

基于改进HMM的物联网入侵路径检测方法

何洋军

(浙江省公众信息产业有限公司, 浙江 杭州 310012)

摘要: 针对物联网多步入侵路径隐蔽性强、标注数据稀缺及加密流量占比高等难题, 对基于改进隐马尔可夫模型(HMM)的入侵路径检测进行了研究。构建融合网络拓扑、配置与漏洞信息的四元组入侵图模型, 利用逻辑斯蒂函数量化状态转移概率, 形成动态概率矩阵。针对传统HMM的零概率缺陷, 设计基于语义相似性的最大似然概率重分配策略。引入状态转移修正系数与自适应动态阈值, 构建多步迭代推断框架, 实现告警序列智能筛选与入侵意图鲁棒识别。实验结果表明, 该方法在复杂攻击场景下路径检测准确率达94%以上, 具备良好的鲁棒性与可扩展性。

关键词: 改进HMM; 物联网; 入侵路径检测; 四元组网络; 节点转移概率; 漏洞信息

中图分类号: TP115

文献标志码: A

doi:10.11959/j.issn.2096-109x.AQ26047

Internet of things intrusion path detection method based on improved HMM

He Yangjun

Zhejiang Public Information Industry Co., Ltd., Hangzhou 310012, China

Abstract: A study was conducted on intrusion path detection based on an improved hidden Markov model (HMM) to address the challenges of strong concealment of multi-step intrusion paths in the Internet of Things, scarce annotated data, and high proportion of encrypted traffic. A quadruple intrusion graph model was constructed that integrated network topology, configuration, and vulnerability information, state transition probability was quantified using logistic functions, a dynamic probability matrix was formed, and a maximum likelihood probability reallocation strategy was designed based on semantic similarity for zero probability defects in traditional HMM. Introducing state transition correction coefficients and adaptive dynamic thresholds, a multi-step iterative inference framework was constructed to achieve intelligent screening of alarm sequences and robust recognition of intrusion intentions. The experimental results show that the method has an accuracy of over 94% in path detection in complex attack scenarios, and has good robustness and scalability.

Key words: improve HMM, internet of things, intrusion path detection, quadruple network, state transition probability, vulnerability information

0 引言

随着物联网(internet of things, IoT)技术的快速发展与广泛应用, 网络边界正从传统的信息系统向物理世界深度延伸。智能家居、工业互联

网、车联网、智慧城市等物联网场景的普及, 在提升生产效率与生活便利性的同时, 也带来了前所未有的安全挑战。据权威机构统计, 物联网设备已成为网络攻击的首要目标, 针对物联网的入

收稿日期: 2026-02-27; 修回日期: 2026-05-16

通信作者: 何洋军, heyangjun369@163.com

侵事件年均增长率超过 30%。与传统 IT 环境不同, 物联网环境具有设备资源受限、协议栈多样、海量异构节点泛在互联、攻击面广泛等显著特征, 这使传统网络安全防护手段难以直接移植应用。现阶段物联网入侵正从单点向多步骤路径化演变, 入侵者经侦察、入侵、控制、扩散到破坏阶段实施完整链路入侵, 隐蔽性强。现有检测手段弊端明显, 其依赖预设特征, 难以应对未知和变异入侵, 只能检测单个行为, 无法还原入侵路径。并且还需大量标注样本训练, 模型参数固定, 难以适应动态环境, 缺乏对时序性入侵路径的实时识别能力。

针对上述难点, 伍均玺等^[1]提出基于改进深度学习的主动式通信网络入侵行为自适应识别方法, 该方法从流量数据中提取空间特征, 并通过长短期记忆网络捕捉这些特征随时间变化的时序依赖关系; 引入注意力机制, 为特征赋予不同权重, 以此突出关键特征对入侵行为识别的作用。但是主动学习所选样本若不能有效代表数据分布特征, 不仅无法提升模型性能, 反而会引入噪声数据。刘强等^[2]提出自适应类增量学习的物联网入侵检测方法, 当出现新的入侵类别数据时, 启动类增量学习机制, 使新模型在保留对旧类别知识记忆的同时学习新类别特征; 引入自适应策略, 依据新数据的分布特征和模型性能动态调整学习率参数, 将实时采集的物联网流量数据输入已训练好的模型中, 模型根据所学知识判断是否存在入侵行为及入侵类别。但是自适应调整参数的策略依赖于对数据分布和模型性能的准确评估, 而物联网环境复杂多变, 数据分布难以精准预测, 导致参数调整难以达到最优, 进而影响模型对新类别的学习效果。Mansahaj 等^[3]提出网络安全和网络入侵检测的联合方法, 构建基于深度学习的联合模型, 运用卷积神经网络对流量数据空间特征进行提取, 挖掘其中随时间演变的入侵迹象; 通过训练优化模型参数, 输出最终的入侵检测结果; 深度学习模型结构复杂, 训练时需大量标注数据。然而, 网络入侵场景多样, 获取全面且准确的标注数据成本高、难度大, 这将导致模型过拟合, 泛化能力受限。Lu 等^[4]采用基于堆叠学习的分布式网络入侵检测方法, 划分分

布式网络节点模型, 将基础模型的输出作为新的特征, 与原始数据中的部分关键特征进行融合, 构建一个元特征集; 将元特征集输入到神经网络元学习器中进行二次训练与学习, 由元学习器给出最终的入侵检测判定结果。但基础检测模型的选择和配置对整体性能影响极大, 当选择不当时, 各节点输出的初步检测结果质量不高, 会导致融合后的元特征集缺乏有效信息, 降低准确率。

当前, 图神经网络 (graph neural network, GNN) 及其变体凭借强大的图结构学习和表征能力, 在网络入侵检测领域取得了显著进展, 尤其在捕获节点间复杂空间关联和层级结构方面展现出优势。然而, GNN 方法通常依赖大量标注数据进行端到端训练以学习有效表征, 在网络入侵检测这一特定场景下, 获取完备、高质量且标注准确的攻击路径序列数据成本高昂且困难, 导致小样本数据环境普遍存在。在此条件下, GNN 模型容易因参数过多、数据不足而产生过拟合, 其性能可能受限。相比之下, 本文采用的改进隐马尔可夫模型 (HMM) 方法, 在处理小样本时序数据方面具有天然优势。HMM 作为一种概率图模型, 其参数具有明确的统计意义和强可解释性, 这使模型能够通过相对较少的样本, 有效地从时序依赖关系中学习到状态转移规律。特别地, 通过期望最大化 (expectation-maximization, EM) 等算法, HMM 可以在小样本甚至数据不平衡的场景下进行稳健的参数估计, 避免了复杂深度学习模型对海量数据的依赖。物联网入侵路径与传统网络入侵路径的核心区别在于前者受限于设备资源异构性与协议多样性, 攻击常以横向扩散为主, 路径节点多且状态跳转依赖物理环境与网络配置的耦合; 后者则集中于服务器与终端间的纵向渗透, 路径相对规整且依赖操作系统与应用层漏洞。本文算法基于改进 HMM 构建, 其输入为状态转移概率与观测序列, 不依赖特定网络协议或设备类型, 因此理论上可推广至传统网络、工业控制网络等环境的入侵路径检测, 但需根据目标网络的拓扑结构与漏洞特征重新训练或调整状态转移概率矩阵。因此, 针对物联网入侵路径检测中攻击步骤的时序性、状态转移的隐蔽性及标

注数据稀疏的特点，改进的HMM框架在数据效率、模型收敛速度和时序动态刻画上提供了更契合的解决方案，为在有限数据条件下实现高精度、自适应的入侵路径检测奠定了理论基础。

1 物联网入侵路径状态转移概率矩阵生成

从入侵者的角度出发，建立网络入侵图，描述各个路径和节点的入侵状态转移过程，实现对物联网入侵路径的概率化处理，提高该方法对网络状态变化的自适应能力。

设定网络入侵图为 G ，采用四元组表示为 $G = (S, A, E, \Delta)$ ，其中， S 为状态节点集合， $S = \{S_i | i = 1, 2, \dots, n\}$ ； A 为原子入侵节点集合； E 为状态节点间的有向边集合^[5]， $E \subseteq S_i \times S_j, \forall e_{ij} \in E$ ， e_{ij} 为连接节点 S_i 和 S_j 的边， S_i 为起始状态节点， S_j 为目标状态节点， Δe_{ij} 等于原子入侵 a 的发生概率^[6]； Δ 为状态转移概率集合， $\forall \Delta(e_{ij}) \in \Delta$ ，假设 $P(S_j | S_i, a)$ 为入侵者由节点状态 S_i 经原子入侵 a 转移到节点 S_j 的概率。为了更精确地刻画复杂网络环境中各因素间的非线性交互作用，放弃了传统的线性加权假设，引入一个基于逻辑斯蒂函数的非线性概率聚合模型。该模型不仅考虑了单一特征的影响，还通过引入特征交叉项来捕捉因素间的协同效应，从而提升状态转移概率建模的准确性与鲁棒性。 $P(S_j | S_i, a)$ 的计算式为：

$$P(S_j | S_i, a) = \sigma(\psi_0 + \phi_1 E_v + \phi_2 P_h + \phi_3 (1 - W_t) + \phi_4 C_t + \phi_5 (E_v \times P_h) + \phi_6 (E_v \times C_t)) \quad (1)$$

其中， E_v 为漏洞可利用性评分，该值采用通用漏洞评分系统（common vulnerability scoring system, CVSS）基准分数，结合公开渗透测试库中该漏洞的利用代码成熟度，进行归一化处理得到； P_h 为攻击成功历史概率，主要基于企业内部或公开攻击数据集中同类攻击在相似环境下的成功次数统计，计算成功频率作为先验概率； W_t 为网络配置与策略权重，结合防火墙规则、访问控制列

表（access control list, ACL）、入侵防御系统（intrusion prevention system, IPS）日志，判断当前节点间通信是否受控，赋予权重，该值越低表示转移越困难； C_t 为实时威胁情报置信度，即通过集成外部威胁情报源，获取当前攻击手法活跃度与置信评分； $\sigma\left(\frac{1}{1 + e^{-x}}\right)$ 为逻辑斯蒂（Sigmoid）函数，确保概率值落在 $[0, 1]$ 区间内； $\phi = [\phi_1, \phi_2, \dots, \phi_6]$ 为模型的权重参数集合，本模型的权重参数 ϕ 通过一个独立的监督学习过程自动获得； ψ_0 为偏置项； ϕ_1 、 ϕ_2 、 ϕ_3 、 ϕ_4 分别为 E_v 、 P_h 、 W_t 、 C_t 的权重参数； ϕ_5 为漏洞可利用性与历史成功率的交互项权重； ϕ_6 为漏洞可利用性与实时威胁情报的交互项权重。

式(1)中权重参数 ϕ 的确定，是保证模型精度的关键，为此，设计了一个离线参数学习阶段。具体而言，从历史安全事件库和渗透测试报告中，提取大量的原子攻击实例。对于每个实例，记录其执行前的网络状态特征（即 E_v, P_h, W_t, C_t ）以及攻击的最终结果（成功或失败，作为二元标签 $y \in \{0, 1\}$ ）。这样，就构建了一个包含 M 个样本的训练数据集 $\{(\mathbf{x}^{(k)}, y^{(k)})\}_{k=1}^M$ ，其中 $\mathbf{x}^{(k)}$ 是由四元组特征及其交叉项构成的特征向量。随后，采用批量梯度下降算法，以最小化交叉熵损失函数为目标，对参数 ϕ 进行优化。该优化过程不仅能够自动学习各特征及其交互项的最优权重，还从理论上保证了模型能够最大限度地拟合历史数据中的攻击成败规律。相比于单纯依赖敏感性分析的手动调参，该方法具有更坚实的统计学基础和更强的泛化能力。

节点状态转移概率表示如图1所示，计算节点指向自身边的状态转移概率，以图1中的 S_0 为例， $k_0 = 2$ ， S_0 可以向 S_1 或 S_2 进行转移。若向 S_1 转移得到状态概率为 $P_d(V_1)$ ，失败概率^[7]则为 $1 - P_d(V_1)$ ；若向 S_2 转移得到状态概率为 $P_d(V_2)$ ，失败概率则为 $1 - P_d(V_2)$ 。将失败概率记作节点经过入侵后状态保持不变的概率，则 S_0 的转移概率为：

$$d_0 = \min\{2 - P_d(V_1) - P_d(V_2)\} \quad (2)$$

其中， d_0 为节点 S_0 保持状态不变的概率。

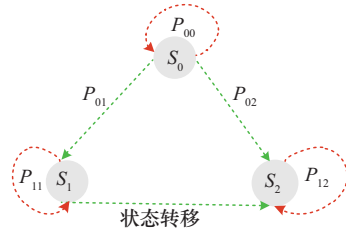


图1 节点状态转移概率表示

对图1进行量化表示, 计算入侵者实施路径入侵后状态转移的成功概率为:

$$P_{00} = \frac{d_0}{d_0 + P_d(V_1) + P_d(V_2)} \quad (3)$$

$$P_{01} = \frac{P_d(V_1)}{P_d(V_1) + P_d(V_2)} \quad (4)$$

$$P_{02} = \frac{P_d(V_2)}{P_d(V_1) + P_d(V_2)} \quad (5)$$

$$P_{12} = P_d(V_2), P_{11} = 1 - P_d(V_2) \quad (6)$$

其中, P_{00} 为节点 S_0 转移到自身的状态转移概率, P_{01} 为 S_0 转移到 S_1 的状态转移概率, P_{02} 为 S_0 转移到 S_2 的状态转移概率, P_{11} 为 S_1 转移到 S_0 的状态转移概率, P_{12} 为 S_1 转移到 S_2 的状态转移概率。

状态转移概率的量化基于网络配置、漏洞可利用性、攻击成功历史数据与实时威胁情报4类信息, 其具体计算依据如下。

2 联合改进HMM的入侵路径自适应识别

联合改进HMM的入侵路径自适应识别方法适用于加密流量, 其核心依据在于: 模型不依赖于对加密载荷的解密, 而是基于网络层和传输层的元数据、上下文关联以及漏洞信息进行推理。在加密流量环境下, 虽然应用层载荷不可见, 但仍可从以下维度获取有效的告警元数据。

(1) 流特征元数据

五元组信息(源IP地址、目的IP地址、源端口、目的端口、传输层协议)在加密流量中明文传输, 可用于建立通信对之间的关联关系; 流统计特征(数据包长度分布、包到达间隔时间、流持续时间、上下行流量比例等)由于加密算法不会改变流量的统计特性, 使不同应用和攻击行为的流模式仍具有可区分性; 传输层安全(transport layer security, TLS)协议握手特征(证书信息、加密套件协议、服务器名称指示(server name indi-

cation, SNI)字段)中, TLS协议在TLS握手中以明文形式传输, 可揭示访问的目标域名^[8]。

(2) 告警日志元数据

入侵检测系统(intrusion detection system, IDS)/入侵防御系统(intrusion prevention system, IPS)告警不依赖解密, 即使流量加密, 部署在边界的入侵检测系统仍可基于流特征和异常行为模式(如异常大的数据包、非标准的端口组合、与已知C2服务器的通信模式)产生告警; 防火墙日志记录了连接尝试、阻断记录和策略命中日志; 身份认证日志则包含多次失败的认证尝试以及异常时间的登录行为^[9]。

(3) 上下文关联信息

漏洞信息包括目标资产的已知漏洞, 攻击者可能利用这些漏洞进行入侵, 无论后续流量是否加密; 网络拓扑与访问关系表明, 加密改变了内容, 但未改变资产间的暴露关系。

将上述元数据与漏洞信息、网络配置信息融合, 构建状态转移概率矩阵。当加密流量中出现异常流模式或触发基于行为的告警时, 模型将信息融合结果作为观测序列的输入, 结合漏洞信息推断攻击路径。

HMM^[10]是一种描述时序状态的概率模型, 核心是刻画不可直接观测的隐藏路径^[11]状态的动态演变规律^[12]。但其在实际应用中存在一定弊端, 若网络入侵状态的真实概率为0, 而出现数量相同的入侵路径信息概率并不一定一样。针对传统HMM在处理稀疏数据时面临的零概率问题, 提出一种基于状态语义相似性的自适应概率重分配策略。该策略的核心思想是: 当一个状态转移在训练数据中未出现时, 不应简单地为其分配均匀概率, 而应借鉴与其语义相似的高频状态转移的概率分布。改进后的HMM概率计算式为:

$$P[x_{t+1} = O_j | q_t = s_i] = \begin{cases} \hat{P}[O_j | O_i]_t - \delta, \hat{P}[O_j | O_i]_d > 0 \\ \frac{\sum_{k \in K} \text{sim}(O_j, O_k) \hat{P}[O_k | O_i]}{\sum_{k \in K} \text{sim}(O_j, O_k)} \cdot \eta, \text{其他} \end{cases} \quad (7)$$

其中, t 为当前时刻的观测序列索引; q_t 为入侵状态变量^[13]; x_{t+1} 为告警信息变量; s_i 为入侵节

点； O_i 、 O_j 为路径 i 、 j 上的具体告警信息； $\hat{P}[O_j|O_i]$ 为最大似然估计得到的原始转移概率^[14-15]； δ 为从高频状态中抽取的概率总量； K 为所有高频状态（即 $\hat{P}[O_k|O_i] > 0$ 的状态）的集合； $\text{sim}(O_j, O_k) \in [0, 1]$ 为告警状态 O_j 与 O_k 之间的语义相似度，基于告警类型、源目的IP地址、端口等元数据特征计算得到； η 为归一化因子，确保概率总和为1。

为深入理解本文基于语义相似性的概率重分配策略的理论特性，将其与经典的平滑技术进行对比分析，并讨论其理论边界。

(1) Add-one 平滑。作为最简单的平滑方法，Add-one 平滑假设所有未观测事件具有相同的先验概率。其优点是实现简单，缺点是会显著低估高频事件的概率，且未考虑事件间的内在关联。本文方法通过引入语义相似性度量，实现了对未观测事件的差异化概率分配，避免了均匀分配可能引入的噪声。

(2) Good-Turing 估计。Good-Turing 估计是一种基于频率的频率（frequency of frequencies）的非参数平滑方法，其核心思想是利用出现 r 次的事件总数来估计未观测事件的概率。该方法在自然语言处理领域取得了巨大成功，但其理论假设是事件的出现服从泊松分布，且未考虑事件间的语义关联。相比之下，本文方法虽然缺乏Good-Turing那样的严格理论证明，但通过引入领域知识，在网络安全这一特定场景下具有更强的可解释性和针对性。

(3) Jelinek-Mercer 平滑。该方法将高阶模型与低阶模型进行线性插值。本文语义相似性加权可视为一种特殊的插值方法，其中插值权重由状态间的相似度动态决定，而非固定的超参数。

本文概率重分配策略的理论边界主要体现在以下3个方面。

(1) 相似性度量的有效性边界。该策略的有效性高度依赖于语义相似度 $\text{sim}(O_j, O_k)$ 的计算准确性。当相似性度量能够准确反映告警状态间的真实关联时，重分配能有效提升模型的泛化能力；反之，若相似性度量引入噪声，则可能加剧概率失真，导致错误的状态关联。因此，相似性度量的设计是该策略的理论核心与实际应用瓶颈^[16]。

(2) 数据稀疏性与高频状态集合的稳定性。当训练数据极度稀疏时，高频状态集合 K 本身可能已不具备统计显著性，此时基于这些状态的相似性加权可能无法提供可靠的概率估计，甚至可能放大采样偏差。极端情况下，若 $K = \emptyset$ ，该策略退化为均匀分配。这表明，在数据稀疏度过高的场景中，需引入外部知识库或先验分布对相似性加权进行约束。

(3) 时间序列的动态性适配。当前策略主要针对静态状态空间进行平滑，未显式考虑攻击路径的时间动态性。在多步攻击场景中，某些零概率转移虽然在静态统计中未出现，但在特定攻击时序上下文具有较高的先验概率。

基于式(3)~式(6)过程求得的路径状态转移概率关系，形成状态转移序列 $\{P_1, P_2, \dots, P_T\}$ ， P_T 为基于时间的状态转移系数^[17]。基于式(7)建立HMM，推测 $\{P_1, P_2, \dots, P_T\}$ 对应的入侵意图序列 $\{S_1, S_2, \dots, S_T\}$ 。

为有效区分有效入侵路径与噪声告警，本文在每个入侵状态转移时引入自适应动态阈值^[18]。若告警序列 $\{O_1, O_2, \dots, O_T\}$ 在时刻 t 对应某入侵意图的概率低于阈值 $\theta_t(j)$ ，则判定该序列中不存在真实的入侵行为，将其作为噪声或误报告警予以丢弃，不参与后续路径推断。若告警序列对应某入侵意图的概率高于或等于阈值 $\theta_t(j)$ ，则判定存在真实的入侵行为，保留该序列并输出入侵路径自适应识别结果。具体的实现步骤如下。

假设 t 时刻 $\{O_1, O_2, \dots, O_T\}$ 对应某入侵意图的概率 $\delta_t(j)$ ，计算式为：

$$\delta_t(j) = \left(\max_{1 \leq i \leq N_m} [\delta_{t-1}(j) a_{ij}] P_{t-1,i} \right) b_j^t(O_T) \quad (8)$$

其中， $\delta_{t-1}(j)$ 为下一时刻入侵意图概率， $P_{t-1,i}$ 为入侵状态转移概率， a_{ij} 为攻击意图的状态转移修正系数， $\theta_t(j)$ 为识别阈值， b_j^t 为 t 时刻下路径 j 的实时状态，其计算式为：

$$b_j^t = P(O_t | q_t = s_j) = \frac{\text{count}(O_t, s_j) + \varepsilon}{\sum_{k=1}^r \text{count}(O_k, s_j) + \tau \varepsilon} w_j^t \quad (9)$$

其中， $\text{count}(O_t, s_j)$ 表示在训练数据中，当隐藏状

态为 s_j 时观测到告警 O_i 的次数; τ 为告警类型的总数; ε 为拉普拉斯平滑系数, 用于避免零概率问题; w_j' 为告警可信度权重, 基于告警源的可靠性和实时威胁情报动态调整。

为有效区分有效入侵路径与噪声告警序列, 设计了一种基于概率分布统计特性的自适应动态阈值机制。该识别阈值 $\theta_T(j)$ 不是一个固定经验值, 而是根据历史告警序列的概率分布特征动态计算得出的, 其定义如下:

$$\theta_T(j) = \mu(\delta_{1:T}(j)) - \kappa\sigma(\delta_{1:T}(j)) \quad (10)$$

其中, $\delta_{1:T}(j) = \{\delta_1(j), \delta_2(j), \dots, \delta_T(j)\}$ 为路径 j 在时刻 1 到 T 的概率值序列; μ 和 σ 分别为该序列的均值和标准差; κ 为可调节的灵敏度系数。该公式的物理意义是: 将阈值设定在概率均值以下 κ 倍标准差的位置, 只有那些概率显著高于历史平均水平的告警序列才被判定为有效入侵路径。

识别阈值 $\theta_T(j)$ 的动态自适应更新遵循以下规则。

(1) 初始化阶段: 在模型部署初期, 若无历史数据可用, 采用保守的初始阈值 $\theta_0 = 0.3$ 。

(2) 滑动窗口更新机制: 随着系统运行, 模型维护一个长度为 L 的滑动窗口 $O_t = \{\xi_{t-L+1}(j), \dots, \xi_t(j)\}$, 记录最近 L 个时间点的概率值。每接收到新的告警序列并完成概率计算后, 窗口向前滑动, 利用窗口内的最新数据重新计算 μ 和 σ , 并依据式(10)更新阈值为:

$$\begin{cases} \mu_t = \frac{1}{L} \sum_{k=t-L+1}^t \delta_k(j) \\ \sigma_t = \sqrt{\frac{1}{L} \sum_{k=t-L+1}^t (\delta_k(j) - \mu_t)^2} \end{cases} \quad (11)$$

(3) 冷启动处理: 当滑动窗口内数据不足 L 时, 采用指数加权移动平均 (exponentially weighted moving average, EWMA) 方法, 使阈值能够快速收敛, 即:

$$\begin{cases} \mu_t = \alpha\delta_t(j) + (1 - \alpha)\mu_{t-1} \\ \sigma_t^2 = \alpha(\delta_t(j) - \mu_t)^2 + (1 - \alpha)\sigma_{t-1}^2 \end{cases} \quad (12)$$

其中, $\alpha \in (0, 1)$ 为衰减因子, 控制历史数据的衰减速度。

完成概率计算后, 将当前时刻路径 j 的概率 $\delta_t(j)$ 与动态阈值 $\theta_t(j)$ 进行比较。

若 $\delta_{t-1}(j) < \theta_T(j)$, 表明当前告警序列的概率显著低于识别阈值, 该序列中包含真实入侵行为的可能性较低, 判定为噪声或误报, 予以丢弃, 不参与后续的路径检测过程。

若 $\delta_{t-1}(j) \geq \theta_T(j)$, 表明当前告警序列的概率达到或超过阈值, 存在真实的入侵行为, 保留该序列并输出自适应识别结果 $\delta_t(j)$ 为:

$$\delta_t(j) = \left(\max_{1 \leq i \leq N_m} [\delta_{t-1}(j) a_{ij}] P_{t-1,i} \right) b_j^{t+1} \quad (13)$$

其中, b_j^{t+1} 为 $t+1$ 时刻下路径 j 的实时状态。

路径检测算法动态更新过程如图2所示, 展示了模型在时序维度上对入侵路径状态概率的迭代优化机制。该过程通过多轮状态转移系数修正与入侵意图概率计算, 实现了对告警序列的持续筛选与路径置信度的动态演化分析。由图2可知, 随着迭代次数增加, 有效入侵路径的概率逐步收敛于高置信区间, 而无效或干扰路径的概率则被快速抑制, 体现了模型对零概率问题的优化效果, 即通过最大似然概率重新分配, 增强了非零概率路径的区分度。同时, 入侵状态转移修正系数与动态阈值的协同作用, 确保了算法在复杂告警环境下对隐藏攻击节点的解耦能力, 显著降低了路径检测的模糊性。该动态更新机制是本方法实现高精度、自适应入侵路径检测的核心, 有效支撑了模型在多步攻击场景下对时序依赖关系的捕捉与意图推断。

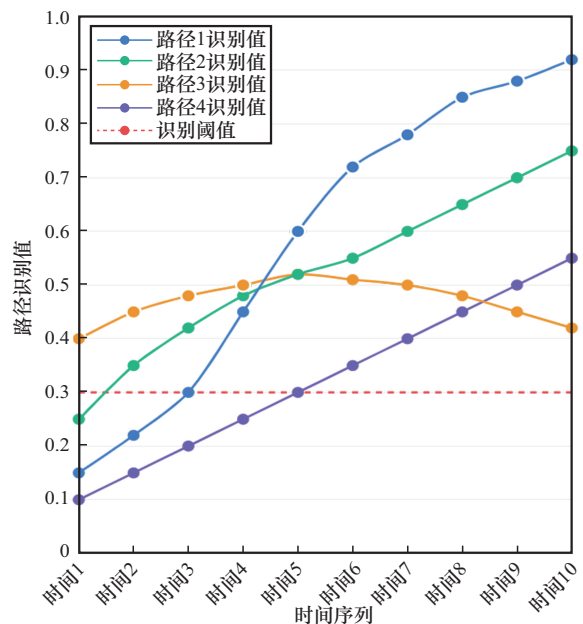


图2 路径检测算法动态更新过程

综合上述分析, 动态阈值驱动的入侵路径推断模块的关键伪代码实现过程如算法1所示。

算法1 关键伪代码实现过程

输入 告警序列 obs_seq , 状态转移概率矩阵 A , 观测概率矩阵 B , 滑动窗口长度 L

输出 入侵路径 $path$, 判定结果 $is_intrusion$

function PathInference(obs_seq, A, B, L):

1) // Viterbi 算法求解最大概率隐藏状态序列

$path = \text{Viterbi}(obs_seq, A, B)$

2) // 计算当前路径的概率

3) $prob = \text{ComputePathProbability}(path, obs_seq, A, B)$

// 滑动窗口动态阈值计算

4) $window_probs = \text{get_recent_probs}(L)$ // 获取最近 L 个序列的概率

5) $threshold = \text{mean}(window_probs) - \lambda * \text{std}(window_probs)$

// 判定是否为有效入侵

6) if $prob \geq threshold$:

$is_intrusion = \text{True}$

output $path$

7) else:

$is_intrusion = \text{False}$

discard sequence

8) return $path, is_intrusion$

该函数首先采用 Viterbi 算法对输入告警序列进行解码, 得到最大概率对应的隐藏状态序列作为候选入侵路径; 然后计算该路径的联合概率; 接着基于滑动窗口内历史概率序列的均值和标准差动态计算判定阈值; 最后将当前路径概率与阈值比较, 高于阈值则判定为有效入侵并输出路径, 否则作为噪声丢弃。

3 性能测试

3.1 测试环境

为验证本文联合改进隐马尔可夫模型的物联网入侵路径自适应识别研究的实用价值, 需搭建包含服务器、终端、路由器等目标网络设备的实验环境。借助漏洞扫描设备获取各设备的漏洞信息与网络配置数据, 利用流量采集设备实时捕获设备间的通信流量及告警日志。为确保实验对比的公平性, 所有参与对比的方法均在完全相同的

数据集、数据划分和预处理条件下进行评估。具体设置如下。

(1) 数据集划分

实验采用 5 折交叉验证策略, 将原始数据集随机划分为 5 个互斥的子集。每次实验使用其中 4 个子集作为训练集, 剩余 1 个子集作为测试集, 轮换 5 次直至每个子集均被用作测试集。最终结果为 5 次实验的平均值, 以消除数据随机划分带来的偏差。

(2) 数据预处理流程

所有输入数据均经过相同的预处理流程: 首先进行流量清洗, 去除无效连接和广播流量; 然后采用 Z-score 方法对数值型特征进行标准化处理; 接着进行告警聚合, 将短时间内来自同一源 IP 地址的同类告警聚合为单条告警以减少冗余; 之后采用长度为 10 的滑动窗口将告警流转换为固定长度的序列样本; 最后对少数类样本采用合成少数类过采样技术 (SMOTE), 以缓解类别不平衡问题。

(3) 对比方法的重现

为确保对比的公平性, 所有对照方法的代码均从作者官方仓库获取, 并在上述统一的数据划分和预处理条件下重新训练和测试。超参数设置: 隐状态空间维度为 8~30 个, 入侵类型覆盖多路径并发攻击、隐节点跳板攻击、拓扑变异攻击, 节点数量为 37 个, 观测噪声强度为 0.1 (表示在数据中添加了 10% 的随机干扰), 跨场景迁移参数为 0.2~0.8。实验环境架构如图 3 所示。

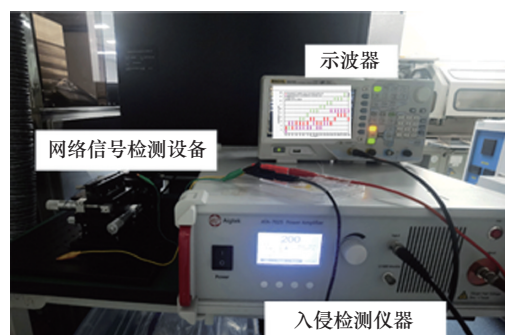


图3 实验环境架构

实验所用设备型号: 服务器采用 Dell Power-Edge R740 (CPU 为 Intel Xeon Gold 5220, 64 GB RAM), 终端设备采用 Raspberry Pi 4B (4 GB

RAM) 及 Intel NUC10i7FNH, 路由器采用 Mikro-Tik RB4011iGS+RM; 漏洞扫描设备使用 Nessus Professional 10.5.1, 流量采集工具使用 Tcpdump 4.99.1 及 Wireshark 3.6.5。软件版本: 服务器操作系统为 Ubuntu 20.04.6 LTS, 终端操作系统为 Raspberry Pi OS 11 (bullseye) 及 Windows 10 IoT Enterprise LTSC 2021; 攻击模拟工具包括 Metasploit Framework 6.3.0 及 Nmap 7.94。

3.2 算法时间复杂度分析

为验证算法有效性, 进行了算法时间复杂度与扩展性量化分析, 算法时间复杂度与运行时性能分析结果如表 1 所示。

从表 1 的实际运行时数据可以看出, 本算法的推断时延和内存开销与理论分析基本一致, 随状态数和序列长度增加呈平方级增长。在中小规模网络 (状态数 $N \leq 100$ 个, 序列长度 $T \leq 50$ 个时序步长) 中, 推断时延控制在 25 ms 以内, 满足实时检测需求。当状态数达到 300 个且序列长度达到 200 个时序步长时, 推断时延接近 900 ms, 已难以满足实时性要求。这表明, 在大规模网络直接应用原始算法存在性能瓶颈, 必须通过状态抽象和滑动窗口等优化策略实现扩展。

3.3 最大似然概率重分配策略有效性评估

为评估最大似然概率重分配策略在数据稀疏或高噪声场景下的鲁棒性, 本文设计了增加与经典平滑方法的对比实验, 不同数据条件下的对比实验如表 2 所示。

由表 2 可知, 在基准场景下, 3 种平滑方法的性能差异不大, 本文方法略优于经典方法。但在更具挑战性的数据稀疏和高噪声场景中, 本文

基于语义相似性的重分配策略展现出显著优势。相比 Good-Turing 估计, 本文方法在数据稀疏场景下的准确率提升了 2.4 个百分点, 状态混淆度降低了 0.04。这表明, 通过引入领域知识对概率重分配过程进行约束, 能够有效缓解经典平滑方法在面对非均匀分布状态空间时的局限性, 增强模型在复杂网络环境下的鲁棒性。

表 2 不同数据条件下的对比实验

场景类型	平滑方法	路径检测准确率	置信谱数均值	状态混淆度
基准场景	Add-one 平滑	92.8%	0.15	0.18
	Good-Turing	93.5%	0.14	0.17
	本文方法	94.2%	0.12	0.15
数据稀疏场景	Add-one 平滑	78.4%	0.35	0.48
	Good-Turing	80.2%	0.31	0.45
	本文方法	82.6%	0.28	0.41
高噪声场景	Add-one 平滑	75.6%	0.42	0.59
	Good-Turing	77.1%	0.38	0.55
	本文方法	79.8%	0.35	0.52

3.4 大规模网络可扩展性分析

为验证本文方法在大规模网络环境下的可扩展性, 设计了不同网络规模下的性能对比实验。实验在保持攻击路径数 (5 条) 和噪声强度 (0.1) 不变的条件下, 逐步增加网络节点数, 评估各项性能指标的变化趋势。大规模网络可扩展性结果如表 3 所示。

由表 3 可知, 当节点数从 50 个增至 200 个时, 准确率从 94.8% 小幅下降至 91.2%, 降幅约 3.6%, 表明本文方法在中等规模网络中仍能保持

表 1 算法时间复杂度与运行时性能分析结果

状态数/个	序列长度/个时序步长	理论计算量相对倍数	推断时延/ms	内存开销/MB	扩展性等级与可行性评估
30	50	1.00 (基准)	2.3	85	基准可行: 实验已验证, 计算效率高
50	50	2.78	5.8	142	轻度负担: 计算量增长约 2.8 倍, 仍可高效运行
100	50	11.11	24.6	312	显著负担: 计算量增长超 10 倍, 需开始关注性能优化
30	200	4.00	9.1	98	线性负担: 因序列增长, 计算量增至 4 倍, 实时性可能受影响
100	200	44.44	97.3	356	较重负担: 计算量增长显著, 接近实时处理能力的临界点
300	200	400.00	892.5	1 245	扩展瓶颈: 计算量激增 400 倍, 直接扩展不可行
50	500	9.26	21.4	178	中等负担: 主要由长序列导致, 适合离线深度分析
300	500	1 000.00	2 356.8	1 560	不可行: 计算负荷达到基准的 1 000 倍, 必须引入优化策略

较高的识别精度。当节点数达到1 000个时,准确率降至82.3%,降幅扩大,主要原因是状态空间膨胀导致参数估计难度增加;推断时延随节点数增加呈超线性增长。当节点数从50个增至200个时,时延从3.2 ms增至28.6 ms,仍能满足实时检测需求。当节点数达到1 000时,时延超过500 ms,已难以支持毫秒级实时响应。内存开销和CPU使用率随节点数增加而显著上升,当节点数达到1 000个时,内存消耗接近3 GB, CPU使用率接近饱和,表明在资源受限环境中可能面临挑战。

表3 大规模网络可扩展性结果

节点数/个	路径检测准确率	推断时延/ms	内存开销/MB	CPU使用率
50	94.8%	3.2	92	18%
100	93.5%	7.8	187	32%
200	91.2%	28.6	412	58%
500	87.6%	124.5	1 120	89%
1 000	82.3%	512.7	2 840	97%

3.5 物联网入侵路径自适应识别结果

为精准、直观地验证物联网入侵路径自适应识别的实际有效性,给出包含38个节点的网络拓扑结构以及入侵路径。设置改进深度学习的主动式通信网络入侵行为自适应识别方法(简称改进深度学习方法)、自适应类增量学习的物联网入侵检测方法(简称自适应类增量学习方法)为对照组,3种方法物联网入侵路径自适应识别结果如图4所示。

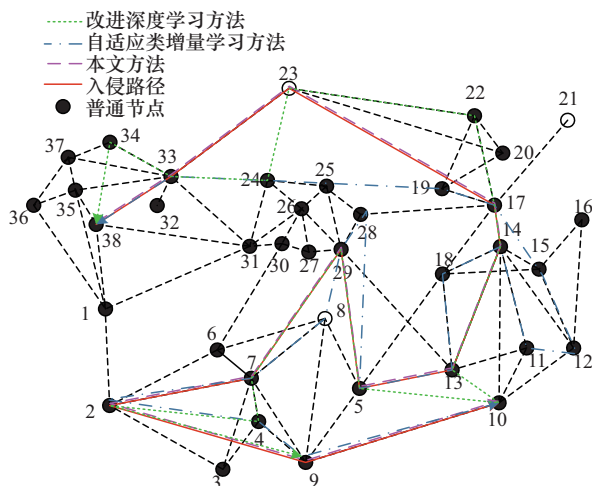


图4 3种方法物联网入侵路径自适应识别结果

从图4可以看出,本文方法在入侵路径的节点关联匹配、动态状态转移轨迹刻画方面表现出显著优势。其识别轨迹与真实入侵路径的重叠度显著高于对照方法,能够精准覆盖入侵行为涉及的关键节点;而其余方法的轨迹则未能完整捕捉多步入侵的状态转移逻辑。这表明,本文方法在多节点网络场景下具备更优的识别精度与动态适应性,为复杂网络环境中的多步入侵行为检测提供了更可靠的技术支撑。

3.6 入侵路径概率置信谱数

入侵路径概率置信谱数指识别方法输出的入侵路径概率分布,借助熵值来衡量算法对入侵路径的聚焦能力,以避免多路径模糊识别。数值越高,意味着算法求得的入侵概率分布分散于多条路径,难以确定哪条是最终入侵路径。3种方法的入侵路径概率置信谱数对比的实验结果如图5所示。

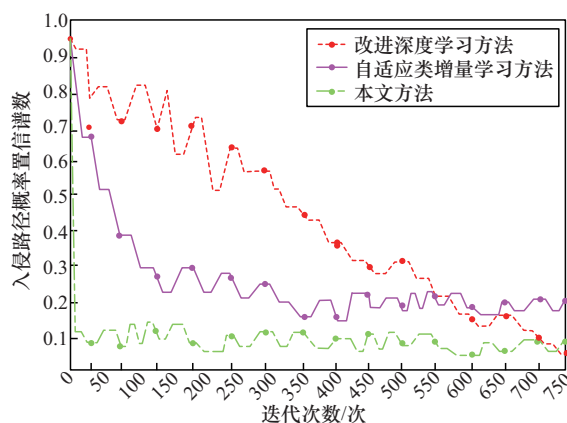


图5 3种方法的入侵路径概率置信谱数对比

由图5可以看出,本文方法的入侵路径概率置信谱数始终维持在0.1的低水平区间,且随迭代次数增加呈现平稳收敛趋势;而另外两种方法不仅置信谱数分布较高且收敛性能较差,需要多次迭代才能达到较低水平。这表明,本文方法的零概率参数优化策略能够有效聚合入侵路径的概率分布,大幅降低多路径检测的模糊性,在入侵路径概率分布的聚焦性和收敛稳定性方面均显著优于对照方法。

3.7 隐藏攻击节点解耦性

在实际网络环境中,恶意攻击者常通过伪造合法行为来误导攻击路径追踪,使有效识别变得

困难。为此,将隐藏攻击节点解耦混淆熵作为评价指标,以量化解耦决策的模糊性。熵值越低,说明该方法解耦攻击与正常节点边界的能力越强,边界越清晰。为增加真实性,设置伪造强度分级:低强度(伪造行为占比 $\leq 30\%$)、中强度($30\% < \text{占比} \leq 60\%$)、高强度(占比 $> 60\%$),伪造强度等级通过人为调整合法/攻击行为的时间占比实现。3种方法隐藏攻击节点解耦性如表4所示。

表4 3种方法隐藏攻击节点解耦性

方法	伪造强度等级对应的隐藏攻击节点解耦混淆熵		
	高强度	中强度	低强度
改进深度学习 学习方法	0.598±0.371	0.452±0.233	0.399±0.147
自适应类增量 学习方法	0.551±0.423	0.457±0.291	0.232±0.115
本文方法	0.225±0.179	0.154±0.099	0.196±0.048

由表4可以看出,针对高、中、低不同伪造强度的隐藏攻击节点场景,本文方法的解耦混淆熵均显著低于另外两种对比方法。这表明,本文方法对隐藏攻击节点伪造合法行为的特征具有更强的区分能力,能够有效降低解耦决策的模糊程度。即使在伪造行为占比更高、识别难度更大的高强度场景中,仍可清晰界定攻击节点与正常节点的边界。

3.8 加密流量环境下的性能评估

为验证本文方法在加密流量环境下的有效性,设计了加密场景对比实验。实验采用CIC-IDS-2017数据集中的加密流量部分,并引入OpenSSL工具对部分攻击流量进行TLS加密处理,构建混合加密环境。本文方法与对照方法在加密流量环境下的性能对比如表5所示。

表5 加密流量环境下的性能对比

方法	流量类型	路径检测准确率	置信谱数均值	推断时延/ms	误报率
改进深度学习 学习方法	明文	88.3%	0.24	18.5	7.2%
	加密	67.5%	0.42	19.2	15.8%
自适应类增量 学习方法	明文	86.7%	0.26	15.3	8.1%
	加密	64.2%	0.45	16.1	17.3%
本文方法	明文	94.2%	0.12	3.2	3.5%
	加密	89.3%	0.19	3.8	5.2%

由表5可知,加密流量对所有方法的性能均造成一定程度的下降,但下降幅度存在显著差异。其中,改进深度学习方法和自适应类增量学习方法在加密环境下的准确率下降幅度分别为20.8%和22.5%,误报率大幅上升。这些方法高度依赖从明文载荷中提取的深度特征,加密后关键特征丢失,导致模型失效。本文方法在加密环境下的准确率为89.3%,相比明文环境仅下降4.9%,误报率从明文环境的3.5%上升至5.2%。性能下降幅度远小于对照方法,验证了本文方法对加密流量的适应性。本文方法在加密环境下仍能保持较高性能,得益于模型主要基于流特征、告警元数据、漏洞信息进行推理,这些信息在加密环境下仍然可用;模型建模攻击行为的时序逻辑,加密仅隐藏内容,不改变攻击的阶段特征;漏洞信息和网络配置信息提供了攻击路径的先验知识,补偿了加密导致的信息损失。

为深入理解加密环境下告警生成机制的变化,分析了加密前后各类告警的分布变化,加密前后告警类型分布变化如表6所示。

表6 加密前后告警类型分布变化

告警类型	明文环境 告警量/个	加密环境 告警量/个	变化率	是否依赖解密
基于签名的SQL注入告警	1 245	86	-93.1%	是
基于行为分析的异常流量告警	342	578	+69.0%	否
基于TLS握手的异常告警	0	213	新增	否
基于连接频率的扫描告警	567	601	+6.0%	否
基于端口的非标准服务告警	123	145	+17.9%	否

由表6可以看出,加密环境下,依赖载荷解析的签名类告警数量锐减,而基于行为分析的告警数量显著增加或保持稳定。这表明,加密环境下的告警生成机制确实发生了变化,从“内容匹配”转向“行为分析”。本文方法恰好适应了这一变化——模型输入正是基于行为的告警元数据、流特征和上下文信息,而非签名匹配结果。因此,加密非但未削弱方法的有效性,反而使方法的优势更加凸显。

4 结束语

为提高网络运行的稳定性和安全性，提出联合改进隐马尔可夫模型的物联网入侵路径自适应识别方法。通过四元组定义网络入侵图，整合目标网络配置与漏洞信息，量化状态节点间的转移概率，构建动态更新的状态转移概率矩阵，实现了入侵路径的概率化描述，可适配漏洞有效性与利用成功率的实时波动。针对传统隐马尔可夫模型的零概率缺陷，通过最大似然概率分配优化模型，提升了对非零概率路径的刻画精度。引入入侵状态转移修正系数与识别阈值，设计多步骤迭代计算逻辑，有效过滤无效告警序列，强化了模型对入侵意图的判别能力。实验数据证明，本文方法在提升识别准确率的同时，保障了对复杂入侵行为的快速收敛的高效性。面对未知攻击，模型通过识别告警事件在状态图中形成的低概率异常转移路径进行检测，其基础是对网络正常行为与漏洞关系的全面描述。对于加密流量，模型不解析内容，而是依据告警元数据、资产漏洞及拓扑关系进行推理，这些信息在加密环境下仍可获得。加密改变了内容但未改变资产间的暴露关系和访问逻辑，因此模型在此环境下依然有效，这种从具体特征向行为逻辑的抽象，正是该方法应对新型威胁与复杂环境的关键优势所在。

参考文献：

- [1] 伍均玺, 林峰, 高红云. 基于改进深度学习的主动式通信网络入侵行为自适应识别算法[J]. 微型电脑应用, 2024, 40(4): 9-12.
Wu J X, Lin F, Gao H Y. An Adaptive recognition algorithm for intrusion behavior in active communication networks based on improved deep learning[J]. Microcomputer Applications, 2024, 40(4): 9-12.
- [2] 刘强, 张颖, 周卫祥, 等. 自适应类增量学习的物联网入侵检测系统[J]. 计算机工程, 2023, 49(2): 169-174.
Liu Q, Zhang Y, Zhou W X, et al. Adaptive class incremental learning-based IoT intrusion detection system[J]. Computer Engineering, 2023, 49(2): 169-174.
- [3] Mansahaj S P, Rudra P S, Navneet K P, et al. A federated learning framework for enhanced data security and cyber intrusion detection in distributed network of underwater drones[J]. IEEE Access, 2025, 13: 12634-12646.
- [4] Lu Q Y, Gao Q Z, Li J N, et al. Distributed cyber-physical intrusion detection using stacking learning for wide-area protection system[J]. Computer Communications, 2024, 215(6): 9151-1023.
- [5] 李志辉, 王德军, 孙贝尔. 基于知识图谱与HMM的入侵路径分析方法[J]. 中南民族大学学报(自然科学版), 2025, 44(2): 226-236.
Li Z H, Wang D J, Sun B R. Intrusion path analysis method based on knowledge graph and HMM[J]. Journal of South-Central Minzu University(Natural Science Edition), 2025, 44(2): 226-236.
- [6] Bouzaachane K, Guarmah E M E, Alnajim A M, et al. Addressing modern cybersecurity challenges: a hybrid machine learning and deep learning approach for network intrusion detection[J]. Computers, Materials & Continua, 2025, 84(2): 2391-2410.
- [7] 杨定坤, 尹徐珊, 孙澄宇, 等. 隐马尔可夫模型下电力信息网络入侵检测方法研究[J]. 自动化仪表, 2025, 46(5): 94-99.
Yang D K, Yin X S, Sun C Y, et al. Research on intrusion detection method in power information network under hidden Markov model[J]. Process Automation Instrumentation, 2025, 46(5): 94-99.
- [8] 于洪, 朱正彬, 魏帅, 等. 基于异构执行体同步控制的入侵检测模型研究[J]. 网络与信息安全学报, 2025, 11(1): 54-65.
Yu H, Zhu Z B, Wei S H, et al. Research on intrusion detection model based on synchronization of heterogenous executer[J]. Chinese Journal of Network and Information Security, 2025, 11(1): 54-65.
- [9] 蒋玉长, 蒋利兵, 黄凯. FL-IoV: 基于联邦学习的综合车联网入侵检测方法[J]. 网络与信息安全学报, 2025, 11(4): 127-140.
Jiang Y C, Jiang L B, Huang K. FL-IoV: comprehensive intrusion detection method for internet of vehicles based on federated learning[J]. Chinese Journal of Network and Information Security, 2025, 11(4): 127-140.
- [10] Nandhini U, Kumar S V. An enhanced whale optimizer based on feature selection technique with effective ensemble classifier for network intrusion detection system[J]. Peer-to-Peer Networking and Applications, 2025, 18(2): 1-28.
- [11] 朱静雯, 许静, 陈亮, 等. 基于隐马尔可夫模型的SQL注入攻击行为检测[J]. 计算机应用与软件, 2023, 40(2): 331-338, 344.
Zhu J W, Xu J, Chen L, et al. SQL injection attack detection based on hidden Markov model[J]. Computer Applications and Software, 2023, 40(2): 331-338, 344.
- [12] Naeem H, Ullah F, Srivastava G. Classification of intrusion cyber-attacks in smart power grids using deep ensemble learning with metaheuristic-based optimization[J]. Expert Systems, 2025, 42(1): 2354-8531.
- [13] Bozorg C A, Zarakovitis C C, Ting T O, et al. Novel modeling and optimization for joint cybersecurity-vs-QoS intrusion detection mechanisms in 5G networks[J]. Computer Networks, 2023, 237(2): 1-11.
- [14] 熊文定, 罗凯伦, 李睿. 基于异常的终端级入侵检测[J]. 计算机系统应用, 2023, 32(2): 181-189.
Xiong W D, Luo K L, Li R. Anomaly-based terminal-level intrusion detection[J]. Computer Systems & Applications, 2023, 32(2): 181-189.
- [15] Zhao J X, Zheng W. Cybersecurity entity recognition method combined with feedforward neural network enhancement[C]//Proceedings of the 3rd International Conference on Machine Learning, Cloud Computing and Intelligent Mining (MLCCIM2024). Berlin: Springer, 2025: 95-104.

- [16] Wang W Y, Li Q, Mu H L. Exploring malware complexities: a behavior and characteristic analysis approach for robust and accurate cybersecurity[J]. Cluster Computing, 2024, 28(2): 82.
- [17] Zhukabayeva T, Ahmad Z, Adamova A, et al. Penetration testing and machine learning-driven cybersecurity framework for IoT and smart city wireless networks[J]. IEEE Access, 2025, 13(9): 86144-86166.
- [18] 胡威, 庞九凤, 张亚昊, 等. 一种基于隐马尔可夫链与分类赋权的网络安全态势评估模型研究[J]. 网络安全技术与应用, 2023(9): 29-33.
- Hu W, Pang J F, Zhang Y H, et al. Research on a network security situation assessment model based on hidden Markov chain and clas-

sification weighting[J]. Network Security Technology & Application, 2023(9): 29-33.

[作者简介]



何洋军 (1988-), 男, 浙江省公众信息产业有限公司高级工程师, 主要研究方向为网络安全、数据安全。