文章编号:1671-4598(2025)12-0042-09 DOI:10.16526/j.cnki.11-4762/tp.2025.12.006 中图分类号:TP393 文献标识码:A

基于 FL 优化联合学习的 WSN 隐私数据入侵节点时空图异常检测方法

王培春[1]，任晓龙[2]
(1. 中电科网络安全科技股份有限公司，成都 610095；
2. 中国科学院 计算技术研究所，北京 100190)

摘要：跨 WSN 子网的时空数据呈现强非独立同分布特性，且子网间容易遭受针对性攻击，如标签翻转攻击、后门攻击，降低 WSN 的安全性；为了解决这些问题，提出了基于 FL 优化联合学习的异常检测方法；各 WSN 子网基于本地数据生成隐私化时空图训练集，并采用 FL 优化联合时空滑动平均法对数据进行平滑处理，以消除传感器噪声干扰，抑制数据突变与异常抖动；对滑动时空窗口内的数据进行最大最小值归一化处理，确保数据分布均匀性，从而提升后续异常检测的准确性；构建基于 FL 优化联合学习的异常检测框架，聚合 WSN 中边缘节点的联邦平均参数，建立异常检测 FL 优化联合学习目标，通过联邦子域微调机制和加密参数共享，结合差分隐私与动态权重聚合，在适配非独立同分布时空数据的同时抑制针对性攻击，实现安全精准的跨域异常检测；实验结果表明，该方法在标签翻转攻击、后门攻击模式下节点空间分布熵最大值分别为 0.8 和 0.6，射频信号标识模拟误差分别为 0.7、0.3，与实验指标一致，说明使用该方法检测结果精准，能够有效保障 WSN 隐私数据的传输与存储。

关键词：FL 优化联合学习；WSN 隐私数据；入侵节点；时空图；异常检测

Spatiotemporal Graph Anomaly Detection Method for WSN Privacy Data Intrusion Nodes Based on FL Optimization Joint Learning

WANG Peichun[1], REN Xiaolong[2]
(1. Cyberspace Security Technology Co., Ltd., CETC, Chengdu 610095, China;
2. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: The spatiotemporal data across wireless sensor network (WSN) subnets exhibits the characteristics of strong non-independent and identical distribution, which are susceptible to targeted attacks such as label flipping attacks and backdoor attacks, thus reducing the security of WSN. To address these issues, an anomaly detection method based on federated learning (FL) optimization joint learning is proposed. Each WSN subnet generates a private spatiotemporal graph training set based on local data, and the FL optimization joint spatiotemporal sliding average method is used to smooth the data, thereby eliminating the noise interference of sensors and suppressing data mutations and abnormal fluctuations. Normalize the maximum and minimum values of sliding spatiotemporal window data to ensure uniform data distribution, thus improving the accuracy of subsequent anomaly detection. Build an anomaly detection framework based on FL optimization joint learning, aggregate the federated average parameters of edge nodes in WSN, establish the FL optimization joint learning objective for anomaly detection, use the federated subdomain fine-tuning mechanism and encrypted parameter sharing, and combine differential privacy with dynamic weight aggregation to adapt to non-independent and identically distributed spatiotemporal data while suppressing targeted attacks, thereby achieving secure and accurate cross-domain anomaly detection. Experimental results show that under the label flipping attack and backdoor attack modes, the maximum values of node spatial distribution entropy are 0.8 and 0.6 and the simulation errors of the RF signal identification are 0.7 and 0.3, respectively, which

收稿日期:2025-05-13； **修回日期**:2025-07-01。
基金项目:浙江省纪律检查委员会浙江省公权力大数据监督应用(二期)项目(E432059)。
作者简介:王培春(1973-),男,硕士,高级工程师。
通讯作者:任晓龙(1988-),男,博士,工程师。
引用格式:王培春,任晓龙.基于 FL 优化联合学习的 WSN 隐私数据入侵节点时空图异常检测方法[J].计算机测量与控制,2025,33(12):42-50.

are consistent with experimental indicators. This indicates that the detection results by this method are accurate and can effectively guarantee the transmission and storage of WSN privacy data.

Keywords: FL optimization joint learning; WSN privacy data; intrusion nodes; spatiotemporal diagram; abnormal detection

0　引言

WSN 无线传感网络在健康监测、科研数据采集、环境监测、军事侦察等方面具有广阔应用前景，然而在这种情况下，WSN 隐私数据相关传感器节点极易成为入侵节点，导致时空图异常数据注入。现有研究大多采用加密、认证等传统防御手段，但在入侵节点被攻破时，隐私数据仍可能泄露[1]。目前仅有极少数检测技术采用了加密、认证及密钥管理等传统防御手段作为辅助，在某个入侵节点被攻击突破时，其存储或关联的 WSN 隐私数据全部机密会被窃取，从而导致现有的防范措施无法奏效[2]。为了应对这种情况，将入侵检测技术作为网络安全领域的二道防线应用于 WSN 隐私数据安全防护体系，可以有效地发现针对隐私数据的恶意攻击。目前，有关入侵节点引发的时空图异常检测技术研究较少。

文献［3］提出了基于边残差注意力机制的检测方法，将网络流量转化成一系列图快照，使用边注意力层从每个离散快照中提取空间信息，给予高相似性节点权重，强化其空间特征。利用 BiGRU 捕获 IP 对之间通信演变，实现入侵检测。该方法在入侵检测过程中容易受到动态攻击影响，无法充分捕捉时空特征，影响检测精度；文献［4］提出了融合稀疏图注意力的检测方法，采用卷积神经网络提取时间戳上下文信息，使用全局时间戳编码和 Transformer 位置编码增强序列之间联系。利用稀疏自注意力关注重要的时间戳与特征，通过自注意力蒸馏突出重要特征，提升表示学习质量。通过构建基于预测和重构的综合损失函数，对模型参数进行优化。由于该方法过分依赖学习时间和特征两个维度复杂关系，无法实现对损失误差的精准判定，导致检测结果精准度不高；文献［5］提出了基于时空依赖关系和特征融合的检测方法，使用视频段之间的索引距离和特征相似程度拟合视频段的时间和空间依赖关系，构建视频段的关系特征，通过融合原始特征和关系特征表达入侵节点时空图动态特性和时序关系，以此实现异常检测。然而，该方法抗干扰性不高，影响检测精度；文献［6］提出了基于注意力特征融合的检测方法，采用轻量级注意力特征融合模块构筑融合机制进行不同空间特征的融合，在增强融合后特征表达能力同时减少网络参数量，提高异常检测算法性能。由于该方法跨域融合能力有限，容易导致不同特征域间存在一定差异，导致时空感知能力不足，影响检测效果。

针对以上问题，提出了基于 FL 优化联合学习的 WSN 隐私数据入侵节点时空图异常检测方法。将 FL 优化联合学习平均法应用于非独立同分布的时空图异常数据检测中，通过联邦子域微调机制使全局模型适配各子域的时空特征，有效解决了多域时空异构性问题。并采用联邦学习方式实现分布式加密模型训练与跨域共享，利用时空注意力加权联邦聚合算法对齐各节点模型，通过动态权重衰减联邦学习算法平衡各节点数据异构性，结合差分隐私时空扰动确保隐私安全隔离，在保护用户隐私的同时，能有效处理不同维度数据，获得可靠检测结果。

1　WSN 隐私数据入侵节点时空图数据 FL 联合优化处理

1.1　数据采集与预处理

由于 WSN 边缘域的入侵节点时空图数据是非独立同分布的，这种统计异构性导致很难训练出一个对所有边缘子域均适用的单一全局入侵检测模型，所以为了解决这种多域时空异构性的挑战，将基于 FL 优化联合学习平均法应用于非独立同分布的时空图异常数据检测之中[7]。

在 WSN 隐私数据采集及时空图预处理阶段，每周期动态采样一次本地入侵监测节点的多模态时空传感器数据，生成 t 时刻的时空图特征张量 $u_t \in U_{M\times L\times C}$，具体地，$M$ 为本地子域内可疑节点与邻域节点数量总和，L 为时空图窗口深度，C 为多模态特征维度[8]。动态标注当前时空图特征张量 u_t 的标签 Υ_{u_t}，判断时空图异常行为如公式（1）所示：

$$\Upsilon_{u_t}=\begin{cases}0 & \text{正常}\\ 1 & \text{异常}\end{cases} \tag{1}$$

若时空图异常校验判定为无攻击行为，则正常；若检测到跨子域攻击模式，则异常。连续采集时空图攻击事件流，直至满足分布式攻击模式收敛准则，即异常时空图张量 $\Upsilon_{u_t}=1$ 的隐私化累积计数超过动态阈值，或 FL 优化联合学习轮次达到自适应迭代上限[9-10]。本地子域生成隐私化时空图训练集，可表示为：

$$S(i)=\{[u_t(i),\Upsilon_{u_t}(i)]\}\cdot\omega_i \tag{2}$$

式中，样本权重 ω_i 由 i 个 FL 优化联合协同贡献熵与时空数据分布熵动态加权聚合确定[11]。在 FL 优化联合时空图入侵检测模型训练阶段，FL 优化联合中心首先将全局时空基模型分发至各边缘入侵检测子域[12]。各子域接收来自 FL 优化联合中心的相关参数，并使用本地

子域生成隐私化时空图训练集在相关参数基础上进行分布式时空图迭代优化，直至达到子域最大本地迭代轮次。

1.2 模型训练与参数聚合

将本地迭代完成的隐私化时空模型参数通过安全多方聚合协议上传至 FL 优化联合中心，该中心作为全局隐私协调方，需满足容忍最多 20% 的恶意子域[13]。FL 优化联合中心接收来自各边缘入侵检测子域的隐私化时空模型参数集 $\{S_{11},S_{12},\cdots,S_{1N}\}$，其中 N 为参与本次 FL 优化联合训练的子域总数[14]。FL 优化联合中心采用时空 FL 优化联合平均法完成全局时空模型参数聚合，生成本轮隐私化全局时空模型参数 S_{i+1}，并通过同态加密安全信道将 S_{i+1} 分发至各子域，供分布式时空图入侵检测优化使用。

1.3 数据平滑与归一化

WSN 无线传感器网络中的节点容易受到电磁干扰，环境温度的变化，以及硬件的老化等因素的影响，使得采集到的数据中含有大量的高频噪音[15]。因此，为了消除传感器噪声干扰，抑制数据突变与异常抖动，对参数集 $\{S_{11},S_{12},\cdots S_{1N}\}$ 使用 FL 优化联合时空滑动平均法对数据集平滑处理，计算式为：

$$S'_{ij}=\begin{cases}S(i)\sum_{j}^{LN}S_{ij} & N<K\\ \sum_{j}^{aN}S_{ij} & \\ \dfrac{}{M-L(K-1)} & N=K\end{cases} \tag{3}$$

式中，j 为边缘入侵检测子域捕获的时空图数据[16]；K 为滑动窗口个数[17]。当数据维度不一致时，会影响到检测准确性和稳定性。因此，为了使滑动时空窗口中的数据更均匀地分布，避免由于维数不同而造成窗宽不合理问题，需要对数据进行最大最小值归一化操作，公式为：

$$S''_{ij}=\frac{S'_{ij}-min(S'_{ij})}{max(S'_{ij})-min(S'_{ij})} \tag{4}$$

在 FL 优化联合中，归一化操作可以在不泄露原始数据的情况下进行，并且归一化后的数据使得异常数据点在数值上更加突出，方便对时空图异常数据检测。

2 基于 FL 优化联合处理的时空图异常检测

2.1 联邦子域微调机制

由于全局时空图模型与 WSN 边缘子域的本地时空图数据存在显著分布差异，且网络容易受到标签翻转攻击、后门攻击等针对性攻击，因此，通过联邦子域微调机制使全局模型适配各子域的时空特征。为使全局模型适配各子域的时空特征，在全局模型分发至各边缘子域后，子域利用本地独有的时空图数据对模型进行微调。本地数据包含子域特定的节点分布、通信模式等时空信息，通过反向传播算法，依据本地数据标签与模型预测结果的损失，更新模型参数，使模型学习到子域的独特特征。

平衡全局模型泛化能力和子域模型个性化性能方面，在微调过程中，设置较小的学习率，避免模型过度偏离全局模型，从而保留全局模型从多子域数据中学习到的通用知识，保证泛化能力。同时，限制微调轮次，防止子域模型过度拟合本地数据，失去对其他子域异常模式的识别能力。此外，采用正则化技术，约束模型参数变化幅度，使模型在适配子域特征的同时，仍能保持对全局时空模式的敏感度，实现泛化与个性化的平衡，提升异常检测的准确性和可靠性。

该框架采用中央聚合节点与边缘子域节点的双层架构，即参与的 WSN 设备分为两种角色：联邦时空客户端与联邦聚合服务端[18]。其中，主要参与节点为边缘子域客户端，任何部署多模态传感器的 WSN 节点均可成为客户端，主要职责是独立训练基于时空图特征的入侵检测模型，并将本地模型更新加密上传至聚合服务端。而聚合服务端通常部署于边缘计算网关或云端，主要职责是调度边缘子域的时空计算任务、聚合多源时空图模型参数，并生成全局隐私化入侵检测模型。边缘子域客户端使用本地采集的时空图数据协同训练入侵检测模型，并服从聚合服务端的分布式训练指令与参数同步策略[19]。聚合服务端通过动态任务分配机制调度各边缘子域的模型训练，负责全局时空模型的分发、参数隐私化聚合以及异常检测规则的跨域更新[20]。基于 FL 优化联合学习的异常检测框架，如图 1 所示。

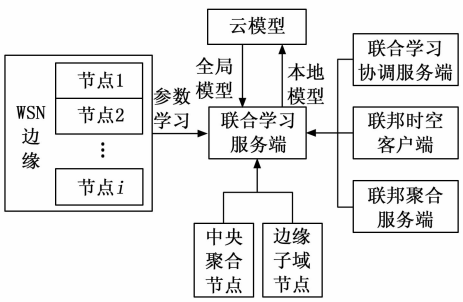


图 1 基于 FL 优化联合学习的异常检测框架

检测初始化阶段通过基于 FL 优化联合学习的异常检测框架，将初始化信息从联合学习协调服务端发送到各个 WSN 边缘节点客户端，并且本轮次所有参与异常数据聚合的候选 WSN 节点客户端将确定下来。等到参与聚合的 WSN 节点客户端训练到预设的本地时空特征迭代轮次后，客户端上传各自训练完成的入侵检测模型梯度更新量及时空图结构增量信息至服务端，服务端将接收到的多节点时空特征更新通过加权时空注意力聚合后再次分发给各个客户端，开启下一轮时空异常模式挖

掘训练，如此反复进行，直到入侵检测模型的时空异常识别准确率达到收敛阈值。联合学习框架中的各个 WSN 异构节点可呈现多模态计算架构差异，使其能够充分适配海量异构无线传感终端，共同训练基于 FL 优化联合学习的异常检测框架。该框架能够高效整合无线传感器网络（WSN）全域设备的多源时空数据流与分布式边缘计算资源，同时依托差分隐私时空扰动确保各节点原始时空采样数据与局部异常检测特征的隐私安全隔离。

在基于 FL 优化联合学习的异常检测框架中，差分隐私通过向模型参数更新过程添加噪声实现。具体而言，每个边缘子域在完成本地模型训练后，不直接上传原始的模型参数更新，而是先对参数更新进行归一化处理，再根据设定的隐私预算计算所需添加的噪声量，采用拉普拉斯噪声或高斯噪声。添加噪声后的参数更新被发送至中心服务器。这样，即使攻击者获取了部分子域的参数更新，也无法准确推断出单个数据点的信息，有效抑制针对性攻击，保护用户隐私。

动态权重聚合依据各子域数据质量、模型性能等因素计算权重。以子域数据的新鲜度、完整性以及本地模型在验证集上的准确率为指标，通过加权平均方式聚合模型，以提升检测精度，因为权重高的子域模型对全局模型的贡献更大，使聚合后的模型更贴合高质量数据特征，从而更精准地识别异常。

2.2　云模型聚合与目标构建

云模型聚合是联邦学习框架中全局模型更新的核心环节，旨在整合多个参与方（如不同机构或边缘设备）的本地模型更新。对所有 WSN 边缘节点模型进行本地时空特征训练后，将其上传至 FL 优化联合学习云平台进行聚合。使用时空注意力加权联邦聚合算法对齐各节点模型，在每轮训练中对 WSN 边缘节点模型进行平均，得到的平均模型为：

$$f(S''_{ij}) = \frac{1}{G}\sum_{g=1}^{G} S''_{ij} \tag{5}$$

式中，$g(g = 1,2,\cdots,G)$ 表示机构数量，在多次迭代之后，云端模型具有更好泛化能力[20]。WSN 隐私数据入侵节点时空图异常全局数据联邦平均参数聚合，如公式（6）所示：

$$S''_{i} = S''_{i-1} - \frac{g}{G}\sum_{i=1}^{n} S''_{i-1} \tag{6}$$

基于此，构建的异常检测 FL 优化联合学习目标，如公式（7）所示：

$$\underset{\omega,e}{\operatorname{argmin}} Q = \sum_{i=1}^{n} l[S''_{i}, f(S''_{ij})] \tag{7}$$

式中，e 为参数学习偏差；l 为损失函数；n 为数据集的大小。在时空图联邦训练阶段，各个 WSN 边缘节点客户端将本地采集的多模态时空数据流中的入侵行为样本分离出来，构成时空异常样本集合。具体而言，异常样本定义为本地数据中满足时空异常模式标签的样本，如节点流量突增、拓扑结构异常波动等。各客户端将本地异常样本集合通过安全聚合通道上传至联合学习云平台。云平台接收来自各节点的异常样本集合，并通过归一化处理打乱样本顺序，生成全局时空异常训练集。云平台基于该数据集对全局时空图神经网络模型进行分布式迭代训练，采用动态权重衰减联邦学习算法平衡各节点数据异构性，直至模型达到预设的最大迭代次数。

在联合异常检测阶段，各 WSN 节点时空采样窗口获取本地传感器的多维度时空数据，生成待检测时空图样本。将待检测样本输入训练好的全局时空异常检测模型，输出对当前样本的时空维度异常率 P_1 和时序维度异常率 P_2，并通过时空注意力加权融合计算综合异常检测分数：

$$P = \frac{\lambda P_1 + (1-\lambda)P_2}{2} \tag{8}$$

式中，λ 为模型置信度。由于全局时空图异常检测模型对分布式入侵行为具有高敏感度，设置模型置信度小于 0.5。当综合时空异常检测分数超过预设的模型置信度，则判定当前 WSN 子网存在时空协同异常，否则判定为正常状态。

基于联邦优化联合学习框架，动态聚合各 WSN 边缘节点（客户端）中分布式时空图异常样本，通过联邦隐私化时空图融合机制将分散的时空图特征编码至中心服务器，完成跨子网时空图联邦模型的训练。同时，在本地节点侧构建时空注意力增强模型，利用各节点独立标注的异常时空图样本进行强化学习，强化模型对非独立同分布时空异常模式的捕获能力。通过动态加权融合策略整合的全局泛化能力，计算综合异常检测分数，实现高精度隐私化时空图异常检测。

3　实验

3.1　实验装置

以 LG-WSN02 型无线传感网络实验系统为研究对象，系统平台采用模块化设计，包括以 ZigBee 片上系统 CC2530 为核心 IEEE802.15.4 网络，基于 IEEE802.11B/G 协议 Wi-Fi 核心网络、蓝牙网络、高集成度 GPRS/GSM 电讯网络和 100 M 以太网等在同一平台。LG-WSN02 型无线传感网络体系结构如图 2 所示。

该系统采用模块化设计，集成多种网络协议，能模拟多种通信环境，一定程度上可反映实际 WSN 应用场景的多样性。然而，实际 WSN 应用场景更为复杂，如节点部署环境差异大、通信干扰因素多等。该实验系统在节点数量、网络规模等方面可能无法完全覆盖实际场景。不过，通过精心设计的实验参数方案，如设置不同

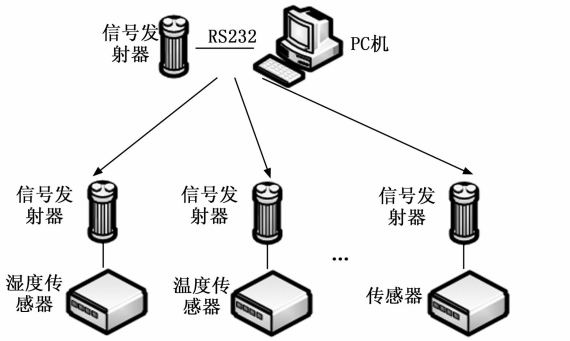


图 2 LG-WSN02 型无线传感网络体系结构

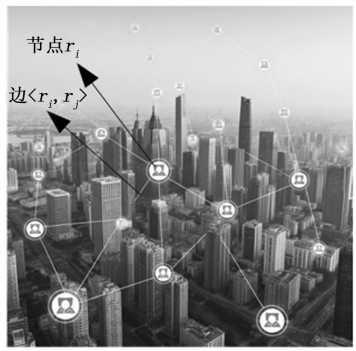


图 3 WSN 入侵节点拓扑分布图

节点数量、网络半径等，可在一定程度上模拟实际场景的复杂情况。围绕 WSN 隐私数据入侵检测需求设计的网络参数方案，如表 1 所示。

表 1 实验参数

参数类型	参数值
节点数量	20～50 个
网络半径	5～35 m
部署模式	混合网格＋随机分布
通信协议	LoRaWAN
传输频率	2.4 GHz
数据包长度	64～256 B
传输速率	50 kbps
隐私数据类型	工业传感器
数据注入率	5%～15%
节点传输频率	10 min/次

一个由无线传感器节点组成的自组织分布式无线传感网络，其核心功能是通过多类型传感器模块对环境中的物理/化学参数进行实时感知、本地化数据融合处理，并经由多跳无线通信协议将结构化环境信息传输至汇聚节点或远程监控终端。

3.2 实验环境

用贝叶斯攻击溯因网络来模拟 WSN 隐私数据入侵场景的空间相关性，贝叶斯入侵关联网络结构图是一个有向无环图 G，其节点表示隐私数据泄露风险变量 r_i，连接两个节点的有向边表示风险变量间存在入侵因果传导关系。考虑到实际 WSN 在中小型监测区域，网络节点数量设定为 50 个，设传感器节点均匀部署，通信范围为 20 m，通过计算节点间距离并结合实际通信干扰因素，将节点间的连接概率设置为 0.35，任意两个节点之间存在直接通信链路并进行数据交互的概率约为 35%。面向隐私数据的 WSN 入侵节点拓扑分布如图 3 所示。

图中节点 r_i 为部署在隐私监测域的高风险数据泄露节点，假设 WSN 隐私入侵场景的拓扑结构可以用连通图 $G=<V, E>$ 表示，定义节点边数作为节点的隐私泄露中继风险度。V 表示部署在监测域的隐私敏感节点集合，Ψ 表示连接节点的动态入侵传播链路，连接敏感节点的单跳通信链路代表有向图的边 $(r_i, r_j) \in \Psi$。高风险隐私节点部署在流域中以监测流域隐私数据链，每个节点有一个动态抗逆向加密标识符，并按隐私泄露威胁指数动态编码。

节点能感知多维隐私入侵特征属性，例如：数据包时空图异常篡改特征向量，即节点位置与时间戳的跨跳攻击路径出现偏移；设备标识伪造漂移量，即射频信号标识模拟误差，判断 WSN 隐私数据入侵节点时空图是否异常。

滑动窗口大小的选择需综合多方面因素，为避免过度平滑掩盖局部特征，窗口不宜过大；若数据存在明显周期性，窗口大小应与周期匹配；从异常检测目标出发，若异常特征尺度大、持续时间长，大窗口能更好检测；若异常是瞬间突变，小窗口可避免平滑掉异常特征。由此可通过均方误差衡量实验，在不同窗口大小下评估平滑效果，选取效果最佳且资源消耗合理的窗口大小，并采用加权平均法确定窗口内数据权重。

3.3 实验攻击设置

为了方便实验，设置了投毒攻击模式。在实验模拟的投毒攻击场景中，恶意节点可通过本地数据篡改模块对传感器采集的原始时序特征集实施选择性样本删除或恶意样本注入，以此改变正常数据初始分布，并降低 WSN 性能。常用的投毒攻击方法主要有两种，分别是标签翻转攻击和后门攻击。

3.3.1 标签翻转攻击

标签翻转攻击即攻击者将 WSN 隐私数据存储中的数据标签改变，并往存储数据中植入提前预备好的恶意数据点，这样可以使得攻击者可自由设定错误标签诱导 WSN 隐私数据存储模型。恶意攻击者通过多轮数据交互从服务端获取全局配置参数，并用其更新本地存储处理参数，恶意攻击者接下来对本地存储处理参数进行数据篡改训练。其中部分或全部隐私数据样本的标签被人为地翻转，最后将训练完成得到的篡改后参数加密打包

后回传到存储服务端。当存储服务端根据最近更新的受攻击的参数进行数据聚合后，WSN 隐私数据存储系统会受到恶意攻击者的破坏并在之后的数据交互中进一步破坏正常用户的数据处理结果。将特定数据错误分类为错误标签对应类别标签翻转方法进行投毒攻击的流程，如图 4 所示。

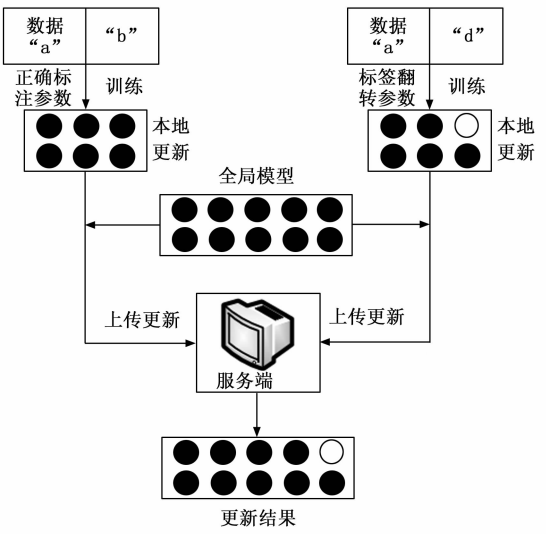


图 4　标签翻转投毒攻击流程

因为在 WSN 运行过程中无法确定恶意攻击目标，所以恶意攻击者获取与回传隐私数据存储篡改参数、恶意数据样本等内容都不会被存储服务端拒绝。并且出于隐私保护的角度，所有隐私数据处理都是在设备本地进行的，存储服务端没有办法也不能去干涉审查，所以无法检查到被篡改标签的隐私数据样本。同时，恶意攻击者能够调整攻击的误导标签。

3.3.2　后门攻击

后门投毒攻击使用了一些特殊的隐藏模式，这些模式被叫做恶意存储后门触发器。它们由恶意攻击者在 WSN 隐私数据样本上专门设计，能够干预隐私数据存储模型的训练阶段，并且在后续数据预测阶段产生的结果与实际正确结果完全不一致。设定攻击频率为每 100 个数据传输周期进行一次攻击操作，数据传输周期为 1 分钟，在一个包含 1 000 个样本的数据集中，每次攻击会翻转 200 个样本的标签，后门攻击的触发频率设定为在模型预测过程中，每 500 次预测尝试触发一次后门。在训练数据中，注入带有后门触发模式的恶意样本比例为 15%。例如，在一个包含 2 000 个训练样本的数据集中，会注入 300 个带有后门触发模式的恶意样本。通过调整这些参数，可以模拟不同强度的后门攻击，以全面评估所提方法对后门攻击的抵御能力。后门投毒攻击流程，如图 5 所示。

对于一个 WSN 隐私数据存储场景下的数据分类任

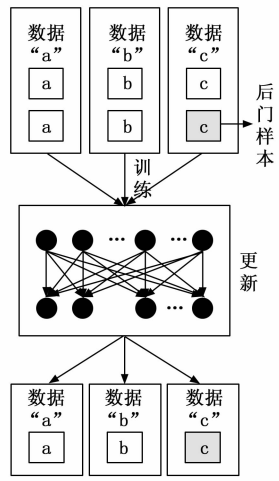


图 5　后门投毒攻击流程

务，恶意攻击者的后门投毒攻击目标是诱导模型将特定数据类别错误判定为预设的异常风险标签，恶意存储后门触发器为植入投毒训练数据的特定区域的特殊加密冗余编码块。因此在隐私数据查询分析阶段时，被植入后门的存储管理模型将能够识别出嵌入该恶意存储后门触发器的数据样本，并且将错误判定输出为投毒设定的异常风险标签，与此同时未携带恶意存储后门触发器的正常数据样本依然可以被正确识别输出。

3.4　实验指标

数据包时空图异常篡改特征向量的表征量是节点空间分布熵，设备标识伪造漂移量的表征量是射频信号标识模拟误差，将其作为实验指标，如图 6 所示。

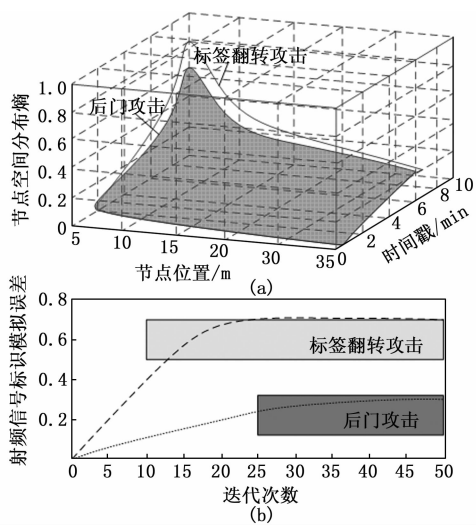


图 6　实验指标

由图 6（a）可知，X 轴节点位置代表了某个特定区域内节点的物理位置分布。Y 轴时间显示了时间维度上的变化。Z 轴节点空间分布概率反映了在特定时间和位置下节点存在的概率。图中有一个三维曲面，显示了节点空间分布熵随节点位置和时间的变化情况。从图中

可以看出，节点空间分布概率是节点位置和时间的函数，呈现出复杂的三维变化趋势。标签翻转攻击情况下节点空间分布熵比后门攻击情况下的分布熵要高，两种攻击下的最大值分别为 0.8 和 0.6，节点空间分布熵值越大，说明数据包时空图异常篡改特征向量的异常程度越高。

由图 6（b）可知，标签翻转攻击下的模型性能指标在开始时较高，但随着迭代次数的增加，性能指标也逐渐下降，不过下降速度相对较慢。这说明标签翻转攻击同样会对模型性能产生负面影响，但影响程度相对后门攻击较小。随着迭代次数的增加，后门攻击下的模型性能指标逐渐下降。这表明后门攻击对模型的性能有显著的负面影响，且随着迭代次数的增多，这种影响愈发明显。在迭代次数为 0 时，两种攻击类型下的模型性能指标相近，但随着迭代次数的增加，两者的性能差异逐渐增大。在标签翻转攻击情况下射频信号标识模拟误差在迭代次数为 10 次时，就达到了 0.6。随着迭代次数继续增加，误差稳定在 0.7。在后门攻击下射频信号标识模拟误差在迭代次数为 25 次时逐渐稳定，最终保持 0.3 不变。

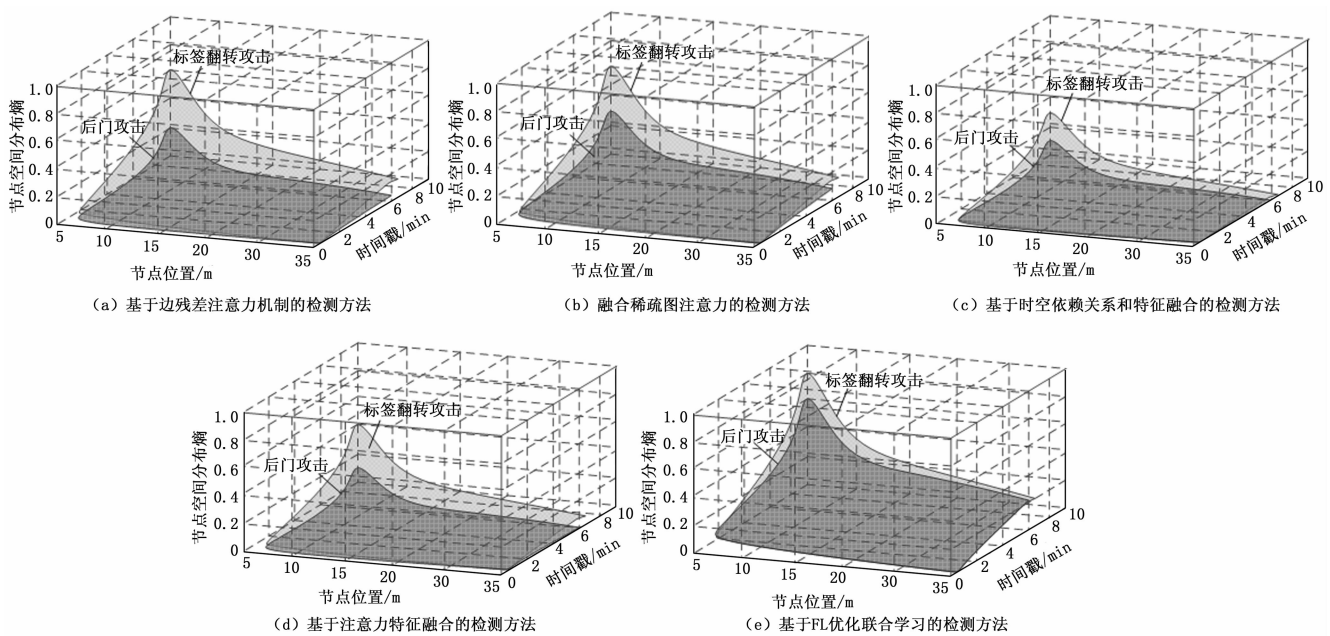
3.5 实验结果与分析

使用基于边残差注意力机制的检测方法、融合稀疏图注意力的检测方法、基于时空依赖关系和特征融合的检测方法、基于注意力特征融合的检测方法、基于 FL 优化联合学习的检测方法，对比分析数据包时空图异常篡改特征向量，分析结果如图 7 所示。

由图 7 可知，使用基于边残差注意力机制的检测方法动态攻击自适应能力弱，难以对复杂的时空特性进行有效捕获，导致对异常篡改特征的敏感性不够，在标签翻转攻击和后门攻击情况下，节点空间分布熵分别为 0.6 和 0.2；使用融合稀疏图注意力的检测方法过度依赖时间与特征维度的复杂关系建模，导致损失误差判定不精准，在标签翻转攻击和后门攻击情况下，节点空间分布熵分别为 0.6 和 0.3；使用基于时空依赖关系和特征融合的检测方法抗干扰能力不足，难以拟合入侵节点的时空动态特性，易受噪声影响，在标签翻转攻击和后门攻击情况下，节点空间分布熵分别为 0.3 和 0.1；使用基于注意力特征融合的检测方法跨域融合能力有限，特征域差异导致时空感知能力不足，无法有效区分异常模式，在标签翻转攻击和后门攻击情况下，节点空间分布熵分别为 0.4 和 0.1；使用基于 FL 优化联合学习的检测方法，通过 FL 优化联合框架，结合多客户端的异常样本训练，显著提升 FL 优化联合鲁棒性，在标签翻转攻击和后门攻击情况下，节点空间分布熵分别为 0.8 和 0.6。只有使用基于 FL 优化联合学习的检测方法数据包时空图异常篡改特征向量分析结果与实验指标一致，其余方法均不一致。

使用不同方法对比分析射频信号标识模拟误差，对比结果如图 8 所示。

由图 8（a）可知，基于 FL 优化联合学习的检测方法误差最大值为 0.7，与实验指标一致，而使用基于边残差注意力机制的检测方法、融合稀疏图注意力的检测方法、基于时空依赖关系和特征融合的检测方法、基于注意力特征融合的检测方法对应的误差最大值分别为 0.5、0.5、0.54、0.6，与实验指标不一致。这表明基于 FL 优化联合学习的检测方法在模拟误差最大值这一



（a）基于边残差注意力机制的检测方法　（b）融合稀疏图注意力的检测方法　（c）基于时空依赖关系和特征融合的检测方法

（d）基于注意力特征融合的检测方法　（e）基于FL优化联合学习的检测方法

图 7　不同方法数据包时空图异常篡改特征向量对比结果

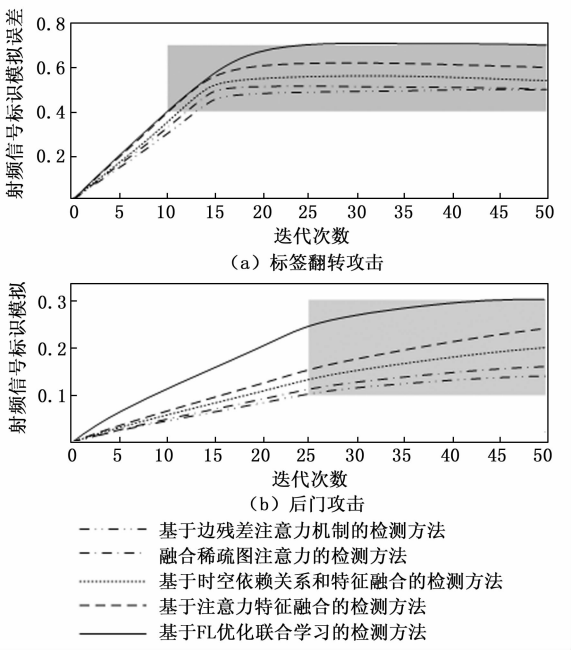


图 8　不同方法射频信号标识模拟误差对比结果

指标上更贴合实验情况，说明其在处理数据时能够更好地反映真实场景下的误差范围。

由图 8（b）可知，基于边残差注意力机制的检测方法、融合稀疏图注意力的检测方法、基于时空依赖关系和特征融合的检测方法、基于注意力特征融合的检测方法、基于 FL 优化联合学习的检测方法的误差最大值分别为 0.14、0.16、0.2、0.24、0.3，只有使用基于 FL 优化联合学习的检测方法模拟误差与实验指标一致。这进一步说明基于 FL 优化联合学习的检测方法在误差最大值的模拟上更符合实验要求，能够更精准地反映检测过程中的误差情况，而其他方法在这方面存在不足，进而影响整体的检测精准度。

由上述结果可知，基于边残差注意力机制的检测方法的局限性在于无法有效处理复杂时空特性的攻击，导致对异常篡改特征的敏感性不足，使其在面对标签翻转攻击和后门攻击时，无法准确识别攻击行为，从而影响了检测结果的准确性。融合稀疏图注意力的检测方法过度依赖复杂关系建模，容易受到数据噪声和干扰的影响，导致损失误差判定不精准。在实际应用中，由于 WSN 环境的复杂性，数据中存在大量噪声，这进一步降低了该方法的检测性能。基于时空依赖关系和特征融合的检测方法抗干扰能力不足，难以拟合入侵节点的时空动态特性。在复杂攻击场景下，无法准确捕捉攻击的时空特征，导致检测结果出现偏差。基于注意力特征融合的检测方法跨域融合能力有限，无法有效区分异常模式。在面对特征域差异较大的攻击时，无法准确识别攻击行为，从而影响了检测结果的可靠性。相较于其他方法，在不同攻击场景下，性能提升幅度较大。在标签翻转攻击和后门攻击情况下，节点空间分布熵和射频信号标识模拟误差等指标均明显优于其他方法，能够有效提高 WSN 隐私数据入侵检测的准确性和可靠性。

4　结束语

由于 WSN 隐私数据入侵节点时空图数据集的异构性，使用传统异常检测方法容易出现误报率增加或模型鲁棒性降低的问题。FL 优化联合学习框架的应用能够解决这些问题，使用时空滑动平均法对数据集平滑处理，利用数据最大最小值归一化操作，使滑动时空窗口中的数据更均匀地分布。通过计算综合异常检测分数，实现异常情况检测。为了使实验环境更接近真实环境，设置了标签翻转、后门攻击模式，在这两种攻击模式下的节点空间分布熵最大值、射频信号标识模拟误差均与实验指标一致，说明联邦学习（FL）优化联合学习的方式可以利用分布在各个客户端中的数据，通过聚合各客户端的模型更新来提升整体模型的性能。这种方式能够整合多源数据，充分利用不同客户端数据的多样性，使得模型在面对复杂的 WSN 隐私数据入侵节点时空图异常检测场景时，具有更好的泛化能力和适应性，从而更精准地检测异常情况，使用该检测方法也能够为下一代网络空间主动防御技术提供理论支撑与工程化路径。

参考文献：

[1] 金　明，丁　蓉. 一种联合时域和空域残差的网络异常检测与节点定位方法 [J]. 电子学报，2023，51（5）：1172－1178.

[2] 袁子淇，孙庆赟，周号益，等. MNDetector：基于多层网络的异常访问检测方法 [J]. 计算机研究与发展，2025，62（3）：765－778.

[3] 闫雷鸣，张定一，陈先意，等. 基于边残差注意力机制的动态图神经网络入侵检测方法 [J]. 中国电子科学研究院学报，2025，20（1）：10－18.

[4] 衡红军，代栋炜. 融合稀疏图注意力的多元时间序列异常检测方法 [J]. 计算机工程与设计，2025，46（3）：841－849.

[5] 柳德云，李　莹，周　震，等. 基于时空依赖关系和特征融合的弱监督视频异常检测 [J]. 数据采集与处理，2024，39（1）：204－214.

[6] 吴沛宸，袁立宁，胡　皓，等. 基于注意力特征融合的视频异常行为检测 [J]. 图学学报，2024，45（5）：922－929.

[7] 袁红春，张文凤. 融合 SimAM 注意力机制和双向 ConvLSTM 的异常检测方法 [J]. 小型微型计算机系统，2023，44（8）：1777－1784.

[8] 赵建军，汪旭童，崔　翔，等. 基于登录行为分析的失陷邮箱检测技术研究 [J]. 西安电子科技大学学报，2023，50（4）：34－44.

[9] 周 丹，凌 捷. 结合对比学习的双分支多维时间序列异常检测方法 [J]. 计算机应用研究，2025，42 (2)：507 - 513.

[10] 贾文雅，杨红菊. 融合判决门限和信任过滤机制的 WSN 异常节点定位方法 [J]. 传感技术学报，2024，37 (10)：1820 - 1826.

[11] 王 静，何苗苗，丁建立，等. 面向多维时间序列异常检测的时空图卷积网络 [J]. 西安电子科技大学学报，2024，51 (3)：170 - 181.

[12] 肖警续，郭渊博，常朝稳，等. 基于 SDN 的物联网边缘节点间数据流零信任管理 [J]. 通信学报，2024，45 (7)：101 - 116.

[13] 梁 硕，韩翔宇，李 慧，等. 分布式网络异常节点挖掘检测方法仿真 [J]. 计算机仿真，2023，40 (7)：409 - 413.

[14] 刘禹含，吉根林，张红苹. 基于骨架图与混合注意力的视频行人异常检测方法 [J]. 计算机应用，2024，44 (8)：2551 - 2557.

[15] 车丽娜，任秀丽. 基于滑动窗口和置信度的无线传感器网络异常检测算法 [J]. 传感技术学报，2023，36 (11)：1801 - 1807.

[16] 王 俊，赖会霞，万 玥，等. 基于角度的图神经网络高维数据异常检测方法 [J]. 计算机工程，2024，50 (3)：156 - 165.

[17] 袁 野，陈 明，吴安彪，等. 基于个性化 PageRank 和对比学习的图异常检测模型 [J]. 计算机科学，2025，52 (2)：80 - 90.

[18] 杨小龙，唐 婷，李兆玉，等. 基于图卷积神经网络的室内穿墙无源目标检测算法 [J]. 电子学报，2024，52 (2)：614 - 625.

[19] 李聪宇，赵利辉，安 洋. 基于图神经网络的物联网入侵检测研究 [J]. 中北大学学报 (自然科学版)，2024，45 (2)：194 - 204.

[20] 肖 迪，余柱阳，李 敏，等. 基于差分隐私与模型聚类的安全联邦学习方案 [J]. 计算机工程与科学，2024，46 (9)：1606 - 1615.

(上接第 27 页)

[8] LI K Y，LI L，WANG P，et al. A fast and non-destructive method to evaluate yield strength of cold-rolled steel via incremental permeability [J]. Journal of Magnetism and Magnetic Materials，2020，498：166087.

[9] 李丽娟，解社娟，陈洪恩，等. 碳钢塑性变形对增量磁导率信号的影响 [J]. 中国机械工程，2018，29 (14)：1653 - 1660.

[10] 何存富，丁冬冬，刘秀成，等. 65Mn 钢板电镀镍层厚度的增量磁导率检测方法 [J]. 北京工业大学学报，2020，46 (7)：727 - 733.

[11] 武伟康，陈剑云，刘思然，等. 基于增量磁导率的钢轨频变阻抗曲线计算 [J]. 铁道学报，2021，43 (12)：79 - 84.

[12] 宫彦双，完小康，王宏军，等. 基于增量磁导率提取的油气管道正交差分涡流内检测技术 [J]. 油气储运，2024，43 (9)：1002 - 1011.

[13] 涂洪铭，伍剑波，柯 瑞，等. 激励参数对 Q235 钢增量磁导率蝶形图的影响研究 [J]. 机械工程学报，2024，60 (2)：27 - 35.

[14] GUPTA B，UCHIMOTO T，DUCHARNE B，et al. Magnetic incremental permeability non-destructive evaluation of 12 Cr-Mo-W-V steel creep test samples with varied ageing levels and thermal treatments [J]. NDT & E International，2019，104 (6)：42 - 50.

[15] 傅 萍，王钰珏，何存富，等. 用于 DH590 钢塑性变形和残余奥氏体表征的多维微磁参量综合评价指标 [J]. 实验力学，2024，39 (3)：261 - 277.

[16] LIK Y，LIU Y R，WANGP，et al. Research on yield strength estimation and lift-off suppression method based on incremental permeability features [J]. Journal of Nondestructive Evaluation，2023，42 (2)：36.

[17] 边 闯，王海涛，刘向兵，等. 针对 RPV 钢磁巴克豪森噪声检测的传感器设计 [J]. 计算机测量与控制，2024，32 (2)：332 - 338.

[18] 杨理践，孙靖萌，高松巍，等. 激励信号对铁磁性材料矫顽力检测的影响 [J]. 无损探伤，2016，40 (3)：10 - 13.

[19] GRIMBERG R，LEITOIU S，BRADU B E，et al. Magnetic sensor used for the determination of fatigue state in ferromagnetic steels [J]. Sensors & Actuators A Physical，2000，81 (1/2/3)：371 - 373.

[20] STASHKOV A N，SCHAPOVA E A，NICHIPURUK A P，et al. Magnetic incremental permeability as indicator of compression stress in low-carbon steel [J]. NDT&EInternational，2021，118：102398.

[21] CHEN H E，XIE S，CHEN Z，et al. Quantitative nondestructive evaluation of plastic deformation in carbon steel based on electromagnetic methods [J]. Materials Transactions，2014，55 (12)：1806 - 1815.

[22] LIK Y，ZHANG Z D，WANG P，et al. A novel 3D simulation prediction model of mechanical properties of ferromagnetic materials via incremental permeability method [J]. Journal of Magnetism and Magnetic Materials，2021，536：168137.

[23] 张政达. 基于增量磁导率的铁磁性材料机械性能检测仿真方法研究 [D]. 南京：南京航空航天大学，2022.

[24] 李 超，张志杰，陈昊泽，等. 基于 ZYNQ 的涡流无损检测系统的设计 [J]. 计算机测量与控制，2021，29 (8)：56 - 61.