基于 ALSTM 的网络流量异常检测模型

曹旭泽[1] 卢胜男[1]
CAO Xuze LU Shengnan

摘 要 随着全球数据流量呈现爆发式增长态势，伴随其产生了诸多由流量数据泄漏所引发的网络安全事件。传统的网络流量检测方法通常都是依赖于已知的规则和签名，这些方法不能有效捕捉到时序特征数据中的周期特点，且无法提取数据点之间的关联特征，因此在面对异常数据流量的检测能力不足。文章针对传统方法在提取具有时序特征的数据点关联性上的局限性，提出了 ALSTM 网络流量异常检测模型。该模型融合了注意力机制（Attention）与长短期记忆神经网络（LSTM），其中注意力机制在模型处理数据时可以关注更有价值的信息，减少后续计算复杂度，长短期记忆神经网络通过引入记忆单元（Cell State）来解决长序列数据处理中的梯度消失问题，为提取数据点之间的关联性带来便利。最后，该方法在 CICIDS2017 公开数据集上对所提出的模型进行验证，结果显示其准确率达到 98%。与传统方法相比，所提模型的检测精准度有所提升。

关键词 流量异常检测；注意力机制；长短期记忆网络

doi：10.3969/j.issn.1672-9528.2025.10.029

0 引言

随着网络使用率的日益增长，网络流量也呈指数性开始增长，由网络安全问题所引起的个人信息泄露和企业系统瘫痪等网络攻击问题频频发生。网络攻击 [1] 的形式包括但不限于恶意软件攻击、网络钓鱼、拒绝服务攻击（DoS/DDoS）和零日攻击等，并且网络攻击的形式也在持续升级，攻击模式愈发多样。因此，检测异常网络流量，保障网络安全一直以来都是研究人员重要的研究方向。同时，关注数据点之间的关联性也是提升检测准确率的关键因素。

传统的网络流量异常检测方法受限于网络规模扩大和攻击手段复杂化，基于机器学习（SVM[2]、随机森林 [3]）和深度学习（LSTM[4]、ARMA[5]、SARIMA[6]）的方法成为主流。深度学习的方法能自动提取高维特征，例如 LSTM 能有效捕捉流量模式变化，同时结合多种模型进一步提升对具有时序特征的网络流量异常值检测的效率和精度。

国外学者 Shafi 等人 [7] 提出了一种全面的行为分析解决方案，以解决当前入侵检测方法在识别零日攻击和新型恶意行为方面的局限性。Ding 等人 [8] 提出了一种针对网络流量分类的通用和可迁移对抗性攻击方法。该方法利用深度学习分类模型的几何特性设计目标损失函数，并迭代优化通用扰动的生成。这使得对网络流量分类进行通用对抗性攻击成为可能，将对抗性网络流量的平均生成时间缩短至约 0.1 ms。Olabanjo 等人 [9] 提出了一种新颖的图卷积网络（GCN）模型，用于智能网络流量分析和分类。Hu 等人 [10] 设计了一种外部注意力和卷积混合学习模型（ECM），利用低复杂度的外部注意力和卷积提取字节级和数据包级特征，实现准确和实时的流量分类。Cai 等人 [11] 提出了一种称为对比原型网络（CPN）的新方法，用于增量式加密流量分类。主要解决的关键挑战是流量歧义性和流量不平衡，这些限制了现有的增量学习方法无法达到令人满意的结果。Marziyeh 等人 [12] 提出了一个名为 ITC-Net-blend-60 的综合数据集，该数据集旨在克服先前仅在单一网络环境中捕获的数据集的局限性，从而提高网络流量分类模型的鲁棒性和兼容性评估。Jung 等人 [13] 提出了一种分类架构，可以使用比现有基于 GNN 的 ETA 算法更少的特征来分类加密的正常 / 恶意网络流量。Altaf 等人 [14] 介绍了一个新颖的框架，利用专门为物联网（IoT）网络环境中僵尸网络检测设计的顺序门控图卷积神经网络（GGCN）。

本文主要是利用深度学习算法构建出更准确和高效的网络流量异常检测模型。通过应用先进的数据分析和预测建模，本文目标是增强对各种网络入侵的检测和分类。工作聚焦于利用 CICIDS2017 数据集进行异常检测，该数据集是业界公认的评估基准。具体步骤涵盖数据集的预处理、特征提取以及深度学习模型的应用。并且，创新性地提出了 ALSTM 模型。该模型巧妙地将注意力机制提取的主要特征作为输入，传递给长短期记忆神经网络（LSTM）。这种做法不仅降低了计算复杂度，还高效地维持了数据点之间的关联性。

1. 西安石油大学 陕西西安 710065

1 基于 ALSTM 的网络入侵检测模型

1.1 长短期记忆神经网络（long short-term memory, LSTM）

由于深度学习的不断发展，长序列问题处理方面已经得到很好地解决，其中递归神经网络（recursive neural networks, RNN）表现得尤其出色。但是，RNN 经常会出现梯度消失或梯度爆炸等问题。为了提高深度学习网络的效率，研究人员致力于在梯度下降的过程中提高模型的收敛性。长短期记忆神经网络的出现便解决了这个问题。LSTM 是一种由单位单元组成的神经网络，每个单位单元使用指定的门控制机制分析学习序列并保存序列特征，这在预测具有时序特征的数据时尤为重要 [16]。

LSTM 的独特之处在于引入了门控机制：输入门、遗忘门和输出门。此外还有一个记忆单元，用来保存时间步长的信息。LSTM 的门控制机制能够有效地保存和删除信息，从而解决 RNN 在处理长序列数据中的梯度消失或梯度爆炸的问题。网络结构如图 1 所示。

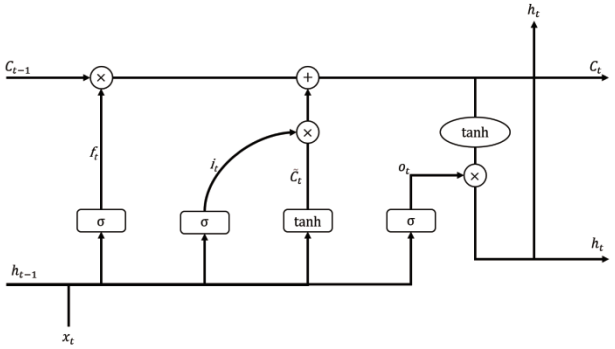

图 1 LSTM 网络结构

LSTM 的第一步是遗忘门，用来判断记忆单元需要丢弃什么信息，计算公式为：

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_f \cdot [h_{t-1}, x_t] + \boldsymbol{b}_f\right) \tag{1}$$

式中：$\boldsymbol{f}_t$ 表示遗忘门的激活值，是一个 0 到 1 之间的向量。接近 1 表示保留信息，接近 0 表示丢弃信息；σ 表示 Sigmoid 激活函数。Sigmoid 函数将输入映射到 0 和 1 之间；h_{t-1} 表示前一时刻的隐藏状态；x_t 表示当前输入；$\boldsymbol{W}_f$ 和 $\boldsymbol{b}_f$ 表示遗忘门的权重和偏置。

第二步是输入门，决定在记忆单元中需要存储哪些新信息。计算主要分为两个部分：

$$\boldsymbol{i}_t = \sigma(\boldsymbol{W}_i \cdot [h_{t-1}, x_t] + \boldsymbol{b}_i) \tag{2}$$

$$\tilde{\boldsymbol{C}}_t = \tanh(\boldsymbol{W}_C \cdot [h_{t-1}, x_t] + \boldsymbol{b}_C) \tag{3}$$

式中：$\boldsymbol{i}_t$ 表示输入门的激活值，是一个 0 到 1 之间的向量。接近 1 表示允许更新信息，接近 0 表示阻止更新；tanh 表示双曲正切激活函数。将输入映射到 -1 和 1 之间；$\tilde{\boldsymbol{C}}_t$ 表示候选单元状态，包含新的信息，根据当前输入和前一个隐藏状态计算出的；$\boldsymbol{W}_i$ 和 $\boldsymbol{b}_i$ 表示输入门的权重和偏置。

第三步是单元状态更新，通过遗忘门选择遗忘旧细胞信息 $\boldsymbol{C}_{t-1}$ 的一部分，输入门选择候选细胞信息 $\boldsymbol{C}_t$ 的一部分，将两者结合得到新的细胞信息 $\boldsymbol{C}_t$，计算公式为：

$$\boldsymbol{C}_t = \boldsymbol{f}_t \cdot \boldsymbol{C}_{t-1} + \boldsymbol{i}_t \cdot \tilde{\boldsymbol{C}}_t \tag{4}$$

第四步是输出门，控制了细胞状态中输出多少信息到隐藏状态，而隐藏状态是 LSTM 在当前时间步的输出，计算公式为：

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_o \cdot [h_{t-1}, x_t] + \boldsymbol{b}_o) \tag{5}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \cdot \tanh(\boldsymbol{C}_t) \tag{6}$$

式中：$\boldsymbol{o}_t$ 表示输出门的激活值，是一个 0 到 1 之间的向量。接近 1 表示允许输出信息，接近 0 表示阻止输出；h_t 表示新的隐藏状态，作为当前时间步的输出，由输出门和单元状态共同决定。

1.2 注意力机制（Attention）

注意力机制是一种源于人类视觉研究的信号处理技术，是机器学习模型，尤其是深度学习模型中一种关键结构。其核心作用在于自动学习并量化输入数据中不同部分对输出结果的影响 [15]。换句话说，注意力机制能够评估输入数据的各个组成部分的重要性。通过这种方式，模型在处理数据时，可以像人类一样，将注意力集中于更有价值、更相关的部分，而忽略或减少对结果影响不大的信息。在深度学习模型中应用注意力机制，相当于引入了类似人脑的“专注”能力。这使得模型能够更有效地处理复杂数据，因为它能够有选择地关注关键信息，从而提升学习效率和预测准确性。简而言之，注意力机制帮助模型聚焦于最关键的输入特征，从而提高整体性能。

注意力机制中的主要参数为 $\boldsymbol{e}_t$、a_t 和 C_t。其中 $\boldsymbol{e}_t$ 为 t 时刻不同特征对应的权重得分，计算公式为：

$$\boldsymbol{e}_t = v \tanh(\boldsymbol{W}_e \boldsymbol{h}_t + \boldsymbol{b}_e) \tag{7}$$

式中：v 和 $\boldsymbol{b}_e$ 为多层感知器计算注意力权重时的权重，$\boldsymbol{b}_e$ 为多层感知器计算注意力权重时的偏置；h_t 为隐藏层在时间 t 的输出。

a_t 为在时间 t 时不同特征对应的注意力权重，计算公式为：

$$\alpha_t = \frac{\exp \boldsymbol{e}_t}{\sum_{j=1}^{n} \boldsymbol{e}_j} \tag{8}$$

式中：e_j 为不同特征在时刻 j 对应的权重得分。

$\boldsymbol{C}_t$ 为整个注意力机制在 t 时刻的输出，计算公式为：

$$\boldsymbol{C}_t = \sum_{j=1}^{n} \alpha_j h_j \tag{9}$$

1.3 ALSTM 模型

网络流量数据具有显著的时序特性，传统方法难以充分捕捉。长短期记忆网络（LSTM）是一种擅长处理时间序列的循环神经网络，能够有效记忆长期依赖关系。因此，利用 LSTM 可以更好地理解网络流量的变化趋势，从而提高异常检测的性能。但是，网络流量的异常特征并非直接显现，而是通过传输数据包的长度、速率、时间间隔等因素体现。有效的异常检测方法需要深入分析这些特征，提取区分正常流量和异常流量的关键信息。显然单一的 LSTM 模型无法准确分析出这些重要特征，因此通过引入注意力机制，可以为模型分配不同的注意力，提高模型对重要因素的自动处理能力，从而较好地解决了这一难题。综上所述，本文提出了一种新的网络流量异常检测模型 ALSTM，该模型利用了注意力机制的特性，在异常检测中更加关注重要的影响因素，提高检测精度。模型架构如图 2 所示。

ALSTM 异常检测模型，由输入层、全连接层、注意力层、LSTM 层、归一化（BN）层、全连接层、输出层组成，具体模型参数如下：输入层的维数设置为 3，全连接层的隐藏单元数为 50，LSTM 层的神经元数为 50，同时使用 tanh 作为激活函数，输出层的最终维数为 1。

构建的模型能够充分利用 LSTM 和注意力层的特性来提高对时间序列数据检测的准确度，BN 层和全连接层的引入也有助于优化模型的训练和整体性能。整个工作流程如下：

首先对网络流量数据进行数据预处理，并对数据进行标准化、归一化后输入预测模型，通过全连接层进行特征映射，并引入注意力机制计算分配给每个输入的注意力权重，加权后输入到 LSTM 层，同时引入 BN 层防止梯度爆炸或消失的问题，最后，LSTM 输出通过全连接层进行集成，输出预测结果。这种层次结构可以更好地适应不同网络环境下的流量数据检测。

2 实验设计与结果分析

2.1 数据集

CICIDS2017 数据集是 Sharafaldin 等人开发的基于流的现代网络入侵数据集。该数据集中使用 B-Profile 系统来描述人类交互的抽象行为，并生成自然的良性背景流量。数据集的生成跨越了 5 天，使用了 14 台机器，包括正常和恶意流量。正常流量是通过 B-Profile 系统从一组 25 人的良性行为中得出，而恶意流量是通过在特定的时间窗口执行现有的攻击工具生成的。

数据采集从 2017 年 7 月 3 日星期一上午 9 点开始，到 2017 年 7 月 7 日星期五下午 5 点结束，共计 5 天。其中包括：良性流量（Benign）、僵尸网络活动（Bot）、端口扫描攻击（PortScan）、分布式拒绝服务攻击（DDoS）、拒绝服务攻击（DoS）、入侵攻击（Infiltration）、针对 Web 应用程序的攻（Web Attack）等不同样本。

对原始 CICIDS 2017 数据集进行去重、删除无关特征等操作后，共 2 812 897 条数据，其中良性流量 2 260 360 条，异常流量 552 537 条。由于整个数据集相当不平衡，良性数据非常庞大，而异常流量数据却相当少，数据集严重不平衡，因此对数据集进行平衡处理。处理后良性流量有 2 260 360 条，异常流量 587 939 条。表 1 列出各类样本的样本数量以及平衡处理后的各样本数量。

表 1 CICIDS 2017 数据集样本数量

类型	原始样本数量	平衡处理后样本数量
Benign	2 260 360	2 260 360
DoS Hulk	229 198	229 198
PortScan	157 703	157 703
DDoS	127 082	127 082
DoS GoldenEye	10 289	10 289
FTP-Patator	7 894	7 894
SSH-Patator	5 861	5 861
DoS slowloris	5 771	5 771
DoS Slowhttptest	5 485	5 485
Bot	1 043	7 772
Web Attack-Brute Force	1 497	5 988
Web Attack-XSS	648	5 184
Infiltration	34	8 704
Web Attack-Sql Injection	21	5 376
Heartbleed	11	5 632

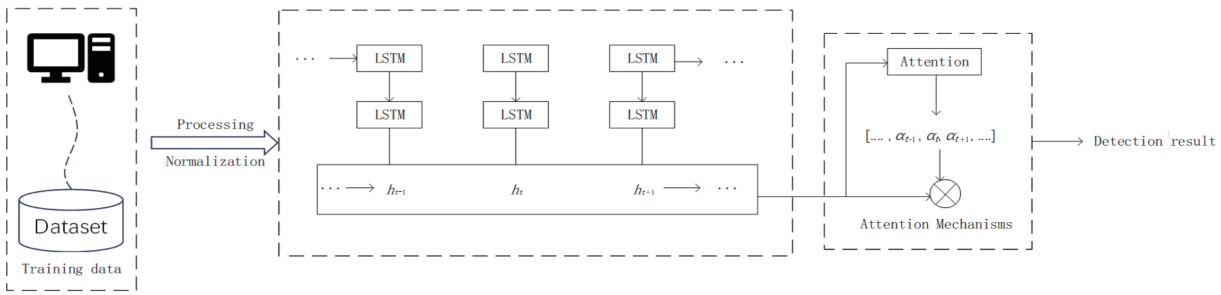

图 2 模型架构图

2.2 评价指标

为了对模型性能进行一般性评估，选用准确率（Accuracy）、精确率（Precision）、召回率（Recall）和 F_1 值（F_1-score）4 个指标评估模型性能。4 个指标分别依赖于数据实例的 4 个不同的类别：真阳性（true positive, TP）、假阳性（false positive, FP）、真阴性（true negative, TN）和假阴性（false negative, FN）。具体释义如表 2 所示。

表 2 依赖指标释义

名称	简称	释义
真阳性	TP	模型正确预测为正的样本数量
假阳性	FP	模型错误预测为正的样本数量（实际是负类）
真阴性	TN	模型正确预测为负类的样本数量
假阴性	FN	模型错误预测为负的样本数量（实际是正类）

基于上述定义，可以计算评估指标：

准确率（Accuracy）：

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{10}$$

精确率（Precision）：

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{11}$$

召回率（Recall）：

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{12}$$

F_1-score:

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$

2.3 实验环境

本实验在一台搭载 Intel Core i5-13400F CPU（10 核），NVIDIA GeForce RTX 4060 GPU（8 GB 显存），内存容量为 32 GB，运行 Windows11 操作系统的环境下实现的。代码实现基于 Python 3.11，深度学习框架采用 TensorFlow2.15.0，时序特征分析使用 pandas 和 sklearn 库。

2.4 实验结果与分析

为了更清楚地展现异常检测模型性能，将 ALSTM 模型与最先进的入侵检测方法在 CICIDS2017 公开数据集上进行了比较。所有的研究方法都使用监督学习方法训练模型。DeepGFL 是一个可以从线性网络流图中提取深度特征的框架。LSTM 和 1D-CNN 是执行二进制分类的深度学习模型。1D-CNN+LSTM 和 LUCID 使用 CNN 来检测 DDoS 攻击。同时，为避免不同研究指标因计算方法不同而带来数据偏差，本文采用准确率（Accuracy）、精确率（Precision）、召回率（Recall）和 F_1 值 4 个指标作为评估指标。表 3 总结了上述模型的性能结果，并将其与本文模型进行了比较。

表 3 各模型性能结果

模型	Accuracy	Precision	Recall	F_1-score
DeepGFL	0.971	0.948	0.448	0.531
MLP	0.705	0.884	0.862	0.872
LSTM	0.975	0.984	0.898	0.895
1D-CNN	0.971	0.981	0.901	0.939
1D-CNN+LSTM	0.978	0.974	0.991	0.982
LUCID	0.955	0.993	0.999	0.996
ALSTM	**0.983**	**0.988**	**0.980**	**0.983**

实验结果表明，本文提出 ALSTM 方法在网络流量异常检测方面表现优异。ALSTM 在召回率和 F_1-score 等方面明显优于 DeepGFL、MLP、1D-CNN 和 LSTM 等现有方法，其中，较 1D-CNN+LSTM 模型准确率提升 0.5%，精确率提升 1.4%。尽管 LUCID 和 1D-CNN+LSTM 在某些指标上略有优势，但它们仅限于二分类场景。相比之下，ALSTM 方法能够在执行多分类任务的同时保持较高的召回率，在实际网络安全应用中更具实用价值。

2.5 实验结果与分析

为了更直观地展示各模型在 CICIDS2017 数据集上的性能表现，本文对检测结果的各项指标（例如：准确率、召回率、F_1-score 等）进行了可视化，结果如图 3 所示。

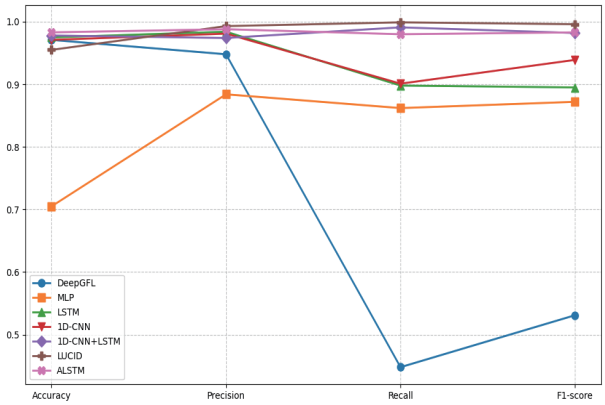

图 3 性能指标

如图 3 所示，ALSTM、1D-CNN+LSTM 和 LUCID 模型在 CICIDS2017 数据集上均表现出良好的异常检测性能，而 DeepGFL 模型的性能则相对较差。与其他模型相比，本文提出的 ALSTM 模型在准确率、精确率、召回率和 F_1 值上均有所提升；尤其是在召回率和 F_1 值上，相较于 DeepGFL 模型分别提升了 53.2% 和 45.2%，提升效果显著。尽管相比于其他模型，各项指标仅提升约 1%，但这表明本文提出的 ALSTM 模型性能优于传统方法，并展现出更强的泛化能力，能够有效分类八种不同类型的网络攻击。

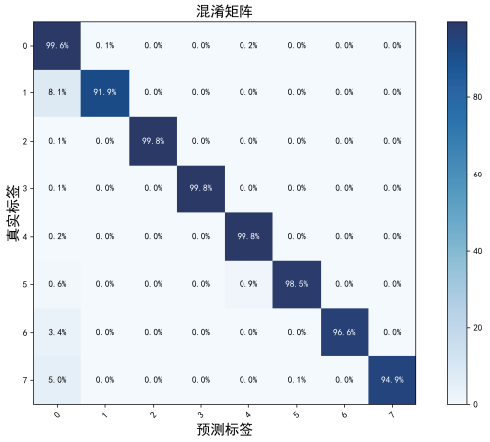

图 4 预测结果混淆矩阵

如图 4 所示，对 CICIDS2017 数据集进行 ALSTM 模型训练后的混淆矩阵分析显示，模型的整体分类准确率较高。然而，非对角线元素的存在表明存在类别混淆。值得关注的是，正常流量（BENIGN）与 Bot、Infiltration 和 Web Attack-XSS 的混淆现象较为明显，这提示正常流量的特征与这些攻击类型的特征可能存在相似性，导致模型难以有效区分。虽然其他类别间的混淆程度较低，但这一发现强调了进一步优化模型以提高对特定类别区分度的必要性。

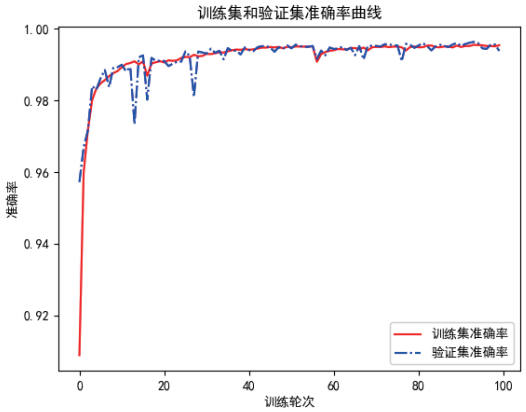

图 5 训练集和验证集准确率曲线

如图 5 所示 CICIDS2017 数据集上 ALSTM 模型的训练集和验证集准确率曲线显示模型表现出色。曲线快速收敛，在 20 个 epoch 左右稳定在高准确率水平。尽管存在少量波动，但训练集和验证集准确率总体趋势一致，表明模型泛化性能良好，可以忽略这些波动。

3 结论

为了提升网络流量异常检测的准确性和效率，本文提出了一种基于注意力机制的 ALSTM 网络流量异常检测模型。该模型通过关注数据间的交互，整合有价值的影响信息，从而在 LSTM 预测的基础上提高了预测精度。为了进一步提高模型的性能，本文还对 CICIDS2017 数据集进行了预处理。实验结果表明，与 DeepGFL、MLP 和 1D-CNN 等传统模型相比，本文提出的 ALSTM 模型在相同实验环境下表现更优。未来，将考虑网络设备的物理位置、使用人群以及使用时间等因素，进一步改进模型，以应对更复杂的网络环境。

参考文献：

[1] 刘新，常英贤，孙莉莉，等．基于深度学习的电网网络攻击检测研究 [J]. 自动化仪表 ,2022,43(12):81-85.

[2] HUANG B, ZHU Y M, WANG Z Z, et al. Imbalanced data classification algorithm based on clustering and SVM[J]. Journal of circuits, systems and computers, 2021, 30(2): 2150036.

[3]ZHANG T, HOU Z M, LI X Q, et al.A novel multivariable prognostic approach for PEMFC degradation and remaining useful life prediction using random forest and temporal convolutional network[J].Applied energy,2025,385:125569.

[4]ER M B, KURAN U, İLHAN N .A novel approach for bird sound classification with cross correlation by denoising with complementary ensemble empirical mode decomposition using B-spline and LSTM features[J].Applied acoustics,2025,233:110601.

[5]LI X, WANG T, GUO J, et al. Corrigendum to "Identification of ARMA models with binary-valued observations" [Automatica 149(2023) 110832][J].Automatica,2025,175:112177.

[6]XIAO H C, SHEN X Y, LI J L, et al. A method for filling traffic data based on feature-based combination prediction model[J]. Scientific reports, 2025, 15(1):8441.

[7]SHAFI M, LASHKARI A H, ROUDSARI A H. NTLFlowLyzer: towards generating an intrusion detection dataset and intruders behavior profiling through network and transport layers traffic analysis and pattern extraction[J]. Computers & security, 2025,148:104160.

[8]DING R Y, SUN L, ZANG W F, et al. Towards universal and transferable adversarial attacks against network traffic classification[J].Computer networks,2024,254:110790.

[9]OLABANJO O, WUSU A, AIGBOKHAN E, et al. A novel graph convolutional networks model for an intelligent network traffic analysis and classification[J/OL].International journal of information technology,2024[2025-05-26].https://link.springer.com/article/10.1007/s41870-024-02032-4.

[10]HU Y H, ZENG Z Q, SONG J P, et al. Online network traffic classification based on external attention and convolution by IP packet header[J].Computer networks,2024,252:110656.

[11]CAI W, HOU C S, CUI M X, et al. Incremental encrypted traffic classification via contrastive prototype networks[J]. Computer networks,2024,250:110591.

基于蜜罐技术的通信网络虚假数据注入攻击防御方法

李志刚[1]
LI Zhigang

摘 要 虚假数据注入攻击会破坏通信信息时间序列，引起攻防对抗演化态势的变化，导致基于原始空间进行的特征映射难以有效过滤虚假数据，降低防御效果。为此，文章开展了一种基于蜜罐技术的通信网络虚假数据注入攻击防御方法研究。通过分析原始网络流量数据的时序特性，构建观测值与时间戳的关联关系。采用不变时间步生成注意力参数，并集成数据权重，生成具有时序基准的通信网络数据时间序列。利用蜜罐技术对当前网络空间中因数据时间序列动态变化引起的攻防对抗演化态势进行表征，输出相应的蜜罐架构，构建出能够适应攻防对抗演化的时间序列映射空间。引入拉普拉斯机制，将包含虚假数据的通信网络时间序列映射至所构建的空间中。利用动态隐私预算调整注入虚假数据幅度以平衡攻击防御与数据质量关系。在构建的映射空间中精准识别并拦截虚假数据，有效防止其在通信网络中传播，实现对虚假数据注入攻击的多层次防御。测试结果表明，该防御方法对攻击的防御遗漏率能够稳定控制在 2.0% 以下，攻击成功率可稳定在 3.0% 以下，具有较为出色的防御效果。

关键词 蜜罐技术；通信网络；虚假数据注入攻击；时间序列注意力参数；攻防对抗演化；映射空间

doi：10.3969/j.issn.1672-9528.2025.10.030

0 引言

虚假数据注入攻击（false data injection attack, FDIA）作为通信网络中极具破坏性的攻击手段之一，其不仅会干扰通信网络的正常运行，还会破坏原始数据的完整性和真实性[1]，降低网络的安全性。因此，针对虚假数据注入攻击设计有效的防御技术对于通信网络安全具有重要的现实意义。

在相关研究中，张贻泉等人[2]提出一种基于深度确定性策略梯度（deep deterministic policy gradient, DDPG）的对抗样本防御框架，通过构建“攻击 - 防御”博弈环境训练智能体动态调整网络参数。在 CIFAR-10 数据集上测试，针对 FGSM（fast gradient sign method）攻击的模型鲁棒性提升 23%，推理速度仅降低 8%。然而，由于虚假数

1. 山西工程科技职业大学 山西晋中 030619

[12]BAYAT M, GARSHASBI J, MEHDIZADEH M, et al.ITC-Net-blend-60: a comprehensive dataset for robust network traffic classification in diverse environments[J/OL].BMC research notes,2024[2025-03-23].https://bmcresnotes.biomedcentral.com/articles/10.1186/s13104-024-06817-5.

[13]JUNG I S, SONG Y R, JILCHA L A, et al. Enhanced encrypted traffic analysis leveraging graph neural networks and optimized feature dimensionality reduction[J]. Symmetry,2024,16(6):733-733.

[14] ALTAF T, WANG X, NI W, et al. GNN-based network traffic analysis for the detection of sequential attacks in IoT[J]. Electronics,2024,13(12):2274.

[15] 孟飞 , 郑卓然 , 黄文聪 , 等 . 基于 CNN-BiLSTM-Attention 的深基坑变形预测方法 [J]. 地下空间与工程学报 , 2025, 21(S1): 87-94.

[16]LAWRYŃCZUK M , ZARZYCKI K .LSTM and GRU type recurrent neural networks in model predictive control: a review[J]. Neurocomputing, 2025, 632l: 129712.

【作者简介】

曹旭泽（1999—），男，陕西西安人，硕士研究生，研究方向：机器学习、大数据技术与应用。

卢胜男（1982—），女，江苏徐州人，博士，副教授、硕士生导师，研究方向：图像处理及机器学习、大数据技术应用。

（收稿日期：2025-05-08 修回日期：2025-09-23）