基于联邦学习与动态注意力的网络运行大数据异常识别

吴晓冰 [1]
WU Xiaobing

摘 要 网络运行大数据异常识别是保障网络系统安全稳定运行的关键环节。受网络环境动态多变以及异常模式随场景快速演变等因素的影响，现有方法难以对时序动态特征进行有效捕捉，因此对于突发与渐进式的识别精度也较低。对此，文章提出一种基于联邦学习与动态注意力的网络运行大数据异常识别方法。该方法采用层次聚类与差分隐私保护相结合的客户端选择策略，通过添加高斯噪声扰动数据特征并计算客户端间特征相似性，利用凝聚式层次聚类的方式筛选出异常样本丰富且分布多样的客户端，在保护隐私的同时提升数据代表性。引入动态多头注意力模块，将客户端数据映射到多个子空间，并利用动态权重矩阵依据时间步调整注意力权重，以此来捕捉数据的时序动态特征，提升对动态异常的感知能力。借助联邦学习算法，各客户端利用捕捉的时序动态特征结合本地数据更新模型参数，再由中央服务器将各客户端更新后的参数按数据量加权聚合形成全局模型，利用该模型计算新样本的异常概率并与设定阈值进行比较，实现异常判定。实验结果表明，所提方法在真实网络数据集上异常识别覆盖率均值达 85%，且在不同异常类型上均表现出更高的稳定性与鲁棒性，有效验证了其在提升识别精度与泛化能力方面的优势，为动态网络环境下实现高精度、隐私安全的异常识别提供了可行路径。

关键词 联邦学习；动态注意力；网络运行大数据；异常识别

doi：10.3969/j.issn.1672-9528.2026.05.024

0 引言

网络运行状态检测环节能够有效保障系统的安全和稳定，该环节主要通过在海量数据中识别出异常行为，进而预防潜在的故障和攻击。但现有的异常识别方法大多为集中式的检测模式，由于网络动态性在逐渐变强，其异常模式也会随着时间以及场景的变化而发生快速演变。常规的集中性检测模式难以适应日益复杂的网络环境，因此有必要研究出一种智能方法，通过搭建出自适应异常识别框架来保障网络运行的安全性。

此前诸多学者尝试将不同算法融入网络异常识别场景当中。毛华彬等 [1] 构建了一个包含多个隐藏层的深度神经网络，通过非线性激活函数来对传感器网络的深层特征进行逐层提取，但是无线传感器网络的数据具有比较强的时序依赖性，深度神经网络模型缺乏对于时序特征的显式建模，导致算法实际识别精度有待提高。韦霞等 [2] 利用集成学习策略训练多棵孤立树，每棵树分别对数据点的隔离路径长度进行计算并通过平均路径长度来判定异常概率，但异常模式在动态网络环境下通常会随着时间而呈现出特定的趋势或周期性的变化，孤立树由于独立处理数据点而难以捕捉这些时序动态特征，导致该算法对渐进式异常的检测能力不足。刘帅等 [3] 选择利用双通道注意力机制分别捕捉局部时序依赖与全局上下文信息，并融合双通道以及全连接层的输出结果完成异常分类，但是由于该算法难以动态调整注意力范围以适应不断变化的异常模式，导致无法准确聚焦关键信息，对异常特征的提取不全面。模型在面对新型异常时的泛化能力容易出现下降的情况，导致误判或漏判。高斌等 [4] 采用在线学习策略逐批次更新模型参数，以此来适应物联网数据的动态变化，同时还设计了基于阈值的动态分类规则，根据正常数据分布的统计特性对判别边界进行自动调整。但是由于在线学习模型对长期时序信息利用还不够充足，分类规则对异常时序特征响应存在一定的滞后性，该方法可能无法及时捕捉到异常模式的动态变化，导致对突发异常等模式出现检测延迟或失效的情况。算法的异常识别及时性不仅会因此降低，还可能会使网络系统在异常发生时无法及时采取应对措施，增加安全风险。为此，本文提出一种基于联邦学习与动态注意力的网络运行大数据异常识别方法，通过引入动态多头注意力模块来自适应捕捉数据流中的动态特性，以此提升对突发异常与渐进式异常的检测灵敏度。

1. 山西工程科技职业大学 山西晋中 030619

1 基于联邦学习与动态注意力的网络运行大数据异常识别方法

1.1 基于层次聚类与差分隐私保护的网络客户端选择

在网络运行大数据异常识别场景中，网络环境动态多变且异常模式随场景快速演变。不同客户端因特定业务场景或设备特性，其网络运行数据中的异常模式分布不均匀，部分客户端会频繁出现特定类型异常。若不进行客户端选择，直接使用所有客户端数据训练模型，模型可能会被异常样本分布不均的情况干扰，难以全面且精准地学习到各类关键异常模式，进而影响异常识别的准确性和有效性。因此，需从众多客户端中挑选出异常样本丰富的客户端，为模型提供更具代表性和多样性的数据，增强模型对关键异常模式的学习能力，因此需进行网络客户端选择。而层次聚类与差分隐私保护相结合的方式，能在保护客户端数据隐私的同时，有效实现客户端的选择。

假设每个客户端有对应的网络运行数据集，b_i 代表客户端 i 的初始数据特征分布。为保护客户端数据隐私，同时为后续聚类提供有效特征，本文选择对每个客户端的特征分布添加差分隐私保护 [5]。设添加差分隐私后的客户端 i 的数据特征分布为 $b_i^{'} = b_i + N(0, \sigma_i^2)$，其中 $N(0, \sigma_i^2)$ 代表均值为 0 且方差为 σ_i^2 的高斯噪声。通过上述操作，可以在保护数据隐私的前提下，对客户端数据特征进行一定程度的扰动，为后续安全地共享特征信息做好准备 [6]。各客户端将添加差分隐私后的数据特征分布 $b_i^{'}$ 上传至中央服务器。服务器接收到所有客户端上传的特征后，对客户端之间的特征相似性进行计算。具体计算公式为：

$$s_{ij} = \frac{1}{1 + d_{ij}} \tag{1}$$

式中：s_{ij} 代表客户端 i 和客户端 j 之间的特征相似性，d_{ij} 代表两个客户端特征分布之间的距离。具体计算公式为：

$$d_{ij} = \sqrt{\sum_{k=1}^{M} \left(b_{ik}^{'} - b_{jk}^{'} \right)^2} \tag{2}$$

式中：M 代表网络运行大数据特征向量的维度，$b_{ik}^{'}$ 和 $b_{jk}^{'}$ 分别代表客户端 i 和 j 的第 k 维特征值。通过计算所有客户端两两之间的相似性，可以得到相似性矩阵 $\boldsymbol{S}$。然后采用凝聚式层次聚类的方式对客户端进行选择 [7]。初始时将每个客户端视为一个单独的簇。按照相似性从高到低的顺序逐步合并相似的簇。当两个簇之间的平均相似性大于阈值 θ 时则可以进行合并。具体表达式为：

$$\frac{1}{n_p n_q} \sum_{i \in C_p} \sum_{j \in C_q} \boldsymbol{s}_{ij} \geq \theta \tag{3}$$

式中：C_p 和 C_q 分别代表两个不同的簇，n_p 和 n_q 分别代表簇 C_p 和 C_q 对应的客户端数量。通过不断对簇进行合并直到达到预设的终止条件并选择簇内特征分布最接近簇中心的客户端。结合上述操作选出的客户端涵盖不同特征类型的异常数据，用于帮助模型全面捕捉各类异常。

1.2 基于动态注意力的客户端网络运行数据时序动态性捕捉

在网络运行大数据异常识别过程中，网络环境动态多变且异常模式随场景快速演变。经过客户端选择后，虽然已经筛选出了异常样本丰富且数据隐私得到保护的客户端，但这些客户端上传的网络运行数据仍具有复杂的时序特性。不同时间步的数据对于异常识别的重要性不同，且异常模式可能在不同的时间尺度上表现出来。传统静态注意力机制由于无法适应时间步变化的动态特性，在捕捉数据时序特征方面存在明显局限性。因此，需要采用动态注意力机制来更好地处理客户端网络运行数据的时序信息，提高异常识别的准确性。

经过客户端选择后，各选定的客户端将本地处理后的数据，即添加差分隐私后的网络特征数据上传至中央服务器进行汇总 [8]。设汇总后的网络数据表示为 $\boldsymbol{X} \in \mathbf{R}^{T \times N \times D}$，其中 T 表示时间步长；N 表示选定客户端数量；D 表示各客户端网络数据的特征维度。

为了更好地处理这些具有复杂时序特性的数据，引入动态多头注意力模块以从不同角度捕捉数据的时序动态特征。首先，需要将输入的网络数据映射到多个不同的子空间，对于第 h 个注意力头，其对应的查询向量、键向量以及值向量分别为 $\boldsymbol{Q}_h$、$\boldsymbol{K}_h$、$\boldsymbol{V}_h$。为有效捕捉时序动态特征，该方法引入动态权重矩阵 $\boldsymbol{W}_{t,h}$，通过时间步自适应的权重调整机制，实现对网络异常模式快速变化的精确建模。第 h 个注意力头的动态注意力得分 $\boldsymbol{A}_h(t)$ 计算公式为：

$$\boldsymbol{A}_h(t) = \text{soft}\max\left(\frac{\boldsymbol{Q}_h(t) \boldsymbol{K}_h(t)^{\mathrm{T}}}{\sqrt{d_k}} + \boldsymbol{W}_{t\,h}(t) \right) \tag{4}$$

式中：d_k 代表查询向量和键向量的维度。以此可以得到第 h 个注意力头的输出结果，具体表达式为：

$$\boldsymbol{O}_h(t) = \boldsymbol{A}_h(t) \boldsymbol{V}_h(t) \tag{5}$$

式中：$\boldsymbol{O}_h(t)$ 代表第 h 个注意力头在时间步为 t 时的输出，即网络运行数据特征的提取结果。通过这种方式，动态多头注意力模块能够根据不同时间步的数据特性动态调整注意力权重，从而更有效地捕捉数据的时序动态特征，为后续的异常识别提供更准确的信息。

1.3 基于联邦聚合与动态阈值的异常判定模型

在通过动态注意力机制捕捉到网络数据的时序动态特征之后，如何利用这些特征实现高精度、可泛化的异常识别是本节所要解决的核心问题。鉴于不同客户端数据分布存在差异，若仅依赖本地模型将导致识别结果缺乏一致性，而传统集中训练模式又无法满足数据隐私保护要求。为此，本节引入联邦学习框架，建立客户端本地更新与服务器全局聚合相

结合的协同训练机制，并设计动态阈值判定规则，实现对网络运行异常的精准识别。

完成时序特征提取后，各客户端基于本地数据对模型进行更新。设客户端接收到的时序动态特征为 O，客户端 k 的本地模式参数为 θ_k，本地损失函数为 $L_k(\theta_k)$，则采用梯度下降法进行参数更新：

$$\theta_k^{\text{new}} = \theta_k - \eta \nabla L_k\left(\theta_k\right) \tag{6}$$

式中：η 代表学习率，$\nabla L_k(\theta_k)$ 代表本地损失函数关于模型参数的梯度。中央服务器收集所有客户端的模型参数后进行加权聚合 [9]，以数据量为权重，构建泛化能力更强的全局模型：

$$\theta_{\text{global}} = \sum_{k=1}^{K} \frac{n_k}{N_{\text{sum}}} \theta_k^{\text{new}} \tag{7}$$

式中：n_k 代表客户端 k 的本地网络运行数据总量，N_{sum} 代表所有客户端产生的网络运行数据总量。

在推断阶段，对于新输入样本的时序动态特征，全局模型的输出其异常概率为 $p(x)$。设定判断阈值 τ 来实现异常识别，如果 $p(x) \geq \tau$ 则代表该样本为异常样本。该阈值可根据验证集上的性能进行动态调整，以平衡误报与漏报比例。

本文所提的联邦聚合与动态阈值判定机制，不仅保障了各客户端数据的本地隐私性，还通过加权融合提升了模型对多样化异常模式的识别能力，为实现分布协作式网络异常检测提供了可行路径。

2 实验环节

2.1 实验准备

本次实验构建的网络运行数据集涵盖了多个网络客户端在连续 30 天内产生的数据。共选取了 50 个网络客户端作为数据来源，每个客户端每天生成的数据量平均为 1 000 条记录，总记录数量为 1 500 000 条。数据以分钟为时间粒度进行采集，每条记录都带有精确到分钟的时间戳，用于分析网络运行的时序特征。每条记录包含流量特征、连接特征以及设备状态特征在内的 10 个特征。其中流量特征主要包括上传流量以及下载流量，用于反映网络传输的数据量情况。

为了模拟真实的网络运行环境，本次实验选择在数据集中采用人为的方式注入一定比例的异常数据。不同类型的异常数据分布情况如表 1 所示。

表 1 不同类型的异常数据分布情况

异常类型	异常记录数量	占比 /%	具体特征参数
突发流量异常	22 500	30	上传流量峰值：超出均值三倍标准差范围，最高达 3 500 MB 下载流量峰值：超出均值三倍标准差范围，最高达 4 000 MB
连接异常	18 750	25	连接持续时间：部分记录超过 5 400 s 连接建立次数：部分记录超过 2 000 次
设备状态异常	15 000	20	CPU 利用率：持续 15~30 min 超过 95% 内存利用率：持续 10 ~ 25 min 超过 92%
传输错误异常	18 750	25	数据包丢失率：部分记录超过 8%

实验选取的对比方法分别为基于孤立森林的传感器网络非线性异常数据点检测方法以及基于 C2-GRU 模型的网络数据流异常识别方法。通过记录三种方法下的异常结果，并与数据集标注的真实异常标签进行比较，以此验证方法在识别精度方面的差异。

2.2 网络运行大数据异常识别结果

本文方法所得到的网络运行大数据异常识别结果如图 1 所示。

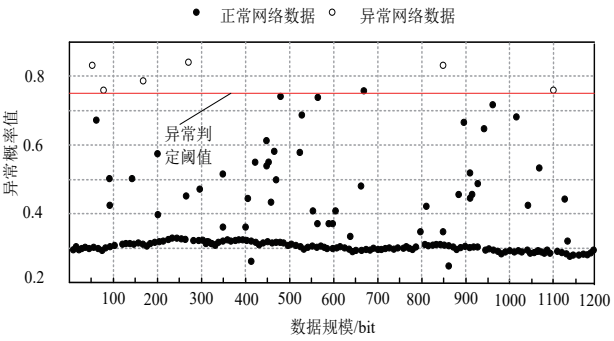

图 1 网络运行大数据异常识别结果

通过图 1 所示的实验结果可以看出，正常网络数据样本的异常概率值大多分布在 0.4 以下，集中分布于低概率区间，表现出较好的聚类特性与稳定性；而异常样本的异常概率值则主要分布在 0.6 以上，与正常样本存在明显的区分度。本文设定异常判定阈值为 0.75，图中可见多数异常样本的异常概率超过该阈值，表明本文所提方法能有效区分正常与异常网络行为。该结果验证了基于联邦学习与动态注意力机制的模型在异常概率估计方面的有效性，为后续异常识别提供了可靠依据。

2.3 识别精度对比

实验以异常标签覆盖率作为对比指标，对方法的实际识别精度进行衡量。具体对比结果如图 2 所示。

通过上述实验结果可以看出，本文方法在各类异常上的覆盖率均显著高于对比方法，整体均值达 85%，且波动范围小，表现出更优的稳定性与泛化能力。相比之下，孤立森林方法在时序动态特征捕捉方面存在不足，C2-GRU 方法虽具备一定的时序建模能力，但难以融合多客户端数据分布，因此在部分异常类型上识别覆盖率较低。实验结果表明，本文所提方法在复杂动态网络环境中具有更优的异常识别精度与鲁棒性。

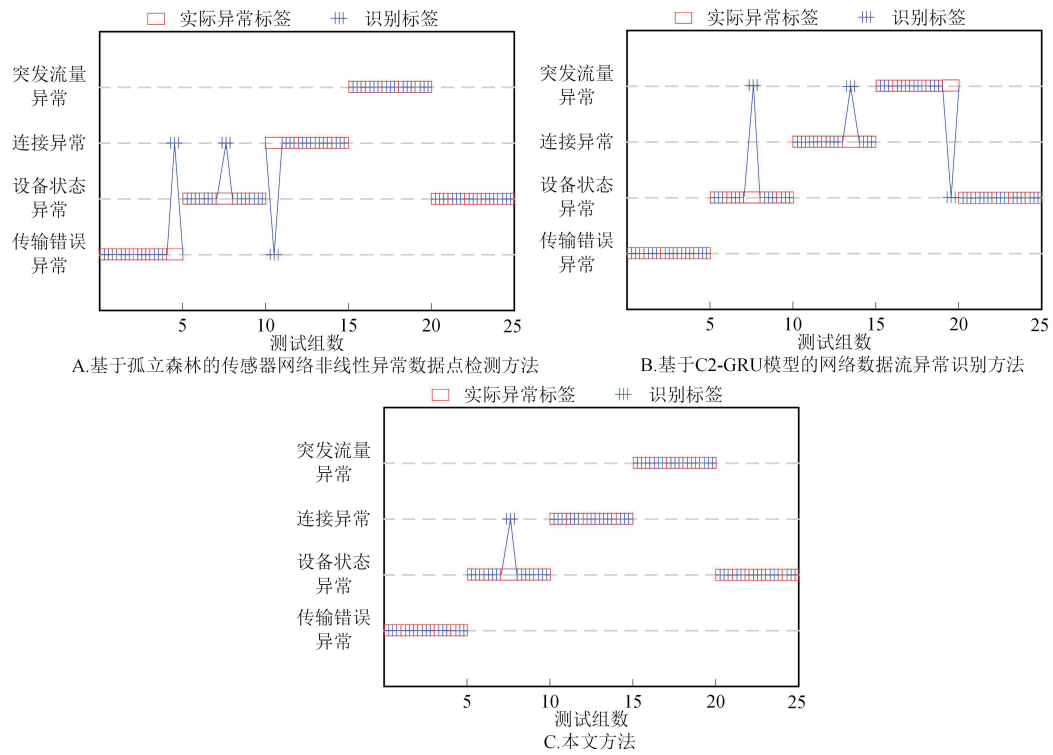

图 2 不同识别方法下的异常标签覆盖率对比结果

3 结论

网络运行大数据异常识别是保障复杂动态网络环境安全稳定运行的关键技术。本文针对现有方法在时序动态特征捕捉、多客户端数据隐私保护及异常识别泛化能力方面的不足，提出了一种融合联邦学习与动态注意力的网络运行大数据异常识别方法。该方法通过层次聚类与差分隐私保护实现客户端的安全筛选与数据代表性增强，利用动态多头注意力机制自适应提取网络数据流中的时序动态特征，并借助联邦聚合机制构建具有强泛化能力的全局异常识别模型。

实验结果表明，本文方法在包含多类异常的真实网络数据集上异常识别覆盖率均值达到 85%，且在突发流量异常、连接异常、设备状态异常和传输错误异常等不同类型异常上均表现出稳定的识别性能。该方法有效克服了传统集中式建模对时序动态响应不足、客户端数据分布差异引发的模型偏差，以及数据隐私泄露风险等问题。

本研究的主要贡献在于构建了一个兼顾隐私保护与动态感知的分布式异常识别框架，为动态网络环境下实现高精度、高安全性的异常检测提供了一条可行的技术路径。未来工作将进一步探索注意力权重的自适应优化策略，并在更大规模的异构网络环境中验证方法的扩展性与实用性。

参考文献：

[1] 毛华彬，熊志文，傅彦铭．基于深度神经网络的无线传感器网络数据异常识别方法 [J]. 传感技术学报，2025, 38 (8): 1491-1498.

[2] 韦霞，熊志文，傅彦铭．基于孤立森林的传感器网络非线性异常数据点检测方法 [J]. 传感技术学报，2025, 38 (8): 1499-1504.

[3] 刘帅，杨锦辉，欧思程，等．基于 C^2-GRU 模型的网络数据流异常识别方法 [J]. 沈阳工业大学学报，2025, 47 (4): 486-492.

[4] 高斌，王以忠，杨国威．基于 BP 神经网络的物联网节点异常数据检测 [J]. 计算机仿真，2025, 42 (7): 267-272.

[5] 黄朝霞，潘云燕．基于弱监督学习的计算机无线网络异常数据流智能辨识方法 [J]. 无线互联科技，2025, 22 (13): 104-107.

[6] 冼德锟．基于无监督学习的通信网络异常数据流智能识别方法 [J]. 长江信息通信，2025, 38 (6): 216-218.

[7] 向思羽，刘才铭．结合混合特征选择和 Transformer 的网络数据流异常检测 [J]. 电子科技大学学报，2025, 54 (3): 442-454.

[8] 肖赣州．基于深度学习的计算机通信工程网络异常数据流辨识 [J]. 长江信息通信，2025, 38 (4): 120-122.

[9] 李晟，李嘉泽楷，任昊．高维空间聚类下多信道光纤网络大数据异常检测 [J]. 激光杂志，2025, 46 (3): 169-174.

【作者简介】

吴晓冰（1984—），女，山西平陆人，本科，实验师，研究方向：计算机应用、职业教育等。

（收稿日期：2025-09-24 修回日期：2026-05-11）