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）

据注入攻击会引发攻防对抗的演化变化，一旦博弈环境构建完成，其规则和参数便相对固定，难以快速适应这种动态变化。当攻击者采用新的攻击策略或模式时，博弈环境可能无法及时更新和调整，导致训练出的智能体在面对新的攻击情况时，防御效果欠佳。王文杰等人[3]开展了基于KNN分类算法的电力系统网络虚假数据注入攻击防御方法研究，利用KNN算法检测PMU量测数据的空间异常特征，并通过动态调整k值权重，在IEEE 14节点系统中的检测准确率达91.7%，有效防御率控制在94.3%以上。但是该方法仅利用PMU量测数据的空间异常特征进行检测，特征维度相对单一，无法全面刻画虚假数据的本质特征。因此，对于一些复杂、隐蔽的虚假数据攻击模式，其检测能力不足，容易出现漏检情况。吴立钊等人[4]设计了一种梯度可信度评估机制（GTEM），结合鲁棒聚合算法防御半异步联邦学习中的投毒攻击。在Non-IID数据场景下，相比传统FedAvg算法，模型准确率受攻击影响降低12%。但虚假数据注入攻击可能采用复杂的对抗样本生成技术，这些攻击手段能够生成高度逼真的虚假数据，难以通过简单的梯度可信度评估机制进行有效区分。GTEM在面对这些复杂对抗样本时，评估准确性会下降，导致虚假数据被误判为可信数据，进而影响模型训练的质量和效果。陈虹等人[5]提出以MACC（multi-layer access control contract）框架为基础，集成状态锁、调用深度验证和资源配额限制三重防护机制的攻击防御方法研究。在以太坊测试网上实现，可100%阻断经典重入攻击，Gas消耗增加不足15%。然而，其验证规则仅基于历史攻击模式制定。当攻击者采用新的对抗样本生成技术时，该方法可能将包含虚假数据的梯度误判为合法数据。同时，复杂的验证计算过程会延长模型聚合时间，降低训练效率，使得系统在面对快速变化的攻击时处于被动防御状态。为此，本文基于蜜罐技术，研究了通信网络中的虚假数据注入攻击防御方法，并通过仿真测试验证了其防御性能。

1 通信网络虚假数据注入攻击防御方法设计

1.1 通信网络数据时间序列生成

在通信网络遭受虚假数据注入攻击时，原始通信网络数据的时间序列特征会遭到破坏[6]，使得数据原有的时序规律与特性被干扰，无法精准反映通信网络的真实运行状态。对此，为有效防御虚假数据注入攻击，本文首先结合原始网络流量数据中包含的时序特性，生成通信网络数据时间序列。

通信网络数据时间序列本身具有时间相关性和特定的时序模式[7]。基于此，本文在将原始数据按时间顺序进行排列处理后，构建了每个观测值与其对应时间戳之间的关联关系[8]。通过重新建模数据的时间相关性，恢复被攻击破坏的时序规律，为后续防御工作提供可信赖的时序基准。

针对原始通信网络数据$X=\{x_1,x_2,\cdots,x_T\}$，本文采用不变时间步生成通信网络数据在时间序列内的注意力参数，其可以表示为：

$$\alpha(x_i,k)=\frac{x_i k_i}{\sqrt{d}} \tag{1}$$

式中：$\alpha(x_i,k)$表示x_i通信网络数据在k时间步内的注意力参数；k_i表示x_i通信网络数据对应的时间戳；d表示原始通信网络数据的时间尺度参数。

根据注意力参数$\alpha(x_i,k)$，设置其在通信网络数据时间序列中的注意力权重参数，确保x_i通信网络数据在时间序列中的特征向量为定值[9]，其计算方式可以表示为：

$$W_{x_i,k}=\frac{\exp(\alpha(x_i,k))}{\sum_{i=1}^{T}\exp(\alpha(x_i,k))} \tag{2}$$

式中：$W_{x_i,k}$表示x_i通信网络数据在时间序列中的权重参数。随后，集成原始通信网络数据$X=\{x_1,x_2,\cdots,x_T\}$中的各数据权重，生成通信网络数据时间序列可以表示为：

$$H=\sum_{i=1}^{T}W_{x_i,k}x_i \tag{3}$$

式中：H表示通信网络数据时间序列。

利用生成的通信网络数据时间序列，能够明确通信网络数据在原特征空间不同位置的信息，通过与潜在特征空间之间建立映射关系，有效避免虚假数据的注入。

1.2 蜜罐技术构建时间序列映射空间

在完成通信网络数据时间序列生成后，虽然恢复了被攻击破坏的时序规律，但虚假数据注入攻击具有动态性和复杂性，单纯的时间序列难以直接准确识别和抵御这些攻击。网络空间中数据处于动态变化状态，攻防对抗也在不断演化，攻击者会不断调整攻击策略以绕过防御机制[10]。因此，需要构建时间序列映射空间，将通信网络数据时间序列特征差分进行映射处理，以更好地适应这种动态变化的攻防环境，实现对虚假数据注入攻击的有效防御。

本文在构建该映射空间的具体过程中，引入了蜜罐技术。利用蜜罐技术，能够对通信网络数据时间序列表现出的长期趋势进行分析，并结合当前网络空间中数据动态变化所体现的攻防对抗演化特征，从而提高对通信网络虚假数据注入攻击攻防行为分析的可信度和准确性[11]。结合演化博弈理论构建的蜜罐架构可以表示为：

$$D=\left(N_H,\phi,S_H,P_H,U_H\right) \tag{4}$$

式中：D表示考虑当前网络空间数据动态变化攻防对抗演化的蜜罐架构，即通信网络数据时间序列映射空间；N_H表示网络空间中，响应通信网络数据时间序列动态变化演化的对象；ϕ表示拟态空间蜜罐架构的状态集合，即拟态空间蜜罐的组

合方式集合；S_H 表示拟态空间蜜罐架构针对通信网络数据时间序列的演化策略；P_H 表示拟态空间蜜罐架构中演化策略的概率参数；U_H 表示拟态空间蜜罐架构对通信网络数据时间序列的防御收益。由此构建出适应攻击场景下攻防对抗演化的时间序列映射空间。

1.3 基于差分映射机制的虚假数据攻击防御方法

在通信网络环境中，攻击者常构造虚假数据以绕过传统检测机制，且这些虚假数据的形式与特征会随时间和攻击策略动态变化。当通信网络数据的原始序列遭受虚假数据注入攻击时，攻击者还可能利用虚假数据干扰蜜罐诱骗逻辑、污染收集数据，而蜜罐技术主要侧重攻击行为监测，数据层面防御存在局限。对此，本文通过建立通信网络数据时间序列特征的差分映射机制来避免虚假数据在通信网络中的传递。

在构建该差分映射机制对通信网络数据时间序列特征进行处理的过程中，本文引入了拉普拉斯机制。考虑到不同虚假数据具有不同的敏感性，同时通信网络在安全管理需求上也存在差异，利用动态隐私预算对注入虚假数据的幅度进行动态调整，以此平衡通信网络攻击防御与通信数据质量之间的关系，实现对攻击的有效防御。其具体的映射方式可以表示为：

$$\Delta(y,x,x')=\max_{X\in D(H)}\left|\log_2\frac{P(H(x')\in X)}{P(H(x)\in X)}\right| \tag{5}$$

$$P(H(x)\in X)=e^{\frac{\varepsilon(y)}{2}}P(H(x')\in X) \tag{6}$$

式中：$\varepsilon(y)$ 表示注入虚假数据 y 后，通信网络数据的隐私预算；$\Delta(y,x,x')$ 表示注入虚假数据后，通信网络数据时间序列的隐私损失。考虑到通信网络数据传递为原特征空间与潜在特征空间的双向传递，因此存在 $\varepsilon(y)=\frac{\Delta(y,x,x')}{2}$；

$P(H(x')\in X)$ 表示注入虚假数据后，通信网络数据时间序列的映射敏感系数；$P(H(x)\in X)$ 表示通信网络数据时间序列的实际映射敏感系数。

按照上述方式，通过建立通信网络数据时间序列特征的差分映射机制，可有效识别并拦截虚假数据，避免其在通信网络中传播，最终实现对虚假数据注入攻击的有效防御。

2 测试分析

2.1 实验数据集

在分析设计方法的防御性能时，分别选取常用的网络入侵检测数据集 NSL-KDD，以及较新的入侵检测领域常用数据集 CICIDS2017 组成混合测试数据集。其具体的实验数据集构成如表 1 所示。

表 1 实验数据集构成

数据集	数据类型	数据规模 / 条	特征数量	覆盖虚假数据注入攻击	训练集划分占比 /%	验证集划分占比 /%	测试集划分占比 /%
NSL-KDD 数据集	网络连接通信数据	104 525	41	DoS, Probe, U2R, R2L	70.0	20.0	10.0
CICIDS2017 数据集	流量通信数据	1 254 620	85	DDoS, Bot, Heartbleed, Portscan, Web attack	70.0	20.0	10.0

对于实验数据集中以离散值形式存在的参量（如 Protocol_type、Service、Flag 等），利用独热编码的方式将其转化到时序空间。通过这样的方式避免生成的通信网络数据时间序列以及时间序列映射空间存在偏差。

2.2 实验数据集时间序列映射空间构建

在针对实验数据集构建时间序列映射空间这一阶段，将实验数据的时间尺度参数 d 值设定为 2.0。在此拟态空间内，蜜罐组合所生成的时间序列映射空间如图 1 所示。

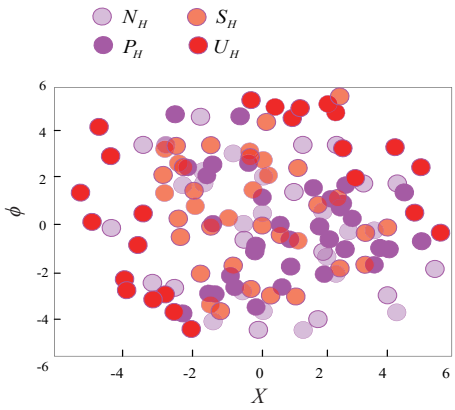

图 1 实验数据集时间序列映射空间

结合图 1 所示的实验数据集时间序列映射空间分布情况进行分析，该映射空间能够表征实验数据时间序列的趋势。从数据动态变化引发的攻防对抗演化角度分析，实现了对趋势的精准呈现。映射空间利用蜜罐架构体现了不同对抗元素之间的攻防关系（具体表现为时间序列映射空间中的叠加态）。这种特性正是其能够针对虚假数据注入攻击中的攻防行为开展高可信度分析的基础，为后续的防御工作提供了坚实保障。

2.3 测试结果与分析

在分析设计方法的防御性能阶段，将基于 KNN 分类算法的电力系统网络虚假数据注入攻击防御方法，以及面向半异步联邦学习的防御投毒攻击方法作为测试的对照组。对于防御性能的分析，分别选取防御遗漏率以及攻击

成功率作为具体的评价指标，以此客观分析设计方法的防御效果。

首先，对比分析了 3 种不同防御方法下，对于测试数据集中不同虚假数据注入攻击的防御遗漏率，结果如图 2 所示。

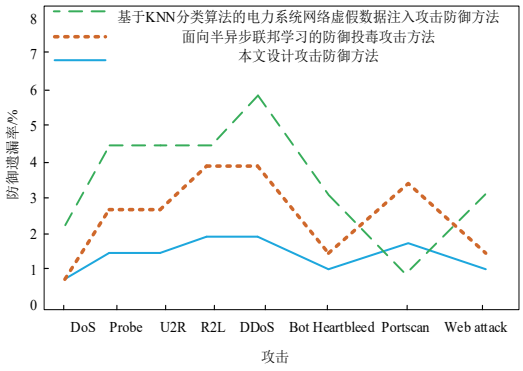

图 2 防御遗漏率

结合图 2 所示的测试结果分析三种方法对于虚假信息注入攻击的防御遗漏率，其中，本文设计方法的遗漏率最低，始终稳定在 2.0% 以下，表明设计防御方法能够实现对攻击的有效防御。这是因为设计方法利用蜜罐技术构建了适应测试数据动态变化对应攻防对抗演化的映射空间，通过映射时间序列差分避免了虚假数据的传递，进而实现了对攻击的有效防御。

其次，对比分析了 3 种不同防御方法下虚假数据攻击的成功率，对应的测试结果如图 3 所示。

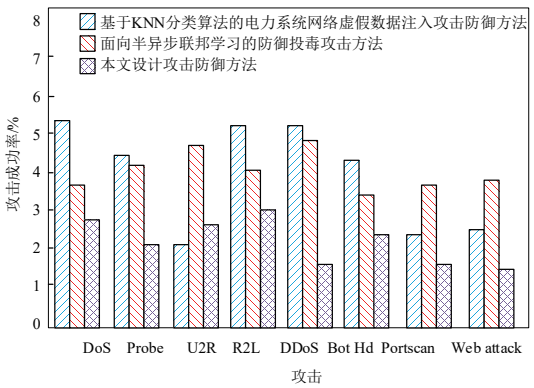

图 3 攻击成功率

结合图 3 所示的测试结果可以看出，在本文设计防御方法下，虚假信息注入攻击的成功率明显低于对照组，整体水平始终稳定在 3.0% 以下，表明其能够实现对虚假信息注入攻击的有效防御。这是因为本文设计方法利用蜜罐架构对不同对抗元素之间的攻防关系进行了表征，使得通过映射测试数据的时间序列特征差分，能够过滤序列中的虚假信息，实现对其对应攻击的有效防御。

3 结语

针对虚假数据注入攻击破坏原始通信信息时间序列引起的攻击成功率较高问题，本文开展了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。结合不变时间步计算得到了通信网络数据的时间序列注意力参数，并生成了通信网络数据的时间序列；在表征当前网络空间中数据时间序列动态变化引起的攻防对抗演化基础上，构建了蜜罐架构形式的时间序列特征映射空间，并将包含虚假数据的通信网络时间序列映射到该空间中，有效过滤虚假信息，实现了遗漏率低于 2.0%、攻击成功率稳定在 3.0% 以下的防御效果。该研究成果可为通信网络安全管理提供参考，后续研究可进一步从防御响应效率角度对方法进行优化。

参考文献：

[1] 滕翠，李容 . 基于文件服务的 Web 攻击实践与防御研究 [J]. 网络安全技术与应用 ,2025(4):19-23.

[2] 张贻泉，张鹏，褚静 . 深度强化学习算法在对抗性攻击防御中的应用探索 [J]. 网络安全和信息化 ,2025(3):52-54.

[3] 王文杰，房海腾，朱成杰，等 . 基于 KNN 分类算法的电力系统网络虚假数据注入攻击防御方法 [J]. 微型电脑应用，2024, 40(10): 130-134.

[4] 吴立钊，汪晓丁，徐恬，等 . 面向半异步联邦学习的防御投毒攻击方法研究 [J]. 信息网络安全，2024,24(10):1578-1585.

[5] 陈虹，谢金彤，金海波，等 . 基于多访问控制的智能合约重入攻击防御方法 [J]. 信息安全研究，2025,11(4): 333-342.

[6] 丁慧霞，薛挺，贾慧贤，等 . 基于攻击图的智能矿山通信立体化网络主动防御技术 [J]. 能源与环保，2024, 46(12): 222-227.

[7] 李志博，蔡英，范艳芳，等 . 面向云边端分层联邦学习中有针对性投毒攻击的鲁棒性防御方案 [J]. 北京信息科技大学学报 (自然科学版), 2024, 39(6): 53-61.

[8] 韩奕昕 . 基于零信任网络的 APT 攻击的分析与防御模型初探 [J]. 长江信息通信，2024, 37(12): 128-131.

[9] 牟大恩，卫志华，孙铭隆，等 . 面向金融领域大语言模型的提示注入攻击防御机制研究及应用 [J]. 网络与信息安全学报，2024, 10(5): 119-133.

[10] 张恒，刘戈威，陈龙文 . 基于多维度攻击模型的网络安全态势感知与防御策略研究 [J]. 中国宽带 ,2025,21(1):58-60.

[11] 刘文钊，郭凯威 . 面向深度神经网络视觉模型对抗鲁棒性的攻击与防御方法研究综述 [J]. 网络安全技术与应用，2025(1): 42-48.

【作者简介】

李志刚（1981—），男，山西临汾人，硕士，讲师，研究方向：软件。

（收稿日期：2025-05-15 修回日期：2025-10-09）