训练中实时诊断学员能力短板，并动态调整训练内容	实现边练边学、动态适配的个性化教学，定位知识或技能短板，提升培训效率与精准度	智能诊断模块，AIGC，闭环学习系统	自适应学习路径规划，专项弱点强化训练
LLM增强的可解释性评估报告	将底层数据转化为包含根因分析与改进建议的深度报告	超越胜负判定、提供深度认知决策支持，提升评估结果的可信度与行动指导价值，将单次对抗转化为深度学习体验	可解释性AI工具（如ICX360），因果推理，自然语言生成	攻防演练复盘分析，生成个性化能力提升计划

构、智能体能力、评估维度三个根本性方面进行系统性对比，如表3所示。这些参照系包括：以自动化渗透测试见长的Metasploit平台[39]、强调防御智能体训练的DeepArmor仿真环境[40]以及用于红蓝对抗演练的CybORG框架[41]等。

可以看出，与现有智能平台相比，本文工作的创新之处并非在于孤立地应用了某项AI技术，而在于系统性构建了一个集成攻防评演、支持智

表3　与现有智能攻防平台的主要特征对比

对比内容	现有典型平台 (Metasploit, DeepArmor, CybORG)	本文提出的AI赋能网络攻防平台	主要区别与进步
主要架构	Metasploit：典型的攻击侧渗透测试框架，提供模块化攻击载荷与利用链，但缺乏智能防御与闭环评估 DeepArmor：侧重于防御侧，研究如何利用深度学习加固终端安全，是一个研究原型而非集成平台 CybORG：面向强化学习算法研究的仿真沙盒，定义了标准化的网络攻防环境接口，但本身不提供成熟的攻防智能体或上层应用	分层模块化集成架构（如图1所示），明确分为智能环境构建、Agent管理、数据分析、用户交互4层。攻击、防御、评测、推演等功能以松耦合的Agent形式在统一管理层中实现闭环集成与数据联动	本架构是一个统一的可配置平台，有机融合了攻击模拟、自适应防御和标准化研究环境，并通过模块化设计支持教育、科研、产业等多场景灵活配置
智能体能力	Metasploit：自动化执行预定义的攻击脚本与利用链，可实现流程自动化，但智能体不具备策略学习与进化能力 DeepArmor：其深度学习模型提供了静态的特征识别功能，但不具备在对抗中自主调整策略的能力 CybORG：提供了智能体训练与测试的标准环境，但智能体策略本身需由研究者外部开发与注入	分层协同与安全可控的智能体体系：构建了攻击、防御、裁判三类智能体协同博弈方式，智能体具备从基础技能到复杂策略、再到持续适应新威胁的自主进化能力。创新性引入智能体行为安全管控机制，确保演练安全可控	本平台的智能体是平台内生的、能够通过对抗持续学习与进化的认知实体，它超越了Metasploit的固定脚本、DeepArmor的静态模型以及CybORG的环境，提供了一个智能体生态
评估方式	Metasploit：评估通常为任务是否成功（如获得Shell），缺乏对过程与策略的量化分析 DeepArmor：评估聚焦于模型本身的分类准确率、误报率等机器学习指标 CybORG：评估主要衡量强化学习智能体在任务中的得分与效率，属于算法级评估	多维度、可解释、闭环评估反馈体系：建立了涵盖技术效能、过程质量、认知能力的综合评估指标体系。评估融入训练全过程，实现基于实时数据的动态反馈与个性化引导。利用LLM生成可解释的评估报告，将演练转化为学习体验	本平台的评估体系不仅能够验证攻击是否成功（如Metasploit）或模型是否准确（如DeepArmor），更侧重于量化策略优劣、诊断能力短板、并提供个性化改进路径

能体安全协同、并具备多维度可解释评估能力的统一平台架构与方法体系。这一体系旨在解决现有平台在架构集成性、智能体演进性、评估发展性等方面存在的不足，为构建真正可持续演进、安全可信的下一代网络攻防训练与研究基础设施提供了新的路径。

3　AI赋能网络攻防平台的构建模式与架构设计

3.1　系统架构设计

为构建高逼真、自适应、可扩展的智能攻防实验平台，本文提出分层模块化架构，如图1所示。

（1）智能环境构建层。集成容器化（Docker/

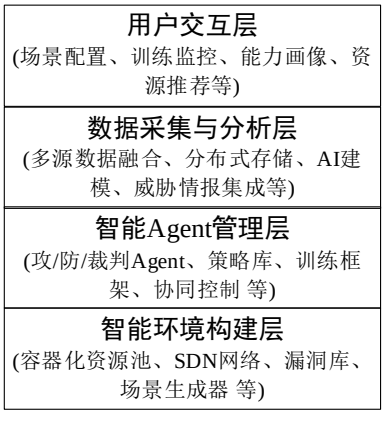

图1　AI赋能网络攻防实验平台系统架构

Kubernetes）与软件定义网络（SDN）等技术，实现拓扑与漏洞环境的动态重构。

（2）智能Agent管理层。作为“核心大脑”，基于多智能体系统（MAS）架构实现攻击、防御、裁判3类Agent的协同博弈与生命周期管理，支持即插即用与自定义策略加载。

（3）数据采集与分析层。通过分布式存储（如Hadoop、Elasticsearch）与Spark计算框架，实现攻防全过程数据的实时采集与深度挖掘，并集成MITRE ATT&CK等威胁情报以保持演进性。

（4）用户交互层。提供Web化界面与角色化视图，支持从场景编排到复盘评估的全流程交互，并通过能力画像与资源推荐实现个性化导学等功能。

3.2 多智能体协同训练方法

为确保持续进化能力，智能攻防平台融合了模仿学习、强化学习、多智能体协同训练及终身学习等方法，构建了一套层次化的训练体系，主要训练方法的核心机制和关键技术与代表性研究如表4所示。这套体系旨在使Agent不仅能掌握基础技能，更能适应动态变化的网络威胁环境。

需要说明的是：这些训练方法并非孤立，而是构成了一个循序渐进的完整体系。模仿学习为智能体提供了可靠的起点，强化学习赋予其自主探索和优化的能力，多智能体协同训练在复杂的博弈中训练其高级策略，终身学习则确保所有能力能够随时间推移而不断进化。同时，为避免智能体在学习新任务时发生灾难性遗忘，平台可采用弹性权重巩固（Elastic Weight Consolidation，EWC）与动态架构扩展相结合的机制。EWC算法通过计算网络参数对旧任务的重要性，在学习新攻击类型时，对重要的旧参数施加“弹性力”，约束其大幅变化，从而将原有攻击策略锚定在参数空间中。同时，对于相关性较低的全新攻击模式，平台允许智能体动态扩展其网络架构的子模块，以实现新旧能力的隔离与共存。通过定期在包含新旧攻击类型的混合场景回放池中进行微调，智能体能够持续巩固并优化其整体策略库，确保新增APT攻击训练后，对原有Web攻击等策略的熟练度保持稳定。

3.3 技术挑战与解决方案

网络攻防平台的建设涉及大量的技术以及不同技术之间的融合，在构建AI赋能的网络攻防平台时，主要面临的挑战和技术应对如下：

（1）环境真实性与资源消耗。为平衡仿真真实性、扩展性与资源成本，平台采用层次化弹性建模与动态资源调度策略。首先，将网络环境划分为高、中、低三个保真度层次，分别对应核心资产、次要节点和背景网络。其次，资源调度器根据实时攻防态势和演练阶段目标，动态调整各层次节点的资源分配与仿真精度：在对抗焦点区域自动提升保真度以确保真实性，在非关键区域则降低消耗。这种按需仿真的策略，辅以容器化部署与混合架构，在保证关键场景真实性的同时实现了资源的集约化利用[46]。

（2）智能体行为的安全性管控。为确保AI智

表4 智能Agent训练方法比较

训练方法	主要机制	关于技术与代表性研究
模仿学习	通过模仿专家示范，快速建立基础行为策略	行为克隆 (Behavior Cloning)，逆强化学习 (Inverse Reinforcement Learning)
强化学习	通过与环境交互和奖励信号，自主优化决策策略	近端策略优化 (PPO)，多目标优化框架 (如MOMA-AC[42])，对抗鲁棒性训练[43]
多智能体协同训练	在对抗或协作中，利用智能体间的相互作用促进整体能力提升	多智能体强化学习 (MARL) [42]，专用多Agent训练框架 (如MarsRL[44])。
终身学习	在不断学习新任务的同时，有效保留对旧知识的记忆	弹性权重巩固 (Elastic Weight Consolidation, EWC) [45]，增量训练 (Incremental Training)

能体在受控范围内进行攻击模拟，需要构建多层次的防护机制[47]。首先是安全沙箱隔离，通过虚拟局域网（VLAN）和容器权限限制，确保攻击行为无法溢出实验环境。其次是行为约束，通过预设的操作白名单和攻击强度阈值，从决策源头规范智能体行为。再者是异常干预，利用异常检测模型实时监控智能体操作，一旦识别到违规行为立即触发暂停和告警。此外，通过建立基于规则优先级和实时推演的冲突仲裁机制，并利用冲突案例驱动评估规则库的持续演进，确保了攻击智能体、防御智能体与裁判智能体在协同过程中的目标一致性与行为协调性。

（3）评估体系的公平性与客观性。一个公平的评估体系需要建立多管齐下、多方协同机制。首先，应根据学员的不同能力等级，匹配不同复杂度的标准化场景，避免出现新手直面专家级攻击的不公平情况。其次，需定期通过基准测试对AI智能体的性能进行校准，确保不同训练批次间的评估基准保持一致。还有，在评估时，可采用类似层次分析法（AHP）的多维度加权评估方法，避免单一指标（如攻击成功率）主导最终结果，从而全面、客观地反映攻防双方的真实能力[48]。

3.4　面向多应用场景的可配置架构

本文提出的分层模块化平台架构（见图1），其本质是一个可根据不同应用场景实现功能调整的可配置框架，能够在不牺牲架构的统一性的基础上，通过调整各层的组件、策略与数据流向，灵活适配教育、科研与产业实践等不同领域的差异化需求。这种可配置架构主要通过以下三个层面实现：

（1）Agent策略库与行为模式的场景化配置。Agent管理层是平台的核心，其内置了可插拔的策略库与可调节的行为参数，以此来实现对不同应用场景的适配。其中，在教育领域，侧重于安全可控与教学引导。在此模式下，攻击Agent主要从“教学案例库”中加载行为，其攻击链的复杂性、隐蔽性和破坏性将被限制在预设范围内，确保实验环境的安全与稳定。防御Agent则更侧重于规则验证与基础响应逻辑。同时，裁判Agent会被赋予更强的教学指导功能，专注于生成详细的、循序渐进的评估报告与操作提示，以辅助学员理解攻防原理；在科研领域，侧重于学术探索与算法验证。平台允许研究人员自定义或导入新的AI模型作为Agent的核心决策引擎，攻击与防御Agent可以有效测试新型算法的有效性。数据采集与分析层会提供更底层、更细粒度的实验过程数据，支持对算法性能的深度分析；在产业实践方面，侧重于高逼真模拟与实战评估。在此模式下，攻击Agent的策略库将紧密集成最新的真实世界威胁情报（如MITRE ATT&CK TTPs），以模拟APT的行为。防御Agent则与企业实际部署的安全产品的API进行对接，演练真实的响应流程。评估体系将可直接评估演练对实际安全运营水平的提升效果。

（2）环境仿真与数据层的差异化构建。智能环境构建层与数据采集分析层可根据场景需求进行调整。其中，在教育与科研领域，可采用轻量级的容器化仿真和开源漏洞库，快速构建标准化或特定拓扑的实验环境，注重可重复性与成本可控；在产业实践领域，则可通过数字孪生技术或由虚拟化与真实设备结合的混合架构构建与企业真实网络高度一致的仿真环境，并导入脱敏后的真实历史流量与日志数据进行训练，强调环境的保真度与针对性。

（3）用户交互与评估反馈的定制化呈现。用户交互层可根据用户需求，提供可视化的交互与评估反馈结果提示。其中，在教育领域，界面可以突出学习过程、知识图谱以及存在的技能短板分析等方面，为教师提供不同个体的能力画像，进而为学生提供个性化练习推荐；在科研领域，界面可提供丰富的算法参数调整、实时训练进程

以及对比实验管理等功能，服务于研究过程的精细控制；而在产业实践方面，界面可以动态场景的视角呈现，聚焦全网实时威胁态势、攻击链还原、团队协作响应效率以及符合行业标准的合规性报告生成等。

可以看出，本平台并非采用死板的一刀切设计，而是通过其内在的模块化、可配置特性以及清晰的管理策略，实现了统一架构下对教育、科研与产业实践等多场景需求的按需适配。

3.5 应用案例与实践成效

AI赋能的网络攻防平台凭借其可配置的架构，已在教育、科研及产业领域取得显著成效，并通过差异化配置精准服务于各场景的核心目标。

在教育领域，平台通过启用教学引导模式和安全约束策略，有效支撑了新型网络安全人才的培养。例如，浙江大学SEED平台集成了AI辅助模块，在受控的教学模式下开展基于深度学习的威胁检测与基于强化学习的攻防对抗实验，帮助学员在安全环境中理解AI在网络安全中的原理与应用[49]。广西科技大学构建的“四元协同”培养模式，则以AI赋能的攻防实验平台为核心支撑，通过调用平台预置的教学案例库与个性化评估反馈机制，整合多智能体对抗、动态场景生成等技术，实现了理论与实践的深度融合[50]。

在科研创新方面，研究者深度利用平台开放的自定义Agent接口和细粒度数据采集功能，加速了前沿技术的探索与验证。例如，厦门市数据安全与区块链技术重点实验室数智安全团队，基于平台灵活的多智能体框架与强化学习训练环境，研发了“基于多智能体协作的Android应用程序漏洞检测系统”与“融合大语言模型与强化学习的多智能体自动化渗透测试系统”，相关成果均获国家发明专利授权[51]。

在产业防护与演练领域，平台通过与企业真实网络拓扑对齐并集成实时威胁情报，实现了高逼真的攻防模拟与实战能力评估。博智安全推出的AI赋能智能攻防推演系统，通过配置产业演练模式，复现针对金融、能源、政府等关键行业的特定攻击链，已服务于多个行业的安全培训、技能考核与应急预案演练，有效提升了组织的主动防御能力[21]。

综合上述案例表明，通过针对性的架构配置，AI赋能平台在人才培养上可缩短技能掌握时间，在科研中能有效推动创新成果转化，在产业应用中可提升威胁检测率，为构建下一代自适应防御体系奠定了坚实的实践基础。

为验证性能提升在不同情境下的稳定性，平台在其模块化架构与自适应机制支持下，展现出了良好的鲁棒性。在教育案例的简单局域网配置与产业案例的复杂混合云拓扑下，平台驱动的威胁检测率得到稳定提升，这表明其智能体决策对网络配置复杂性具备较强的泛化能力。同时，在科研与产业应用中，无论初始防御策略是基于最佳实践（强基线）还是近乎空白（弱基线），防御智能体均能通过与高水准攻击方的对抗实现策略自主进化，最终达到相近的高水平检测率，证明其对初始策略的依赖性较低。此外，平台通过实时集成外部威胁情报与仅依赖内部感知两种模式下的对比显示，其核心检测率波动明显较小，这得益于多源融合感知与强化学习探索的互补机制，降低了对单一时效性情报输入的绝对依赖。为此，平台的性能提升源于其内在的协同进化与持续学习能力，而非对特定理想环境的过拟合，因此在多变的应用情境中保持了稳定的核心效能。

4 挑战与未来发展方向

4.1 技术层面的挑战

（1）模型可解释性与鲁棒性缺失。攻击模拟、威胁检测等核心AI模型（如深度学习、强化学习等）存在“黑箱特性”[52]，决策逻辑透明度

不足，制约教学过程中的原理传导。同时模型鲁棒性薄弱，易受对抗样本攻击（如流量数据扰动导致异常流量分类错误、恶意代码特征修改规避Transformer模型识别），且智能体长期对抗中易出现策略退化，导致训练效果衰减。

（2）环境仿真适配性不足。大规模异构网络环境（如工业互联网、车联网等）仿真面临真实性与扩展性之间的矛盾，纯软件仿真难以复现物理层攻击，与专用设备集成的成本高、扩展性差[53]。对元宇宙、量子网络等新兴场景仿真能力缺失，且大规模演练的资源消耗过高，限制中小机构应用。

（3）数据安全与质量失衡。模型训练与仿真需大量真实攻防数据（如攻击日志、漏洞利用过程等），但此类数据多涉及敏感信息，采集过程中的合规风险高。数据脱敏技术存在过度脱敏降质、脱敏不彻底泄密的双重困境，直接影响模型训练效果。

4.2　伦理与合规性挑战

（1）技术滥用风险。平台高逼真攻击模拟能力可能被恶意利用。为系统性地管控此风险，平台设计并实施了基于角色与场景的权限分级管控体系，核心是对攻击脚本的生成与执行进行严格约束。其中，平台将用户角色与攻击能力解耦，实行三级权限管理：

普通学员/用户：仅能访问和运行经过严格审核的“教学案例库”中的攻击脚本。该库脚本已被剥离真实敏感信息，且其执行被限制在预设的、隔离的标准化实验环境中，无法进行自定义攻击生成或对外部真实目标进行扫描。

专业红队/研究人员：在通过严格资质审核并处于特定“高级研究模式”下，可启用平台的智能攻击生成功能（如LLM生成脚本、强化学习探索）。但其所生成的动作序列必须通过安全策略引擎的实时校验（如禁止对非演练目标IP的扫描、禁止使用未经脱敏的真实漏洞利用代码等），且所有操作均被全链路审计。

平台管理员/裁判：拥有最高权限，负责定义和维护前述各角色的能力边界、审核教学案例库、监督“高级研究模式”的启用与审计日志。

（2）人才培养路径依赖。过度依赖AI辅助易导致学员基础技能退化[54]，形成被动防御思维，且AI模型训练数据偏见可能引发训练不均衡，导致特定领域（如工业控制）学员能力评估失真。

（3）标准化与合规性缺失。缺乏统一的技术标准（如攻击真实性、评估指标等量化规范）与伦理规范，平台建设质量参差不齐，训练成果可比性不足；用户资质审核、操作日志留存等使用规范缺失，易触碰《网络安全法》《数据安全法》等合规性红线[55]。

4.3　未来发展方向

（1）模型可解释性与鲁棒性协同优化。融合可解释AI（XAI）技术（如注意力机制可视化、决策逻辑提取等），实现模型决策过程透明化；采用联邦学习、差分隐私技术增强模型抗对抗能力，构建对抗样本库与持续优化之间的闭环机制，缓解策略退化问题[56]。

（2）数字孪生赋能场景仿真升级。引入数字孪生技术构建物理网络高保真虚拟映射，复现物理层攻击；扩展工业互联网、车联网、元宇宙等多领域场景仿真[57]，结合边缘计算架构优化资源调度，降低大规模演练的资源消耗。

（3）隐私计算驱动数据合规共享。依托联邦学习、同态加密等隐私计算技术，构建跨机构脱敏数据共享平台[58]；建立数据全生命周期管理规范，结合区块链技术实现数据溯源，参考欧盟“数据空间”模式，平衡数据安全与训练需求。

（4）人机协同训练模式创新。构建从AI辅助到人工主导、再到创新提升的三阶训练体系，基础阶段依托AI完成场景生成、威胁预警等基础任务，进阶阶段强化学员主导权，创新阶段鼓励攻

防技术研发；结合知识图谱与智能导师系统，实现个性化精准教学。

（5）标准化与行业协同治理。推动制定技术标准（场景真实性、模型性能、评估指标等量化规范）与伦理规范（使用范围限制、用户资质审核、操作日志留存等），通过高校、企业、科研机构协同，实现平台技术兼容与成果互认。

另外，本平台采用的终身学习（Lifelong Learning）[59-60]与联邦遗忘学习（Federated Unlearning）[61]在目标上存在本质区别。终身学习的核心目标是持续积累与整合知识，在不断接入新任务（攻击/防御技术）时，尽可能保留并优化旧有知识，避免遗忘，其关键挑战是稳定性与可塑性困境。而联邦遗忘学习主要应用于隐私合规场景，其核心目标是根据用户请求，从已训练的联合模型中选择性、可验证地删除特定用户或数据点的贡献信息，其关键挑战是精确擦除。简言之，前者旨在防止非预期的知识流失，后者旨在实现可控的知识移除。两者技术路径不同，但均为构建可持续、可信AI系统的重要研究方向。

5 结论

本文系统论述了人工智能驱动网络攻防平台从静态自动化向动态智能化演进的核心路径与构建方法。通过创新性地融合多智能体协同博弈、大语言模型语义理解及强化学习自主优化等技术，并设计分层可配置的系统架构，本研究提出的平台方案在智能攻击模拟、自适应防御决策与多维度可解释评估等关键环节实现了集成性突破。该平台能够灵活适配教学、科研与产业演练等多重场景，显著提升了攻防训练的逼真度、适应性与评估深度，已成为支撑网络安全人才培养和技术创新的重要基础设施。然而，AI模型自身的可靠性及其应用的治理挑战依然突出，主要体现在决策可解释性不足、数据隐私与安全风险、以及技术可能被滥用的潜在威胁等方面。展望未来，构建可持续发展的智能攻防平台生态，关键在于推进技术创新与治理框架的协同：一方面，需持续攻坚模型可解释性、鲁棒性及仿真保真度等技术瓶颈；另一方面，必须建立融技术标准、伦理规范与安全合规于一体的综合治理体系。网络空间的竞争归根结底是人才的竞争，因此，必须深化技术研发、教育实践与产业应用之间的协同联动，推动平台从一项先进工具向开放、进化的融合性生态升级，从而为构筑坚实可靠的国家网络安全防线提供核心能力支撑。

参考文献：

[1] 王群,李馥娟,郭向民,等.网络靶场实训平台的规划与实践[J].火力与指挥控制,2021,46(07):136-141.
WANG Q,LI F J,GUO X M. Planningand Practiceof Training Platform of Cyber Range [J]. Fire Control & Command Control, 2021,46(07):136-141.

[2] 杨丽,朱凌波,于越明,等.联邦学习与攻防对抗综述[J].信息网络安全,2023,23(12):69-90.
YANG L, ZHU L, YU Y, et al. Review of Federal Learning and Offensive Defensive Confrontation[J]. Netinfo Security, 2023, 23(12): 69-90

[3] 张靖如,方志耕,孙云柯,等.多阶段攻防对抗体系效能评估MS-GERT模型[J].系统工程与电子技术,2025,47(07):2237-2245.
ZHANG J R,FANG Z G,SUN Y K,et al.MS-GERT model for effectiveness evaluation of multi-stage offensive and defensive adversarial system-of-systems[J]. Systems Engineering and Electronics,2025,47(7):2237-2245.

[4] CHEN X, LIU Y, WANG J. Dynamic Attack Simulation for Cyber Security Training Based on Multi-Agent Systems[J]. IEEE Transactions on Education, 2022, 65(3): 217-226.

[5] DU W, WANG H, LIU C. SEED: A Suite of Hands-On Laboratory Exercises for Computer Security Education[J]. ACM Transactions on Computing Education,2020,20(2):1-28.

[6] 朱兆梁,沈建京,郭晓峰,等.基于复杂网络-灰靶理论的网络空间攻防方案评估[J].火力与指挥控制,2022,47(04):90-95+103.
ZHU Z L,SHEN J J,GUO X F,et al.Evaluation of cyberspace attack and defense scheme based on complex network and grey target theory[J].Fire Control & Command Control, 2022,47(4): 90-95.

[7] 博智安全科技股份有限公司.AI攻防演练系统技术白皮书

[R]. 南京：博智安全科技股份有限公司,2024.Bozhi Security Technology Co., Ltd. Technical White Paper on AI Attack-Defense Exercise System [R]. Nanjing: Bozhi Security Technology Co., Ltd., 2024..

[8] 马兆丰,彭海朋,陈秀波,等. 新形势下网络空间安全创新型专业人才培养体系研究[J]. 信息安全研究,2025,11(04): 385-391.
MA Z F, PENG H P, CHEN X B, et al. Innovative and professional talent education architecture of cyberspace security in new situation[J].Journal of Information Security Research,2025, 11(04):385-391.

[9] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[J]. Advances in neural information processing systems,2014,27:1-9.

[10] BROWN T, MANN B, RYDER N, et al. Language models are few-shot learners[J]. Advances in neural information processing systems,2020,33:1877-1901.

[11] 刘艾杉,郭骏,李思民,等. 面向深度强化学习的对抗攻防综述[J]. 计算机学报,2023,46(08):1553-1576.
LIU A S, GUO J, LI S M, et al. A survey on Adversarial Attacks and Defenses for Deep Reinforcement Learning[J]. Chinese Journal of Computers,2023,46(08):1553-1576.

[12] 王立夫,高聪,郭戈,等. 异构多智能体网络拓扑可辨识性[J]. 自动化学报,2025,51(03):559-569.
WANG L F, GAO C, GUO G, et al. Discernibility of Heterogeneous Multi-agent Networks Topology[J].ACTA Automatica Sinica,2025,51(03):559-569.

[13] 张学旺,卢荟,谢昊飞. 基于节点中心性和大模型的漏洞检测数据增强方法[J]. 信息网络安全,2025,25(04):550-563.
ZHANG X W, LU H, XIE H F. A Data Augmentation Method Based on Graph Node Centrality and Large Model for Vulnerability Detection[J]. Netinfo Security, 2025,25(4): 550-563.

[14] MADDIREDDY B R, MADDIREDDY B R. The role of reinforcement learning in dynamic cyber defense strategies[J]. International Journal of Advanced Engineering Technologies and Innovations, 2024, 2(1): 267-292.

[15] SARHAN M, LAYEGHY S, MOUSTAFA, et al. Cyber threat intelligence sharing scheme based on federated learning for network intrusion detection[J]. Journal of Network and Systems Management,2023,31(1):3.

[16] LANKA P, GUPTA K, VAROL C. Intelligent threat detection-AI-driven analysis of honeypot data to counter cyber threats[J]. Electronics,2024,13(13):2465.

[17] NGUYEN T, RNDI N, NETH A. Noise Contrastive Estimation-based Matching Framework for Low-Resource Security Attack Pattern Recognition[J]. arXiv:2401.10337,2024.

[18] JABER A, FRITSCH L. Towards ai-powered cybersecurity attack modeling with simulation tools: Review of attack simulators[C]//International Conference on P2P, Parallel, Grid, Cloud and Internet Computing. Cham: Springer International Publishing,2022:249-257.

[19] SHEN G, ZHAO D, FENG L, et al. PANDAGUARD: Systematic Evaluation of LLM Safety against Jailbreaking Attacks[J]. arXiv preprint arXiv:2505.13862, 2025.

[20] HIRSCHOWITZ E, RAMOS F. Harnessing Bounded-Support Evolution Strategies for Policy Refinement[J]. arXiv preprint arXiv:2511.09923,2025.

[21] YAO Y, DUAN J, XU K, et al. A survey on large language model (llm) security and privacy: The good, the bad, and the ugly[J]. High-Confidence Computing, 2024, 4(2): 100211.

[22] 万维易源. AI 赋能实战网络靶场：构筑智能防御新篇章[EB/OL]. (2025-10-10) [2025-11-11]. https://www. showapi. com/news/article/68e7e60e4ddd79d13511fa1f. Wanwei Yiyuan. AI-Empowered Live-Fire Cyber Ranges: Forging a New Chapter in Intelligent Defense[EB/OL]. (2025-10-10) [2025-11-11]. https://www. showapi. com/news/ article/68e7e60e4ddd79d13511fa1f.

[23] ADAWADKAR A M K, KULKARNI N. Cyber-security and reinforcement learning-a brief survey[J]. Engineering Applications of Artificial Intelligence,2022,114: 105116.

[24] NGUYEN T T, REDDI V J. Deep reinforcement learning for cyber security[J]. IEEE Transactions on Neural Networks and Learning Systems,2021,34(8):3779-3795.

[25] BATES E, HICKS C, MAVROUDIS V. Less is more? Rewards in RL for Cyber Defence[J]. arXiv preprint arXiv:2503.03245, 2025.

[26] YANG Z, FU M, QU H, et al. Incremental model-based reinforcement learning with model constraint[J]. Neural Networks, 2025,185:107245.

[27] ZHANG H, CHEN Z, DENG H, et al. LazyAct: Lazy actor with dynamic state skip based on constrained MDP[J].PloS one, 2025,20(2): e0318778.

[28] PURVES T, KYRIAKOPOULOS K G, JENKINS S, et al. Causally aware reinforcement learning agents for autonomous cyber defence[J]. Knowledge-Based Systems,2024,304:112521.

[29] OH J, FARQUHAR G, KEMAEV I, et al. Discovering state-of-the-art reinforcement learning algorithms[J].Nature,2025:1-2.

[30] REN M, CHEN N, QIU H. Human-machine collaborative decision-making: An evolutionary roadmap based on cognitive intelligence[J]. International Journal of Social Robotics, 2023, 15(7):1101-1114.

[31] WANG Y, LIU J, LIAO X. Preference Construction: A Bayesian Interactive Preference Elicitation Framework Based on Monte

Carlo Tree Search[J]. arXiv preprint arXiv:2503.15150, 2025.

[32] TORI A R, HASAN K F. An Evaluation Framework for Network IDS/IPS Datasets: Leveraging MITRE ATT&CK and Industry Relevance Metrics[J]. arXiv preprint arXiv: 2511.12743, 2025.

Security Geek. AIRTBench: Measuring Autonomous AI Red Teaming Capabilities in Language Models[EB/OL]. (2025-06-26)[2025-11-01].https://www.secrss.com/articles/80190?app=1.

[33] 安全极客.AIRTBench：衡量大语言模型的自主AI红队能力[EB/OL]. (2025-06-26) [2025-11-01]. https://www. secrss. com/articles/80190?app=1.

KOUREMETIS M, DOTTER M, BYRNE A, et al. Occult: Evaluating large language models for offensive cyber operation capabilities[J]. arXiv preprint arXiv:2502.15797, 2025.

[34] POIREAULT K. MITRE: Russian APT28's LameHug, a Pilot for Future AI Cyber-Attacks[EB/OL]. (2025-08-12) [2026-03-02]. https://www. infosecurity-magazine. com/news/mitre-russian-apt28-lamehug/.

[35] 启明星辰.一个MANUS化智能体集群的安全监测体系架构方案:构建“感知-分析-验证-预警”自主闭环[EB/OL].(2025-03-20)[2025-11-02].https://www.venustech. com.cn/new_type/cpdt/20250320/28455.html.

VenusTech. A Security Monitoring Architecture Scheme for MANUS-based Intelligent Agent Clusters: Constructing an Autonomous Closed Loop of "Perception-Analysis-Verification-Early Warning" [EB/OL]. (2025-03-20) [2025-11-02]. https://www.venustech.com.cn/new_type/cpdt/ 20250320/ 28455.html.

[36] WEI D,LUSS R,HU X,et al.ICX360:In-Context eXplainability 360 Toolki[J].arXiv preprint arXiv:2511.10879, 2025.

[37] CHEN G, LIU D, LUO T, et al. Beyond External Monitors: Enhancing Transparency of Large Language Models for Easier Monitoring[J]. arXiv preprint arXiv:2502.05242, 2025.

[38] RASTOGI N, DHANUKA D, SZXENA A, et al. Survey Perspective: The Role of Explainable AI in Threat Intelligence[J]. arXiv preprint arXiv:2503.02065, 2025.

[39] BODKHE S, JADHAV M, WARKAR S. Metasploit for Exploit Automation and Threat Detection on Linux[J]. International Journal of Advanced Research in Science, Communication and Technology，2025,5(9):482-490.

[40] KATIYAR N, TTRIPATHI M S, KUMAR M P, et al. AI and Cyber-Security: Enhancing threat detection and response with machine learning[J]. Educational Administration: Theory and Practice, 2024, 30(4): 6273-6282.

[41] EMERSON H, BATES L, HICKS C, et al. Cyborg++: An enhanced gym for the development of autonomous cyber agents [J]. arXiv preprint arXiv:2410.16324,2024.

[42] CALLAGHAN A, MASON K, MANNION P. MOMA-AC: A preference-driven actor-critic framework for continuous multi-objective multi-agent reinforcement learning[J]. Neurocomputing, 2025: 132032.

[43] CHEN Q, DING K, ZHANG X, et al. Improving robustness by action correction via multi-step maximum risk estimation[J]. Neural Networks, 2025, 184: 107045.

[44] LIU S, DU D, YANG T, et al. MarsRL: Advancing Multi-Agent Reasoning System via Reinforcement Learning with Agentic Pipeline Parallelism[J].arXiv preprint arXiv:2511.11373, 2025.

[45] LIU L, KUANG Z, CHEN Y, et al. Incdet: In defense of elastic weight consolidation for incremental object detection[J]. IEEE transactions on neural networks and learning systems, 2020, 32 (6): 2306-2319.

[46] SAMEH A, SELIM S. Adaptive Dual-Layer Web Application Firewall (ADL-WAF) Leveraging Machine Learning for Enhanced Anomaly and Threat Detection[J].arXiv preprint arXiv: 2511.12643,2025.

[47] 周诣.生成式人工智能驱动的网络安全攻防博弈演化及防御对策研究[J].中国信息界,2025,(07):144-146.

ZHOU Y. A Study on the Evolution of Cyber Attack-Defense Games and Defensive Countermeasures Driven by Generative AI[J].Information China,2025,(07):144-146.

[48] TANG Y, LIU Y, LAN J, et al. Security of LLM-based Agents Regarding Attacks, Defenses, and Applications: A Comprehensive Survey[J]. Information Fusion, 2025: 103941.

[49] SEED Project. SEED User Survey Report 2024[EB/OL]. (2024-06-15)[2025-11-06]. https://seedsecuritylabs.org/.

[50] 广西科技大学.网络空间安全“四元协同”人才培养模式实践报告[R]. 柳州：广西科技大学计算机科学与通信工程学院，2024.Guangxi University of Science and Technology. Practice Report on the "Four-Element Collaboration" Talent Cultivation Model in Cyberspace Security [R]. Liuzhou: School of Computer Science and Telecommunication Engineering, Guangxi University of Science and Technology, 2024.

[51] 厦门市数据安全与区块链技术重点实验室.融合大模型与强化学习的多智能体自动化渗透测试系统 [P]. 中国发明专利: ZL202310567890.1,2024-03-15.

Xiamen Key Laboratory of Data Security and Blockchain Technology. Multi-Agent Automated Penetration Testing System Integrating Large Models and Reinforcement Learning [P]. Chinese Patent: ZL202310567890.1, 2024-03-15.

[52] HASSIJA V, CHAMOLA V, MAHAPATRA A, et al. Interpreting black-box models: a review on explainable artificial intelligence[J].Cognitive Computation,2024,16(1): 45-74.

[53] MILLER E, MINK D, SPELLINGS P, et al. Classifying cyber

ranges: A case-based analysis using the UWF cyber range[J]. Encyclopedia,2025,5(4): 162..

[54] BROWN T B, MANN B, RYDER N, et al. Language Models are Few-Shot Learners[C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver: MIT Press, 2020: 1877-1901.

[55] 中国信息安全测评中心.网络安全实验平台伦理规范（草案）[R].北京：中国信息安全测评中心,2024.China Information Technology Security Evaluation Center. Cybersecurity Experimentation Platform Ethical Code (Draft) [R]. Beijing: China Information Technology Security Evaluation Center, 2024.

[56] LI T, SAHU A K, TALWALKAR A, et al. Federated learning: Challenges, methods, and future directions[J]. IEEE signal processing magazine, 2020, 37(3): 50-60.

[57] TAO F, ZHAGN M, LIU A, et al. Digital Twin in Industry: State-of-the-Art[J]. IEEE Transactions on Industrial Informatics, 2019, 15(4): 2405-2415.

[58] BRAUD A, FROMENTOUX G, RADIER B, et al. The road to European digital sovereignty with Gaia-X and IDSA[J]. IEEE network, 2021, 35(2): 4-5.

[59] CARLSON A, BETTERIDGE J, KISIEL B, et al. Toward an architecture for never-ending language learning[C]//Proceedings of the AAAI conference on artificial intelligence,Atlanta,USA, July 11–15, 2010.USA:AAAI Press,2010, 24(1): 1306-1313.

[60] RUVOLO P, EATON E. ELLA: An efficient lifelong learning algorithm[C]//International conference on machine learning. PMLR,2013: 507-515.

[61] ROMANDINI N, MORA A, MAZZOCCA C, et al. Federated unlearning: A survey on methods, design guidelines, and evaluation metrics[J]. IEEE Transactions on Neural Networks and Learning Systems, 2025,36(7):11697-11717.

[作者简介]

王群（1971-），男,甘肃天水人，博士,教授，硕士生导师，CCF杰出会员(33021D)，主要研究领域为信息安全，计算机网络体系结构与协议。

李馥娟（1974-），女，陕西西安人，硕士，教授，硕士生导师，主要研究领域为计算机网络技术与应用，信息安全。

高光亮（1989-），男，山东临朐人，博士，副教授，硕士生导师，江苏高校“青蓝工程”优秀青年骨干教师培养对象。主要研究领域为公安大数据挖掘，社会网络安全等方向。