· 专题一：双清论坛“智能算网的基础理论与核心技术”
DOI：10.16262/j.cnki.1000-8217.20250428.002

面向多模态网络的威胁防御技术*

任 奎[1,2**] 王荐芃[1,2] 杨文元[3] 王鑫磊[1,2] 卓书果[1,2] 巴钟杰[1,2]

1. 浙江大学 区块链与数据安全全国重点实验室，杭州 310027
2. 浙江大学 网络空间安全学院，杭州 310027
3. 中山大学 网络空间安全学院，深圳 518107

[摘 要] 在虚拟现实、工业互联网、智慧城市等新兴网络服务的驱动下，算力网络、传感器网络、软件定义网络、SCION 网络等新兴架构不断涌现，推动网络走向多模态共存共管的新格局。在全流程可自定义的多模态网络架构下，威胁防御机制也被深度植入网络框架，使得安全策略可随业务需求的变化进行动态部署。本文主要从威胁检测和威胁应对两个方面出发，以技术发展脉络为主线，梳理现有的网络威胁防御技术。在回顾已有威胁防御手段的设计思想与技术路线的同时，探讨在多模态网络的环境下，各类技术所面临的局限性。随着多模态网络技术的进一步普及，研究与多模态网络环境相适应的威胁防御技术迫在眉睫。亟需探索大模型、内生安全、软件定义网络等新兴技术，推动威胁防御技术体系的发展，为多模态网络当中的各类业务提供坚实的网络基座。

[关键词] 多模态网络；流量检测；主动防御；人工智能；跨模态攻击；内生安全

随着虚拟现实、工业互联网、智慧城市等新兴业务形态的迅猛发展，上层应用对网络服务高可靠、低时延、高带宽等方面的个性化需求日益凸显。网络系统需要根据业务形态的需求侧重，针对性地提供个性化网络服务。然而，传统 IP 网络因其固定的网络结构、单一的路由方式以及“尽力而为”的设计理念，已越来越难以满足新兴技术的个性化需求。

为缓解网络上层应用与底层架构的矛盾，研究人员正积极探索软件定义网络[1]、传感器网络[2]、算力网络[3]等新兴网络技术，旨在构建一个全维可定义的多模态网络支撑环境，以实现异构接入设备、多样化交互协议及多模态路由方式的和谐共存。郑州已联合三大运营商，接入网络真实用户与现网流量，成为多模态网络环境现网落地的试点城市。内生安全多模态网络控制器也已应用于多模态网络环境，为网络管理控制提供关键支撑。多模态网络不仅在传统的数据转发与路由方面表现优异，还具备了动态适应和实时决策的能力。利用全维可自定义的网络硬件，可以将网络功能与专用硬件解耦，实现网络拓扑结构的动态变化与资源重构。相比于以往需要依赖外挂式专用安全设备的网络架构，多模态网络硬件能够直接承载多样化的安全功能，为动态防御策略的高效部署提供技术保障[4]。

然而，多模态网络在满足个性化业务需求的同时，也给网络安全技术带来了前所未有的挑战。接入设备、通信协议、交互方式等要素从单一同构向动态异构的转变，极大地增加了网络环境的复杂性，造成更加多样化且难以防范的网络安全威胁。现有威胁防御方法在未知威胁覆盖能力、跨模态泛化能力、感知与应变能力等方面均存在明显不足，难以适应

收稿日期：2024-10-30；修回日期：2025-03-16
* 本文根据国家自然科学基金委员会第 371 期“双清论坛”讨论的内容整理。
** 通信作者，Email：kuiren@zju.edu.cn
本文受到科技部国家重点研发计划专项（2023YFB2904000，2023YFB2904001）和国家自然科学基金项目（U23A20306）的资助。

引用格式：任奎，王荐芃，杨文元，等. 面向多模态网络的威胁防御技术. 中国科学基金，2025，39(2)：229—239.
Ren K，Wang JP，Yang TW，et al. Threat defense technology for polymorphic networks. Bulletin of National Natural Science Foundation of China，2025，39(2)：229-239. (in Chinese)

多模态网络当中复杂动态的网络环境。

因此，如何有效应对网络环境变化所带来的新型安全风险，同时发挥其安全效益，是多模态网络技术在发展过程中所必须面对的重要课题。亟需探索与多模态网络架构相适应的新型威胁防御技术体系，以确保多模态网络技术的稳健发展，为新兴业务形态提供坚实可靠的网络支撑。

1 多模态网络安全的新挑战

多模态网络代表着多种标识体系、交互协议、路由方式共管共治的新一代网络架构[5]。相比于传统单一的 IP 网络，多模态网络中接入设备种类繁多、通信协议彼此不同、网络环境复杂多变。现有的网络威胁防御技术在面对多模态网络环境时存在较大局限，具体可以表现为以下几方面：

(1) 海量未知威胁难以覆盖

现有的网络威胁防御措施大多根据已知威胁的特征建立起专家知识库，通过匹配网络流量与知识库中的威胁特征，进而针对性防御特定威胁。然而在多模态网络当中，网络接入设备、通信协议、访问方式等要素均呈现出明显的多元化趋势。多种网络模态所引入的多元异构软硬件将带来众多未知威胁形式。传统威胁知识库的建立主要依靠安全人员对威胁特征的认知或高质量的威胁样本，其扩充速度与迅速增长的未知威胁数量不相适应，导致海量未知威胁难以有效处理。

(2) 跨模态泛化能力不足

现有的威胁流量检测方案大多依赖协议相关的报文头信息，以及通信双方交互过程中所产生的时间戳、报文长度等侧信道信息。然而，在多模态网络中，每种网络模态对应不同的通信协议。这使得即便是同一种威胁，在不同网络模态报文当中的呈现形式也具有明显差异。威胁检测方法需要具有跨模态泛化能力，才能更准确地解析多模态网络中异构报文的流量特征。已有威胁检测方案往往根据特定网络协议的语义或交互行为进行特征提取，在应对异构呈现形式的网络流量时难以有效迁移威胁特征，在跨模态情景下泛化能力受限。

(3) 感知与应变能力缺乏

现有的威胁防御方法往往仅依据单一网络模态的专家知识，收集当前网络模态下的网络情报信息，搭建相对静态的防御机制。然而，在多模态网络环境当中，标识信息、连接关系、访问方式等因素会随着设备在不同网络模态中的切换而不断重构。这就要求威胁防御方案在感知多种模态网络环境变化的同时，能够自适应调整防御策略。已有的威胁防御方法大多缺乏对多模态环境的感知和适应能力，难以在不断重构的网络环境中及时调整安全策略，无法保障防御效果。

(4) 跨模态攻击难以防范

相比于传统的网络架构，多模态网络当中的不同网络模态之间协议规范有所不同。若攻击者构造跨协议边界的恶意数据包，协议转换网关可能无法正确解析，导致其行为与预期有所偏差。另一方面，当网络信息流在不同网络模态间动态切换时，原有的安全上下文（如会话密钥、访问控制策略等）可能无法正确跨模态继承。攻击者可能在网络模态切换期间，进行跨模态会话劫持。已有的威胁防御方法仅能关注单网络模态下的安全问题，对于跨模态攻击的防范能力明显不足。

现有威胁防御方法在未知威胁覆盖率、跨模态泛化能力、感知与应变能力、跨模态攻击防范能力等方面存在诸多局限，难以应对各种新型安全挑战。建立与多模态网络相适应的新一代威胁防御技术体系，正成为网络技术研究当中的一项关键性问题。

2 网络威胁防御技术发展现状

网络威胁防御技术大致可分为威胁检测技术和威胁应对技术两个方面。为了有效防御网络威胁，一方面要及时准确地检测网络流量当中潜在的威胁隐患，另一方面要高效灵活地对攻击行为做出响应。本章将以图 1 中所展示的发展脉络为主线，分别梳理威胁检测技术和威胁应对技术的发展现状，阐述各种技术路线的优势及局限性，探讨现有技术在多模态网络当中的应用难点。

2.1 威胁检测技术

网络威胁检测技术是一种关键的网络安全手段，旨在通过监控和分析网络中的数据包识别异常流量，从而发现潜在的攻击行为或恶意活动。近年来，网络威胁检测技术从规则驱动为主转向数据驱动为主，智能化程度不断加深。本节主要梳理网络威胁检测技术的发展沿革，并总结每种技术路线的基本思想与典型实例。

2.1.1 规则驱动的检测技术

在网络威胁检测技术的发展初期，研究人员多采用一系列固定规则来审查数据包内容，这种方法也被称为深度包检测（Deep Packet Inspection）。Fernandes 等[6]构建了一个包含 60 款应用流量特征

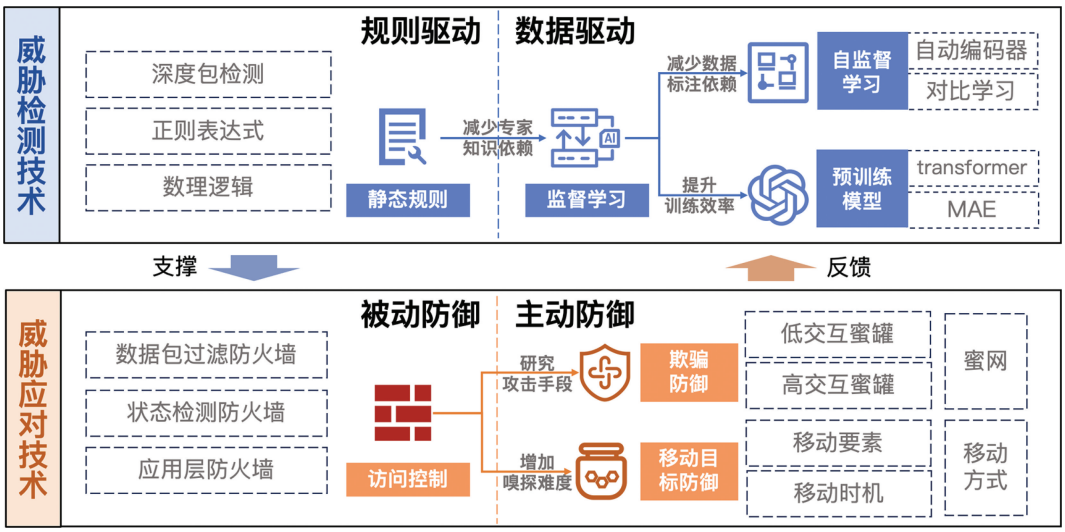

图 1　网络威胁防御技术发展脉络

Fig. 1　The Development Trend of Network Threat Defense Technology

的规则匹配数据库，用于实现流量分类。Keralapura 等[7]构建了一种基于规则的多阶段分类算法，以识别高风险 P2P 流量。

虽然基于规则的检测方法对于已知威胁有一定效果，但检测规则的构建非常依赖于安全人员的专业知识。从一种新的攻击的出现到编写出有效的检测规则，常常存在数以月计的时延[8]，从而给攻击者可乘之机。与此同时，越来越多的网络应用出于对用户隐私的考量，采用加密技术来加密数据载荷[9]。对于这些加密流量，主要依赖数据包载荷中明文信息的深度包检测方法近乎无能为力。综上所述，规则驱动的检测方法在威胁日趋复杂多样且加密流量为主的网络环境中越来越力不从心，其种种缺陷促使着威胁检测技术从规则驱动到数据驱动的转变。

2.1.2　基于监督学习的检测技术

为了有效处理加密流量，并解决规则驱动的流量检测方法灵活性差、依赖专家知识、对于新攻击手段响应不及时等问题，研究人员开始将随机森林、卷积神经网络、图神经网络等人工智能技术引入威胁检测领域。利用人工智能算法的特征提取能力，安全人员得以在不解密数据包载荷内容的情况下，非侵入性地分析网络交互行为，识别潜在风险。

随机森林是网络威胁检测当中的一种常用算法。它通过构建多个决策树，并将这些决策树的预测结果进行组合，从而提高模型的准确性。Hayes 等[10]和 Taylor 等[11]利用随机森林技术，在网络连接加密的情况下，分别识别用户当前浏览的网站和安卓 APP。

卷积神经网络也是一种威胁检测技术当中常用的人工智能算法。它通过多次进行卷积运算，提取不同层次的特征，来完成流量分析任务。Xi 等[12]通过分析挖矿流量的数据包特点，构建用于训练卷积神经网络的鲁棒性特征，能比相关商业应用的入侵检测系统提前数天发现挖矿行为。中国移动公司[13]开发了基于卷积神经网络和多头注意力的流量检测模型，可在 5G 网络和物联网环境中进行流量扫描，实时检测 DDoS 攻击、端口扫描等异常行为。中国电信公司[14]融合了小波变换和深度学习技术，可在金融交易和工业互联网场景中实时监测数据流，提升异常流量监测的准确性。

图神经网络是一种适合处理结构多变数据的人工智能模型。它能将图结构数据转化为统一的矩阵形式数据，以方便输入到神经网络中进行处理。Zhang 等[15]利用图神经网络技术，对一个数据包的头部和载荷分别进行字节级别的特征提取，并通过门控机制将这两个特征进行融合，提取出信息损失更小、更加细粒度的特征用于威胁检测任务。

综上所述，人工智能技术在多种网络流量分类任务中大幅领先于规则驱动的方法，将许多之前难以完成的任务变为可能。然而，这些监督式人工智能算法必须使用大量人工标注的高质量数据集进行训练，在实际应用当中成本较高。另一方面，监督式人工智能算法虽然不再需要构建完备的规则，但特征提取过程还未能摆脱人工介入。很可能由于忽略一些关键性特征，导致模型训练效果无法达到最优。为此，研究人员把目光转向不依赖于人工标注数据的自监督式和无监督式算法，以期进一步提升威胁检测任务的自动化程度。

2.1.3 基于自监督学习的检测技术

为解决监督式人工智能算法特征提取困难，高质量标注数据集获取成本高的问题，研究人员尝试将自监督式人工智能算法引入威胁检测领域。常见的自监督式人工智能算法有对比学习（Contrastive Learning）、自动编码器（Autoencoders）等，这些算法均利用数据本身来构造“伪标签”，设计相应的训练任务。

对比学习通过构造“正样本对”和“负样本对”，让模型自动学习如何缩减正样本对之间的距离，同时拉大负样本对之间的距离。在网络威胁检测任务中，正样本对常常由同一个节点在不同条件下所采集的数据组合而成，负样本对则是由不同节点所采集的数据组合而成。Han 等[16]采用本地行为和全局交互两种特征构造正样本对和负样本对进行对比学习。在网络运行过程中，检测算法需观测正样本和负样本的距离分布情况，若出现相对于历史记录的大幅度偏移，则认为该连接为恶意连接。Liu 等[17]对于网络当中的目标节点进行采样生成正样本，对与目标节点没有直接连接的局部子图采样生成负样本，通过对比学习评估每个节点的异常程度。Wang 等[18]结合了对比学习和联邦学习的思想，协调网络内物联网设备共同学习入侵检测模型，以进行多模态网络中的异构设备入侵检测。

自动编码器的实现思路是学习一个可重构输入数据的神经网络。对于网络流量样本而言，良性的网络流量在训练集当中占绝大多数。因此，自动编码器能较好学习良性网络流量的重建方法，重建误差一般较小。而对于恶意流量，则会产生较大的重建误差。Mirsky 等[19]利用正常流量训练自动编码器，并在网络运行时将流量输入自动编码器并观测其重建误差。若误差异常大，则认为该流量为恶意流量。Ding 等[20]使用图卷积网络来融合网络拓扑和节点属性信息，再使用自动编码器测量节点的重构误差来检测异常流量。

可见自监督学习技术的一大优势是不需要人工对数据集进行标注，即可自行学习网络中正常和异常流量的区别，对于发现多模态网络中复杂未知的攻击手段有一定的效果。然而，自监督式学习算法对数据量的要求更高，需要使用大量的正常流量，来构造正负样本对或者学习正常流量的模式。除此之外，若攻击者在数据集采集之前就已经潜入系统，将使得数据集存在噪声，训练出的模型也难以对类似的攻击进行防护。与监督学习方法相比，自监督学习模型的可解释性也更差，安全人员更加难以理解其决策过程[21]。

2.1.4 基于大型预训练模型的检测技术

前文所述的各类威胁检测方法，往往针对每一种威胁检测任务从零开始训练检测模型，需要一定量面向目标任务的数据集，才能达成较好的训练效果。较差的跨任务迁移性导致训练数据的利用率低下，成为了威胁检测模型进一步提升性能的瓶颈。因此，研究人员开始探索将预训练模型技术引入流量分析领域，希望用更小规模的训练集，达到相同甚至更强的分析性能。预训练模型技术的基本原理是，首先利用大量未标记的数据进行预训练，以学习通用的特征表示；然后使用面向特定下游任务的小规模数据集进行模型微调。

学术界目前已有多个将预训练模型引入流量分析领域的案例。Lin 等[22]和 Lin 等[23]都采用自然语言处理任务中常用的 transformer 模型为基础，前者[22]以字节为单位进行训练，后者[23]结合了数据包和数据流的多种特征进行训练。得到的预训练模型可以用于威胁检测、流量分类资源调度、协议分析等各类下游任务。Wang 等[24]采用另一种自然语言处理模型 T5，并设计三种不同数据分割方式的子任务对流量分析模型进行训练。Wang 等[25]使用序列数据处理模型 Mamba 进行预训练任务，该模型相对于 Transformer 模型计算复杂度更低，可以减少流量分类任务的推理时间和内存占用。Zhao 等[26]采用了计算机视觉任务中常用的 MAE（Masked Autoencoders）模型，设计多层流表示矩阵来提取原始流量数据中的不同层次流量特征。Zhou 等[27]利用 Transformer 模型，设计了掩码预测任务和细粒度多分类任务，并使用定制化的数据增强策略，以便有效捕获流量数据的本质特征，提升模型对网络协议的理解能力。Ferrag 等[28]提出了一种轻量级的 BERT 模型变体，优化自注意力机制与层结构，大幅减少了推理时间和模型参数量，适合在 IoT 设备上部署。

总之，在处理多模态网络当中异构复杂的流量特征方面，预训练模型技术具有相当大的潜力。然而，目前基于预训练模型的威胁检测技术尚处于发展的早期阶段，如何高效处理非结构化的流量数据还有待于实践探索[29]。同时，预训练模型对入侵检测设备的算力资源也提出了较高的要求，亟需研究在软件和硬件层面提高运行效率的方法，以降低部署成本。

2.1.5　基于强化学习的检测技术

强化学习是深度学习领域当中的一种特殊的训练范式。它通过构建一个智能体(Agent)，使其在环境当中进行交互式的自主学习，以便获得动态决策能力。强化学习的要点是对环境进行合适的建模、构建动作选择策略以及设定合理的奖励机制。通过奖励机制的引导，智能体可以不断优化动作选择策略，使其在与环境的交互当中得到最大的收益[30]。

学术界已有将强化学习应用到流量威胁检测领域的案例。Phan 等[31]利用强化学习引导知识迁移过程，将传统网络下采集的数据集迁移到 SDN(Software Defined Network)网络中，为数据稀缺的多模态网络威胁检测场景提供了新思路。Venturi 等[32]利用强化学习自动化生成对抗样本，为验证威胁检测系统的鲁棒性提供了测试基准。Sethi 等[33]提出了一种基于强化学习的威胁检测系统框架，通过差异化奖励策略，引导模型重点关注恶意样本，以缓解训练样本不平衡的问题。

近年来，强化学习技术在学术界得到了广泛关注。尤其是生成式人工智能领域，强化学习技术能在减少训练数据标注需求的同时，显著增强模型的推理能力与泛化能力[34]。然而在网络威胁检测领域，训练样本规模难以满足强化学习算法的需要，深度学习模型的高算力需求与现有网络设备的计算资源尚存在矛盾。仍需进一步解决数据需求、计算开销、环境仿真、可解释性等问题，以推动强化学习在网络安全领域的实际应用。

2.2　威胁应对技术

为保障网络系统的安全运行，一方面要准确检出潜在流量威胁，另一方面要根据威胁形式采取对应的防御策略，保障网络系统免受恶意行为干扰。传统的访问控制措施仅能依据固定的策略对威胁流量进行过滤，在应对未知攻击形式时防御能力明显不足。为提升网络系统免疫未知攻击的能力，研究人员提出了欺骗防御、移动目标防御两种主动防御思路。本节将梳理已有的网络威胁应对方案，并阐述在多模态网络背景下如何应用这些技术，有效加固网络系统。

2.2.1　利用访问控制技术，被动过滤网络流量

以防火墙为代表的访问控制技术，是威胁应对技术中发展历史最久，应用也最为广泛的一类。其基本原理是通过设置网络流量的访问控制策略，监控和过滤进出网络的数据流，使得网络内部的设备免受来自网络外部的潜在威胁。防火墙技术依据其流量过滤机制，大致经历了数据包过滤防火墙、状态检测防火墙、应用层防火墙三个发展阶段。

数据包过滤防火墙是最经典且最基础的网络防火墙类型，它仅依据单个数据包的报文头信息来决定是否允许数据包通过。iptables[35]是 Linux 操作系统中普遍采用的一种数据包过滤防火墙。它能够通过手动配置规则，拦截指定源 IP 地址、源端口号或协议的流量。

状态检测防火墙不再仅依赖单个数据包的信息，而是以网络连接的历史状态为依据，能更准确高效地过滤数据包。NetScreen[36]是 Juniper Networks 推出的一种状态检测防火墙，它通过跟踪 TCP 连接的状态，对属于同一 TCP 连接的一系列数据包进行聚合分析，提升检测效率。

应用层防火墙在利用网络层报文头信息的同时分析数据包的载荷内容，通过识别应用层协议的特征，将不同应用层服务所产生的流量进行差异化分析，以提升流量过滤规则的针对性。TLSFilter[37]通过分析加密协议握手阶段的元数据，防止攻击者通过混淆手段逃避防火墙的检测。

访问控制技术通过对网络流量进行过滤，可一定程度上阻断恶意流量，从而保护系统免受攻击者入侵。然而，为了部署有效的流量过滤规则，需要安全人员对攻击者的攻击手段有比较深入的了解，这一要求在实际操作层面面临较大挑战。此外，访问控制技术需要在一个特定的网关设备上部署，以将网络区隔为内网和外网两个信任程度不同的区域，这与多模态网络中节点间灵活访问的特性相矛盾。

2.2.2　利用欺骗防御技术，主动捕获恶意流量

在欺骗防御技术出现之前，网络攻防双方处于严重信息不对称的局面。具体来说，攻击方可以通过长期监控和探测目标网络，从而全面了解攻击目标的软硬件部署情况，而防御方却对攻击者可能采用的手段一无所知。为了扭转这种被动局面，研究人员提出了欺骗防御技术。其核心思想是建立诱饵系统，让攻击者误认为诱饵系统为攻击目标，从而主动捕获并监控其攻击行为。诱饵系统一般可按其拟真程度，分为低交互蜜罐、高交互蜜罐和蜜网三种。

低交互蜜罐仅能对系统的行为做出简单的模

拟，攻击者比较容易识别出蜜罐行为和真实行为的不同。但低交互蜜罐性能开销小，部署成本低，适合在资源受限的场景中部署。Dionaea[38]是一种专注于捕获恶意软件的低交互蜜罐，支持模拟 SMB、HTTP、FTP、MySQL 等多种服务。Honeyd[39]是一种可以模拟任意 TCP 或 UDP 网络服务的低交互蜜罐，支持模拟 IIS、Telnet、pop3 等网络服务。

高交互蜜罐往往利用虚拟化技术，提供更加贴近真实生产环境的交互体验，迷惑性更强。攻击者在高交互蜜罐中的行为还能被有效记录，安全人员可以据此了解攻击者的惯用手段。Shadow Deamon[40]是一种专门针对 Web 应用的高交互蜜罐，可以有效捕获 SQL 注入，跨站脚本攻击等攻击形式。Lyrebird[41]是一种用途广泛的高交互蜜罐，能模拟文件操作、命令执行等多种系统操作，诱导攻击者发起入侵行为。然而，高交互蜜罐允许攻击者访问更多的系统资源，若实现不当反而可能成为攻击者入侵网络系统的跳板[42]。

蜜网往往利用软件定义网络技术，将多个蜜罐组成高可控的蜜罐网络，甚至根据被保护系统的网络拓扑结构进行定制化模拟，以进一步提升迷惑性。Zarca 等[43]利用网络功能虚拟化技术，实现了一种高可扩展性的虚拟物联网蜜罐，以模拟真实的传感器网络。Fan 等[44]将蜜网当中的诱饵模块、协调模块、捕获模块解耦，协调组织多种异构的蜜罐，实现不同模块的独立开发和更新。

总而言之，欺骗防御技术能够主动吸引攻击者的火力，并采集高质量的网络威胁样本，以支撑防御措施的进一步迭代。然而，蜜罐和蜜网的部署也增加了网络当中的暴露点，如果配置不当，可能会引入新的潜在威胁形式[45]。

2.2.3 利用移动目标防御技术，提升攻击者嗅探难度

攻击者遵循预设的入侵路径，是欺骗防御技术能够发挥作用的基本前提。然而在多模态网络当中，攻击者很可能通过非预期的途径侵入系统，致使部署的欺骗防御策略失效。为解决上述问题，研究人员提出了移动目标防御技术[46]（Moving Target Defense），旨在从系统本身的设计出发，不断改变攻击者可访问的攻击面，从而增加针对系统进行侦察和攻击的难度。为有效达成这一防御目标，需要从移动要素（What to Move）、移动方式（How to Move）、移动时机（When to Move）这三个维度出发，结合网络系统的具体情况，设定合理的防御策略。

“移动要素”是系统中需要动态改变的组件或资源。Jafarian 等[47]将 IP 地址作为随机化要素，周期性进行跳变，使得目标主机难以追踪。Jia 等[48]将服务器节点与客户端之间的对应关系作为随机化要素，通过动态变化客户端所接入的代理节点，防范分布式拒绝服务攻击导致的服务中断。“移动方式”是系统改变移动要素的具体方式，Van 等[49]采用简单随机化的方式改变 XML 命名空间前缀，以此来识别中间人注入的恶意网页代码。Feng 等[50]基于贝叶斯博弈推导最佳迁移策略，将系统中的关键资源地址动态迁移，使攻击者难以确定攻击目标的实际位置。“移动时机”是系统改变移动要素的时间间隔。Javadpour 等[51]通过长短间隔交错的策略实现主机网络地址的迁移，确保系统在不牺牲稳定性的同时具备一定的动态调整能力。Zhang 等[52]通过自适应调整策略，持续评估网络威胁形势，动态调整迁移周期。Wu 等[4]在云原生环境下综合了移动目标防御的三要素。将异构执行体作为移动要素，通过随机或者预设的移动方式，根据多模裁决结果决定移动时机，以干扰攻击链。仿真结果表明，在云原生系统中引入异构执行体机制，仅引入 5% 左右的性能开销，即可达成 MTD（Moving Target Defense）的典型安全效益。若同时引入联合裁决机制，可实现更低的入侵概率，同时使开销在线性可控范围内。

移动目标防御技术通过持续不断地调整系统内的组件或资源，为攻击者在入侵过程中制造了高度不确定的攻击环境，从而显著增强了系统对未知威胁的防御效能。若想有效发挥移动目标防御技术的防御效果，需统筹安排移动要素、移动方式和移动时机，实现系统开销与防御效果的动态平衡。

2.3 小结

表 1 总结了本节所述网络威胁防御技术的关键方法、优势及局限性。总体而言，在威胁检测领域，技术发展经历了从依赖专家经验的静态规则匹配，到基于监督学习的流量分析，再到无需标注的自监督学习和通用化预训练模型的跃迁。威胁应对技术则从被动过滤发展到主动防御，从静态规则转向动态重构。两类技术共同面临着多模态网络环境下协议异构、环境动态、跨模态攻击等核心挑战。

表 1　网络威胁防御技术发展脉络

Table 1　The Development Trend of Network Threat Defense Technology

类别	子类	核心方法	优势	局限性
威胁检测技术	规则驱动	深度包检测、正则表达式	可解释性强、实现简单	依赖专家知识、无法处理加密流量、规则更新滞后
	监督学习	随机森林、卷积神经网、图神经网	可处理加密流量、可自动提取特征	依赖标注数据、特征提取需人工介入、泛化能力差
	自监督学习	对比学习、自动编码器	无需标注数据、能发现未知威胁	数据需求量大、可解释性差
	大型预训练模型	Transformer、MAE、Mamba	跨任务迁移能力强、高效利用训练样本	算力需求高、数据处理 难、部署成本高
	强化学习	环境建模、动作选择策略、奖励机制	容易适应动态环境、人工介入少	数据需求量大、环境仿真困难、奖励机制难以设定
威胁应对技术	访问控制	数据包过滤、状态检测、应用层分析	技术成熟、实现简单	策略静态僵化、依赖先验知识
	欺骗防御	低交互蜜罐、高交互蜜罐、蜜网	主动捕获威胁、获取攻击情报	增加攻击面、资源消耗大
	移动目标防御	移动要素、移动时机、移动方式	增加攻击成本、能抵御未知威胁	系统开销大、安全与性能难以平衡

3　多模态网络安全技术的前景展望

现有的威胁检测和应对技术多以传统的 IP 网络架构为依托，存在推理泛化能力不足、依赖特定协议语义、部署方式固定僵化等缺陷。在面对多模态网络当中海量多元的接入设备、彼此异构的通信协议、动态变化的网络环境时，往往效果不佳。

为应对上述挑战，需要积极探索大模型、软件定义网络、内生安全等技术，建立和多模态网络发展趋势相适应的新型网络安全防护体系，以便有效处理多模态网络发展过程中所涌现出的新问题，消除层出不穷的新型安全风险。具体来说，可针对以下技术要点进行科研攻关：

(1) 发展生成式样本扩充技术，助力检测模型有效加固

多模态网络当中的威胁种类复杂多样，高质量威胁样本的稀缺性更为凸显，难以满足智能化威胁检测模型的训练样本量需求。可以利用生成式人工智能模型的强大生成能力，精准拟合多模态网络中复杂的流量数据分布，以自动化扩充威胁样本，助力威胁检测模型的有效加固。

(2) 发展异构联合裁决技术，强化防御体系可靠性

单一的流量检测方法往往不能做到面面俱到，可靠性难以保证。攻击者很可能利用某一检测方法的弱点，通过探测、穷举、混淆等手段针对性击破，使得威胁防御手段失效。可以利用多模态网络的智能化硬件优势，融合内生安全技术，建立自动化联合决策的检测模型与应对方法体系，防范“单点失效”，提升防御体系的可靠性。

(3) 发展形式化安全度量技术，支撑防御体系持续优化

多模态网络基于生成式人工智能、智能化设备、主动防御等技术构建威胁防御屏障，但这些技术的决策过程往往难以理解，其安全机理难以得到证明与分析。可以利用形式化分析方法、认知科学解释模型等技术，建立防御效果与防御方案之间的关联关系，帮助安全人员理解模型的决策理由，基于防御机理反馈、优化威胁防御体系，实现一定参数配置下攻击不可达的系统级可证明安全。

(4) 发展跨模态信息迁移技术，实现威胁要素标准化

多模态网络当中存在多种异构的网络协议，导致不同网络模态之间存在语义鸿沟，阻碍了信息在跨模态场景下的有效迁移。可研究信息要素的标准化映射技术，构建多模态统一威胁样本库，提取协议无关的深层攻击模式，建立跨模态的协同防御决策链，以解决威胁情报难以在异构协议间迁移应用的难题。

(5) 发展分布式信任评估技术，实现跨模态信任传递

现有网络当中的信任评估机制往往较为单一，难以满足设备的跨模态灵活接入需求。需构建分布式、高可靠的信任评估体系。通过跨模态身份

转换协议，实现信任要素在模态间的透明转换，从而为跨模态的信任传递提供条件。依托跨模态的连续认证技术，构建多维信任图谱，实现信任等级在网络模态间的无损传递和动态调整。

总体而言，如图 2 所示，维护多模态网络的安全是一项复杂的系统性任务。需要综合大模型、软件定义网络、内生安全等多项前沿技术，在威胁检测和威胁应对两方面双管齐下，推动现有威胁防御技术体系适应多模态网络的新变化，以应对多模态网络在发展过程当中所面临的新问题、新挑战。

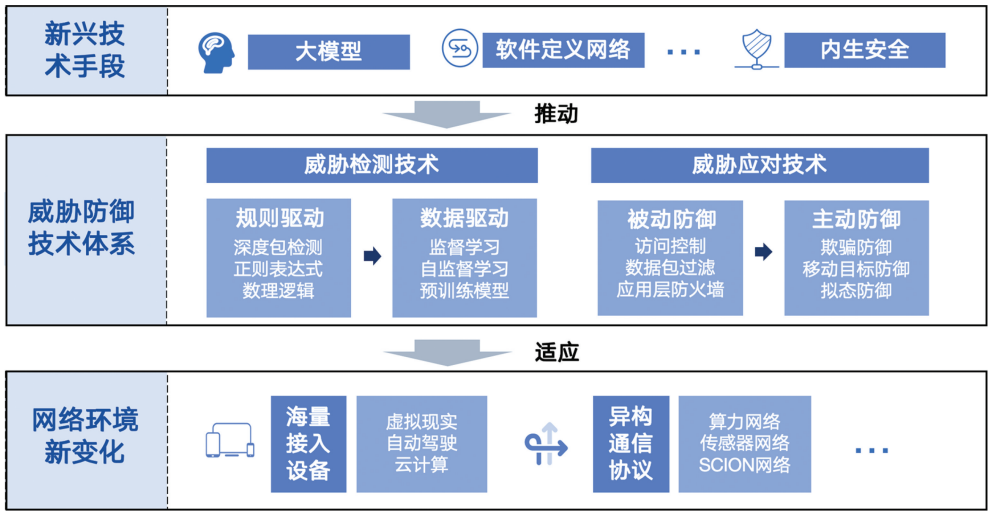

图 2 多模态网络安全未来发展方向示意图

Fig. 2 Schematic Diagram of Future Development Direction of Polymorphic Network Security

4 结 语

多模态网络技术在其发展过程中机遇与风险并存，在满足新兴业态个性化用网需求的同时，也面临着一系列全新的安全挑战。为顺应多模态网络带来的新变化，应对多模态网络安全的新挑战，本文提出如下几点关键研究方向：一是要在关注已知威胁的同时，用智能化手段努力提升对海量未知威胁的覆盖率。二是要提升防御方法的跨模态泛化能力，准确提取深层次威胁特征，适应多模态网络中异构信息的呈现形式。三是要提升防御方法的感知与应变能力，适应复杂动态的多模态网络环境，实现防御策略的按需部署。为实现上述发展愿景，亟需探索大模型、软件定义网络、内生安全等多项前沿技术，通过在威胁检测与威胁应对两方面不断地探索和创新，构建起与多模态网络架构相适应的新型网络安全防御体系，助力多模态网络技术稳健发展。

参 考 文 献

[1] Ferguson AD, Gribble S, Hong CY, et al. Orion: Google's software-defined networking control plane// Proceedings of the 18th USENIX Symposium on Networked Systems Design and Implementation (NSDI 21). Boston, MA: USENIX Association, 2021: 83—98.

[2] Poovendran R. Cyber-physical systems: Close encounters between two parallel worlds [Point of View]. Proceedings of the IEEE, 2010, 98(8): 1363—1366.

[3] Moon SJ, Sekar V, Reiter MK. Nomad: Mitigating arbitrary cloud side channels via provider-assisted migration// Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. Denver, Colorado, USA. ACM, 2015: 1595—1606.

[4] Wu Q, Wang R, Yan XC, et al. Intrinsic security: A robust framework for cloud-native network slicing *via* a proactive defense paradigm. IEEE Wireless Communications, 2022, 29(2): 146—153.

[5] 胡宇翔，伊鹏，孙鹏浩，等. 全维可定义的多模态智慧网络体系研究. 通信学报，2019，40(8)：1—12.
Hu XY, Yi P, Sun PH, et al. Research on the full-dimensional defined polymorphic smart network. Journal on Communications, 2019, 40(8): 1—12. (in Chinese)

[6] Fernandes S, Antonello R, Lacerda T, et al. Slimming down deep packet inspection systems// IEEE INFOCOM Workshops 2009. Rio de Janeiro, Brazil: IEEE, 2009: 1—6.

[7] Keralapura R, Nucci A, Chuah CN. Self-learning peer-to-peer traffic classifier// Proceedings of 18th International Conference on Computer Communications and Networks. San Francisco, CA, USA. IEEE, 2009: 1—8.

[8] Bilge L, Dumitras T. Before we knew it: An empirical study of zero-day attacks in the real world// Proceedings of the 2012 ACM Conference on Computer and Communications Security. Raleigh, North Carolina, USA. ACM, 2012: 833—844.

[9] Zscaler. 2023 threatlabz state of encrypted attacks report. (2023-12-05)/[2024-10-26]. https://www.zscaler.com/resources/2023-threatlabz-state-of-encrypted-attacks-report.

[10] Hayes J, Danezis G. k-fingerprinting: A robust scalable website fingerprinting technique// 25th USENIX Security Symposium (USENIX Security 16). Austin, TX: USENIX Association, 2016, 1187—1203.

[11] Taylor VF, Spolaor R, Conti M, et al. Robust smartphone app identification *via* encrypted network traffic analysis. IEEE Transactions on Information Forensics and Security, 2018, 13(1): 63—78.

[12] Xi SK, Fu TY, Bu K, et al. MineShark: Cryptomining traffic detection at scale// Proceedings of 2025 Network and Distributed System Security Symposium. San Diego, CA, USA. Internet Society, 2025: 1—20.

[13] 中国移动通信集团浙江有限公司，中国移动通信集团有限公司，中国移动(浙江)创新研究院有限公司，等. 网络异常流量检测模型训练方法、装置及电子设备：中国，CN202410522307.1. 2024-10-18.

[14] 天翼安全科技有限公司，异常网络流量检测方法及相关硬件：中国，CN119383006A. 2025.01.28.

[15] Zhang HZ, Yu L, Xiao X, et al. TFE-GNN: A temporal fusion encoder using graph neural networks for fine-grained encrypted traffic classification// Proceedings of the ACM Web Conference 2023. Austin TX, USA. ACM, 2023: 2066—2075.

[16] Han XY, Cui SS, Qin J, et al. ContraMTD: An unsupervised malicious network traffic detection method based on contrastive learning// Proceedings of the ACM Web Conference 2024. Singapore, Singapore. ACM, 2024: 1680—1689.

[17] Liu YX, Li Z, Pan SR, et al. Anomaly detection on attributed networks *via* contrastive self-supervised learning. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(6): 2378—2392.

[18] Wang N, Chen YM, Hu Y, et al. FeCo: Boosting intrusion detection capability in IoT networks *via* contrastive learning// IEEE INFOCOM 2022-IEEE Conference on Computer Communications. London, United Kingdom: IEEE, 2022: 1409—1418.

[19] Mirsky Y, Doitshman T, Elovici Y, et al. Kitsune: An ensemble of autoencoders for online network intrusion detection// Proceedings of 2018 Network and Distributed System Security Symposium. San Diego, CA: Internet Society, 2018: 1—15.

[20] Ding KZ, Li JD, Bhanushali R, et al. Deep anomaly detection on attributed networks// Proceedings of the 2019 SIAM International Conference on Data Mining. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2019: 594—602.

[21] Jacobs AS, Beltiukov R, Willinger W, et al. AI/ML for network security: The emperor has no clothes// Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security. Los Angeles CA, USA. ACM, 2022: 1537—1551.

[22] Lin XJ, Xiong G, Gou GP, et al. ET-BERT: A contextualized datagram representation with pre-training transformers for encrypted traffic classification// Proceedings of the ACM Web Conference 2022. Virtual Event, Lyon France. ACM, 2022: 633—642.

[23] Lin P, Ye KJ, Hu YS, et al. A novel multimodal deep learning framework for encrypted traffic classification. IEEE/ACM Transactions on Networking, 2023, 31(3): 1369—1384.

[24] Wang QN, Qian C, Li XC, et al. Lens: A foundation model for network traffic. [2024-10-26]. https://arxiv.org/abs/2402.03646v4.

[25] Wang TZ, Xie XH, Wang WD, et al. NetMamba: Efficient network traffic classification *via* pre-training unidirectional mamba. [2024-10-26]. https://arxiv.org/abs/2405.11449v4.

[26] Zhao RJ, Zhan MW, Deng XW, et al. Yet another traffic classifier: A masked autoencoder based traffic transformer with multi-level flow representation. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(4): 5420—5427.

[27] Zhou G, Guo X, Liu Z, et al. Trafficformer: An efficient pre-trained model for traffic data// Proceedings of the 2025 IEEE Symposium on Security and Privacy (SP). San Francisco, CA: IEEE Computer Society, 2024: 102—102.

[28] Ferrag MA, Ndhlovu M, Tihanyi N, et al. Revolutionizing cyber threat detection with large language models: A privacy-preserving BERT-based lightweight model for IoT/IIoT devices. IEEE Access, 2024, 12: 23733—23750.

[29] Zhang J, Bu HY, Wen H, et al. When LLMs meet cybersecurity: A systematic literature review. Cybersecurity, 2025, 8(1): 55.

[30] Sabeel U, Heydari SS, El-Khatib K, et al. Unknown, atypical and polymorphic network intrusion detection: A systematic survey. IEEE Transactions on Network and Service Management, 2024, 21(1): 1190—1212.

[31] Phan TV, Sultana S, Nguyen TG, et al. $Q $-TRANSFER: A novel framework for efficient deep transfer learning in networking// 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC). Fukuoka, Japan: IEEE, 2020: 146—151.

[32] Venturi A, Apruzzese G, Andreolini M, et al. DReLAB-deep REinforcement learning adversarial botnet: A benchmark dataset for adversarial attacks against botnet intrusion detection systems. Data in Brief, 2021, 34: 106631.

[33] Sethi K, Madhav YV, Kumar R, et al. Attention based multi-agent intrusion detection systems using reinforcement learning. Journal of Information Security and Applications, 2021, 61: 102923.

[34] DeepSeek-AI, Guo DY, Yang DJ, et al. DeepSeek-R1: Incentivizing reasoning capability in LLMs *via* reinforcement learning. [2024-10-26]. https://arxiv.org/abs/2501.12948v1.

[35] Kerrisk M. iptables(8)—Linux manual page. (2024-06-26)/[2024-10-26]. https://www.man7.org/linux/man-pages/man8/iptables.8.html.

[36] Chris Roeckl. Stateful inspection firewalls. Sunnyvale, CA: Juniper Networks, 2004.

[37] Phillips M. TLS filter: An application-level firewall for transport layer security. [2024-10-26]. https://kevincurran.org/papers/Thesis%20-%20Application%20Level%20Firewall.pdf.

[38] Phibos. Dionaea. (2020-11-30)/[2024-10-26]. https://github.com/DinoTools/dionaea.

[39] Honeyd. Developments of the honeyd virtual honeypot. (2007-05-27)/[2024-10-26]. https://www.honeyd.org/.

[40] Zecure information technologies. Shadow daemon. (2023-07-23)/[2024-10-26]. https://github.com/zecure/shadowd.zecure.org/.

[41] Meituan-Dianping. Lyrebird. (2018-08-03)/[2024-10-26]. https://github.com/Meituan-Dianping/lyrebird.

[42] Javadpour A, Ja'fari F, Taleb T, et al. A comprehensive survey on cyber deception techniques to improve honeypot performance. Computers & Security, 2024, 140: 103792.

[43] Zarca AM, Bernabe JB, Skarmeta A, et al. Virtual IoT HoneyNets to mitigate cyberattacks in SDN/NFV-enabled IoT networks. IEEE Journal on Selected Areas in Communications, 2020, 38(6): 1262—1277.

[44] Fan WJ, Du ZH, Smith-Creasey M, et al. HoneyDOC: An efficient honeypot architecture enabling all-round design. IEEE Journal on Selected Areas in Communications, 2019, 37(3): 683—697.

[45] Manadhata PK, Wing JM. An attack surface metric. IEEE Transactions on Software Engineering, 2011, 37(3): 371—386.

[46] Cho JH, Sharma DP, Alavizadeh H, et al. Toward proactive, adaptive defense: A survey on moving target defense. IEEE Communications Surveys & Tutorials, 2020, 22(1): 709—745.

[47] Jafarian JH, Al-Shaer E, Duan Q. An effective address mutation approach for disrupting reconnaissance attacks. IEEE Transactions on Information Forensics and Security, 2015, 10(12): 2562—2577.

[48] Jia Q, Sun K, Stavrou A. MOTAG: Moving target defense against Internet denial of service attacks// 2013 22nd International Conference on Computer Communication and Networks (ICCCN). Nassau, Bahamas: IEEE, 2013: 1—9.

[49] Van Gundy M, Chen H. Noncespaces: Using randomization to enforce information flow tracking and thwart cross-site scripting attacks. [2024-10-26]. https://haochen.org/publications/vangundy2009noncespaces.pdf.

[50] Feng XT, Zheng ZZ, Cansever D, et al. A signaling game model for moving target defense// IEEE INFOCOM 2017-IEEE Conference on Computer Communications. Atlanta, GA, USA. IEEE, 2017: 1—9.

[51] Javadpour A, Ja'fari F, Taleb T, et al. SCEMA: An SDN-oriented cost-effective edge-based MTD approach. IEEE Transactions on Information Forensics and Security, 2022, 18: 667—682.

[52] Zhang T, Xu CQ, Zhang BC, et al. Towards attack-resistant service function chain migration: A model-based adaptive proximal policy optimization approach. IEEE Transactions on Dependable and Secure Computing, 2023, 20(6): 4913—4927.

Threat Defense Technology for Polymorphic Networks

Kui Ren[1,2*] Jianpeng Wang[1,2] Wenyuan Yang[3] Xinlei Wang[1,2] Shuguo Zhuo[1,2] Zhongjie Ba[1,2]

1. *The State Key Laboratory of Blockchain and Data Security, Zhejiang University, Hangzhou 310027, China*
2. *The School of Cyber Science and Technology, Zhejiang University, Hangzhou 310027, China*
3. *The School of Cyber Science and Technology, Sun Yat-sen University, Shenzhen 518107, China*

Abstract Driven by emerging network services such as virtual reality, industrial Internet, and smart cities, new architectures including computing power networks, sensor networks, software-defined networks, and SCION networks are propelling the network towards a new paradigm of coexistence and co-governance of multiple modalities. The polymorphic network architecture, with its fully customizable processes, enables deep integration of threat defense mechanisms within the architecture. Consequently, security policies can be dynamically deployed as service requirements evolve. This paper investigates both threat detection and response strategies. Following the timeline of technological development, we systematically review existing network threat defense technologies. While revisiting their design philosophies and technical approaches, we particularly examine the limitations faced by these technologies in polymorphic network environments. With the increasing adoption of polymorphic network technologies, research on threat defense mechanisms compatible with such environments has become critically urgent. There is also a pressing need to explore emerging technologies such as large models, endogenous security, and software-defined networks to promote the development of threat defense technical framework, thereby providing a solid network foundation for diverse services in polymorphic networks.

Keywords polymorphic network; traffic analysis; active defense; artificial intelligence; cross-modal attack; endogenous security

任 奎 教授，博士生导师，浙江大学计算机科学与技术学院院长，区块链与数据安全全国重点实验室常务副主任，浙江大学学术委员会委员，AAAS、ACM、CCF 和 IEEE 会士。主要从事网络与通信安全、人工智能安全、云安全与数据安全等领域的研究。

（责任编辑 陈 鹤 张 强）

* Corresponding Author, Email: kuiren@zju.edu.cn