网络行为孪生驱动的物联网异常流量检测

何高峰[1]，田健峥[1]，李亚文[1]，徐丙凤[2]，朱海婷[1]，张璐[3]，郭乃瑄[4]
（1.南京邮电大学物联网学院，江苏 南京 210003；2.南京林业大学信息科学技术学院、人工智能学院，江苏 南京 210042；
3.南京审计大学计算机学院、统计金融联合实验室，江苏 南京 211815；4.盐城工学院信息工程学院，江苏 盐城 224007）

摘　要：针对现有物联网异常流量检测主要依赖于机器学习或深度学习算法，不仅资源消耗高，还易产生大量误报的问题，提出一种基于网络行为孪生的异常流量检测方法。该方法利用大语言模型从设备源码中自动提取网络规则，构建物联网设备的网络行为数字孪生模型，并以此为基础实时模拟设备的正常网络行为，实现对异常流量的高效检测。实验结果表明，所提方法在拒绝服务攻击、命令与控制通信以及内网扫描等典型场景下的检测任务中，检测效果均显著优于现有检测方法。与最新的预训练模型 TrafficFormer 相比，模型大小由 682 MB 降至 17 KB，计算和存储资源消耗分别降低 85.44% 和 94.06%。所提方法兼具高检测精度与边缘部署能力，适用于资源受限的物联网环境，为物联网网络安全提供了虚实结合的动态防护新思路。

关键词：数字孪生；大语言模型；异常流量检测；物联网安全；轻量化模型
中图分类号：TP391
文献标志码：A
DOI：10.11959/j.issn.1000–436x.2026040

Network behavior twin-driven traffic anomaly detection for the Internet of things

He Gaofeng[1], Tian Jianzheng[1], Li Yawen[1], Xu Bingfeng[2], Zhu Haiting[1], Zhang Lu[3], Guo Naixuan[4]
1. School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
2. College of Information Science and Technology & Artificial Intelligence, Nanjing Forestry University, Nanjing 210042, China
3. School of Computer Science and Technology and Joint Lab for Statistics and Finance, Nanjing Audit University, Nanjing 211815, China
4. School of Information Engineering, Yancheng Institute of Technology, Yancheng 224007, China

Abstract: To overcome the limitations of existing Internet of things (IoT) traffic anomaly detection methods, which predominantly rely on machine or deep learning algorithms and thus incur high resource consumption and frequent false positives, a novel detection framework based on the network behavioral twin was proposed. The proposed method harnessed large language model (LLM) to automatically extract network interaction rules from device source code, thereby constructing a digital twin that accurately mirrors the IoT device's network behavior. This digital twin was employed to simulate the device's legitimate network activities in real time, enabling precise and efficient detection of anomalous traffic. Experimental results demonstrate that the proposed method significantly outperforms existing detection methods in detection tasks under typical scenarios, such as denial of service (DoS) attacks, command and control (C&C) communication, and intranet scanning. Meanwhile, compared with the latest pre-trained model TrafficFormer, the model size is reduced from 682 MB to 17 KB, and the computation and storage resource consumption are reduced by 85.44% and 94.06%, respectively. By combining high detection accuracy with exceptional computational efficiency, the proposed method is well-suited for resource-constrained IoT environments and establishes a new approach of dynamic, cyber-physical protection in IoT network security.

Keywords: digital twin, large language model, traffic anomaly detection, Internet of things security, lightweight model

收稿日期：2025–12–02；修回日期：2026–02–06
通信作者：徐丙凤，bingfengxu@njfu.edu.cn
基金项目：国家自然科学基金资助项目（No.62572252, No.62372240）；江苏省高等学校自然科学研究重大基金资助项目（No.22KJA520005）；统计金融联合实验室开放课题基金资助项目（No.2025JLSF302）
Foundation Items: The National Natural Science Foundation of China (No.62572252, No.62372240), Key Project of Natural Science Research in Jiangsu Provincial Colleges and Universities (No.22KJA520005), The Open Project of Joint Lab for Statistics and Finance (No.2025JLSF302)

0　引言

随着硬件芯片、网络通信以及人工智能技术的快速发展，物联网（Internet of things, IoT）已经成为连接物理世界与数字世界的桥梁。从智能家居到工业自动化，从健康医疗到环境监测，物联网的应用无处不在，极大地提升了人们的生活质量和工作效率。然而，物联网的安全问题也日益凸显，这是因为：物联网设备通常计算能力有限、存储空间较小，难以安装和运行复杂的安全防护软件；在工业制造等领域，为保证生产过程的稳定性，物联网设备大多不具备在线升级功能，从而难以修正系统漏洞[1]；物联网设备普遍存在弱口令问题，一般用户对这些口令的存在及其修改方式往往缺乏了解[2]。同时，随着加密传输成为网络主流，黑客可进一步利用加密通道隐藏攻击特征，导致传统基于特征匹配的恶意攻击检测失效。这些均为黑客发动远程网络攻击提供了可乘之机，给物联网安全防护带来了巨大挑战。

为保障物联网安全，在网络中进行异常流量检测是一种有效方法。该方法可部署于物联网边缘网关处，运行时不需要或仅需终端设备的少量参与，且能够在攻击初期提供预警，有效防止实质性危害的发生。因此，物联网环境中的异常流量检测已成为当前研究热点。对现有异常流量检测方法进行概括分类，可以分为基于机器学习[3]（machine learning, ML）、深度学习[4]（deep learning, DL）以及预训练模型[5]等方法。但在实际应用中，这些方法均存在一个共同挑战，即误报数量大，例如，在一个中等规模的大学校园网络中，一天的误告警数量就多达 24 000 条[6]。大量误报使网络安全管理人员疲于应对，制约了此类方法的实际部署。

分析误报产生的原因，一个重要因素是现有方法都是基于网络流量数据集训练检测模型，而训练数据集与实际应用环境之间通常存在显著的分布漂移，这违背了许多机器学习算法所依赖的独立同分布（independent and identically distributed, IID）前提[7]。具体地，分布漂移可体现为数据漂移[8]和概念漂移[9]。数据漂移即真实网络流量的统计特征（如报文长度分布、流量大小等）会随时间、业务场景的变化而改变，与静态的训练数据产生差异。概念漂移是指网络攻击或正常行为的定义本身发生了变化。例如，当一种新型攻击手法将其行为特征伪装成正常物联网应用时，数据特征及其背后代表的真实概念（即“恶意”或“正常”）之间的对应关系便发生了改变。这使模型基于旧概念学到的知识失效，进而产生误报。此外，训练数据集本身往往存在覆盖不全的问题，如样本类别不均衡、攻击类型和正常行为模式的代表性不足，这进一步加剧了模型在真实环境中泛化能力的不足，最终导致较高的误报率。

基于上述分析，本文提出以下研究问题：不依赖网络流量数据集能否实现有效的（加密）流量异常检测？对此，本文证实了其可行性，提出了一种基于网络行为孪生的异常流量检测方法。该方法受物联网数字孪生技术[10]启发，其核心思路是构建设备的网络行为数字孪生模型，从根源上精确刻画合法流量产生的时序规律与行为动机，进而为识别异常流量提供依据。具体执行过程为：组合利用多个大语言模型（large language model, LLM）从设备代码中提取网络交互规则，并通过交叉验证确保规则提取的准确性；将提取的网络交互规则转换为有限状态自动机，构建物联网设备的网络行为数字孪生模型；基于有限状态自动机，并结合物联网设备当前状态，实时模拟网络流量生成；最后执行异常流量检测，若实际流量与模拟流量不匹配，则判断为异常流量。所提方法不需要任何训练数据集，同时将异常流量检测转变为网络流量的模拟与匹配，运行简单、高效，且检测结果易于解释——在当前条件下不应该出现的流量即异常流量。

本文的主要贡献如下。

1) 利用 LLM 从物联网设备的源代码中自动提取网络交互规则并分析不同规则间的关联关系，为网络行为数字孪生模型的构建奠定基础。此外，通过多个 LLM 的交叉验证，提高网络交互规则提取和分析的准确性。

2) 基于网络交互规则，构建物联网设备网络行为的数字孪生模型，当检测到与模型不符的流量时，即可视为异常流量。检测过程不需要依赖复杂的机器/深度学习算法，而是通过精确的仿真和比对，快速识别异常行为，提高检测的准确性和效率。

3) 对常见的物联网设备包括智能监控摄像头、智能环境监测终端、智能物流追踪器构建网络行为数字孪生模型，分别使用拒绝服务（distributed denial of service, DDoS）攻击、命令与控制（com-

mand and control, C&C）通信和内网扫描对物联网设备进行恶意攻击。实验结果表明，所提方法的F1分数、误报率等指标均优于现有方法。与TrafficFormer[11]等最新的预训练模型相比，所提方法的模型大小由682 MB降至17 KB，计算和存储资源消耗分别降低85.44%和94.06%。

1 相关工作

本节按照从机器学习、深度学习到最新预训练方法的技术演进路线，对物联网恶意攻击和异常流量检测领域的代表性工作进行介绍。在研究初期，研究人员广泛探索了基于机器学习的检测方法[12]。整体上，此类方法的准确性依赖于选择特征的有效性。例如，文献[13]采用PCA（principal component analysis）、RFE （recursive feature elimination）等特征选择方法并结合DT（decision tree）、NB（naive Bayes）、KNN（K nearest neighbor）分类器进行研究，以提升物联网异常流量攻击检测性能。实验结果表明，利用关键特征能够显著降低计算成本，有效提升检测性能。然而，该类方法高度依赖特征选择与专家经验，导致其鲁棒性不足，且难以适应多样化的网络环境与应用场景，通用性有限。

为减少对特征选择的依赖，研究人员进一步提出基于深度学习的检测方法。文献[14]采用含长短期记忆网络（long short-term memory, LSTM）层的递归神经网络（recurrent neural network, RNN）架构，结合NAdam优化算法，通过分析网络流量特征实时识别潜在威胁，检测准确率高达99%。传统的深度学习方法依赖于充足算力，难以适配资源受限的物联网应用场景。针对该问题，研究人员持续探索轻量级深度学习检测方法。文献[15]提出轻量级知识蒸馏模型，基于神经网络通过知识蒸馏机制将复杂模型的检测知识迁移至轻量级学生模型，并结合可分离卷积降低计算复杂度，实现对物联网异常流量的高效检测与模型体积压缩。文献[16]提出TSCRNN（time-segmented convolutional RNN）方案，利用堆叠双向LSTM捕捉时序依赖关系，并通过长时流量降采样策略优化时序建模，降低模型复杂度。类似地，文献[17]通过主成分分析，结合扩展压缩结构、逆残差模块与通道混洗操作降低计算开销，并引入损失函数优化样本，在保证恶意攻击检测精度的同时实现模型轻量化。文献[18]提出LKD-STNN （lightweight knowledge distillation space time neural network）模型，基于轻量级残差块模块，融合ShuffleNetV2的跨层连接思想与Ghost模块替代传统卷积，以低成本生成丰富特征表示，大幅减少模型参数、缩短推理时间。

上述方法虽能轻量化运行，但模型的训练需要大量的带标签训练数据（如标记为恶意的网络流量），这些数据在实际网络中却难以获得并准确标注[11]，限制了此类方法的实际应用。为此，研究人员提出预训练方法，充分利用无标签网络流量数据来训练基础模型，然后使用少量标签数据进行微调以适应具体检测任务。文献[19]提出一种以字节对为基本单元的分词方案，并利用未标注的流量数据对Transformer编码器模型进行预训练。随后，他们使用少量标注的加密流量对预训练模型进行微调，以完成下游分类任务。这一思路也被ET-BERT[20]、YaTC[21]和TrafficFormer[11]采用。ET-BERT将突发（即同一方向上连续的数据包序列）视为一个句子，并通过掩码突发建模和下一句预测任务来学习流量模式。YaTC将每个流表示为一幅图像，并在预训练阶段采用掩码图像建模任务。TrafficFormer在预训练阶段保留了掩码建模任务，同时引入了一种细粒度的多分类任务——同源方向流任务，以捕获流量中的方向性和序列信息。但小样本微调仍难以全面反映物联网设备的全部流量特点，因而仍易产生较多误报。

本文研究工作受数字孪生技术启发。数字孪生技术是一种通过创建物理对象或系统的虚拟模型来实现对其状态、行为和性能的全面理解和优化的技术[22]。这项技术最初应用于航空航天领域，随着物联网的发展，数字孪生的应用范围迅速扩展到了制造业、医疗保健、智慧城市等多个领域。在数字孪生的辅助下，可以实现预防性维护[23]、智能海上运输[24]以及构建弹性供应链[25]等多种智能服务。在物联网安全领域，数字孪生主要应用于网络安全态势感知[26]和策略优化[27]等，但是将数字孪生技术应用于网络流量异常检测的相关研究较少。

2 网络行为建模与异常流量检测

本文旨在解决不依赖网络流量数据集的异常流量检测难题，并为此提出了一种基于网络行为数字孪生的检测方法。方法的整体架构如图1所示。

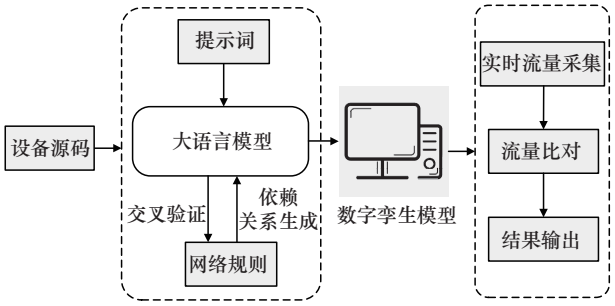


图1　方法的整体架构

所提方法不需要采集任何网络流量作为训练样本，而是以设备源码为输入，利用LLM强大的代码分析能力，从源码中直接提取网络交互规则。网络交互规则即代表了设备的网络行为逻辑。对提取的网络交互规则进行归纳分类，再次利用LLM，分析交互规则之间可能存在的关联关系，为物联网设备的网络行为数字孪生模型构建奠定基础。在此过程中，通过多LLM组合交叉验证策略，保障了网络交互规则提取以及孪生模型构建的准确性。利用构建的网络行为数字孪生模型，采用流量比对实现异常流量检测，即将实际网络流量与数字孪生模型中的预期流量进行实时对比，不符合流量预期的则被标记为异常流量，方法运行高效。

所提方法需以设备源码为输入，在实际应用中可由物联网设备制造商完成孪生模型的构建。具体应用模式为：设备制造商负责生成并更新物联网设备对应的网络行为数字孪生模型，用户以插件形式下载和应用网络行为数字孪生模型。此应用模式与现有网络入侵检测规则（如Cisco Talos提供的商业版Snort入侵检测规则）的购买使用类似，因而具备实际可行性。本文设定攻击者已占据物联网设备，并利用该设备发动进一步的恶意攻击。与现有工作类似[10,20-21]，攻击者无法控制或干扰检测模型的具体运行，即检测设备是安全的。

2.1　网络交互规则提取

目前，从源码中提取网络交互规则的一种常见实现方式是构建控制流图（control flow graph, CFG）[28]，如在CFG中定位出网络访问应用程序接口（API）以及相关联的节点，利用图的遍历确定在何种事件触发下会导致网络交互的发生。但此类方法需针对不同编程语言进行单独设计实现，实现难度大。同时，由于图中关联节点众多，遍历所需的时间复杂度高且难以准确确定具体的触发事件。为此，本文提出一种基于LLM的网络交互规则提取方法，利用LLM的强大代码分析能力，从设备源码中定位网络访问功能，并回溯出具体的触发事件。定义大语言模型集合 $L = \{M_1, M_2, M_3, \cdots, M_n\}$。将代码片段标记为$S$，根据$S$的特性，如编程语言、协议类型、设备特性和功能等，生成提示语集合 PromptSet = $\{P_{ij}\}$，其中，i 代表模型集合中第 i 个大语言模型，j 代表为该模型定制的第 j 条特定提示语。网络交互规则提取的具体实现流程如算法1所示。

算法1　网络规则提取算法

输入　设备源码S，LLM模型集合L，提示语集合PromptSet

输出　验证后的网络交互规则集合V_{final}

1)初始化V_{final}为空
2)for each模型$M_i \in L$ do
3)　初始化模型结果列表V_i为空
4)　for each提示语P_{ij} ∈PromptSet do
5)　　V_{ij} ←使用模型M_i和提示语P_{ij} 分析S
6)　　将分析结果V_{ij} 添加到V_i
7)　end for
8)end for
9)将所有模型的分析结果V_i标准化为统一表示形式
10)统计各规则的支持模型数量
11)设定阈值th
12)for each规则 do
13)　if该规则的支持模型数量>th then
14)　　将规则加入V_{final}
15)　else
16)　　使用所有模型重新验证该规则，若支持模型数量>th，则加入V_{final}
17)　end if
18)end for
19)返回验证后的规则集合V_{final}

在算法1中，提示语的生成会综合多方面因素。在本文工作中，对于不同类型的流量，设计不同的提示词。例如，若代码涉及HTTP流量，会生成类似“假如你是一个Python工程师，这是一段智能摄像头的Python源代码，该设备使用HTTP，请分析其网络行为，并提取其网络规则，包括触发条件以及触发流量特征、IP”等提示语。若涉及

TCP流量，则会设计针对TCP特性的提示语，如“请分析这段代码中使用TCP的网络交互逻辑，明确其连接建立、数据传输和连接断开的规则及触发条件”。同时，考虑不同大模型的特点来设计提示词，如提示词同时准备中文和英文版本，以适应不同模型的需求。提示语的不同设计可以引导模型更准确地分析网络流量交互相关的逻辑。

对于每个模型$M_i \in L$，使用提示语P_{ij}对代码S逐条分析，生成分析结果V_{ij}，例如该智能摄像头通过HTTP与服务器（172.20.10.11:5000）进行周期性交互，行为规则示例如下。定时轮询：每秒发起两次随机间隔的请求。固件检查：GET/firmware/check，User-Agent包含设备IP（172.20.10.10），若响应中new_version_available为true，则触发。位置上报：POST/location/update发送随机生成的经纬度。随后将V_{ij}添加到该模型的结果列表V_i。为了克服LLM幻觉问题[29]并确保分析结果的准确性，采用多模型交叉验证方法。具体过程如算法1中步骤9)~步骤18)所示。首先，对所有模型的分析结果V_i进行处理。由于不同大模型生成的规则样式可能不同，因此需要先对每个模型生成的规则进行标准化处理，将其转换为统一的表示形式。统一表现形式包括触发条件、流量特征（如IP、协议、端口号、请求方法、URL等）以及行为描述3个部分。触发条件和流量特征高度一致的规则算作同一条规则。具体地，规则一致性使用Jaccard相似度进行衡量，计算式为

$$\mathrm{Sim}\left(R_1,R_2\right)=\alpha\left|\frac{T_1\cap T_2}{T_1\cup T_2}\right|+\beta\left|\frac{F_1\cap F_2}{F_1\cup F_2}\right| \tag{1}$$

其中，R_1和R_2代表经过标准化处理后的规则，包含触发条件T和流量特征F，对T和F进行分词处理后，分别计算对应的交集和并集，从而计算出相似度；α和β是触发条件在综合相似度计算中的权重系数，取值范围为[0, 1]，表示不同条件对判定两条规则是否高度一致的重要程度，若相似度大于一定阈值，则被判定为一致。

最后，设定阈值th，th的取值可根据实际需求和模型数量来灵活确定。若某个规则对应的支持模型数量大于设定的阈值th，则认为该规则具有较高的可信度，将其加入验证结果集合V_{final}中；若支持模型数量小于或等于阈值th，则丢弃该规则。经过上述筛选后，验证后的结果集合V_{final}会整合为一个完整的流量触发规则集，这个规则集具有高准确性和可靠性，为后续网络行为数字孪生模型的构建提供基础。

2.2 网络行为数字孪生模型构建

为了从网络交互规则中构建出一个能准确反映流量行为逻辑的数字孪生模型，本文对不同物联网设备的网络交互规则进行深入分析，并将物联网流量触发条件归纳为自启动式、状态触发式以及访问触发式。自启动式流量触发规则指的是设备在启动时自动触发的流量，例如设备上电连接到云服务器或发送初始化状态。状态触发式流量触发规则由设备内部状态变化引发，比如检测到新固件版本后发起下载请求，定位服务启用时进行位置数据上传，或者当传感器检测到环境参数（如温度、湿度、光照强度等）超出预设阈值时，设备触发警报或通知。访问触发式流量触发规则是指在外部实体（如用户或其他设备）向物联网设备发出请求时产生的流量，包括合法的API调用和配置更改指令等。在网络行为数字孪生模型构建过程中，设定自启动式>状态触发式>访问触发式的优先级顺序，以反映网络交互规则间的依赖关系。优先级的设定是依据物联网设备的内在运行逻辑：自启动式流量是设备启动后的基础操作，之后设备会检查内部和外部状态，然后再接收用户请求。触发条件的归纳和优先级的定义可以确保模型能够高效、准确地模拟设备网络行为，避免规则冲突。基于提取的网络交互规则，物联网设备的网络行为数字孪生模型构建流程如算法2所示。

算法2 网络行为数字孪生模型构建算法

输入 网络交互规则集合V_{final}，LLM模型集合L

输出 数字孪生模型M

1)初始化状态集合S，初始化状态机集合SM、数字孪生模型M为空

2)将V_{final}按规则类型划分为自启动式集合V_{auto}、状态触发式集合V_{state}、访问触发式集合V_{access}

3)for each 规则$r \in V_{\text{final}}$ do

4) 提取r的核心要素：触发条件$C(r)$、网络行为$B(r)$、目标地址$A(r)$、优先级$P(r)$

5) 添加$C(r)$中的状态到去重集合S中

6)end for

7)for each 模型$M_i \in L$ do

8) 并行处理V_{final}，生成规则依赖关系$G_i = M_i(V_{\text{final}})$

9)end for

10)对$\{G_1, G_2, \cdots, G_n\}$进行交叉比对，生成最终依赖关系$G_{\text{final}} = \text{HC}(\{G_i\})$

11)以规则集合$R = V_{\text{final}}$、状态集合S，调用算法3

12)通过DFA转换算法的NFA构建、子集构造、最小化流程，得到确定性有限自动机SM（含状态集Q、转移函数δ、初始状态q_0、接收状态动作标记B）

13)for each 状态$q \in \text{SM}.Q$ do

14) 找到q对应的规则r（通过$\text{SM.B}[q] = B(r)$匹配）

15) 根据G_{final}确定r的前置条件$\text{Pre}(r)$，更新SM的转移函数δ：仅当$\text{Pre}(r)$满足时，才执行从当前状态经输入符号到目标状态的转移

16)end for

17)for each $r \in \text{SM}$ do

18) if $r \in V_{\text{auto}}$，封装$\{A(r)$, 固定触发时间$T_0\}$为自动组件

19) else if $r \in V_{\text{state}}$，绑定传感器阈值为触发条件

20) else if $r \in V_{\text{access}}$，添加用户IP校验

21) end if

22) 将r加入M中

23)end for

24)在M中添加实时数据接口，用于获取外部访问信息和设备状态变化

25)返回数字孪生模型M

在孪生模型构建的过程中，需结合不同规则间的关联关系，识别可能的流量依赖性或冲突性，包括不同规则可能共享相同的触发条件以及流量的产生依赖于其他规则的前置行为。为此，本文提出双层验证机制，如算法2中步骤1)~步骤9)所示。首先由多个LLM并行处理规则集，利用模型间的认知差异捕捉潜在解析盲区。然后通过LLM建立规则交互关系，对各模型输出的依赖关系进行交叉比对，重点校验规则优先级排序、条件触发阈值、行为执行前置条件等核心要素。例如，在分析工业传感器网络规则时，某模型可能遗漏“设备认证成功”对“数据上传规则”的触发约束，另一模型则能准确识别该依赖关系。通过综合研判，能够有效规避单一模型的局限性，提升分析的准确性与全面性。

解析后的规则集向确定有限状态自动机（deterministic finite automaton, DFA）的转化，是连接规则解析与流量预测的关键环节，该过程通过算法2中步骤11)~步骤16)与算法3（DFA转换算法）协同实现，形成兼具形式化严谨性与实际运行逻辑的状态机模型。具体而言，确定有限状态自动机的构建分为两个核心阶段。

第一阶段为基础DFA生成，由算法3完成。该算法以规则集合$R=V_{\text{final}}$（经分类验证的网络交互规则，每条规则包含触发状态条件C、网络行为B及优先级P）和状态集合S（从规则触发条件中抽象的离散设备状态，如“设备上电”“传感器阈值超标”等，由算法2中步骤3)~步骤6)提取）为输入，通过三步核心操作生成确定性有限自动机DFA：首先为每条规则构建NFA子图（将触发条件分解为状态序列的正则表达式，通过Thompson算法[30]转换为NFA）；随后合并所有NFA并通过子集构造法生成DFA，明确状态集Q、转移函数δ（描述“当前状态 + 输入符号→下一状态”的映射）及初始状态q_0；最终通过Hopcroft算法[31]最小化DFA，并按优先级为每个状态绑定唯一网络行为$B[q]$。此时生成的DFA已实现规则到“状态–转换–行为”的形式化映射，构成状态机SM的基础框架。

第二阶段为依赖关系融合，由算法2中步骤14)~步骤16)实现。算法3生成的基础DFA仅体现单条规则的触发逻辑，未考虑规则间的依赖关系，如固件下载需依赖固件检查结果。因此需基于算法2中步骤7)~步骤10)得到的最终依赖关系G_{final}，即通过多LLM交叉验证确定的规则前置条件集合，对DFA的转移函数δ进行优化：遍历DFA的每个状态q，通过行为标记$B[q]$匹配对应规则r，从G_{final}中提取r的前置条件$\text{Pre}(r)$，并更新转移逻辑，即仅当$\text{Pre}(r)$满足时，才允许执行当前状态经输入符号到目标状态的转移。

基于前置关系G_{final}和分类处理过的规则集V_{final}，状态机SM可形式化定义为

$$\text{SM} \doteq (S, P, R) \rightarrow E \tag{2}$$

其中，S为状态集合（由规则触发条件抽象而来，即算法2中构建的状态集），P为优先级关系（确保自启动式>状态触发式>访问触发式规则的执行顺序，与算法3中按优先级绑定行为逻辑一致），R为分类后的规则集合V_{final}，E为触发事件集（对应算法3中状

态序列分解得到的触发条件）。基于该状态机，可精确描述不同设备状态下由特定事件触发的网络交互行为，为后续数字孪生模型的构建提供支撑。

算法3 DFA转换算法

输入 规则集合R（每条规则含触发状态条件C、动作B、优先级P），状态集合S（输入的状态信息集合）

输出 确定性有限自动机DFA（状态集Q，转移函数δ，初始状态q_0，接受状态动作标记B）

1)Q = [], δ = {}, q_0 = None, B = {}

2)for each 规则 r in R do

3) 分解$C(r)$为状态S的序列组合成的正则表达式

4) 使用Thompson算法将规则的正则表达式转换为NFA(r)

5)end for

6)合并NFA，新增总起始状态q_0

7)通过ε-转移连接所有规则NFA(r)的起始状态

8)初始状态q_0_DFA= ε-闭包 (q_0)（即从q_0出发仅通过ε-转移可达的所有状态）

9)对每个状态和输入符号，计算可达状态集合

10)生成DFA状态转换表

11)for each $q \in Q$_DFA do

12) 若q包含多个NFA接收状态，按$P(r)$选择最高优先级动作$B(r)$

13) 标记q的动作$B[q] = B(r)$

14)end for

15)使用Hopcroft算法合并等价状态

16)更新Q_DFA和δ_DFA

17) Q = Q_DFA, δ = δ_DFA, q_0= q_0_DFA, B = B_DFA

18)return (Q, S, δ, q_0, B)

基于状态机SM，可生成物联网设备对应的网络行为数字孪生模型。如算法2中步骤17)~步骤23)所示，将触发阈值和IP地址等具体信息添加至对应规则，并添加与设备交互的数据接口，形成描述型的网络行为数字孪生模型。最后，采用代码构建可执行的网络行为数字孪生模型。通过设计标准化的代码结构，封装触发条件、IP地址、时间信息等属性，实现模型的可执行化。具体地，对于自启动式流量，直接输出对应网络流量信息；对于状态触发和访问触发式流量，在代码中通过接口获取物联网设备输入状态信息中的用户访问信息，并与流量事件、时间信息进行关联整合，借助函数调用实现数据的动态更新与流量生成。通过代码层面的结构化处理，将大模型输出转化为可稳定运行的网络行为数字孪生模型，确保其能准确反映不同触发机制下的网络行为模式。

2.3 异常流量检测

基于构建的网络行为数字孪生模型，通过实时流量比对，即将实际网络流量与数字孪生模型中的预期流量进行实时对比，能够及时检测出异常流量。为了更准确地描述异常流量检测过程，首先定义一些核心数据结构。

DeviceState表示原设备的当前状态，包括运行状况、已安装固件版本、启用的服务列表、接收到的用户请求的IP地址等信息。

TrafficModel表示网络行为数字孪生模型，包括流量触发条件的匹配逻辑和流量预测的基本结构。

TrafficEvent表示模型输出的预测结果，包括流量的时间序列、目标地址等。

在检测时，物联网设备向网关发送当前状态信息DeviceState。网关接收DeviceState后，依托部署于其中的网络行为数字孪生模型TrafficModel，根据流量触发条件的匹配逻辑与流量预测结果，持续生成预期流量记录数据TrafficEvent。与此同时，网关实时采集物联网设备的实际网络流量，获取源IP、目标IP、协议、时间戳、数据包大小等信息。随后进入流量比对环节，执行如算法4所示的双重条件联合决策策略。第一重判别策略是目标IP一致性验证，即核查实际流量的目标IP与预期流量预测结果TrafficEvent中的目标IP匹配度，检测实际流量目标地址是否偏离模型预测的地址白名单。第二重判别策略是时间点一致性验证，重点关注设备状态变更时刻与实际流量触发时刻的时序偏差，要求实际流量在Δt容差窗口内产生。若时序偏差超过阈值或目标地址偏离白名单，则触发异常标记机制，将该流量标记为异常流量。

算法4 异常流量检测算法

输入 捕获的网络流量信息集合F，合法目标IP集合$\mathrm{IP}_{\mathrm{expected}}$，时间阈值$\Delta t$

输出 标记为异常流量的集合A

1)初始化异常流量集合A为空

2)获取预期目标IP地址集合$\mathrm{IP}_{\mathrm{expected}}$

3)获取预期流量时间序列 T

4)for each 实际流量事件 $e \in T$ do

5)　提取事件 e 的目标 IP 地址为 IP_{dst}，时间戳为 t_{real}

6)　if IP_{dst} 不在 $IP_{expected}$ 中 then

7)　　将 e 加入异常流量集合 A

8)　　跳过后续时间验证

9)　end if

10)　从 T 中确定 e 所对应的预期时间点 t_{exp}

11)　if $|t_{real} - t_{exp}| > \Delta t$ then

12)　　将 e 加入异常流量集合 A

13)　end if

14)end for

15)返回异常流量集合 A

在算法 4 中，异常流量判定采用双重条件联合决策策略，即目标 IP 一致性和时间点一致性，这两类策略具体解释如下。

第一类策略针对目标 IP 地址的合法性。数字孪生模型基于设备正常行为规则预设“地址白名单”，该名单仅包含设备按规则应通信的合法目标（如云端服务器、授权终端）。若实际流量的目标 IP 不在白名单中，则判定为异常流量。例如，传感器设备正常仅向指定告警服务器发送数据，若出现向未知域名或黑客服务器的流量，则直接触发异常标记，可能表明设备被异常控制或数据遭窃取。

第二类策略聚焦流量触发的时序一致性，核心在于设备状态变更（如传感器阈值突破、用户请求接收）与流量产生的因果关系。合法流量应在状态变更后的合理时间窗口 Δt 内触发，该窗口基于设备正常处理延迟、通信协议耗时等因素设定。若实际流量在状态未变更时提前出现（如伪造的虚假告警）或在状态变更后超出 Δt 仍未出现（如恶意阻塞导致响应延迟），均视为时序异常。例如，温度超阈值后，设备应在数秒内发送告警，若延迟过长或无响应，可能是攻击导致资源占用或通信中断。通过这一策略，可有效识别违反设备行为逻辑的异常操作，如无状态变更的伪造流量或延迟注入的攻击流量。

3　实验

3.1　实验环境与配置

实验中选用多种物联网设备来验证方法的实际运行效果，具体实验设备如下。智能监控摄像头：周期性上传门禁触发视频流（M-JPEG 协议），每日固定时段（00:00、12:00）发起固件更新请求，异常门禁操作时实时上报设备位置与事件日志，传输视频数据等。智能环境监测终端：每 5 min 向云端传输温湿度、PM2.5 数据（HTTP），网络中断后缓存数据并在恢复时补发，接收云端指令调整数据采集频率。智能物流追踪器：移动过程中每 30 s 上报 GPS 位置信息，途经预设节点时上传货物状态（如振动、倾斜数据），支持远程触发固件版本校验与更新。

本文方法组合利用多个大语言模型来构建设备的网络行为数字孪生模型。实验中采用的大语言模型有 GPT-4o、Gemini 2.5 Pro 和通义千问 3.5。LLM 通过自注意力机制实现对海量文本数据的深度语义理解，其核心优势在于通过参数规模扩展（百亿至万亿级）以获得涌现能力，可完成逻辑推理、代码解析等复杂任务。这充分利用了 LLM 对协议语法、网络行为的强表征能力，为物联网设备的网络行为数字孪生系统的构建提供了智能分析引擎。在实验中，首先将设备源码保存至文本文件，然后上传至 LLM 服务器。针对上传的文件，输入提示词，让 LLM 分析提取设备的网络交互规则。在利用 LLM 提取规则和生成规则依赖关系时，将模型的 Temperature 和 Top_p 参数均设置为 0.1，以保证输出结果的确定性与一致性。提示词采用结构化预设模板，包含角色设定、任务描述和输出规范 3 个部分，以引导模型生成标准化的规则描述。不同 LLM 生成的网络交互规则示例如图 2 所示。接着通过算法 2 和算法 3，形成最终的网络行为孪生模型。最后利用 Python 代码实现网络行为孪生模型，即智能监控摄像头、智能环境监测终端以及智能物流追踪器的网络行为数字孪生模型均为运行于边缘网关的 Python 可执行文件，易于部署和应用。

本文采用误报率（false positive rate, FPR）、精确率（Precision）、召回率（Recall）和 F1 分数（F1 score）来衡量和对比不同异常流量检测方法的运行效果。设 TP 为真阳性数量，TN 为真阴性数量，FP 为假阳性数量，FN 为假阴性数量，则有

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \times 100\% \tag{3}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \times 100\% \tag{4}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100\% \tag{5}$$

规则名	触发条件	目标地址	路径/参数	关键字段/内容
固件更新检查	0:00-1;00周期性请求，版本落后时触发下载流量	服务器IP：5000	/firmware/check	User-Agent含设备IP
位置上报	gps开启，每隔一分钟周期性执行	服务器IP：5000	/location/update	经纬度数据
视频流请求	合法用户访问视频流	摄像头IP：8080	/?action=stream	持续返回MJPEG流
快照请求	合法用户访问快照	摄像头IP：8080	/?action=snapshot	返回单张JPEG图像
命令控制	合法用户命令操控	摄像头IP：8080	/?action=command	包含命令ID和参数

(a) GPT-4o生成的网络交互规则

规则名	触发条件	目标地址
固件更新检测	周期性每天 0:00 ~ 0:59 触发，根据版本信息触发下载流量 /firmware/check	服务器IP:5000
位置上报	gps开启时周期性每分钟执行一次 /location/update	服务器IP:5000
视频流传输	携带身份信息访问/?action=stream或/?action=snapshot	摄像头IP:8080

(b) Gemini 2.5 Pro生成的网络交互规则

规则名	触发条件	目标地址	关键字段/内容
固件更新检查	每天 0:00 ~ 0:59 之间首次触发，根据版本信息触发下载流量	服务器IP:5000/firmware/check	用于固件更新检测
位置上报	gps开启时每分钟执行一次	服务器IP:5000/location/update	GPS 上报经纬度
视频流传输（流媒体）	客户端访问视频流	摄像头IP:8080/?action=stream	持续发送视频流
快照请求（快照）	客户端访问快照	摄像头IP:8080/?action=snapshot	返回单帧 JPEG 图像
命令处理	客户端访问命令操控	摄像头IP:8080/?action=command	解析命令操作

(c) 通义千问3.5的网络交互规则

图2 网络交互规则示例

$$\mathrm{F1} = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \times 100\% \tag{6}$$

此外，本文还对比了不同方法的资源消耗情况。具体地，使用CPU占用率和GPU占用率以及内存使用量评估模型部署的资源消耗程度。

1) CPU 占用率：实验期间边缘网关的平均CPU使用率。

2) GPU 占用率：实验期间边缘网关的平均GPU使用率。

3) 内存使用量：实验期间边缘网关运行内存的平均占用量。

方法参数依据实验需求设定。考虑到本文采用3个大模型进行协同验证，故将算法1中的 th 设置为1，即当某条规则获得的支持模型数量大于或等于2时，判定其具备高可信度，纳入最终验证结果集合 V_{final}。将式(1)中的触发条件特征权重 α 设定为0.5，流量特征权重 β 设定为0.5，若计算结果大于0.8，则判断为一致有效的网络行为规则。为验证本文方法对异常流量的检测性能，将算法4中的时间戳阈值 Δt 作为核心变量，通过为其设置一系列不同的数值进行对比实验。

实验通过多种渗透测试工具模拟不同攻击场景，用于验证数字孪生模型的异常流量检测能力，具体介绍如下。

1) DDoS攻击

DDoS通过向目标服务器发送海量请求流量，耗尽目标资源（如带宽、计算能力），导致合法用户无法访问服务。本实验采用轻量级渗透测试工具模拟HTTP Flood攻击，基于Python多线程技术构建并发请求逻辑，模拟僵尸网络的高频访问行为。攻击脚本通过requests库向目标服务器发送构造的HTTP GET请求以形成流量负载。

2) C&C通信

C&C是攻击者通过隐蔽通道远程控制僵尸网络的核心技术。攻击者通过恶意程序感染设备后，僵尸主机会主动连接C&C服务器以下载指令并回传数据。本实验通过Python脚本模拟C&C客户端行为，设定设备向预设的C&C服务器发送HTTP POST请求。请求负载包含设备ID、当前固件版本等信息。脚本同时解析本地预设的“指令文件”，

模拟从DNS或HTTP头部获取加密指令的过程。

3) 内网扫描

通过主动探测目标主机的开放端口和服务，寻找潜在漏洞进行渗透攻击。本实验物联网设备通过渗透测试工具Nmap对子网内设备进行端口探测与漏洞嗅探。扫描脚本对扫描目标IP地址的80、443、22、8080等常用端口发送TCP SYN探测包，通过响应报文的标志位判断端口状态。

3.2　检测效果对比

为全面验证所提网络行为孪生模型的检测性能，实验选取5种主流物联网恶意攻击和异常流量检测方法作为对比基线，分别为卷积神经网络（convolutional neural network，CNN）[15]、时空特征融合循环神经网络（time-segmented convolutional RNN，TSCRNN）[16]、轻量级神经网络（lightweight deep neural network，LNN）[17]、轻量级知识蒸馏时空神经网络（lightweight knowledge distillation space time neural network，LKD-STNN）[18]以及网络流量预训练方法TrafficFormer[11]等。依据对应的原始论文内容，对前4种方法分别采用数据集进行模型训练。

1) CNN采用KDD-CUP99和UNSW-NB15[32]数据集，首先对离散特征进行独热编码，然后对所有特征进行min-max归一化处理，最后对教师网络和学生网络进行50轮训练。

2) TSCRNN采用SCX-Tor2016[33]数据集，通过5元组进行流分割，随机采样后将流量向量化，随后除以255进行归一化。模型首先通过CNN提取空间特征，再输入堆叠双向LSTM学习时间特征，共训练50轮。

3) LNN采用UNSW-NB15和Bot-IoT[34]数据集，在数据预处理阶段，对符号特征进行独热编码，将攻击类别标签数值化并转换为独热编码以适配多分类任务，随后采用min-max归一化，利用PCA算法进行特征降维以降低输入复杂度，最后设置模型参数训练50轮。

4) 对于LKD-STNN，采用Edge-IIoTset[35]、USTC-TFC2016[36]、ToN-IoT[37]、CIC IoT2023[38]数据集进行训练，将原始流量进行整合、分割（基于5元组）、清洗（删除无效文件）、截断与填充以及归一化后，以IDX格式输入模型训练50轮。

由于TrafficFormer公开了其源代码和模型，实验中直接采用其公开发布的预训练模型，不需要额外训练。所有对比方法与本文的网络行为数字孪生模型均针对同一批模拟生成的攻击流量开展检测。具体地，分别对智能监控摄像头、智能环境监测终端以及智能物流追踪器进行攻击并统计流量。针对DDoS、C&C通信和内网扫描攻击，每类攻击流量的总规模均设定为10 GB，通过模拟持续攻击场景，以验证模型在大数据量下的处理性能与统计显著性。实验共持续一个月，动态采集的正常网络流量约为150 GB。在所有对比实验中，均采用10次独立重复实验的平均值作为最终结果数据，以消除随机性带来的误差。检测过程中记录各方法的准确率、召回率、F1分数等关键指标，以实现全方位性能对比。依据实验设备性能，设定数字孪生模型的时间容差Δt值从0.5变化至1.75。对不同设备的检测结果进行综合统计，如表1所示。

表1　网络行为孪生模型异常流量识别结果

评价指标		DDoS攻击	C&C通信	内网扫描
召回率	Δt=0.5	96.31%	96.00%	96.00%
	Δt=0.75	95.52%	95.12%	94.23%
	Δt=1	94.10%	93.10%	92.36%
	Δt=1.25	92.81%	92.00%	88.60%
	Δt=1.5	87.00%	84.40%	80.10%
	Δt=1.75	75.23%	76.00%	70.20%
精确率	Δt=0.5	100.00%	100.00%	100.00%
	Δt=0.75	100.00%	100.00%	100.00%
	Δt=1	100.00%	100.00%	100.00%
	Δt=1.25	100.00%	100.00%	100.00%
	Δt=1.5	100.00%	100.00%	100.00%
	Δt=1.75	100.00%	100.00%	100.00%
误报率	Δt=0.5	0	0	0
	Δt=0.75	0	0	0
	Δt=1	0	0	0
	Δt=1.25	0	0	0
	Δt=1.5	0	0	0
	Δt=1.75	0	0	0
F1分数	Δt=0.5	98.12%	97.96%	97.96%
	Δt=0.75	97.71%	97.49%	97.03%
	Δt=1	96.96%	96.43%	96.03%
	Δt=1.25	96.27%	95.83%	93.96%
	Δt=1.5	93.05%	91.54%	88.95%
	Δt=1.75	85.86%	86.36%	82.49%

从表1中可以看出，孪生模型在面对不同类型的异常流量时表现出良好的检测性能。特别地，在设定的不同时间容差$\Delta t\in[0.50, 1.75]$范围内，本文方法的检测误报率均为0，显示了本文方法的有效性。此外，为验证本文方法的先进性，开展了多项对比实验，整体结果如表2所示。在对比实验中，依据原始论文内容，多次调整CNN、TSCRNN、LNN以及LKD-STNN的配置参数，直到最优分类效果；对于TrafficFormer，直接采用公开的预训练模型，设置Δt=0.5。表2中的实验结果表明，本文方法的召回率（96.10%）、精确率（100.00%）、误报率（0）以及F1分数（98.00%）都显著优于现有方法。

表2　　不同方法检测效果对比

模型	召回率	精确率	误报率	F1分数	模型大小
本文方法	**96.10%**	**100.00%**	**0**	**98.00%**	**17 KB**
CNN	91.51%	91.02%	7.63%	91.26%	925 KB
TSCRNN	93.41%	96.32%	3.91%	94.84%	3 784 KB
LNN	93.10%	95.91%	4.60%	94.48%	655 KB
LKD-STNN	92.20%	96.12%	3.91%	94.12%	78 KB
TrafficFormer	95.63%	97.90%	5.20%	96.75%	682 MB

为确保对比实验的公平性，本文测试了不同参数下各种方法的检测效果，实验结果如图3~图6所示。由于不同方法具有不同参数，为进行有效的对比，在[0,1]内逐步遍历不同判断阈值，并测试在同一阈值、不同参数下各方法的最佳计算结果。具体实现时，对于对比方法，直接以置信度概率作为阈值，统计在不同阈值下的最优检测结果。对于本文方法，将Δt在其实验设置范围$[\Delta t_{\min}, \Delta t_{\max}]$内进行线性映射，计算式为$\dfrac{\Delta t-\Delta t_{\min}}{\Delta t_{\max}-\Delta t_{\min}}$，其中，$\Delta t_{\min}$=0.5，$\Delta t_{\max}$=2.5。实验结果表明，在阈值区间内，本文方法的F1分数、误报率、精确率以及召回率等指标均优于现有方法。特别地，如表2所示，本文方法在保持高检测精度的同时，以17 KB（智能监控摄像头对应的孪生模型大小）的轻量化设计突破了现有方法在资源占用方面的瓶颈，支持边缘设备实时部署与多设备并行检测。本文提出的虚实协同的检测机制不仅解决了传统方法依赖数据集和特征工程的局限性，更具备轻量化运行特性，为大规模物联网的网络安全防护提供了一种有效解决方案。

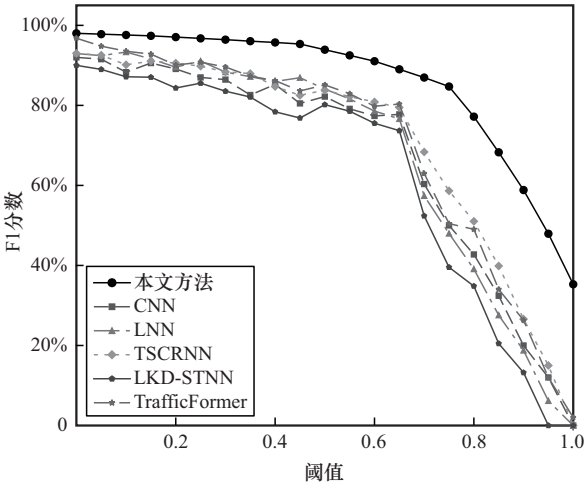


图3　F1分数对比

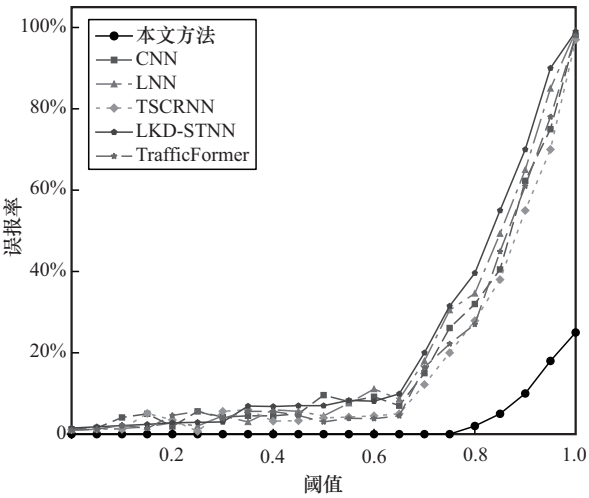


图4　误报率对比

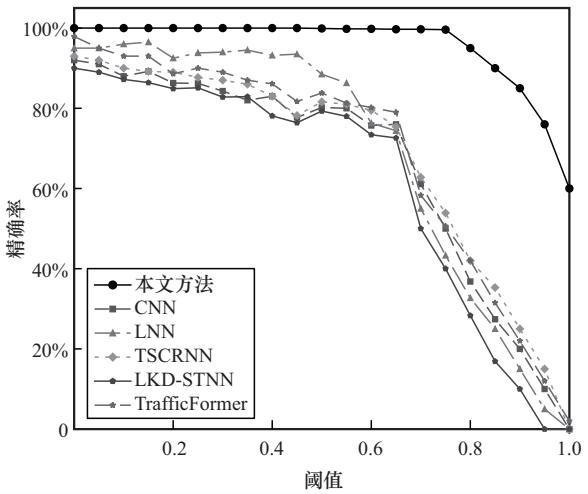


图5　精确率对比

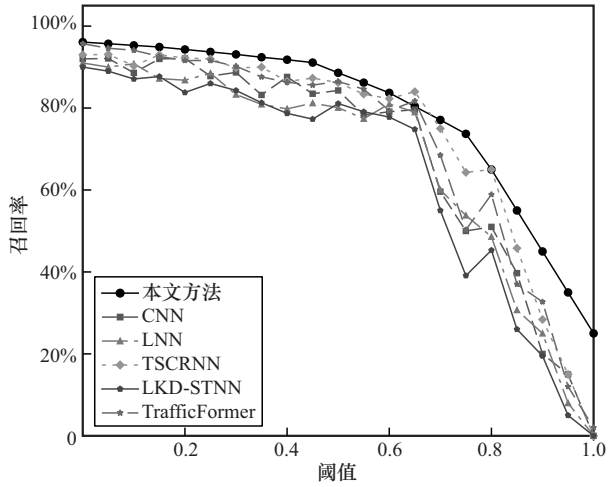


图6　召回率对比

3.3　资源消耗对比

为对比不同方法的资源消耗，本节分别统计本文方法以及CNN、TSCRNN、LNN、LKD-STNN、TrafficFormer在运行时消耗的CPU、GPU和内存资源，从而实现不同方法的资源消耗对比。实验中，不同方法均运行于同一台边缘网关服务器（边缘服务器配置为4 核 CPU，4 GB 内存，NVIDIA GeForce RTX 4060 GPU），对物联网设备产生的网络流量进行持续分析判断，持续时间为1 h。在此过程中，使用htop工具统计方法的CPU、GPU和内存使用量，并以均值为最终统计结果，如图7所示。相较于传统的深度学习方法，基于数字孪生模型的检测方法的CPU使用率与内存使用量均显著更低。相比之下，TrafficFormer的CPU使用率高达10%，是本文方法的6.7倍；内存使用量达1 GB，是本文方法的16.5倍；GPU使用率更是攀升至97%，资源消耗差距尤其明显。在实验中，本文提出的网络行为孪生检测方法以仅1.5%的CPU使用率、0的GPU使用率、60.78 MB的内存使用量实现了高效检测。这种极小的资源代价，使本文方法在资源利用效率上优势突出，适用于物联网环境。

4　结束语

本文针对物联网异常流量检测挑战，提出了一种基于网络行为数字孪生的创新解决方案，以替代传统的机器（深度）学习模型。利用大语言模型自动提取物联网设备的交互规则并构建网络行为数字孪生模型，然后将孪生模型转变为可执行代码，实现了网络流量的实时比对与异常流量的有效检测。为验证方法的有效性，本文以3种物联网设备为例，完整实现了网络交互规则提取、数字孪生模型构建与运行、异常流量检测等过程。实验结果验证了所提方法在典型攻击场景下的高效性，特别是方法的轻量化特性突破了传统方法对算力与存储的依赖，为物联网边缘计算场景提供了可行的安全防护手段。

在研究和实验中，笔者也发现针对实时网络流量分析，数字孪生模型的检测性能受事件同步及时性影响较大。若孪生设备无法及时同步和处理数

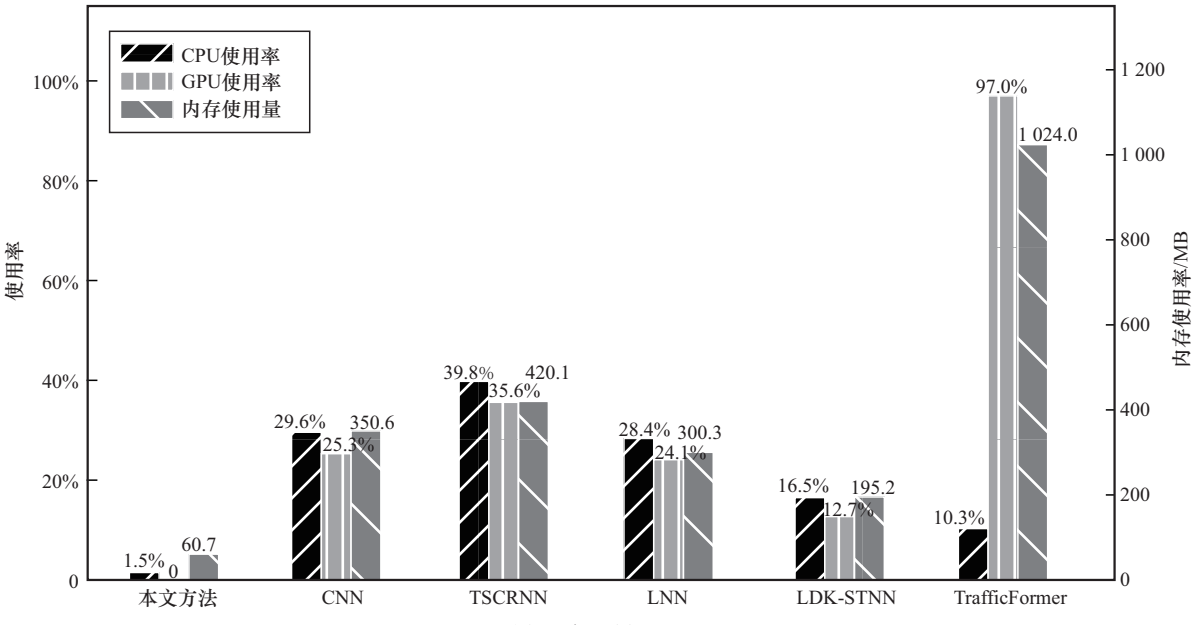


图7　资源消耗对比

据，则期望的事件发生时间同实际的事件发生时间之间会产生较大偏差，从而降低检测准确性。在未来研究中，需要设计物理设备同孪生设备间的快速数据同步和高效处理机制，从而进一步提升方法的鲁棒性。此外，针对互联网服务器等复杂类型设备与应用，如何构建完备的网络行为数字孪生模型，将是未来工作的重点与难点。

参考文献：

[1] 杨毅宇, 周威, 赵尚儒, 等. 物联网安全研究综述: 威胁、检测与防御[J]. 通信学报, 2021, 42(8): 188-205.
Yang Y Y, Zhou W, Zhao S R, et al. Survey of IoT security research: threats, detection and defense[J]. Journal on Communications, 2021, 42(8): 188-205.

[2] He G F, He T Y, Chen R H, et al. Multidevice collaborative authentication for Internet of things[J]. IEEE Internet of Things Journal, 2025, 12(14): 27753-27768.

[3] Wang Z H, Fok K W, Thing V L L. Machine learning for encrypted malicious traffic detection: Approaches, datasets and comparative study[J]. Computers & Security, 2022, 113: 102542.

[4] Lin K D, Xu X L, Xiao F. MFFusion: a multi-level features fusion model for malicious traffic detection based on deep learning[J]. Computer Networks, 2022, 202: 108658.

[5] Fu C P, Li Q, Bertino E, et al. Training with only 1.0 ‰ samples: malicious traffic detection via cross-modality feature fusion[C]//Proceedings of the 2025 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2025: 3930-3944.

[6] 何高峰, 魏千峰, 肖咸财, 等. 支持数据隐私保护的恶意加密流量检测确认方法[J]. 通信学报, 2022, 43(2): 156-170.
He G F, Wei Q F, Xiao X C, et al. Confirmation method for the detection of malicious encrypted traffic with data privacy protection[J]. Journal on Communications, 2022, 43(2): 156-170.

[7] Liu M Y, Wang P, Li Z Y, et al. UDDA-TC: unsupervised real-time drift detection and adaptation for continual traffic classification in mobile edge computing[J]. IEEE Transactions on Consumer Electronics, 2025, 71(3): 8940-8952.

[8] Johari S S, Tornatore M, Boutaba R, et al. Few-shot domain adaptation for effective data drift mitigation in network management[C]//Proceedings of the 2025 IEEE 45th International Conference on Distributed Computing Systems (ICDCS). Piscataway: IEEE Press, 2025: 417-427.

[9] Xu L J, Ding X, Peng H P, et al. ADTCD: an adaptive anomaly detection approach toward concept drift in IoT[J]. IEEE Internet of Things Journal, 2023, 10(18): 15931-15942.

[10] Tao F, Xiao B, Qi Q L, et al. Digital twin modeling[J]. Journal of Manufacturing Systems, 2022, 64: 372-389.

[11] Zhou G M, Guo X W, Liu Z T, et al. TrafficFormer: an efficient pretrained model for traffic data[C]//Proceedings of the 2025 IEEE Symposium on Security and Privacy (SP). Piscataway: IEEE Press, 2025: 1844-1860.

[12] Shafiq M, Tian Z H, Bashir A K, et al. CorrAUC: a malicious bot-IoT traffic detection method in IoT network using machine-learning techniques[J]. IEEE Internet of Things Journal, 2021, 8(5): 3242-3254.

[13] Sah G, Banerjee S, Singh S. Intrusion detection system over real-time data traffic using machine learning methods with feature selection approaches[J]. International Journal of Information Security, 2023, 22(1): 616.

[14] Woźniak M, Siłka J, Wieczorek M, et al. Recurrent neural network model for IoT and networking malware threat detection[J]. IEEE Transactions on Industrial Informatics, 2021, 17(8): 5583-5594.

[15] Zhao R J, Chen Y, Wang Y J, et al. An efficient and lightweight approach for intrusion detection based on knowledge distillation[C]//Proceedings of the ICC 2021 - IEEE International Conference on Communications. Piscataway: IEEE Press, 2021: 1-6.

[16] Lin K D, Xu X L, Gao H H. TSCRNN: a novel classification scheme of encrypted traffic based on flow spatiotemporal features for efficient management of IIoT[J]. Computer Networks, 2021, 190: 107974.

[17] Zhao R J, Gui G, Xue Z, et al. A novel intrusion detection method based on lightweight neural network for Internet of things[J]. IEEE Internet of Things Journal, 2022, 9(12): 9960-9972.

[18] Zhu S Z, Xu X L, Zhao J, et al. LKD-STNN: a lightweight malicious traffic detection method for Internet of Things based on knowledge distillation[J]. IEEE Internet of Things Journal, 2024, 11(4): 6438-6453.

[19] He H Y, Yang Z G, Chen X N. PERT: payload encoding representation from transformer for encrypted traffic classification[C]//Proceedings of the 2020 ITU Kaleidoscope: Industry-Driven Digital Transformation (ITU K). Piscataway: IEEE Press, 2020: 1-8.

[20] Lin X J, Xiong G, Gou G P, et al. ET-BERT: a contextualized datagram representation with pre-training transformers for encrypted traffic classification[C]//Proceedings of the ACM Web Conference 2022. New York: ACM Press, 2022: 633-642.

[21] Zhao R J, Zhan M W, Deng X W, et al. Yet another traffic classifier: a masked autoencoder based traffic transformer with multi-level flow representation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(4): 5420-5427.

[22] Mihai S, Yaqoob M, Hung D V, et al. Digital twins: a survey on enabling technologies, challenges, trends and future prospects[J]. IEEE Communications Surveys & Tutorials, 2022, 24(4): 2255-2291.

[23] Schroeder G N, Steinmetz C, Rodrigues R N, et al. A methodology for digital twin modeling and deployment for industry 4.0[J]. Proceedings of the IEEE, 2021, 109(4): 556-567.

[24] Liu J, Li C L, Bai J P, et al. Security in IoT-enabled digital twins of maritime transportation systems[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(2): 2359-2367.

[25] Lv Z H, Qiao L, Mardani A, et al. Digital twins on the resilience of supply chain under COVID-19 pandemic[J]. IEEE Transactions on Engineering Management, 2024, 71: 10522-10533.

[26] Saad A, Faddel S, Youssef T, et al. On the implementation of IoT-based digital twin for networked microgrids resiliency against cyber attacks[J]. IEEE Transactions on Smart Grid, 2020, 11(6): 5138-5150.

[27] 张会迈, 胡晓娅, 周纯杰. 面向 TSN 工控系统安全策略生成与优化的数字孪生框架[J]. 系统仿真学报, 2025, 37(4): 861-874.
Zhang H M, Hu X Y, Zhou C J. Digital twin framework for the generation and optimization of security policies for TSN industrial control systems[J]. Journal of System Simulation, 2025, 37(4): 861-874.

[28] Liu K Z, Yang M, Ling Z, et al. Samba: detecting SSL/TLS API misuses in IoT binary applications[C]//Proceedings of the IEEE INFOCOM 2024 - IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2024: 2029-2038.

[29] Banerjee S, Agarwal A, Singla S. LLMs will always hallucinate, and we need to live with this[C]//Proceedings of the Intelligent Systems Conference. Berlin: Springer, 2025: 624-648.

[30] Bijoor Y, Kantode R, Chavan O, et al. Intelligent system to convert natural language queries to deterministic finite automata[C]// Proceedings of the International Conference on Data Science and Big Data Analysis. Berlin: Springer, 2024: 147-165.

[31] Leung H. Descriptional complexity of nfa of different ambiguity[J]. International Journal of Foundations of Computer Science, 2005, 16(5): 975-984.

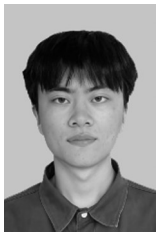
[32] Moustafa N, Slay J. UNSW-NB15: a comprehensive data set for net-

work intrusion detection systems (UNSW-NB15 network data set)[C]// Proceedings of the 2015 Military Communications and Information Systems Conference (MilCIS). Piscataway: IEEE Press, 2015: 1-6.
[33] Habibi Lashkari A, Draper Gil G, Mamun M S I, et al. Characterization of tor traffic using time based features[C]//Proceedings of the 3rd International Conference on Information Systems Security and Privacy. Setúbal: SciTePress - Science and Technology Publications, 2017: 253-262.
[34] Moustafa N, Slay J, Creech G. Novel geometric area analysis technique for anomaly detection using trapezoidal area estimation on large-scale networks[J]. IEEE Transactions on Big Data, 2019, 5(4): 481-494.
[35] Ferrag M A, Friha O, Hamouda D, et al. Edge-IIoTset: a new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning[J]. IEEE Access, 2022, 10: 40281-40306.
[36] Wang W, Zhu M, Zeng X W, et al. Malware traffic classification using convolutional neural network for representation learning[C]//Proceedings of the 2017 International Conference on Information Networking (ICOIN). Piscataway: IEEE Press, 2017: 712-717.
[37] Booij T M, Chiscop I, Meeuwissen E, et al. ToN_IoT: the role of heterogeneity and the need for standardization of features and attack types in IoT network intrusion data sets[J]. IEEE Internet of Things Journal, 2022, 9(1): 485-496.
[38] Neto E C P, Dadkhah S, Ferreira R, et al. CICIoT2023: a real-time dataset and benchmark for large-scale attacks in IoT environment[J]. Sensors, 2023, 23(13): 5941.

[作者简介]

何高峰（1984– ），男，安徽安庆人，博士，南京邮电大学副教授、硕士生导师，主要研究方向为网络安全和物联网安全。

田健峥（2001– ），男，安徽滁州人，南京邮电大学硕士生，主要研究方向为物联网恶意攻击和异常流量检测。



李亚文（2002– ），男，江苏盐城人，南京邮电大学硕士生，主要研究方向为时间序列异常检测。

徐丙凤（1986– ），女，安徽安庆人，博士，南京林业大学副教授、硕士生导师，主要研究方向为网络安全威胁建模与评估。

朱海婷（1983– ），女，江苏如皋人，博士，南京邮电大学副教授、硕士生导师，主要研究方向为网络管理和网络性能优化。

张璐（1983– ），男，江苏盐城人，博士，南京审计大学副教授，主要研究方向为网络与信息安全、信息系统审计、智能审计等。

郭乃瑄（1991– ），男，江苏盐城人，博士，盐城工学院讲师，主要研究方向为网络安全、图像处理等。