多智能体自适应安全体系构建问题研究

Research on Construction of Multi-Agent Adaptive Cybersecurity System

孔令飞[1],孙翊豪[1],沈　元[2](1.中国联通河北分公司,河北 石家庄 050000;2.中国联通数据科学与人工智能研究院,北京 100000)

Kong Lingfei[1],Sun Yihao[1],Shen Yuan[2](1. China Unicom Hebei Branch,Shijiazhuang 050000,China;2.China Unicom Data Science and Artificial Intelligence Research Institute,Beijing 100000,China)

摘　要:

聚焦多智能体自适应安全体系的“鲁棒性—适应性—业务约束动态平衡机制”,针对动态威胁环境下攻防对抗的复杂特征,拆解三维度耦合难题,构建“感知—决策—执行”三层架构,设计分布式协同机制、特征融合算法及动态平衡策略以探讨解决局部感知局限、执行一致性及目标分歧等问题,并展望了工程实践的技术演进方向。

关键词:

多智能体;自适应安全体系;鲁棒性;适应性;动态平衡机制

doi:10.12045/j.issn.1007-3043.2025.09.006

文章编号:1007-3043(2025)09-0032-06

中图分类号:TN915.08

文献标识码:A

开放科学(资源服务)标识码(OSID):

Abstract:

It focuses on the “robustness-adaptability-business constraints dynamic balance mechanism” of multi-agent adaptive security systems. Aiming at the complex characteristics of attack-defense confrontations in dynamic threat environments, it breaks down the three-dimensional coupling problem, builds a three-layer architecture of “perception-decision-making-execution”, designs distributed collaborative mechanisms, feature fusion algorithms, and dynamic balancing strategies to address issues like local perception limitations, execution consistency, and objective divergence, and looks forward to the technological evolution direction of engineering practice.

Keywords:

Multi-Agent systems(MAS);Adaptive security paradigm;Robustness;Adaptability;Dynamic balance mechanism

引用格式:孔令飞,孙翊豪,沈元. 多智能体自适应安全体系构建问题研究[J]. 邮电设计技术,2025(9):32-37.

0 引言

目前,网络威胁并未趋于减少,而是结合新型技术变得更加复杂多样。尽管网络安全防御体系借助最新技术得到显著强化,但攻防双方的态势在未来仍将维持动态平衡。攻击手段的持续变异、威胁路径的动态演化及业务需求的实时变化相互交织,仍是网络安全防御面临的重大挑战。攻击者利用漏洞框架的自动化生成能力与攻击工具的模块化组合特性,持续压缩防御响应时间窗口,导致防御方长期处于“补丁滞后于漏洞、策略落后于变异”的被动状态,应对难度显著提升。在此背景下,基于多智能体技术的自适应安全体系成为研究重点,其动态感知攻防形态并自主调整策略的能力,为有效应对动态攻防问题提供了关键支撑。

收稿日期:2025-08-19

1 研究背景、问题与路径

1.1 研究背景与问题提出

多智能体安全防护体系通过群体智能协同突破传统集中式防御的局限[1]，通过抽象化态势理解实现跨域资源调度[2]。这种架构能解决“物理分散—逻辑集中”的安全管控矛盾，其分布式决策特性与通信网络拓扑分布特征高度适配[3]。通过3个维度的压力模型分析自适应能力适配动态威胁环境的挑战：在时间维度上，攻击技术迭代周期可能呈现指数级缩短的态势，静态策略库的更新速度难以与之匹配；在空间维度上，通信网络异构性持续加剧（如5G网络中同时存在NFV虚拟化环境、传统IP网络、OT专用信道），单一防御策略普适性显著下降；在强度维度上，AI生成式攻击工具使攻击成本呈几何级数降低（如基于大语言模型的漏洞利用代码生成工具可大幅缩短攻击准备时间），传统基于特征匹配的防御机制面临理论失效风险。这3个维度的压力共同要求自适应安全体系突破线性响应模式，强化“感知—决策—执行”的动态协同调整能力。

当前网络空间的攻防对抗呈现3类显著特征：一是攻击面动态扩张，云计算、物联网等场景使防御边界模糊，脆弱点更迭速度大幅提升；二是攻击策略协同化演进，高级持续性威胁（APT）通过多阶段、跨域协同实现攻击，单一节点局部响应难以阻断完整攻击链；三是安全与业务矛盾加深，金融交易、工业控制等关键场景对业务连续性的高要求（如毫秒级时延、99.999%可用性），限制了传统“边界式”防御的应用。这些特征共同指向了核心问题：防御体系如何在动态威胁环境中，同时满足鲁棒性（抗干扰能力）、适应性（环境响应速度）与业务约束（低代价、高可用）要求。

1.2 关键命题与研究目标

现有多智能体安全防护体系在自适应决策上存在理论缺陷，导致在实际部署中出现“防御过松则风险失控，防御过严则业务中断”的困境。第一，通常假设攻击面特征分布为平稳过程，但实际通信网络中攻击面状态满足非齐次马尔可夫特性，导致策略有效性随时间衰减的速率偏离预期；第二，普遍要求智能体间实现全局信息同步，这与通信网络带宽约束、隐私保护等实际条件形成理论冲突，使协同决策在分布式环境下出现逻辑悖论；第三，多数方案从工程角度采用经验化调整参数，未建立调整幅度与系统稳定性的数学映射关系，可能引发防御策略震荡甚至失稳。这些缺陷导致现有体系在动态威胁下的自适应能力有效性和可靠性不足，亟需构建新的理论框架。

本研究聚焦动态威胁环境下多智能体自适应安全体系的“鲁棒性—适应性—业务约束”（Robustness—Adaptability—Business Constraints，RAB）动态平衡机制，探索解决三维度耦合难题的理论路径。研究目标是构建理论体系和技术框架，在多智能体自适应安全体系中实现攻击面演化追踪、应变和业务影响无感化控制，并给出具体实现路径。

1.3 研究思路与创新意义

本研究将RAB问题拆解为3个子问题展开系统研究，其创新点主要体现在以下3个方面。一是提出RAB核心框架，突破既有研究仅在单一维度（如单纯优化鲁棒性或适应性）进行局部优化的局限，实现三维度的协同平衡；二是融合多智能体联邦学习与势博弈理论，通过数学推导证明动态威胁环境下全局最优策略的快速收敛性，为分布式决策提供理论支撑；三是设计场景化平衡调节策略，例如针对通信网络带宽波动、工业互联网时延约束等场景特征，动态调整防御措施的强度与粒度，确保策略适配性。

本研究的理论意义在于丰富动态防御的基础理论体系，为多智能体自适应控制问题提供一套可扩展的分析框架；实践价值则体现在为基础通信网络、工业互联网、云计算与物联网等关键信息基础设施的安全防护提供新的技术路径，助力提升网络空间的整体抗风险能力。

2 相关研究综述

动态防御策略是实现适应性的关键，相关研究主要聚焦于策略空间构建与动态调整机制。早期的策略空间静态建模方法，例如，IBM的移动目标防御（MTD），通过预设策略集合实现防御多样性，但依赖人工经验[4]，缺乏对攻击面演化的动态适配[5]；Lee等人的端口跳变算法[6]，策略切换周期固定，易被破解[7]。Sutton等人基于强化学习方法（RL），将攻防对抗建模为MDP[8]，通过Q-learning求解最优防御策略。但其在复杂场景中存在局限，应考虑业务约束下的多智能体协同存在目标分歧的问题。Olfati-Saber一致性协议提出基于共识算法的协同框架，但其假设智能体目标一致[9]，无法处理现实冲突。联邦强化学习在节点数量过多时通信开销大，决策延迟增加[10]；RL在多智能体场景中不应被视为以同步方式执行原始动作[11]。

鲁棒性是防御系统的基础要求，对抗训练、模型

正则化、集成学习与动态架构调整是提升鲁棒性的关键技术。对抗训练通过添加扰动增强特征学习[12];混合精度训练兼顾容错与资源优化[13];深度模型集成增强数据偏移适应性[14];基于强化学习的架构算法可动态适配环境。然而,现有方法仍存在对抗样本迁移难、动态环境适应不足等问题。虽有研究提出“鲁棒性阈值”的概念,但因阈值无法动态调整,难以实现鲁棒性与适应性的有效平衡。跨场景防御方案的通用性探索,如MITRE的ATT&CK框架标准化攻击战术建模[15],未提供防御策略及动态调整逻辑。

3 多智能体感知层的协同自适应机制

感知层作为多智能体防御系统的“神经末梢网络”,需通过智能体间的协同交互,解决“局部感知局限”“数据异构性”“攻击协同性”等分布式场景特有的问题。

3.1 多智能体感知的分布式协同架构

多智能体的感知层应采用“边缘节点—区域簇头—全局中心”三级架构,通过分工协作实现疑似攻击事件报警消息的汇总和提炼。

3.1.1 智能体角色与协作模式

a)边缘感知节点。部署于网络边缘(如IoT设备、终端主机),负责本地原始数据采集与初步信息提取,受限于资源,应仅运行轻量化模型(如微型CNN),感知范围覆盖单一子网或设备集群。

b)区域簇头节点。由边缘节点选举产生(基于剩余算力与通信质量),负责融合簇内各节点的局部特征,识别跨节点攻击模式(如DDoS攻击的流量协同特征),运行中等复杂度模型(如联邦学习聚合器)。

c)全局中心节点。部署于云端,汇聚各簇头的融合特征,识别跨区域攻击链(如APT攻击的多阶段特征),运行复杂模型(如图神经网络分析攻击传播路径)。

协作模式采用“局部决策—簇内协商—全局校准”的递进式策略。例如,边缘节点保留一定比例的本地感知决策权(如即时阻断单节点攻击),仅将跨节点信息(如异常流量的时空关联性)上传至“簇头”(即结合具体场景定义的有限小范围内的中心节点)。簇头节点在较小的时间切片内与簇内节点同步一次感知结果,通过一致性算法消除局部偏差。全局中心则在稍大的时间切片内(通常与簇头节点的较小时间片成线性倍率,具体参数可通过实验确定)向各簇头推送全局攻击态势,校准区域感知的片面性。

3.1.2 感知任务的动态分配机制

基于智能体能力与攻击特征的匹配度,动态分配感知任务以提升效率。

a)能力向量。定义智能体感知能力$c_i=(c_{i1},c_{i2},c_{i3})$,其中$c_{i1}$为特征提取精度,$c_{i2}$为抗噪声能力,$c_{i3}$为通信延迟。

b)任务要求。定义攻击感知任务需求$t_j=(t_{j1},t_{j2},t_{j3})$,其中$t_{j1}$为特征精度要求,$t_{j2}$为抗噪声要求,$t_{j3}$为实时性要求。

c)匹配度计算。$\mathrm{score}(i,j)=\left(c_i\times t_j\right)\Big/\left(\left\|c_i\right\|\left\|t_j\right\|\right)$,通过匈牙利算法实现智能体的最优匹配的任务。攻击感知任务将会被优先分配给抗干扰能力强的边缘节点,跨区域攻击感知任务将被分配给通信延迟低的簇头节点,以提升整体感知效能。

3.2 多智能体特征融合的协同算法

单一智能体的感知存在局限性(如边缘节点无法获取全局流量),可通过特征融合实现信息互补,解决“局部感知盲区”问题。

3.2.1 异构特征的协同机制

不同类型智能体(如工业传感器与IT服务器)的感知数据存在异构性(如协议格式、特征维度差异),可通过以下方法实现规范化。

a)特征标准化。采用最大—最小归一化将异构特征映射至同一空间[0,1],消除量纲差异。

b)元特征提取。定义跨类型通用元特征(如“异常持续时间”“流量波动频率”),作为异构特征的统一交互载体。

c)联邦字典学习。各智能体在本地训练特征字典,仅上传字典参数至簇头,通过字典对齐实现特征语义统一,避免原始数据泄露。

3.2.2 对抗性场景的协同机制

需设计多智能体协同验证机制识别对抗性攻击行为,以防止攻击者通过欺骗单一感知节点实施定向攻击(如伪造边缘节点的感知数据)。

a)交叉验证规则。对于关键攻击特征,需多个不同位置的智能体独立感知并达成一致(高置信度条件),否则标记为“可疑特征”。

b)时空一致性校验。结合智能体的物理位置与攻击传播速度,验证特征出现的时空合理性。例如,恶意载荷从A节点传播至B节点的时间应≥2个节点

的距离/传播速率。

c）权重动态调整。根据智能体的历史可信度r_i分配投票权重，可信度小于阈值的节点投票权重予以调降。可信度r_i可以由统计（或基线学习）得到的正确感知次数与总感知次数之比确定。

3.3 感知误差的分布式协同抑制

单节点感知误差的传导会因智能体协同而放大（如一个节点的误判可能误导整个簇的决策），需建立分布式误差抑制机制。防御措施与实时威胁的匹配本质上是策略空间与攻击空间的映射问题，其偏差源于两者演化的不同步与特征提取的不完整。构建多智能体偏差传播网络，如式（1）所示。

$$\varepsilon_d^{(g)}(t)=\sum_{i-1}^{n}\omega_i\left(k_i\varepsilon_s^{(i)}(t)+\delta_i\varepsilon_d^{(t)}(t-1)\right) \quad (1)$$

其中，$\varepsilon_d^{(g)}(t)$为全局决策偏差，ω_i为智能体i的权重（与可信度r_i正相关），$\varepsilon_s^{(i)}(t)$为智能体i的局部感知偏差，k_i、δ_i为其偏差放大系数与惯性系数。

局部感知偏差可分解为演化延迟偏差（源于防御措施更新滞后于攻击特征变化）、特征失配偏差（源于攻击特征感知误差）和映射偏差（源于系统设计中最优映射函数的近似或累进误差）三类分量（但三类分量之和可能不小于局部感知偏差）。防御效能应可量化为与偏差有关的幂函数并且与智能体的层级数直接相关。进一步推论则可证明分布式场景下进行误差协同抑制的必要性，可改进一致性算法，建立感知偏差的协同抑制。一种改进步骤如下。

a）由各智能体i在本地计算感知偏差$\varepsilon_s^{(i)}(t)$广播至邻居节点。

b）采用加权平均共识$\varepsilon_s^{(i)}(t+1)=\sum_j a_{ij}\varepsilon_s^{(j)}(t)$，其中$a_{ij}$为信任系数，$a_{ij}\propto r_i r_j/d_{ij}$（$d_{ij}$为通信距离）。

c）若$\varepsilon_s^{(i)}(t)$与共识值偏差超过阈值，则触发本地感知模型重训练。

3.4 多智能体感知的资源协同优化

多智能体感知的协同需消耗通信与计算资源，应避免协同通过开销抵消效能增益的问题。在通信开销的动态控制方面，采用“事件触发通信”替代周期性通信的设计，仅当局部特征变化量超过阈值时，智能体才向簇头发送更新信息，否则保持静默。在智能体计算资源的协同分配方面，采用基于优先级的动态分配算法。例如，高优先级任务采用本地计算优先模式；低优先级任务采用闲时计算模式并且预留一定比例的能力作为弹性资源，由簇头统一调度（例如，临时分配给遭受攻击的节点）。

4 多智能体执行层的协同控制机制

执行层作为多智能体防御系统的“行动终端”，承担将决策层策略转化为具体防御动作的核心功能。与单一节点执行不同，多智能体执行需解决“分布式动作协同”“跨节点执行一致性”“执行效果闭环反馈”等问题。因此，需通过智能体间的动作协调与偏差修正，确保全局防御策略的可靠执行。

4.1 多智能体执行的分布式协同架构

应构建“全局指令分解—局部动作协同—执行效果聚合”的三层架构，实现分布式场景下的动作一致性与效率平衡。

4.1.1 执行智能体的角色分工

a）终端执行节点。部署于网络边缘（例如防火墙、入侵防御设备、终端加固工具），负责具体防御动作的执行（例如端口封堵、进程终止、固件修复）并具备生成局部策略的最小权限。

b）区域协调节点。由终端执行节点选举产生（基于执行成功率与通信质量），负责区域内执行动作的协同（如避免重复封堵同一端口）并处理局部执行冲突。每个区域协调节点仅存储本域及相邻域的执行状态，避免全局信息集中存储。

c）域间协同节点。由区域协调节点动态轮值担任（基于时间窗口或其他判定要素进行轮换），负责跨域执行指令的转发与冲突仲裁，不具备全局控制权限且仅保留本轮次的临时协调权限。

4.1.2 决策指令的分布式分解机制

决策层输出的全局策略仅包含目标约束，不涉及具体执行指令。区域协调节点通过“目标分解博弈”自主协商子目标，博弈收益函数与区域防御能力、攻击强度正相关，确保了子目标分配的公平性与可行性。子目标确定后，由区域节点各自分解为终端执行指令，分解过程通过智能合约验证是否符合全局目标约束，避免局部指令偏离整体策略。

4.2 执行动作的协同一致性控制

考虑拜占庭容错协议建立跨域时序同步机制。各区域节点应具备仅通过交换时间戳与签名信息便可实现同步的能力。应避免对绝对时间源的依赖，当局部网络与外部时间源断开时，仍能保持域内节点的相对同步。

5 多智能体决策层的动态平衡机制

决策层作为多智能体自适应安全体系的“大脑”，需在分布式架构下实现目标协同、资源调度与策略进化的动态平衡。与集中式决策不同，多智能体决策面临“局部信息不对称”“目标分歧”“资源竞争”等特有挑战，需通过博弈论与分布式优化的融合，构建兼顾全局效能与局部利益的决策机制。

5.1 环境—策略的动态映射模型

将攻击面演化特征转化为最优防御策略空间，解决“环境变化—策略调整”的匹配问题。

5.1.1 攻击面演化的特征量化与状态建模

构建攻击面状态向量$A(t)$，该向量包含3类核心特征。

a）攻击强度特征$a_1(t)=(a_{11},a_{12})$，其中a_{11}为攻击流量强度（归一化至[0,1]），a_{12}为攻击工具复杂度（基于漏洞利用链长度量化）。

b）攻击扩散特征$a_2(t)=(a_{21},a_{22})$，其中a_{21}为失效节点比例，a_{22}为跨域传播速度（节点/s）。

c）攻击目标特征$a_3(t)=(a_{31},\cdots,a_{3n})$，其中$a_{3i}$为针对第$i$类核心资产的攻击概率。

通过隐马尔可夫模型（HMM）对$A(t)$进行状态聚类，定义5种典型攻击状态：S_0（正常）、S_1（试探或低强度攻击）、S_2（单点攻击）、S_3（区域扩散）以及S_4（全局爆发）。状态转移概率矩阵$P(S_i\rightarrow S_j)$，可通过历史攻击数据训练得到。

5.1.2 防御策略空间的分布式构建

防御策略空间D由各智能体的局部策略d_i协同构成，包括基准策略和自适应调节2部分。其中，基准策略库包括m类通用防御策略（如端口封堵、流量清洗、固件更新等），每类策略通过参数化描述以保证其灵活性。自适应调节由局部策略向量$d_i=(d_{i1},\cdots,d_{im})$通过分布式张量分解$D(t)=\bigoplus_{i-1}^{n}d_i(t)$提供，其中$d_{ik}$表示智能体$i$对第$k$类基准策略的执行强度。

5.1.3 映射关系的动态学习与优化

可以采用深度Q网络（DQN）与联邦学习结合的方式，训练环境—策略映射函数。局部训练由每个区域协调节点基于本地攻击状态与策略效果数据完成。域间协同节点负责收集各节点的模型参数，通过加权平均（权重与节点决策准确率正相关）聚合为全局模型，再下发至各节点微调。周期性触发模型迭代，采用增量学习方式避免灾难性遗忘。

5.2 多智能体协同的决策的博弈机制

多智能体在决策过程中存在天然的目标分歧（如某节点优先保障自身业务，另一节点侧重全局防御），需通过博弈机制实现目标协同，平衡局部利益与全局效能。

5.2.1 目标协同的势博弈模型

将多智能体决策过程建模为势博弈，确保存在纯策略纳什均衡（PNE），使协同决策收敛至稳定解。

a）局中人集合。$N=\{1,2,\cdots n\}$（区域协调节点集合）。

b）策略空间。每个局中人i的策略集为D_i（局部策略向量空间）。

c）效用函数。$u_i\left(d_i,d_{-i}\right)=\alpha g(D)+(1-\alpha)h_i\left(d_i\right)$，其中$g(D)$为全局防御效能，$h_i\left(d_i\right)$为智能体$i$的局部收益，$\alpha\in[0,1]$为全局—局部权重（动态调节）。

d）势函数。$\Phi(D)=\sum_{i-1}^{n}u_i\left(d_i,d_{-i}\right)$，可证明该博弈是精确势博弈（Exact Potential Game），即$\forall i,\Delta u_i=\Delta\Phi$，保证PNE存在且可通过局部改进收敛。

5.2.2 分布式优化的一致性算法

为求解博弈均衡，采用分布式交替方向乘子法（ADMM），使各智能体在仅交换局部信息的情况下达成全局一致。

a）全局优化目标。求解全局最大化约束收益$\min_D J(D)=\sum\left[\lambda\left(1-g(D)\right)+(1-\lambda)\left(1-h_i\left(d_i\right)\right)\right]$，其中$\lambda$为全局效能权重。

b）局部子问题。每个智能体i求解局部最优，采用中位数提高抗干扰性。

5.2.3 动态联盟的形成与解体机制

针对跨域攻击的协同防御需求，设计多智能体动态联盟机制，实现临时协同决策。当跨域传播速度快时触发联盟条件，由受影响区域的节点发起联盟邀请。基于节点大于阈值的历史协同效率选择联盟成员，并且控制在一个较小的规模，以避免决策效率下降。当攻击扩散特征或联盟决策误差超过阈值时，自动解体并释放资源。阈值由迭代确定。

5.3 鲁棒性与适应性的平衡机制

过度强调鲁棒性会导致策略僵化（无法应对新攻击），过度追求适应性则会增加系统波动（防御效果不稳定），需通过动态调节，实现二者的平衡。

5.3.1 鲁棒性与适应性的量化指标体系

定义2类核心评估指标，分别为鲁棒性指标R和适应性指标A，为平衡决策提供依据。其中，$R=\alpha R_1+(1-\alpha)R_2$，$R_1$为攻击耐受度指标，由攻击成功率和攻击强度决定；$R_2$为策略稳定性指标，由策略执行的波动率决定。类似地，$A=\beta A_1+(1-\beta)A_2$，A_1为环境响应速度，由策略调整延迟和攻击特征变化速度决定；A_2为新攻击覆盖度，由识别的新攻击类型占比决定。

5.3.2 动态权重调节的平衡算法

基于攻击面状态向量$A(t)$，动态调整鲁棒性与适应性的决策权重。

a）权重函数：$\omega_R(t)=\sigma\left(\omega A(t)+b\right)$，其中$\sigma$是S型函数，$\omega$为权重向量（可通过强化学习训练得到）。

b）平衡决策的目标函数：$\max_{D(t)}\left(\omega_R(t)R+\omega_A(t)A\right)$，需同时满足业务约束。

当遭遇超出预期的极端攻击时，触发策略降级以保证核心功能。例如，将降级层级设计为三层：一级降级时保留绝大多数核心决策功能，关闭非必要功能；二级降级时保留部分核心功能，激活预定义应急策略；三级降级时启动最小防护模式。当攻击缓解后再以适当的间隔逐步逆序恢复。

6 结论与未来工作

本研究围绕多智能体自适应安全体系的RAB关键命题，通过理论建模和机制设计，创新性地提出一种分布式自适应安全体系架构，各层通过多智能体协同感知与特征融合、势博弈与联邦学习融合机制、动态轮值协同与闭环反馈提升效能，形成无中心节点的全局协同自适应闭环。本研究在动态平衡机制方面取得理论创新，建立量化平衡框架，提出攻击面动态建模、动态权重平衡算法及极端场景策略降级机制。此外，本研究丰富了动态防御基础理论，为复杂网络系统主动防御提供可扩展分析工具与方法论。尽管理论研究取得了阶段性成果，但在极端环境适应性、长期演化稳定性和跨领域普适性方面仍存在局限性，需进一步深化实用场景验证。未来研究将强化极端环境适应性，拓展跨领域普适性，旨在实现“自组织、自进化、自平衡”的智能防御生态。

参考文献：

[1] RASHID T, SAMVELYAN M, DE WITT C S, et al. QMIX: monotonic value function factorisation for deep multi-agent reinforcement learning[EB/OL]. [2025-06-06]. https://arxiv.org/abs/1803.11485v2.

[2] CHAKRABARTY P K. Swarm intelligence in SaaS ecosystems: multi-agent coordination for threat neutralization [J]. International Journal of Computer Engineering and Technology (IJCET), 2025, 16 (3): 135-149.

[3] BABAHAJI M, FIROUZMAND E, AGHDAM A, et al. Consensus control of multi-agent systems with uncertain communication links [C]//2022 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE). Winnipeg: IEEE, 2022: 93-98.

[4] 陈子涵，程光. 基于Stackelberg-Markov非对等三方博弈模型的移动目标防御技术[J]. 计算机学报，2020，43(3)：512-525.

[5] CONNELL W, MENASCÉ D A, ALBANESE M. Performance modeling of moving target defenses with reconfiguration limits [J]. IEEE Transactions on Dependable and Secure Computing, 2021, 18 (1): 205-219.

[6] LEE H C J, THING V L L. Port hopping for resilient networks [C]// IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. Los Angeles: IEEE, 2004: 3291-3295.

[7] 范晓诗，李成海，王昊. 基于可变时隙与动态同步的端口跳变技术研究[J]. 计算机工程与设计，2013，34(10)：3465-3469.

[8] SUTTON R S, PRECUP D, SINGH S. Between MDPs and semi-MDPs: a framework for temporal abstraction in reinforcement learning [J]. Artificial Intelligence, 1999, 112(1/2): 181-211.

[9] OLFATI-SABER R, MURRAY R M. Consensus problems in networks of agents with switching topology and time-delays [J]. IEEE Transactions on Automatic Control, 2004, 49(9): 1520-1533.

[10] CHANG H H, SONG Y F, DOAN T T, et al. Federated multi-agent deep reinforcement learning (Fed-MADRL) for dynamic spectrum access [J]. IEEE transactions on wireless communications, 2023, 22 (8): 5337-5348.

[11] XIAO Y C, TAN W H, HOFFMAN J, et al. Asynchronous multi-agent deep reinforcement learning under partial observability [J]. The International Journal of Robotics Research, 2025, 44(8): 1257-1286.

[12] MADRY A, MAKELOV A, SCHMIDT L, et al. Towards deep learning models resistant to adversarial attacks [EB/OL]. [2025-06-19]. https://arxiv.org/pdf/1706.06083.

[13] 隋晨红，王奥，周圣文，等. 面向鲁棒学习的对抗训练技术综述[J]. 中国图像图形学报，2023，28(12)：3629-3650.

[14] KARIYAPPA S, QURESHI M K. Improving adversarial robustness via promoting ensemble diversity [EB/OL]. [2025-01-25]. https://arxiv.org/abs/1901.09981.

[15] MITRE. Defend[EB/OL]. [2025-02-03]. https://d3fend.mitre.org/.

作者简介：
孔令飞，高级工程师，学士，主要从事网络空间安全架构设计工作；孙翊豪，助理工程师，硕士，主要从事数据安全管理工作；沈元，工程师，博士，主要从事人工智能研究工作。