深度学习下计算机通信网络动态跟踪异常数据流辨识

卜莹雪[1]
BU Yingxue

摘 要 计算机通信网络入侵攻击往往混合多种协议，攻击者会依据网络防御策略的调整而改变攻击方式，使得异常数据流的特征随之发生变化。依赖固定阈值的辨识方法，其特征提取能力有限，无法充分挖掘数据在不同尺度和频率上的特征信息，导致辨识准确度不高，且在辨识过程中产生脏数据。为此，文章提出了一种深度学习下计算机通信网络动态跟踪异常数据流辨识方法。通过数据预处理环节标准化处理原始数据，利用降噪自编码器进行降维处理。在特征提取阶段，构建了包含卷积层、池化层和全连接层的深度学习模型，结合小波包分析和 Soft 分类器，实现高区分度的特征提取。设计多状态特征辨识矩阵，通过动态调整矩阵结构来适应不同网络环境，基于该矩阵构建异常数据流辨识模型，引入数据包最大丢失率作为关键阈值实现首次辨识；运用动态跟踪方法将辨识结果进行验证，实现对计算机通信网络动态跟踪异常数据流的精准辨识。实验结果表明，所提方法的数据流辨识准确度高达 90%，在辨识过程中产生的脏数据区间为 0 个，能够达到良好的异常数据辨识效果。

关键词 深度学习；计算机；网络入侵；数据流；辨识

doi：10.3969/j.issn.1672-9528.2025.09.040

0 引言

计算机通信网络已成为社会运转和经济发展的关键基础设施。随着网络规模的不断扩大、网络应用的日益复杂以及网络攻击手段的持续升级，计算机通信网络面临着前所未有的安全挑战。其中，异常数据流的产生和传播对网络的安全性和稳定性构成了严重威胁，可能导致数据泄露、服务中断、系统瘫痪等严重后果。因此，如何有效地动态跟踪计算机通信网络中的异常数据流并进行准确辨识，成为当前网络安全领域亟待解决的重要问题。

文献 [1] 通过采样和差分隐私技术，为数据流中的最近数据集快速创建可以安全发布的差分隐私直方图。根据用户的隐私保护强度，通过隐私分配的方式自适应生成最终可发布直方图。文献 [2] 通过主动选择最具信息量的样本进行标注和学习，提高模型对非平衡数据流的分类性能。将多个分类器集成在一起，通过投票方式提高算法对非平衡数据流的适应性。以上两种方法没有充分考虑异常数据流特征随网络环境变化的动态性。当攻击模式发生变化时，无法提取高区分度的特征，而这些特征对于准确辨识异常数据流非常重要，从而影响整体的辨识效果。基于此，本研究以计算机通信网络动态跟踪异常数据流辨识作为研究对象，结合深度学习技术，开展新的研究分析。该方法的创新点如下：

（1）在数据预处理环节，对原始数据进行标准化处理，

1. 江西飞行学院 江西南昌 330088

[8] 刘诗琼，石强，汤继周，等 . 基于球缝模型优化反演的裂缝储层评价新方法 [J]. 地球物理学报 ,2022,65(4):1451-1460.
[9] 李艳霞，柴毅，胡友强，等 . 不平衡数据分类方法综述 [J]. 控制与决策 , 2019,34(4):673-688.
[10]VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. NewYork: ACM, 2017: 6000-6010.
[11] 张曦，杨颖，陈超君，等 . 增强双流 Transformer 的柴油发动机剩余寿命预测模型 [J]. 汽车工程 ,2025,47(2):292-300.

【作者简介】
侯耀晗（2001—），男，河南新乡人，硕士研究生，研究方向：机器学习、图像处理等。

（收稿日期：2025-05-08 修回日期：2025-09-15）

并采用降噪自编码器进行降维，提升数据质量和后续处理效率。

（2）构建包含卷积层、池化层和全连接层的深度学习模型，结合小波包分析和 Soft 分类器，实现高区分度的特征提取，有效挖掘数据中的关键特征。

（3）设计多状态特征辨识矩阵，能够动态调整矩阵结构以适应不同的网络环境，增强模型的适应性和通用性。

（4）基于多状态特征辨识矩阵构建异常数据流辨识模型时，引入数据包最大丢失率作为关键阈值，实现异常数据流的首次精准辨识。

（5）运用动态跟踪方法对首次辨识结果进行二次验证，进一步提高异常数据流辨识的准确性和可靠性，实现对计算机通信网络动态跟踪异常数据流的精准辨识。

1 网络动态跟踪异常数据流辨识

1.1 深度学习下的异常数据流特征提取

1.1.1 异常数据预处理

在网络动态跟踪异常数据流辨识中，原始数据往往包含噪声、冗余信息以及非数值化的标识，这些信息会干扰后续的特征提取和模型训练过程 [3]。因此，在进行特征提取前，需要对异常数据进行预处理，去除干扰因素，提高数据的质量和可用性 [4]。

对于符号值属性数据，进行有效转换。采用空间映射策略，将这类非数值化的标识映射为大于 1 的整数，从而确保它们能够作为有效的数据点参与到后续的分析中。这种转换不仅保留了数据的原始信息，还为其在数值计算中的处理提供分析点 [5]。对于特殊数据流类型，采用 *Z*-score 标准化处理，其公式为：

$$z=\frac{y-y_{\min}}{y_{\max}-y_{\min}} \tag{1}$$

式中：y 为数组的原始值；$y_{\min}$ 和 $y_{\max}$ 分别为数组的最小值和最大值。

针对计算机网络数据流数据包中少数类数据不平衡问题，采用距离度量方法来增强少数类样本的表示能力。随机选取少数类样本作为参照点，利用欧式距离公式来计算各样本之间的距离，以此识别出相似或不同的数据点 [6]。通过合成技术生成新的样本，增加少数类样本的数量。其计算公式为：

$$Y=R\times[z(j)-y+y'] \tag{2}$$

式中：R 为邻近样本数据；$z(j)$ 为样本数据编号；y' 为新生成的样本值。

基于距离度量的结果，获取出少数类样本的邻近样本，利用这些邻近样本的信息，通过集合操作来生成新的样本，以此增加少数类样本的数量。

1.1.2 深度学习特征提取

在网络动态跟踪异常数据流辨识过程中，数据预处理后的原始数据通常包含大量冗余信息、噪声和复杂的非线性关系，直接用于异常辨识可能效果不佳。因此，需要构建一个深度学习模型，以自动学习并捕捉预处理后数据中的深层次特征，有效剔除无关信息，保留对异常检测至关重要的特征。建立深度学习模型的公式为：

$$x=Y\times w\times b(l) \tag{3}$$

式中：w 为数据权重；$b(l)$ 为一定长度下的神经元偏置。

在卷积神经网络模型中，卷积层的输出结果会传递给池化层进行进一步的处理。池化层会执行最大池化操作，基于池化窗口内的数据计算出特征数据的统计值，更接近于平均池化，实现了对特征数据的聚合，以达到数据组的最优化处理效果。

为了进一步提升模型的泛化能力，避免数据过度拟合，可以考虑结合小波包分析来提取更精细的特征，并利用 Soft 分类器来输出用户异常数据流特征的分类结果。分类器的输出表示每个类别的可能性，其公式为：

$$D=\frac{1}{2\varepsilon}\times x\times a \tag{4}$$

式中：ε 为小波包参数；a 为数据拟合次数。

模型能够作出最终的分类决策，从而成功完成异常数据流的特征提取。

1.2 多状态特征异常辨识模型构建

1.2.1 多状态特征辨识矩阵构建

在完成深度学习特征提取后，构建多状态特征辨识矩阵，在不同通信网络运行环境下，根据特征的变化及异常数据流的类型，对数据流的多个状态特征进行量化和编码。通过动态化调整辨识矩阵结构，可使辨识效果更加显著。定义其中关键的状态特征，作为数据流映射数据和信息的采集。表 1 为数据流映射及其范围值，作为数据采集的基础。

表 1 数据流映射数据

数据流映射	范围
数据包频率 /Hz	125~126
传输速度 /(Mbit·s^{-1})	100~150
转换可控差	1.25~1.30
权重	11.20

在数据采集阶段，将根据实时收集的数据流信息，填充至辨识矩阵的对应位置。通过这种方式，多状态特征辨识矩阵不仅能够捕捉数据流的单一特征变化，还能通过特征间的

组合关系，深入挖掘潜在的异常模式。因此，设计多状态特征辨识矩阵的公式为：

$$M=[f(x_1),f(x_2),\cdots,f(x_n)] \tag{5}$$

式中：$f(x_1)$ 为第 1 个数据流的特征值；$f(x_n)$ 为第 n 个数据流的特征值。

在矩阵中导入深度学习算法，结合异常数据流特征的变化，自动调整辨识矩阵的结构，确保矩阵在不同通信网络环境下均能保持优异的辨识性能。

1.2.2 深度学习下的异常数据流辨识模型构建

在数据流辨识过程中，对海量无标签网络数据进行深入剖析和模式识别，建立通信网络异常数据流辨识模型，获取通信网络数据流样本。特征提取环节将异常数据流特征分为基础特征和高级特征。针对数据包丢失问题，计算数据包所能容忍的最大丢失率并设定为关键阈值。将辨识矩阵作为模型输入数据，设计深度学习下的辨识模型，通过训练和测试模型以获得辨识结果，及时发现并处理潜在异常数据流，确保网络安全，并优化模型性能。

基础特征聚焦于数据包频率这一直观指标，而高级特征则深入挖掘数据包之间的内在联系与相互影响。计算出数据包所能容忍的最大丢失率，并将这一数值设定为数据包辨识处理过程中的关键阈值。具体而言，数据包允许出现的最大丢失率可以通过公式进行计算：

$$F=\alpha\times\frac{s}{d\times(1+\rho^2)\times M} \tag{6}$$

式中：α 为数据包总量；s 为辨识范围；ρ 为监督定向差值；d 为数据包大小。

基于当前测量数据，计算网络传输中数据包的最大允许丢失率，并作为辨识处理的阈值标准。在此基础上，构建深度学习下的异常数据流辨识模型。该模型的输入数据为多状态特征辨识矩阵，其结构可根据网络环境动态调整，以优化特征表达。模型训练过程中，通过调整参数提升辨识精度，最终输出异常检测结果。其辨识模型表示为：

$$B=\eta\times\partial(G-\pi) \tag{7}$$

式中：η 为辨识梯度差值；∂ 为多状态辨识均值；G 为辨识数据量。

通过训练和测试，该模型能够有效识别潜在异常数据流，从而增强网络安全防护能力。结合辨识矩阵的反馈，持续优化模型性能，确保检测的准确性与适应性。

1.3 动态跟踪数据流辨识

为提高准确性，采用动态跟踪方法动态跟踪网络数据流状态，对捕获的数据包进行全面分析，并将结果与预设模型输出进行对比验证，实现对通信网络数据流的辨识，显著提升了网络稳定性，适应多变的通信网络环境。动态跟踪数据流辨识结果公式为：

$$M=B\times M(x) \tag{8}$$

式中：$M(x)$ 为对数据包的二次辨识结果。

在验证所有信息一致后，输出最终的辨识结果，从而获取检测到的异常数据流情况，显著提升网络的稳定性，以更好地适应多变的通信网络环境。

2 实验测试与分析

2.1 搭建实验环境

为了验证所提方法在计算机通信网络动态跟踪异常数据流辨识中的效果，本次实验选取了电网公共网络入侵数据集 R1 作为实验数据。该数据集具备模拟实际计算机通信网络中异常状况仿真参数的能力，具体的仿真参数配置详情如表 2 所示。在实验准备阶段，设定计算机数据点为准确监测到异常数据的过程。将数据流划分为 100 个数据节点，定义节点数据的监测结果。为了防止通信网络因测试而过度拟合，制定虚拟的辅助网络异常数据流指令，将测试目标导入辨识程序中。预设数据流辨识的周期，每个周期都需要对数据信息进行采集。针对当前通信网络的覆盖控制区域，设定检测节点。根据异常数据流的捕捉范围，设定检测点允许出现的最大检测差为 1.5。结合异常数据流入侵情况，并依据上述设计模型进行实验测试。

表 2 仿真参数设置

参数	数值
传感器节点数	350 个
控制包	110 bit
初始能量	0.5 J
数据包大小	1 500 bit
射频能耗	55 nJ/bit
阈值	85 m

2.2 结果分析

在本次测试环境下，设置 3 个测试小组。针对异常入侵检测的数据流类型，设计并实施了自动检测。其中运用本文方法的小组为实验组，另外两个小组为对照组（对照 1 组为基于特征补全的无线传感器网络异常数据流检测方法，对照 2 组为非平衡数据流在线主动学习方法）。通过建立的测试平台对所提出的方法的辨识效果进行应用效果验证。在设计的模型中，针对初始构建的异常数据流入侵检测标签，构建了相应的时间序列。最后输出辨识结果，计算对应的辨识准确度，得到结果如图 1 所示。

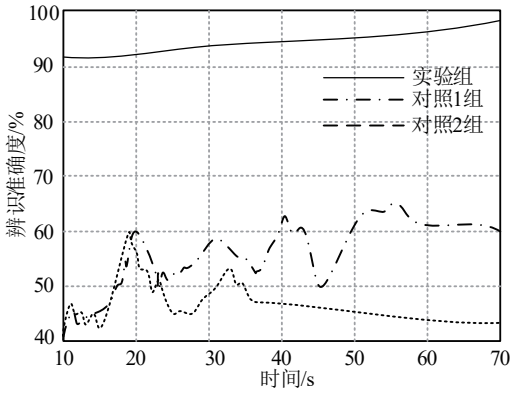

图 1 三种方法异常数据辨识准确度对比结果

由图 1 可知，本文方法在辨识准确度方面表现出显著优势，实验组达到了 90% 以上的高准确率，明显优于两个对照组的测试结果。这一优异的辨识性能不仅验证了所提方法的有效性，更显著提升了电力通信网络的安全防护水平。研究表明，深度学习的引入为电力通信网络安全防护提供了创新性的解决方案，同时通过优化数据处理结构，使系统能够更加灵活地适应复杂多变的网络环境。

在测试过程中，将数据合理划分为 100 个数据节点，并明确了节点数据的监测标准。在确保实验环境稳定的情况下，设定 3 个测试小组对辨识过程中产生的脏数据情况进行分析。当异常数据的值处于 3~4 的区间时，判定该监测方法产生了脏数据。通过对数据流在辨识过程中的脏数据情况进行分析，得到了具体结果，如图 2 所示。

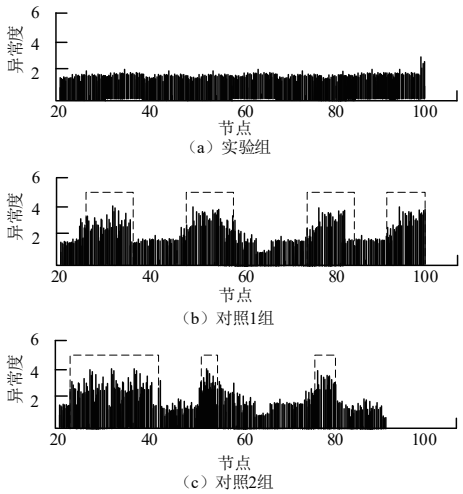

图 2 三种方法的脏数据对比结果

由图 2 可知，在辨识过程中，两个对照组产生了较多脏数据，这些脏数据主要源于实时监测过程中的误判或干扰。实验组表现优异，实现了零脏数据。动态跟踪的二次验证机制，通过将首次辨识结果与实时捕获的数据包特征进行二次比对，有效剔除了首次辨识中可能产生的误判（即脏数据），确保了异常数据监测结果的准确性。

3 结语

本文提出深度学习下计算机通信网络动态跟踪异常数据流辨识方法，该方法在数据预处理阶段，对原始数据开展标准化操作，借助降噪自编码器完成降维工作。在特征提取环节，构建由卷积层、池化层以及全连接层构成的深度学习模型，进行高区分度的特征提取目标。设计了一种多状态特征辨识矩阵，此矩阵能够依据不同网络环境动态调整自身结构。基于该矩阵，构建起异常数据流辨识模型，同时引入数据包最大丢失率作为关键阈值，以此实现初次辨识；运用动态追踪手段对辨识结果进行二次辨识，达到计算机通信网络入侵异常数据流辨识的目的。实验结果表明，该方法检测准确率达 90% 以上，且无脏数据产生，相比传统固定阈值方法能更有效应对混合协议和多变攻击。

当前主要局限在于模型可解释性有待提升，未来研究工作可以将深度学习与区块链、边缘计算等技术的融合，网络异常高效辨识技术将向更智能、高效的方向发展，为构建安全网络环境提供有力支撑。

参考文献：

[1] 王修君，莫磊，郑啸，等．基于采样的数据流差分隐私快速发布算法 [J]. 计算机研究与发展 ,2024,61(10):2433-2447.

[2] 李艳红，任霖，王素格，等．非平衡数据流在线主动学习方法 [J]. 自动化学报 ,2024,50(7):1389-1401.

[3] 郑俊华，魏晋宏．基于特征补全的无线传感器网络异常数据流检测 [J]. 传感技术学报 ,2024,37(6):1061-1066.

[4] 孙丽君，李方方，胡祥培．基于离散数据流分割算法的预测方法 [J]. 管理科学学报，2024, 27(9):48-61.

[5] 代少升，边志奇，袁中明．结合软约束的演化数据流模糊聚类算法 [J]. 重庆邮电大学学报 (自然科学版), 2024, 36(2): 287-298.

[6] 郑万波，李磊．移动边缘计算下矿山安全监控视频数据流自适应传输 [J]. 现代电子技术，2024,47(2):74-78.

【作者简介】

卜莹雪（1981—），女，江西南昌人，硕士，副教授，研究方向：计算机网络、电子商务。

（收稿日期：2025-04-16 修回日期：2025-09-12）