基于 DDPM–RF 的网络安全态势感知算法研究

胡琦涛 [1]
HU Qitao

摘 要 针对网络安全态势感知中普遍存在的数据不平衡问题，文章提出一种基于去噪扩散概率模型（DDPM）与随机森林（RF）的混合算法（DDPM-RF）。该算法首先利用 DDPM 生成高质量、多样化的少数类攻击样本，实现数据集平衡；随后采用 RF 模型完成高效准确的攻击识别；同时结合层次分析法（AHP），构建服务层、主机层、网络层的多层次安全态势量化评估体系。基于 UNSW-NB15 数据集的实验结果表明，DDPM 生成样本的均方误差（MSE）和 Wasserstein 距离均优于传统生成方法；DDPM-RF 模型在准确率、召回率及 F_1-score 等评价指标上均达到 97% 以上，显著提升了少数类攻击识别性能与态势评估可靠性。研究结果证实，所提方法可有效缓解不平衡数据导致的模型偏差，为构建精准、自适应的网络安全防护体系提供了新的技术思路与理论支撑。

关键词 网络安全态势感知；扩散模型；数据平衡；随机森林；攻击检测

doi：10.3969/j.issn.1672-9528.2026.03.027

0 引言

随着信息技术的快速发展，网络攻击手段日趋复杂化、隐蔽化，网络安全威胁已成为影响国家安全和社会稳定的重大挑战 [1]。网络安全态势感知（network security situation awareness, NSSA）作为网络防御体系的核心技术，能够通过对网络环境中多源异构安全要素的采集、融合与分析，实现对安全状态的实时感知、威胁评估与趋势预测，为网络防御决策提供关键支撑 [2]。然而，在实际网络环境中，安全事件的发生具有显著的“长尾分布”特征，正常流量样本数量远远超过各类攻击样本，导致网络安全数据集普遍存在严重的类别不平衡问题。这种数据失衡现象使得传统机器学习模型在训练过程中过度拟合多数类样本，而对少数类攻击样本的识别能力严重不足，最终导致威胁检测漏报率高、态势评估结果出现系统性偏差。

为解决数据不平衡对网络安全态势感知性能的制约，研究者们提出了多种数据增强与重采样技术 [3]。传统的过采样方法通过插值生成合成样本，但易导致特征空间模糊和边界样本过拟合；基于生成对抗网络的方法虽能生成多样样本，却常面临训练不稳定、模式崩溃等挑战，生成样本的质量和真实性难以保证。因此，探索一种能够生成高保真、多样化攻击样本的新型数据平衡方法，成为提升网络安全态势感知模型性能的关键。

近年来，扩散模型（diffusion models）作为一种新兴的深度生成模型，在图像、语音等领域展现出卓越的数据生成能力。其通过模拟数据分布中逐步加噪与去噪的马尔可夫过程，能够学习复杂的数据分布并生成高质量、多样化的样本。这为解决网络安全领域的数据不平衡问题提供了新的思路。本文将去噪扩散概率模型（denoising diffusion probabilistic model, DDPM）引入网络安全态势感知任务，旨在通过其强大的数据生成能力，合成高质量的攻击流量样本，以有效平衡数据集。进一步，本文提出一种融合 DDPM 与随机森林（random forest, RF）的网络安全态势感知算法框架。该框架首先对原始网络流量数据进行预处理与特征增强，利用 DDPM 进行针对性的攻击样本生成以解决类别不平衡问题，然后采用随机森林分类器进行攻击行为判别，最终结合层次分析法（analytic hierarchy process, AHP）实现从服务层、主机层到网络层的多层次安全态势量化评估。本研究以公开数据集 UNSW-NB15 为基础进行实验验证，结果表明，所提出的方法能显著提升模型在失衡数据下的攻击识别准确率与态势评估可靠性，为构建精准、自适应的网络安全防护体系提供了新的技术途径。

1 DDPM-RF 网络安全态势感知算法设计

1.1 算法整体框架

为系统性地解决网络安全态势感知中的数据不平衡问题并实现精准的攻击检测与态势评估，本文提出了基于 DDPM-RF 的网络安全态势感知算法。该算法构建了一个从数据预处理到态势评估的完整处理流程，其整体框架如图 1 所示。该框架以数据平衡为核心，首先对网络公开数据集进行特征分析与预处理，通过 DDPM 模型实现高质量少数类攻击样本的

1. 中国船舶集团有限公司第七一六研究所 江苏连云港 222000

生成，并结合随机欠采样优化多数类样本分布；随后利用包括随机森林在内的多种机器学习模型进行攻击识别；最终基于层次化分析方法，综合考虑服务、主机及网络层级的权重与威胁信息，实现对网络安全态势的量化评估与可视化呈现。该框架的模块化设计确保了各环节的协同运作，为提升在不平衡数据场景下的态势感知性能提供了系统化解决方案。

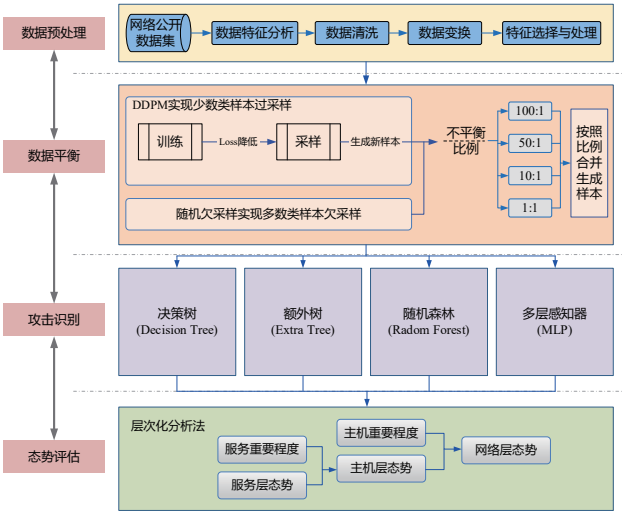

图 1 网络安全态势感知算法整体框图

1.2 数据预处理模块

数据预处理是构建高效网络安全态势感知模型的基础环节，直接影响后续数据平衡与攻击识别的性能[4]。本模块采用系统化的处理流程对原始网络数据集进行规范化与优化，具体包含以下关键步骤：首先，对 UNSW-NB15 公开数据集进行全面的数据特征分析，明确其包含的 42 个特征属性及 9 类攻击流量与正常流量的分布情况；其次，实施严格的数据清洗操作，针对数据中存在的重复值、缺失值与异常值进行处理，以提升数据质量与一致性；随后进行数据变换，通过标签编码与最大最小值归一化等方法，将非数值型特征转换为数值型并统一特征尺度；最后，开展特征选择与处理，基于互信息理论分析各特征与攻击标签的关联性，对低互信息特征施加噪声干扰以增强关键特征与标签间的相关性，为后续的样本生成与模型训练奠定高质量的数据基础。

1.3 基于 DDPM 的数据平衡模块

为有效解决网络安全数据集中普遍存在的类别不平衡问题，本模块创新性地引入了去噪扩散概率模型（denoising diffusion probabilistic model, DDPM）作为核心的数据增强方法，DDPM 模型结构图如图 2 所示。

DDPM 通过模拟数据分布中的前向加噪与反向去噪过程，能够学习并生成与原始攻击样本分布高度一致的高质量新样本。具体而言，该模块首先将经过预处理的少数类攻击样本输入 DDPM 进行训练，使其学习到该类样本的潜在特征分布；随后，根据预设的不平衡比例（如 100:1、50:1、10:1、1:1），通过 DDPM 的采样过程生成指定数量的多样化攻击样本。与此同时，对数量占优的正常类样本采用随机欠采样策略以控制整体数据规模。最终，将生成的攻击样本与下采样后的正常样本合并，构建出类别分布均衡的训练数据集。通过均方误差（MSE）和 Wasserstein 距离（WD）等指标评估表明，DDPM 生成的样本在保真度和多样性上均显著优于传统方法，从而为下游攻击判别模型提供了更优质、更平衡的数据基础，有效缓解了因样本不足导致的模型偏见与性能下降问题。

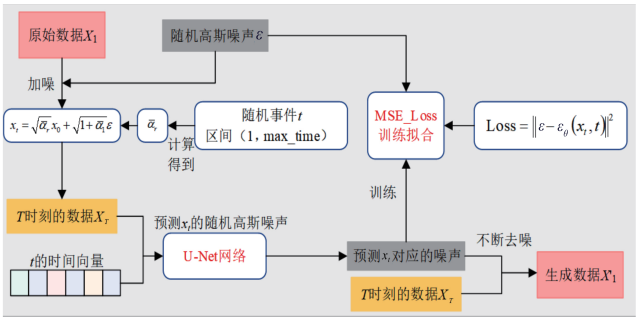

图 2 DDPM 模型结构图

1.4 基于 RF 的攻击判别模块

在完成数据平衡后，攻击判别模块负责对网络流量进行分类，识别其中蕴含的攻击行为。本模块选用随机森林（random forest, RF）作为核心分类器，如图 3 所示，并与其他典型机器学习模型进行性能对比。随机森林是一种集成学习算法，通过构建多棵决策树并综合其预测结果进行分类决策，其核心优势在于能够有效避免过拟合、具备较强的特征选择能力与良好的泛化性能，尤其适用于高维、复杂的网络安全数据[5]。

该模块的具体实现分为三个关键阶段：首先，在特征选择阶段，每棵决策树在构建时随机选取部分特征进行分裂，增强了模型的多样性与鲁棒性；其次，在决策树构建阶段，采用基尼系数作为节点分裂准则，递归地划分数据直至满足停止条件；最后，在集成预测阶段，通过多数投票机制综合所有决策树的分类结果，形成最终的攻击类型判别。实验中将 RF 与决策树（DT）、极端随机树（ET）及多层感知器（MLP）进行对比分析，以全面评估不同模型在不平衡数据场景下的攻击识别能力，为网络安全态势的精准感知提供可靠的分类基础。

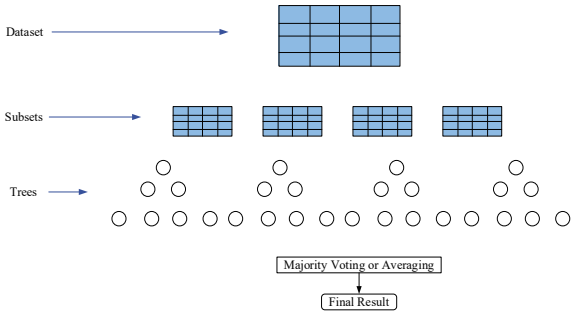

图 3 随机森林结构图

1.5 基于层次分析法的态势评估模块

为实现对网络安全状态的系统化、量化评估，本模块采用层次分析法（analytic hierarchy process, AHP）构建多层级态势评估模型，其层次化网络安全感知架构如图 4 所示。该架构将网络安全态势分解为服务层、主机层和网络层三个层次，通过自底向上的方式逐层聚合威胁信息，最终形成对整体网络安全状况的综合评价。

本模块首先依据 Snort 分级标准与权系数理论对各类攻击行为进行威胁量化，将攻击划分为高、中、低三个威胁等级并分配相应的威胁值。在服务层评估中，根据攻击发生的频率、类型及威胁值，结合服务的重要性和使用频率计算服务态势值；在主机层评估中，综合主机上所有服务的态势值及其权重，计算主机安全态势；最终在网络层，通过聚合所有主机的态势值及其在网络中的重要程度，得到网络整体安全态势值。该模块不仅实现了态势的实时计算，还可通过可视化方式动态展示各层级态势的变化趋势，为网络安全管理与应急响应提供直观、量化的决策依据。

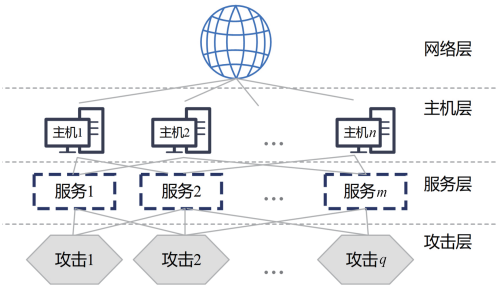

图 4 层级化网络安全感知架构

2 实验与结果分析

2.1 实验环境与数据集

本实验在配置为 NVIDIA V100 GPU、8 核 CPU 与 32 GB 内存的阿里云服务器上进行，采用 TensorFlow 2.14 与 PyTorch 2.1 作为深度学习框架，操作系统为 Ubuntu 22.04。数据集选用澳大利亚新南威尔士大学发布的 UNSW-NB15 公开网络流量数据集，该数据集包含 254 万条记录，涵盖正常流量与 Analysis、Shellcode、Generic、DoS、Fuzzers、Exploits、Worms、Backdoors、Reconnaissance 九类攻击流量[6]。经数据清洗与整合后，最终得到包含 153 684 条样本的有效数据集，每条样本包含 42 个特征属性和 1 个分类标签，其中正常样本占比 55.58%，攻击样本分布极不均衡，如 Worms 类攻击仅占 0.11%，为典型的长尾分布数据集。

2.2 评价指标设定

为全面评估 DDPM 生成样本质量及攻击判别模型的性能，本研究采用以下两类评价指标：

在生成质量评估方面，采用均方误差（MSE）和 Wasserstein 距离（WD）作为核心指标。MSE 通过计算生成样本与真实样本在特征空间中的点对点差异，衡量生成样本的数值保真度；WD 则从概率分布角度衡量生成分布与真实分布之间的整体差异，能更全面地评估生成样本的分布一致性。

在攻击判别性能评估方面，采用准确率（Accuracy）、召回率（Recall）、精确率（Precision）和 F_1-score 四个分类指标，综合评估模型对不同类别攻击的识别能力。其中召回率尤其重要，反映了模型对少数类攻击样本的检出能力，对解决数据不平衡问题具有重要意义。

2.3 DDPM 生成样本质量评估

为验证 DDPM 在网络安全数据生成中的有效性，本研究在 4 种不同不平衡比例（100:1、50:1、10:1、1:1）下，对 DDPM 与传统 SMOTE 方法的生成样本质量进行了对比评估。表 1 展示了在 10:1 不平衡比例下，DDPM 与 SMOTE 生成 9 类攻击样本的 MSE 和 WD 指标对比结果。

表 1 10:1 不平衡比例下 DDPM 与 SMOTE 生成质量对比

攻击类型	样本量	生成方法	MSE	WD
Exploits	45 000	SMOTE	0.08	0.25
		DDPM	0.02	0.13
Fuzzers	30 000	SMOTE	0.09	0.26
		DDPM	0.02	0.13
Reconnaissance	30 000	SMOTE	0.06	0.21
		DDPM	0.03	0.18
Generic	20 000	SMOTE	0.15	0.33
		DDPM	0.08	0.28
DoS	15 000	SMOTE	0.002	0.05
		DDPM	0.001	0.03
Shellcode	10 000	SMOTE	0.002	0.03
		DDPM	0.001	0.02
Analysis	10 000	SMOTE	0.05	0.19
		DDPM	0.008	0.08
Backdoor	10 000	SMOTE	0.003	0.06
		DDPM	0.001	0.02
Worms	8 000	SMOTE	0.01	0.09
		DDPM	0.004	0.06

从表 1 可以看出，在所有 9 类攻击样本生成中，DDPM 在 MSE 和 WD 指标上均优于 SMOTE 方法。特别在 DoS、Shellcode 等少数类攻击样本上，DDPM 的 MSE 值低至 0.001~0.002，WD 值在 0.02~0.03 之间，显著低于 SMOTE 的 0.002~0.06，表明 DDPM 能够生成与真实攻击样本分布更接近的高质量数据。这一结果验证了 DDPM 在处理网络安全不平衡数据时的优越性。

2.4 攻击判别性能对比分析

为评估数据平衡后各分类模型的攻击识别能力，本研究比较了决策树（DT）、极端随机树（ET）、随机森林（RF）

和多层感知器（MLP）在四种不平衡比例下的性能表现。表2展示了10:1不平衡比例下各模型的分类性能指标。

表2 10:1不平衡比例下各模型攻击判别性能

单位：%

模型	准确率	召回率	精确率	F_1-score
DDPM-DT	96.88	96.88	96.88	96.88
DDPM-ET	97.41	97.41	97.41	97.41
DDPM-RF	97.44	97.44	97.45	97.43
DDPM-MLP	88.59	88.59	88.59	88.59

表2显示，在使用DDPM平衡数据后，基于集成学习的模型（ET、RF）表现出更优的攻击判别性能。其中DDPM-RF模型在所有指标上均表现最佳，准确率达到97.44%，召回率和精确率也达到97.44%和97.45%，显著高于DDPM-MLP的88.59%。这表明随机森林在处理网络安全不平衡数据时具有更强的鲁棒性和泛化能力。

为进一步验证DDPM-RF方法的优越性，表3将其与其他先进方法在UNSW-NB15数据集上的性能进行了对比。

表3 不同算法在UNSW-NB15数据集上的性能对比

单位：%

算法	准确率	召回率	精确率	F_1-score
CGAN-DSAE-RESMSCNN	84.2	84.7	84.3	84.5
GMCE-GraphSAGE	89.4	89.4	90.3	89.4
SMOTE-Transformer	93.3	92.1	93.5	92.6
LightGBM	93.8	55.8	59.2	56.2
SMOTE-RF	93.6	93.6	93.5	93.5
DDPM(10:1)-RF	97.4	97.4	97.5	97.4

从表3可以看出，DDPM(10:1)-RF在所有对比算法中表现最优，准确率、召回率、精确率和F_1-score均达到97.4%以上，显著优于传统SMOTE-RF（93.6%）和其他先进方法。特别是在召回率指标上，DDPM(10:1)-RF达到97.4%，而LightGBM仅为55.8%，这充分说明DDPM数据平衡方法能有效提升模型对少数类攻击样本的识别能力。

2.5 态势评估结果可视化

基于层次分析法的态势评估模块能够直观展示网络安全状态的变化趋势。图5展示了一天内网络整体安全态势的变化情况。从图5可以看出，网络态势值在正常状态下接近0，但在第40个时间窗口出现明显峰值，表明该时段网络遭受了集中攻击。通过对服务层和主机层的进一步分析发现，这一异常主要源于特定主机的DNS服务遭受攻击。

图6展示了主机层态势变化情况，反映了网络中三台主机的安全状态波动。图6显示，主机1和主机2在第100个时间窗口附近出现较高的态势值，表明这两台主机在该时段遭受了较频繁的攻击，需要重点关注和安全加固。

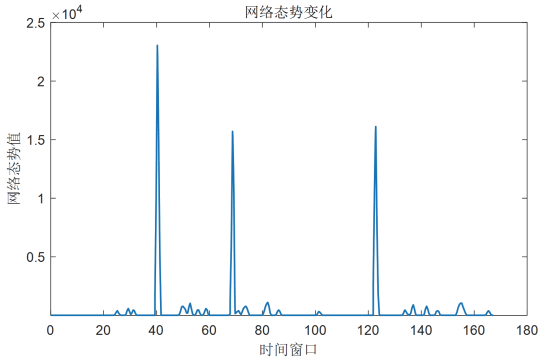

图5 网络整体安全态势变化趋势图

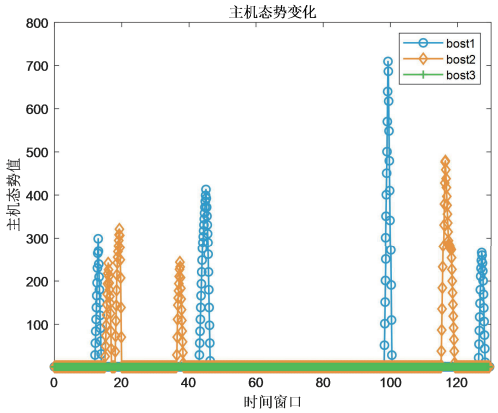

图6 主机层安全态势变化图

这些可视化结果不仅验证了本文提出的DDPM-RF方法在实际网络安全态势感知中的有效性，也为网络安全管理提供了直观的决策支持，有助于安全管理员及时识别威胁、定位问题源头并采取针对性防护措施。

3 结论

针对网络安全态势感知中的数据不平衡问题，本文创新性地提出了基于去噪扩散概率模型（DDPM）与随机森林（RF）的网络安全态势感知算法。通过系统性的实验验证与分析，得出以下主要结论：

首先，DDPM在网络安全数据生成中展现出显著优势。实验结果表明，相较于传统SMOTE方法，DDPM能够生成更高质量、更多样化的攻击样本，其生成样本与真实样本在MSE和Wasserstein距离指标上均表现出更好的匹配度，特别是在DoS、Shellcode等少数类攻击样本上表现尤为突出。这为解决网络安全领域长期存在的数据不平衡问题提供了有效的技术途径。

其次，DDPM与机器学习模型的结合显著提升了攻击识别性能。在10:1不平衡比例下，DDPM-RF模型的准确

基于改进哈希算法的网络信息安全溯源体系研究

康　茜 [1]
KANG Qian

摘　要　针对复杂网络环境下信息安全溯源精度不足、抗篡改能力薄弱，且在大规模数据场景下溯源效率低下的问题，文章对网络信息安全溯源方法进行了深入探索与优化。以深度哈希算法为核心框架，重点优化了哈希码生成、数据一致性检测及节点验证三大关键环节，尤其对 DRH 算法的特征提取层结构与汉明空间映射机制进行了针对性改进，构建了高效可靠的溯源模型。实验结果表明，与现有主流溯源模型相比，所提方法在溯源精度与运行效率上均展现出显著优势：信息溯源检索精度最高可达 92.3%，信息检索一致性稳定在 85% 以上，响应时间仅为 98.67 ms，吞吐量最高达到 215.55 请求 /s。研究结果证实，所提方法在真实复杂网络环境中具备优异的效率与可靠性，可为复杂环境下的网络信息安全溯源提供新的技术路径与解决方案。

关键词　哈希算法；DRH；网络信息；溯源；生成平衡

doi：10.3969/j.issn.1672-9528.2026.03.028

0 引言

随着数字化和智能化的发展，网络信息安全问题日益突出，信息的追溯和篡改检测成为保障数据可信性的重要研究课题 [1]。信息溯源能够有效跟踪数据的来源和流通过程，从而及时发现数据异常，保障信息的完整性和真实性。在此背景下，溯源技术在供应链管理、物联网等领域逐渐受到广泛关注 [2]。李达等 [3] 研究发现，在能源电力项目工程数据生成过程中，数据易遭篡改，导致数据追溯和存证验证难度较高的问题，设计了一种基于区块链的能源电力数据验证与追溯模型。实验结果表明，与传统方法相比，该模型能够有效解

1. 山西工程科技职业大学 山西晋中 030619

率达到 97.44%，召回率和精确率均超过 97%，显著优于传统 SMOTE-RF 模型及其他先进算法。表明 DDPM 生成的高质量平衡数据能够有效缓解分类模型对少数类样本的学习不足，提升模型在实际不平衡网络环境中的攻击检测能力。

同时，基于层次分析法的态势评估模块实现了网络安全状态的多层次量化评估与可视化。通过服务层、主机层到网络层的自底向上态势计算，能够精准定位安全威胁源头，直观展示网络安全状态的动态变化趋势，为网络安全管理和应急响应提供了科学的决策依据。

综上所述，本文提出的 DDPM-RF 网络安全态势感知算法通过高质量数据生成、精准攻击识别和系统化态势评估三个环节的有机结合，有效提升了在不平衡数据场景下的网络安全态势感知性能。该方法不仅为网络安全防护提供了新的技术思路，也为相关领域的数据不平衡问题解决提供了可借鉴的解决方案。未来研究可进一步探索 DDPM 在其他网络安全数据集上的泛化能力，并尝试将深度学习模型与集成学习方法相结合，以应对日益复杂的网络攻击威胁。

参考文献：

[1] 黄健 . AI 在网络安全态势感知中的异常检测技术研究 [J]. 互联网周刊 , 2025(17):34-36.

[2] 耿美月 . 基于深度学习的计算机网络安全态势感知技术 [J]. 信息记录材料 , 2025,26(9):143-145.

[3] 康会娟 , 张腾飞 . 基于机器学习的网络安全态势感知系统研究 [J]. 电脑知识与技术 ,2025,21(20):88-90.

[4] 杨宇 , 谷宇恒 . 网络安全态势感知综述 [J]. 科学技术与工程 , 2022,22(34):15011-15019.

[5] 朱光剑 , 李阳冬 , 钱平 . 基于智能化网络模型的信息安全态势感知算法 [J]. 信息技术 ,2025(5):135-139.

[6] 林广朋 , 李闯 . 入侵攻击下无线网络安全态势感知算法 [J]. 计算机仿真 , 2023,40(12):451-454.

【作者简介】

胡琦涛（1997—），男，黑龙江佳木斯人，本科，研究方向：网络安全、应用电子。

（收稿日期：2025-12-25　修回日期：2026-03-11）