arXiv:2601.12231v1 [cs.LG] 18 Jan 2026

通过偏差调制和分辨率自适应注意力对多通道用户日志进行小波感知异常检测

Kaichuan Kong[1] *DongjieLiu*[1*] *Xiaobo Jin*[2] *ShijieXu*[1] *GuanggangGeng*[1]中国广州暨南大学网络安全学院[2]中国苏州西交利物浦大学先进技术学院

摘要

内部威胁检测是企业安全领域的一项关键挑战，它依赖于能够捕捉丰富且复杂行为模式的用户活动日志。这些日志通常具有多通道、非平稳等特点，且异常事件十分罕见，因此异常检测颇具挑战。为解决这些问题，我们提出了一种新型框架，结合了小波感知调制、多分辨率小波分解和分辨率自适应注意力，以实现稳健的异常检测。我们的方法首先采用偏差感知调制方案，抑制例行行为，同时放大异常偏差。接着，离散小波变换（DWT）将日志信号分解为多分辨率表示，从而捕捉长期趋势和短期异常。最后，通过可学习的注意力机制，动态调整最具判别力的频带权重，以进行检测。在 CERT r4.2 基准测试中，我们的方法在不同时间粒度和场景下的精确率、召回率和 F1分数均持续优于现有基线方法。

索引词— 内部威胁检测、日志异常检测、多分辨率分析、离散小波变换、注意力机制

1. 引言

内部威胁检测是企业环境中的一项关键安全任务，因为经过身份验证的用户可能有意或无意地从事恶意活动。为支持此类检测，记录访问模式、资源交互和系统操作的用户活动日志提供了丰富的行为信号 [1]。这些日志本质上具有多通道、随时间演变的特征，既反映正常的日常活动，也体现潜在威胁 [2–4]。然而，由于以下三个关键因素，从这些日志中检测异常仍极具挑战性：(i) 数据具有多源、高维结构；(ii) 异常行为罕见且隐蔽，常被占主导地位的常规模式所掩盖；(iii) 非平稳的时间动态将长期趋势与短期突发交织在一起。

先前研究采用了统计特征、卷积神经网络以及循环神经网络和 Transformer 等序列模型进行日志建模 [5, 6]。然而，仍存在若干局限：许多模型对冗余操作较为敏感；分布不平衡会导致过拟合；对长距离依赖进行建模的成本通常较高。此外，频域方法（如快速傅里叶变换）能够捕捉全局变化 [7, 8], ，但其粗粒度特性阻碍了对局部异常的检测。

为应对这些挑战，我们提出了一种小波感知多分辨率框架，用于检测多通道用户活动日志中的异常。首先，将原始日志聚合为结构化的行为矩阵，并通过一种偏差感知方案进行调制，以抑制稳定的常规行为，同时突出不规则偏差。其次，通过离散小波变换(DWT) [9], 获得多分辨率表示，从而区分稳定的长期趋势与瞬态突发。最后，分辨率自适应注意力模块对频率分量进行自适应重新加权，增强对异常检测最具信息量的尺度。总体而言，我们的方法兼顾了鲁棒性和计算效率。

本研究的主要贡献如下：

- 我们提出一种偏差感知调制机制，通过自适应地为行为矩阵加权，突出异常偏差。
- 我们结合多分辨率小波分解，从行为矩阵中提取频率感知时序特征。
- 我们设计了一种分辨率自适应注意力机制，用于自适应校准多尺度表示。
- 在 CERT 基准测试上的实验表明，我们的方法在各种场景和粒度级别上均优于最先进的基线方法。

2. 相关工作

早期基于日志的内部威胁检测方法依赖于使用人工特征的经典机器学习模型。研究人员采用了 Log2Vec [10]、信息增益（IG）和基于相关性的特征选择（CFS） [11] 等技术来建模用户的潜在行为。随机森林（RF）、支持向量机（SVM）和极端梯度提升（XGBoost）等模型表现出良好的性能 [12]。在无监督场景中，研究人员使用孤立森林（iForest）、局部异常因子（LOF）和自编码器[13–15]来检测审计日志中的行为偏差。

为解决时序和语义方面的限制，基于注意力机制的长短期记忆

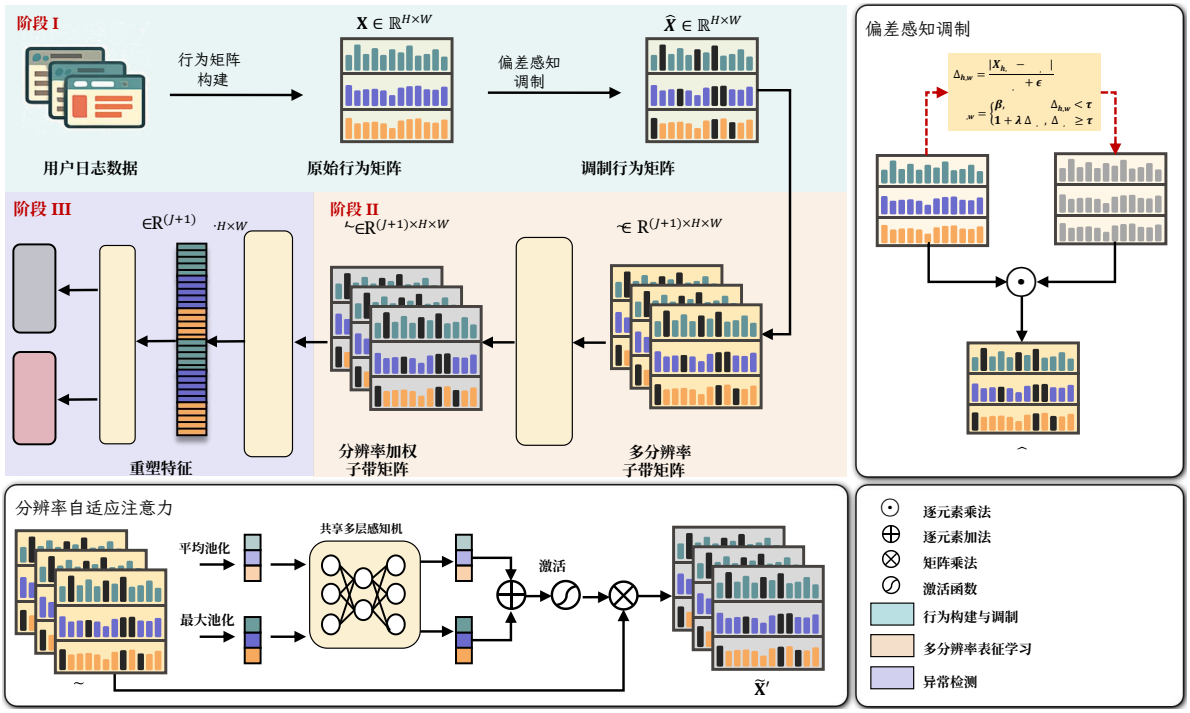

图 1。我们框架的概览。原始用户日志被转换为动作矩阵 $\mathbf{X} \in \mathbb{R}^{H \times W}$，随后由偏差感知权重 $\mathbf{M} \in \mathbb{R}^{H \times W}$ 进行调制，得到调制矩阵 $\hat{\mathbf{X}} = \mathbf{X} \odot \mathbf{M}$（右侧）。使用离散小波变换将调制矩阵分解为 $J+1$ 分辨率子带 $\tilde{\mathbf{X}} \in \mathbb{R}^{(J+1) \times H \times W}$，其中权重由分辨率自适应注意力机制确定。随后将这个由多个分辨率子带组成的矩阵输入检测模型，以生成异常结果。

网络（Att-LSTM） [5], BERT（双向 Transformer 编码器表示）-LSTM [16], 以及基于 Transformer 的架构 [6] 等深度学习方法已得到探索。表示学习还通过生成对抗网络（GANs） [17], 时频分析 [18], 和图神经网络（GNNs） [19,20] 得到了进一步发展，用于用户关系建模。最近，大语言模型（LLMs）通过提示 [21,22] 或微调 [23]实现了对日志序列的零样本/少样本检测。

近年来，信号处理方法已应用于审计日志异常检测，并利用小波变换进行时频分解。例如，离散小波变换（DWT）已用于异常评分（ [24]）；采用多层级小波特征的胶囊网络（CapsNet）取得了较高的准确率（ [25]）；去噪技术则提升了基于聚类的模型的性能（ [8]）。这些方法展现出良好的前景，但通常使用固定基函数和单变量信号，限制了它们适应动态多通道行为的能力。

3. 方法论

图 1 展示了整体框架。我们的方法包含两个主要阶段：(i) 将原始日志转换为结构化且偏差感知的行为矩阵；(ii) 学习能够同时捕捉稳定例行行为和不规则异常的多分辨率表示。以下小节将详细介绍各个阶段。

3.1. 行为构建与调制

用户活动日志本质上具有稀疏、不规则的特点，且以例行行为为主，因此直接建模效果不佳。我们首先构建行为矩阵，将原始日志转换为时间序列，然后应用偏差感知调制，抑制冗余模式，同时增强信息性异常。

行为矩阵构建。事件日志包含异构记录（例如登录、文件访问、电子邮件），每条记录均带有时间戳。设 $\mathcal{E}$ 为已记录事件的集合，其中每个事件 $e \in \mathcal{E}$ 由类型 $e.\text{type} \in \mathcal{H}$ 和时间戳 $e.\text{time}$ 描述。这里 $\mathcal{H}$ 是行为类别集合，其基数为 $H = |\mathcal{H}|$。我们将时间轴划分为 W 个等宽且互不重叠的区间 $\{w_1, \ldots, w_W\}$（例如小时槽）。行为矩阵 $\mathbf{X} \in \mathbb{R}^{H \times W}$ 为：

$$X_{h,w} = \sum_{e \in \mathcal{E}} \mathbb{I}\{e.\text{type} = h,\ e.\text{time} \in w\}, \tag{1}$$

其中 $\mathbb{I}\{\cdot\}$ 是指示函数。由此得到一种结构化表示，其中行对应行为类型，列对应时间区间。

偏差感知调制。日常活动（例如办公时间登录）往往在 $\mathbf{X}$ 中占据主导地位，并掩盖罕见但具有信息性的信息性异常。为了为每个坐标提供总体基准，我们计算$\mu_{h,w}$ 和$\sigma_{h,w}$，分别作为 $X_{h\ w}^{(n)}$,在仅正常训练样本上的均值和标准差；这些值在训练后固定，并在推理时使用。标准化偏差分数定义如下：

$$\Delta_{h,w} = \frac{|X_{h,w} - \mu_{h,w}|}{\sigma_{h,w} + \epsilon}, \tag{2}$$

其中，$\epsilon > 0$是用于保证数值稳定性的一个小常数。然后，我们计算权重矩阵$\mathbf{M}$：

$$M_{h,w} = \begin{cases} \beta, & \Delta_{h,w} < \tau, \\ 1 + \lambda\Delta_{h,w}, & \Delta_{h,w} \geq \tau, \end{cases} \tag{3}$$

其中，$\beta \in (0,1)$用于抑制稳定行为，$\lambda > 0$ 用于放大偏差，$\tau \geq 0$则通过验证数据进行调优。调制后的矩阵为：

$$\hat{\mathbf{X}} = \mathbf{X} \odot \mathbf{M}, \tag{4}$$

其中，$\odot$ 表示哈达玛积。这种重新加权会抑制频繁但缺乏信息的事件，并凸显罕见偏差，从而为后续异常分析提供更清晰的表示。

3.2. 多分辨率表示学习

用户行为跨越多个时间尺度展开，从长期工作规律到突发的异常活动。为同时捕捉稳定模式和瞬态偏差，我们采用小波进行多分辨率分解，并进一步设计一种注意力机制，自适应地突出最具信息量的尺度。

基于小波的多分辨率分解。用户行为很少局限于单一时间尺度。例如，每周电子邮件周期会与突发的异常数据传输并存。为同时捕捉粗粒度规律和细粒度的不规则性，我们采用离散小波变换（DWT）[9], 这一信号处理中的经典工具进行多分辨率分析。

对于每个序列 $\hat{}X_{h,\ :} \in \mathbb{R}^W$，$J$ 层离散小波变换可得：

$$\hat{X}_{h,:} = A_J^{(h)} + \sum_{j=1}^{J} D_j^{(h)}, \quad h = 1, 2, \cdots, H, \tag{5}$$

其中 $A_J^{(h)}$ 表示低频近似（长期趋势），而 $\{D_j^{(h)}\}$ 是细节系数，对应不同尺度的高频波动。

由于这些分量的长度不同，我们将它们插值回原始分辨率 W，并将其堆叠：

$$\tilde{\mathbf{X}} \in \mathbb{R}^{(J+1)\times H\times W}. \tag{6}$$

这种分解丰富了矩阵表示：$A_J^{(h)}$ 反映稳定的用户惯例，而 $D_j^{(h)}$ 突出了可能对应安全违规行为的短暂突发。

分辨率自适应注意力聚合。并非所有分解后的分量都同样有用：低频近似可能会淡化异常，而某些高频带可能大多包含噪声。为了自适应地聚焦于信息丰富的尺度，我们设计了一种跨分辨率的通道注意力机制。经过重新加权后，所有分量被拼接成一个统一嵌入，从而紧凑地表示多分辨率用户行为。

对于每个分辨率 c，我们提取两个描述符：

$$z_c^{\text{avg}} = \frac{1}{HW}\sum_{h,w}\tilde{\mathbf{X}}[c,h,w], \quad z_c^{\max} = \max_{h,w}\tilde{\mathbf{X}}[c,h,w]. \tag{7}$$

将这些描述符拼接后，通过一个两层感知机进行变换：

$$s = \sigma(W_2\,\delta(W_1[z^{\text{avg}}\|z^{\max}])), \quad s \in \mathbb{R}^{J+1}, \tag{8}$$

其中 $\delta(\cdot)$ 是 ReLU，$\sigma(\cdot)$ 是 Sigmoid。权重 s_c 决定分辨率 c的重要性，重新加权后的张量为：

$$\tilde{\mathbf{X}}'[c,h,w] = s_c \cdot \tilde{\mathbf{X}}[c,h,w]. \tag{9}$$

3.3. 异常检测

将重新加权后的各分量拼接并重塑为嵌入：

$$\mathbf{Z} = \text{Reshape}(\tilde{\mathbf{X}}') \in \mathbb{R}^{(J+1)\cdot H\times W}, \tag{10}$$

该嵌入能够联合捕获语义行为类型和频率感知的重要性。

给定异常检测器 $f(\cdot)$，预测结果如下：

$$\hat{y} = f(\mathbf{Z}), \quad \hat{y} \in \{\text{Normal}, \text{Abnormal}\}. \tag{11}$$

4. 实验

为评估我们模型的有效性，我们在广泛使用的基准数据集上与最先进的基线方法进行了对比实验。

4.1. 实验设置

数据集与基线方法。CERT r4.2 [26] 是一个公开基准测试，其中包含三种场景下用户活动日志的异常标签：(i) 导致数据泄露的非工作时间活动；(ii) 因求职引发的数据盗窃；以及 (iii) 心怀不满的管理员实施的内部人员滥用。该基准测试包括信号分解方法（DWT-OCSVM [8], MWCap-sNet [25]）、基于 Transformer 的模型（CATE [6]）以及大语言模型微调方法（ITDLM [23]）。

实现细节。日志被划分为行为矩阵，窗口大小分别为 24 小时（每日）、72 小时（三日）和 168 小时（每周），步长为 24 小时。窗口内的任何异常事件均标记为异常。对于每种场景，我们随机选择三名用户，并报告其平均性能，以减轻用户特定偏差。所有模型均使用 PyTorch 实现，并在四块 NVIDIA L20 GPU 上训练。为验证泛化能力，我们

表1。不同场景下多种算法检测性能的比较

场景粒度		DWT-OCSVM			MWCapsNet			CATE			ITDLM			Ours		
		Pre	Rec	F1	Pre	Rec	F1	Pre	Rec	F1	Pre	Rec	F1	Pre	Rec	F1
场景1	24h	0.88	1.00	0.93	0.93	1.00	0.96	0.99	0.99	0.99	0.89	0.92	0.90	0.98	0.99	0.98
	72h	0.85	1.00	0.92	0.99	0.99	0.99	0.98	0.99	0.99	0.91	0.93	0.93	0.97	0.98	0.97
	168h	0.80	1.00	0.89	0.99	0.99	0.99	0.50	0.98	0.67	0.94	0.98	0.96	0.96	0.98	0.97
场景2	24h	0.90	0.59	0.72	0.79	0.80	0.80	0.77	0.87	0.82	0.83	0.87	0.83	0.96	0.95	0.96
	72h	0.85	0.77	0.81	0.82	0.95	0.88	0.76	0.93	0.84	0.86	0.89	0.90	0.96	0.97	0.96
	168h	0.74	0.93	0.83	0.80	0.97	0.88	0.61	0.96	0.75	0.92	0.93	0.93	0.97	0.98	0.98
情景3	24h	0.89	1.00	0.94	0.93	0.99	0.96	0.98	0.99	0.98	0.87	0.89	0.89	0.98	0.98	0.98
	72h	0.91	1.00	0.95	0.99	0.99	0.99	0.99	0.99	0.99	0.89	0.91	0.90	0.97	0.98	0.97
	168h	0.86	1.00	0.92	0.97	0.99	0.98	0.99	0.99	0.99	0.94	0.96	0.95	0.96	0.97	0.97
Avg.	–	0.85	0.93	0.89	0.93	0.95	0.95	0.84	0.95	0.90	0.90	0.93	0.92	**0.97**	**0.98**	**0.97**
Rank	–	–	–	5	–	–	2	–	–	4	–	–	3	–	–	**1**

表 2。检测模型对比

模型	Pre	Rec	F1
AE	0.49	0.98	0.66
孤立森林	0.83	0.66	0.73
XGBoost	0.98	**1.00**	**0.99**
Transformer	0.87	0.93	0.90
TCN	**1.00**	0.98	**0.99**

实例化五种检测器：三种经典方法（自编码器 [27], 孤立森林和 XGBoost [12]）以及两种深度模型（Transformer [28] 和时序卷积网络 [29]）。我们报告**精确率**、**召回率**和**F1分数**。

4.2. 主要结果与分析

表1报告了三个实验场景下的检测结果，每个场景还在多个时间粒度（24小时、72小时和168小时窗口）上进行了评估。在所有设置下，我们提出的方法均表现出最均衡且稳健的性能，平均精确率、召回率和F1分数分别为0.97、0.98和0.97。与传统的DWT-OCSVM以及MWCapsNet、CATE和ITDLM等近期基线方法相比，我们的方法在不同场景和时间粒度下始终保持较高的精确率和召回率。具体而言，尽管某些基线方法会根据窗口大小提高精确率或召回率，但我们的方法在所有三项指标上均取得稳定提升，并获得了最佳综合排名。这些结果证实了我们的框架在捕捉多尺度用户行为动态以及确保各种使用场景下可靠的异常检测方面的有效性。

4.3. 检测模型选择

为检验不同检测器与我们表征的适用性，我们使用**场景2**中的一名代表性用户进行了模型选择研究。如表 2 所示，自编码器和孤立森林等传统无监督方法表现尚可，而XGBoost和深度序列模型（Transformer、时序卷积网络）则始终取得较高的F1分数。**时序卷积网络和XGBoost尤其表现出色，F1分数达到0.99**，凸显了它们在捕捉时间依赖关系和判别模式方面的有效性。

表 3. 组件消融研究

模型	Pre	Rec	F1
我们的方法（完整模型）	**0.98**	**1.00**	**0.99**
不使用偏差调制	0.96	0.90	0.92
不使用离散小波变换	0.93	0.94	0.94
无注意力机制	0.94	0.92	0.93

该实验验证了我们的框架与各种检测范式的兼容性，后续评估中我们将采用表现最佳的模型。

4.4. 消融研究

表 3 展示了消融实验结果。完整模型取得了最佳性能（F1=0.99）。移除偏置调制后，召回率降低（1.00→0.90），凸显了其在突出异常行为方面的作用。移除离散小波变换后，F1分数降至 0.94，说明多分辨率分解对于捕捉时间动态至关重要。最后，移除注意力机制后，F1分数进一步降至 0.93，证实了自适应调整分辨率权重的有效性。这些结果表明，每个组件都发挥了作用，而它们的组合带来了最佳性能。

5. 结论

在本研究中，我们提出了一种小波感知多分辨率框架，用于检测多通道用户活动日志中的日志异常。我们将原始日志转换为结构化的行为矩阵，通过离散小波变换（DWT）对其进行分解，并应用分辨率自适应注意力机制，从而有效地区分常规操作与异常偏差。实验结果表明，在精确率、召回率和F1分数方面，我们的方法始终优于最先进的机器学习和深度学习基线方法。未来，我们计划将该框架扩展至规模更大、种类更多样的日志数据集（例如云基础设施日志），或通过动态阈值调整增强流式日志分析框架。

6. 参考文献

[1] F. R. Alzaabi 和 A. Mehmood，“使用机器学习方法进行恶意内部威胁检测的最新进展、挑战与机遇综述”，*IEEE Access*，第 12 卷，第 30907–30927 页，2024 年。[2] Chiming Duan、Tong Jia、Yong Yang、Guiyang Liu、Jinbu Liu、Huxing Zhang、Qi Zhou、Ying Li 和 Gang Huang，“Eager-log：用于基于日志的异常检测的主动学习增强型检索增强生成”，载于 *ICASSP 2025-2025 IEEE* 国际声学、语音与信号处理会议 *(ICASSP)*。IEEE，2025 年，第 1–5 页。[3] Lingzhe Zhang、Tong Jia、Mengxi Jia、Yifan Wu、Hongyi Liu 和 Ying Li，“ScalaLog：使用大语言模型进行可扩展的基于日志的故障诊断”，载于*ICASSP 2025-2025 IEEE* 国际声学、语音与信号处理会议 *(ICASSP)*。IEEE，2025 年，第 1–5 页。[4] Mingjie Zhou、Weidong Yang、Lipeng Ma、Sihang Jiang、Bo Xu 和 Yanghua Xiao，“LogSI：一种用于系统增量日志分析的基准测试”，载于*ICASSP 2025-2025 IEEE* 国际声学、语音与信号处理会议 *(ICASSP)*。IEEE，2025 年，第 1–5 页。[5] W. He、X. Wu、J. Wu、X. Xie、L. Qiu 和 L. Sun，“基于用户历史行为和注意力机制的内部威胁检测”，载于*IEEE* 第六届数据科学与网络空间国际会议 *(DSC)* 论文集，2021 年，第 564–569 页。[6] H. Xiao、Y. Zhu、B. Zhang、Z. Lu、D. Du 和 Y. Liu，“揭开阴影：一种基于统计分析与序列分析的全面内部威胁检测框架”，计算机与安全，第 138 卷，第 103665 页，2024 年。[7] T. Yang、Z. Yang、F. Li 和 H. Wang，“一种基于多变量信号分解和变量选择的短期风电功率预测方法”，*Applied Energy*，第 360 卷，第 122759 页，2024 年。[8] D.-W. Kim、G.-Y. Shin 和 M.-M. Han，“基于离散小波变换的内部威胁分类异常检测”，*Computer Systems Science and Engineering*，第 46 卷，第 1 期，第 153–164 页，2023 年。[9] Tomohiko Nakamura 和 Hiroshi Saruwatari，“结合离散小波变换的 Wave-U-Net 时域音频源分离”，载于*ICASSP 2020-2020 IEEE* 国际声学、语音与信号处理会议 *(ICASSP)*。IEEE，2020 年，第 386–390 页。[10] F. Liu、Y. Wen、D. Zhang、X. Jiang、X. Xing 和 D. Meng，“Log2Vec：一种基于异构图嵌入的企业网络威胁检测方法”，载于*2019* 年 *ACM SIGSAC* 计算机与通信安全会议论文集 *(CCS)*，2019 年，第 1777–1794 页。[11] B. Bin Sarhan 和 N. Altwaijry，“使用机器学习方法进行内部威胁检测”，*Applied Sciences*，第 13 卷，第 1 期，第 259 页，2022 年。[12] D. C. Le、N. Zincir-Heywood 和 M. I. Heywood，“使用机器学习分析内部威胁检测的数据粒度级别”，*IEEE*网络与服务管理汇刊，第 17 卷，第 1 期，第 30–44 页，2020 年。[13] F. W. Bartoszewski、M. Just、M. A. Lones 和 O. Mandrychenko，“内部威胁异常检测：机器学习模型与集成方法的客观比较”，载于*IFIP ICT*系统安全与隐私保护国际会议论文集。2021 年，第 367–381 页，Springer。[14] D. C. Le 和 N. Zincir-Heywood，“使用无监督集成方法检测内部威胁异常”，*IEEE*网络与服务管理汇刊，第 18 卷，第 2 期，第 – 页。

[15] R. Yousef、M. Jazzar、A. Eleyan 和 T. Bejaoui，《用于检测内部人员网络犯罪威胁的机器学习框架与开发》，载于*2023*年智能应用、通信与网络国际会议论文集（*SmartNets*）。2023年，第1–6页，IEEE。[16] W. Huang、H. Zhu、C. Li、Q. Lv、Y. Wang 和 H. Yang，《It-dbert：用于内部威胁检测的时序语义表示》，载于*IEEE*计算机与通信研讨会论文集（*ISCC*），希腊雅典，2021年，第1–7页。[17] R. G. Gayathri、A. Sajjanhar 和 Y. Xiang，《采用SPCAGAN增强的内部威胁分析混合深度学习模型》，《专家系统与应用》，第249卷，第123533页，2024年。[18] A. Budˇzys、O. Kurasova 和 V. Medvedev，《面向关键基础设施内部威胁检测的深度学习身份验证》，《人工智能评论》，第57卷，第10期，第272页，2024年。[19] K. C. Roy 和 G. Chen，《Graphch：用于评估内部威胁检测中网络与人类因素的深度框架》，《*IEEE*可靠与安全计算汇刊》，2024年，提前在线发表。[20] X. Cai、Y. Wang、S. Xu、H. Li、Y. Zhang、Z. Liu 和 X. Yuan，《Lan：通过学习自适应邻居实现实时内部威胁检测》，《*IEEE*信息取证与安全汇刊》，2024年。[21] J. Qi、S. Huang、Z. Luan、S. Yang、C. Fung、H. Yang、D. Qian、J. Shang、Z. Xiao 和 Z. Wu，《Loggpt：探索使用ChatGPT进行基于日志的异常检测》，载于*IEEE*高性能计算与通信、数据科学与系统、智慧城市以及传感器、云与大数据系统及应用可靠性国际会议论文集（*HPCC/DSS/SmartCity/DependSys*），2023年，第273–280页。[22] Y. Liu、S. Tao、W. Meng、J. Wang、W. Ma、Y. Chen、Y. Zhao、H. Yang 和 Y. Jiang，《利用大语言模型与提示策略进行可解释的在线日志分析》，载于第*32*届*IEEE/ACM*程序理解国际会议论文集（*ICPC*），2024年，第35–46页。[23] S. Song、Y. Zhang 和 N. Gao，《应对内部威胁：基于大语言模型微调的行为日志精确异常检测》，载于第*31*届计算语言学国际会议论文集（*COLING*），2025年，第8589–8601页。[24] W. Feng、W. Yan、S. Wu 和 N. Liu，《利用小波变换与无监督机器学习检测云文件共享中的内部威胁》，载于*IEEE*情报与安全信息学国际会议论文集（*ISI*），中国北京，2017年，第155–157页。[25] K. D. Randive 和 M. Ramasundaram，《MWCapsNet：一种利用图像表示进行内部威胁检测的新型多层小波胶囊网络》，《神经计算》，第553卷，第126588页，2023年。[26] B. Lindauer，《内部威胁测试数据集》， 2020， [在线]。网址：https://doi.org/10.1184/R1/12841247.v1。[27] Richa Singh、Nidhi Srivastava 和 Ashwani Kumar，《基于自编码器的多数据集网络异常检测：综述》，《工程领域近期专利》，第18卷，第9期，第63–77页，2024年。[28] Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N Gomez、Lukasz Kaiser 和 Illia Polosukhin，《Attention Is All You Need》，《神经信息处理系统进展》，第30卷，2017年。[29] Jin Fan、Ke Zhang、Yipan Huang、Yifei Zhu 和 Baiping Chen，《用于多变量时间序列预测的并行时空注意力时序卷积网络》，《神经计算与应用》，第35卷，第18期，第13109–13118页，2023年。