基于半监督自编码器的网络流量异常检测关键技术研究

易 伟，孙 惠

（郑州科技学院，河南 郑州 450064）

摘 要 随着物联网（IoT）设备数量的爆炸式增长，安全问题日益凸显。设备精准识别与流量异常检测是构建物联网安全体系的核心基础，但实际中异常样本稀缺，传统监督学习因样本不均衡性能受限。本文针对物联网设备识别与异常检测中的关键技术进行系统性研究，提出了一种基于一维卷积神经网络（1D CNN）的设备自动识别方法和半监督深度自编码器异常检测模型，通过编码器—解码器学习正常流量潜在特征，用少量标记样本微调编码器后的分类器，形成重构误差与分类损失的综合目标函数，解决异常样本稀缺问题。

关键词 物联网安全；异常检测；半监督学习；深度自编码器；数据增强

基金项目：河南省科技厅科技攻关项目“基于深度学习的物联网设备识别异常检测关键技术研究”（项目编号：242102210094）；河南省教育厅高等学校重点科研项目“基于大数据采集的罗德相关失效模型在高速切削分析中的应用研究”（项目编号：24B460028）。

中图分类号：TN76 **文献标志码**：A **DOI**：10.3969/j.issn.2097-3365.2025.31.001

0 引言

物联网（Internet of Things，IoT）作为新一代信息技术体系的核心构成单元，近年来呈现高速发展态势。随着物联网设备数量的爆发式增长，网络攻击面呈几何级拓展。受限于设备算力与成本控制等因素，终端安全防护能力普遍薄弱，安全漏洞持续暴露，给物联网生态带来巨大压力。在物联网安全研究领域，设备识别和异常数据检测至关重要。传统方法依赖人工提取特征，效率低且准确性难以保证，无法适应设备数量众多、种类繁杂的特点。深度学习技术在图像分析、语音识别等领域的突破性成果，为解决物联网设备识别和异常检测问题提供了新思路[1]。本文研究基于深度学习的物联网设备识别和异常检测关键技术，旨在为提高物联网设备的安全性和可靠性提供参考。

1 物联网设备识别与异常检测

1.1 物联网流量特征分析

物联网流量是由数据包按时间顺序传输而形成的序列。从网络协议栈的角度来看，物联网流量涵盖了IP、TCP、UDP等基础协议，同时也广泛应用CoAP、MQT等轻量级应用层协议[2]。数据包都包含头部和有效载荷两部分，头部携带丰富的控制信息，构成了设备通信的基础特征。

通过对数据包长度分布的研究发现，不同设备的数据包长度具有显著特征。智能门锁的控制命令数据包通常较短，而智能摄像头的视频数据包则相对较长。传输频率也是重要分析指标，恒温器会周期性地发送温度数据，传输频率稳定；智能门锁在用户操作时会产生突发流量。

时间序列分析则侧重于研究流量随时间的变化规律，通过分析历史流量数据构建设备正常行为的基线。智能音箱流量在一天中的不同时间段会呈现出不同的模式，晚上使用频率较高时流量较大。通过建立时间序列模型，可以准确地描述这种流量变化规律。

1.2 深度学习基础技术

1.2.1 卷积神经网络（CNN）

CNN通过局部连接、权值共享和池化操作，能高效自动提取数据中的空间特征[3]。一维CNN尤其适合处理时序数据，如数据包长度序列。卷积核在序列上滑动可自动提取关联特征。池化层进行下采样，保留显著特征，全连接层最终完成分类。

1.2.2 梯度提升树（LightGBM）

LightGBM作为GBDT框架下的高效机器学习模型，适用于高维度、大样本量的数据建模。通过多轮迭代训练，将多个弱分类器集成为强分类器，实现对复杂数据模式的精准拟合。

1.2.3 深度自编码器（DAE）

深度自编码器采用“编码器—解码器”双模块协同的网络架构，通过数据的“压缩—重构”过程实现特征学习与信息表征[4]。

1.3 性能评估指标体系

设备识别常用指标包括准确率、精确度、召回率、AUC和AUPRC。异常检测常用指标包括AUC、F1分数、误报率和漏报率。这些指标共同构成完整的模型性能评估体系。

2 局域网环境下的物联网设备识别方法

2.1 研究现状与挑战

传统局域网设备识别依赖于MAC地址和设备指纹。MAC地址易被篡改，设备指纹因加密通信和软硬件动态变更而稳定性差。深度学习技术在物联网设备识别应用中存在计算成本高、泛化性差等问题。

2.2 基于一维卷积神经网络的特征提取

2.2.1 数据预处理与特征构建

通过网络抓包工具，在局域网内的关键节点采集物联网设备通信流量数据。预处理阶段提取每个会话的数据包长度序列作为核心特征。将会话数据包长度序列归一化为固定长度，通过滑动窗口生成多个样本，充分利用序列信息。

原始包长序列保留了完整的时序信息，能直接反映物联网设备通信行为模式，避免了特征工程中的信息损失。

2.2.2 网络架构设计

基于一维CNN构建物联网设备识别模型，由输入层、卷积层、池化层和全连接层组成。输入层接收预处理后的数据包长度序列。卷积层采用32个大小为5的卷积核，自动提取局部关联特征。池化层压缩特征维度并保留关键信息。全连接层整合特征并完成设备类型分类。

2.3 实验验证与性能分析

在物联网设备识别研究领域，目前有三个公开可获取的真实设备数据集，它们的具体信息如下：2016年由Aalto大学构建的数据集，包含27台设备；同年发布的UNSW数据集，涵盖28台设备；以及2019年推出的IoTFinder数据集，涉及51台设备。

在实验挑选数据集时，对这三个数据集进行了全面考量，结合本研究对数据规模、特征维度及场景适配性的需求，最终确定采用UNSW数据集，保障模型训练样本分布均衡，采用精确度、召回率和F1分数作为核心评价指标。实验结果显示，该模型的识别准确率均稳定在99%以上，验证了其在物联网设备识别任务中的有效性与泛化能力。

3 网络地址转换后的物联网设备识别方法

3.1 NAT环境下的识别难点

在家庭网络中，NAT设备实现多个内部设备共享一个公网IP地址，无法直接通过MAC地址或IP地址来区分不同物联网设备。更严峻的是数据分布极度不平衡，存在已知漏洞的特定型号设备占比极小。

3.2 基于焦点损失的LightGBM模型优化

3.2.1 数据增强与特征工程

从NAT转换后的流量数据中提取源端口分布、目标IP地域熵、TCP连接时长等30维统计特征。针对数据分布失衡问题，引入SMOTE过采样策略平衡样本分布。通过对少数类样本特征向量进行邻域搜索，构建保留核心特征属性又具备适度特征差异的虚拟样本。

3.2.2 模型训练与参数调优

选择LightGBM作为基础模型，集成焦点损失函数[5]。采用贝叶斯优化算法进行参数调优，确定学习率为0.01，叶子节点数为31，焦点损失函数中的α为0.8，γ为2，此配置使模型更加关注少数类样本，提高对稀有设备型号的识别能力。

3.3 实验部分

与某大型运营商合作，在其家庭网络中进行了实际部署和数据采集。通过在运营商的网络边缘设备上部署流量采集工具，采集NAT出口流量数据。在数据标注阶段，重点标注了含漏洞的摄像头型号，某品牌的V1.0版本智能摄像头存在安全漏洞，将其作为目标类别进行标注。在采集到的10万用户流量数据中，该型号的摄像头样本数量占比仅为0.7%，属于典型的少数类样本，符合实际应用中具有安全漏洞设备型号占比较低的情况。

为系统验证所提方法的性能优势，本文将实验结果与现有相关研究展开对比分析，验证了该方法在真实不平衡场景下的有效性，实验结果对比见表1。

4 基于半监督学习的物联网流量异常检测

4.1 异常检测场景与数据困境

在实际的物联网应用中，异常样本稀缺、传统监督学习因缺乏正样本而表现不佳。无监督方法对复杂、隐蔽的异常模式泛华能力不足，难以检测高级持续威胁攻击。

表1 实验结果对比

设备	AUC（本文）	AUPRC（本文）	AUPRC
webcam.Amcrest.IPM-721W	1.0 000	1.0 000	1.0 000
socket.Wemo.Insight	1.0 000	1.0 000	1.0 000
webcam.Samsung.SNH-1011N	1.0 000	0.9 999	0.9 990
streamer.Amazon.Fire_TV_Stick	0.9 999	0.9 995	0.9 920
light_bulb.TP_Link.LB130	1.0 000	1.0 000	1.0 000
speaker.Sonos.One	0.9 998	0.9 975	0.9 520
doorbell.Amazon.Ring	1.0 000	0.9 999	0.9 890

4.2 深度自编码器与半监督训练策略

4.2.1 模型架构设计

编码器采用3层全连接层，结构为256→128→64。将原始流量特征压缩至64维低维隐空间，解码器与编码器对称，将低维特征还原为原始特征。在编码层后接入Softmax分类器，利用少量标记异常样本进行有监督训练[6]，实现异常流量的精准分类。最终构建半监督目标函数包含重构误差和分数误差两个关键组件，通过权重系数平衡两者的贡献度，提升分类准确率。

4.2.2 训练流程优化

先用10万条筛选后的正常物联网流量数据，通过深度自编码器的编解码结构无监督预训练，最小化输入与重构输出的均方误差（MSE），使模型掌握正常流量的潜在分布与模式。随后引入占总样本4%的标记异常样本进行有监督微调，联合优化重构损失与分类损失：前者保障正常流量特征表达能力，后者借交叉熵引导精准识别异常。在线检测时，实时计算新流量重构误差，超基于正常流量统计特征的动态阈值即判为异常，结合分类器输出概率双重验证，提升检测可靠性。

4.3 数据集与对比实验

为应对样本量不平衡问题，本文提出一种流量数据扩展策略，具体可采用以下单一方法或多种方法组合的形式实现：

1. 时间序列反转：针对流量数据时序依赖性强、模型易过度拟合单一时间流向的问题，采用时间序列逆向处理策略：将24小时历史流量序列沿时间轴翻转，使原始序列的起始片段转化为终止片段、终止片段转化为起始片段。

2. 特征变换：考虑到流量数据特征（如数据包大小、传输速率）常存在分布偏态问题，对核心特征实施针对性数学变换：包括对数转换（用于压缩数值跨度较大的特征）、平方根运算（缓解右偏分布特征的极端值影响）、指数变换（增强低数值特征的区分度）等。

3. 时空变换：若流量数据携带有地理空间标识（如设备接入位置的经纬度坐标），则针对空间维度实施时空变换：对地理坐标施加微小随机扰动（如基于高斯分布的±0.001经纬度偏移），模拟不同地理位置（如同一区域内不同家庭、同一楼栋不同楼层）的设备接入场景。

4. 采样变换：针对时间序列类流量数据的“多分辨率分析需求”，采用多时间尺度重采样策略：一方面通过下采样将分钟级流量数据聚合为小时级（如按60分钟滑动窗口求和），降低时间分辨率以凸显宏观流量趋势；另一方面通过上采样（如线性插值）将小时级数据拆解为分钟级，提升时间粒度以捕捉微观流量波动。

5. 噪声注入：为模拟现实场景中数据不确定性（如物联网设备传感器的采集误差、无线传输链路的电磁干扰），向流量数据中注入符合实际扰动特性的随机噪声。

6. 数据切片：采用时间片段化重组方法，将完整的流量时序数据（如24小时序列）切割为若干等长独立子片段（如每2小时1个片段），再通过随机排列子片段的顺序构建新的样本序列（如将“片段1—片段2—片段3”重组为“片段3—片段1—片段 2”）。

7. 缺失模拟：针对实际应用中“数据传输中断、设备离线导致的流量数据缺失”问题，采用缺失场景模拟策略：基于掩码机制随机掩盖部分时间步（如随机选取10%的时间点）或特征维度（如随机掩盖“数据包重传率”特征）的数值，构建含缺失值的流量样本。表2为训练集和测试集中不同流量类型的样本数量列表。

N-BaIoT数据包含9类物联网设备的真实流量数据，其中7台设备同时受到BASHLITE与Mirai两种僵尸网

表2 训练集和测试集中不同流量类型的样本数量列表

报文类型	协议类型	非恶意报文	恶意报文
ClientHello报文	密码套件	提供密码套件0x002f(TLS_RSA_WITH_AES_128_CBC_SHA)	提供3个过时的密码套件，包括0x0004(TLS_RSA_WITH_RC4_128_MD5)
	扩展块	大多数TLS流中支持0x000d，以及0x0005（状态请求）、0x3374（下一个协议协商）、0xff01	客户端支持的TL扩展块值相同
ClientKeyExchange报文	密钥长度	256位椭圆曲线密码作为公钥	2 048位RSA公钥
ServerHello报文		—	过时的密码套件
Certificate报文		0.1%为自签名	70%是自签名

络的感染，另外2台则仅被BASHLITE单一恶意程序入侵。该数据集提供正常网络流量与异常恶意流量，涵盖100毫秒、500毫秒、1.5秒、10秒和1分钟五种时间窗口的采样数据。其23维特征分为四类：第一类包含8个特征，可捕捉单方向数据包的传输规模特性；第二类涵盖4个特征，聚焦于单位时间内数据包总量的统计需求，形成用于量化数据传输频次的计数特征；第三类涉及3个特征，侧重描述数据包传输过程中时延波动的特性，构建反映数据传输稳定性的抖动特征；第四类包含8个特征，可同时关联入站与出站双向数据包的传输规模，实现对双方向数据交互体量的综合表征。上述所有超参数的具体取值细节已系统汇总于表3，为后续实验复现、结果验证及方法改进提供清晰参考。

表3 超参数的取值

超参数	值
批量大小	128
预训练迭代次数	100
迭代次数	50
预训练学习率	0.001
学习率	0.001
权重衰减系数 λ	1e-6
平衡因子n	1

在半监督学习的研究范式中，存在三个关键调控参数。第一个参数为训练集中标记样本的占比，$\gamma_l=m/(n+m)$，其核心作用是调控标记样本在训练数据中的分布权重：当该占比取值为0时，半监督学习范式将退化为无监督学习；取值为1时，则完全转化为监督学习，二者分别对应不同的样本依赖场景。第二个参数是无标记样本的污染率 γ_p，γ_p 具体指无标记训练样本中异常样本的占比，γ_p 的值越趋近于1，表明无标记样本中异常样本的掺杂程度越高，模型对“正常—异常”特征边界的学习易受干扰，进而导致异常检测任务的难度显著提升。第三个参数为标记样本所含的异常类数量 k_l。

5 结束语

本研究针对物联网安全监控需求，提出设备识别与异常检测完整技术方案。在局域网中，1D-CNN设备识别模型通过提取数据包长度序列深度特征，实现各类设备精准识别，准确率超99%；在NAT网络中，创新用焦点损失优化LightGBM，解决数据不平衡问题，显著提升稀有漏洞设备检测能力。此外，结合深度自编码器与半监督学习，构建适配异常样本稀缺场景的流量检测模型，有效识别高级持续威胁。成果可服务智能家居、工业物联网等场景，保障用户隐私财产安全，助力运营商精准管理设备资产、预警隐患，为物联网安全生态提供技术支撑。

参考文献：

[1] 王海峰,吴军,郑波.物联网设备指纹特征构建与识别方法[J].通信学报,2021,42(04):156-167.
[2] 刘建辉,孙志刚,肖建国.物联网设备流量行为分析与特征提取[J].计算机学报,2019,42(06):1254-1268.
[3] 王鹏,李静,张涛.基于一维卷积神经网络的物联网设备识别方法[J].计算机研究与发展2021,58(03):643-654.
[4] 张强,王敏,周昊,基于深度自编码器的物联网异常检测模型[J].软件学报,2020,31(08):2456-2470.
[5] 刘伟,陈明,赵磊.面向不平衡数据集的物联网设备识别研究[J].计算机工程,2022,48(05):123-130.
[6] 张明,李娜,黄伟.半监督学习在网络异常检测中的应用研究[J].计算机应用研究,2020,37(09):2742-2746.