基于RL-WGAN的5G网络异常数据生成方法

宁兆龙[1]，邹道远[1]，周力[2]，欧阳瑞崎[1]，熊炫睿[1]
（1. 重庆邮电大学通信与信息工程学院，重庆 400065；2. 国防科技大学电子科学学院，湖南 长沙 410073）

摘　要：为了解决5G网络异常检测中数据稀缺性、协议复杂性及动态攻击隐蔽性等难题，提出一种基于强化学习优化的Wasserstein生成对抗网络（RL-WGAN）的异常数据生成方法。通过融合强化学习的动态奖励机制与协议约束条件，构建多阶段联合优化框架：设计分层解析策略破解通用分组无线服务隧道协议用户面（GTP-U）协议封装瓶颈，实现流量特征精准提取；创新协议合规性奖励函数与Wasserstein对抗损失的协同优化机制，确保生成数据在协议语义与统计分布上逼近真实数据；采用双向时序建模增强生成器对流量动态演化规律的捕捉能力。实验表明，该方法显著提升了生成样本的分布保真度与协议合法性，有效缓解了异常检测模型的训练数据匮乏问题，为5G网络动态安全防护提供了可靠的数据增强解决方案。

关键词：强化学习；生成对抗网络；5G网络安全；异常数据生成；数据增强
中图分类号：TP393
文献标志码：A
DOI：10.11959/j.issn.1000–436x.2026011

RL-WGAN based method for 5G network anomalous data generation

Ning Zhaolong[1], Zou Daoyuan[1], Zhou Li[2], Ouyang Ruiqi[1], Xiong Xuanrui[1]
1. School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
2. College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China

Abstract: To address the challenges of data scarcity, protocol complexity, and stealthiness of dynamic attacks in 5G network anomaly detection, an anomalous data generation method based on a reinforcement learning Wasserstein generative adversarial network (RL-WGAN) was proposed. By integrating the dynamic reward mechanism of reinforcement learning with protocol constraints, a multi-stage joint optimization framework was constructed. A hierarchical parsing strategy was designed to resolve the general packet radio service turneling protocol user plane (GTP-U) protocol encapsulation bottleneck, enabling precise extraction of traffic features. An innovative protocol compliance reward function was combined with Wasserstein adversarial loss to ensure that generated data approximates real traffic in both protocol semantics and statistical distribution. Bidirectional temporal modeling was adopted to enhance the generator's capability to capture dynamic traffic evolution patterns. Experimental results demonstrate that both the distributional fidelity and protocol compliance of the generated samples are significantly enhanced. This effectively mitigates the problem of training data scarcity for anomaly detection models, providing a robust data augmentation solution for dynamic security in 5G networks.

Keywords: reinforcement learning, generative adversarial network, 5G network security, anomalous data generation, data augmentation

收稿日期：2025–08–05；**修回日期：**2026–01–07
通信作者：周力，zhouli2035@nudt.edu.cn
基金项目：国家自然科学基金资助项目（No.62171449, No.62272075）；重庆市自然科学基金资助项目（No.CSTB2024NSCQ-JQX0013, No.CSTB2024NSCQ-QCXMX0058, No.CSTB2025NSCQ-LZX0050）
Foundation Items: The National Natural Science Foundation of China (No.62171449 ,No.62272075), The Natural Science Foundation of Chongqing (No.CSTB2024NSCQ-JQX0013, No.CSTB2024NSCQ-QCXMX0058, No.CSTB2025NSCQ-LZX0050)

0　引言

随着 5G 的快速发展，其网络流量在数据分布、协议特征等方面与传统互联网呈现显著差异[1-2]。在此背景下，5G通信网络安全领域面临数据集缺乏的严峻挑战，尤其在网络入侵检测系统（network intrusion detection system, NIDS）方面，高质量训练数据的匮乏严重制约了检测模型的性能与泛化能力[3]。传统的通过模拟真实攻击抓取流量的方式，不仅效率低下，其有效性也难以保证，无法满足5G通信网络对大规模、多样化数据集的需求[4]。因此，5G通信系统亟须有效的异常数据模拟生成技术以应对这些挑战。

近年来，数据生成技术，特别是基于深度学习的生成模型，被认为是解决上述问题的有效途径。该技术旨在通过学习真实数据的分布特性，生成高质量、高保真度的模拟数据，从而有效扩充现有数据集，改善样本数据匮乏问题[5]。最新研究进展进一步深化了生成式方法在通信场景的应用潜力[6-7]。针对5G通信网的特殊性，研究能够模拟其复杂流量模式和多样化异常行为的数据生成方法，对于构建有效的5G安全数据集、提升入侵检测系统的准确性和鲁棒性具有非常重要的理论意义和应用前景。

当前网络异常数据生成方法主要分为3类：基于抽样的生成方法、基于优化的生成方法和基于模仿的生成方法。

基于抽样的生成方法主要依托数据变换规则实现样本扩增。文献[8]通过插值技术缓解类别不平衡问题；文献[9]通过上下文依赖建模增强语义分割的域迁移能力。然而，此类方法受限于原始数据的特征空间分布，制约了生成数据的泛化能力。

在基于优化的生成方法领域，研究者普遍采用进化算法进行对抗样本的迭代生成[10]。但此类方法存在显著缺陷：单样本生成所需的时间成本与5G场景的实时性需求存在量级差异；适应度函数的预设依赖性导致生成数据的真实性验证存在方法论缺陷。

基于模仿的生成方法，尤其是生成对抗网络（generative adversarial network, GAN）及其变体，在数据生成领域展现出巨大潜力[11-14]。在网络入侵检测系统（network intrusion detection system, NIDS）领域，文献[15]基于GAN提出了一种针对实时网络流量学习的入侵检测系统的逃避策略，提高了攻击者的成功率；文献[16]提出基于生成对抗网络的目标检测生成框架，用于生成特征以提高低质量图像上目标检测的鲁棒性。文献[17]提出了基于指数信息度量的生成对抗网络模型，用于无监督学习下的异常检测，通过调整目标函数中的事件概率，优化异常数据生成。同时，基于强化学习（reinforcement learning, RL）的生成对抗模型能够适应多样化网络场景和需求[18-19]，文献[20]提出多自我生成对抗网络（multiself generative adversarial network, MultiselfGAN）算法，引入多控制器和奖励重塑机制，利用判别器的输出作为一种自引导的性能评估指标，为神经架构搜索提供反馈，从而替代如起始分数等传统的外部评价指标，提升神经架构搜索效率。在数据增强领域，文献[21]提出了一种基于多周期模式和增强知识GAN来生成大规模城市蜂窝网络流量的方法。

尽管近年来融合RL与GAN的生成方法[22]与自编码器（VAE）、扩散模型等生成式AI模型[23-24]在通用网络流量生成与NIDS领域取得了显著进展，但在精确刻画5G通信网络特有的用户面流量行为和严格遵循复杂网络协议规范等方面，仍面临诸多挑战。这些挑战直接限制了所生成数据在提升5G NIDS针对性与鲁棒性方面的实际效用。综上所述，如何深度融合强化学习在动态决策与序列控制方面的优势与GAN在复杂分布建模方面的能力，生成一种在协议结构上严格合规，同时在统计分布和行为模式上高保真的异常5G流量数据，具有重要的理论研究意义与广泛的实际应用价值。

本文的主要研究工作如下。

1) 提出了一种基于强化学习优化的Wasserstein生成对抗网络（RL-WGAN）的异常数据生成方法，该方法利用Wasserstein生成对抗网络（Wasserstein generative adversarial network, WGAN）生成统计上逼真的网络流量，再通过RL对生成过程进行策略优化，使其能够生成在协议结构上完全合规，但在流量行为上展现特定异常模式的数据，为构建全面、均衡的5G网络安全数据集提供了关键技术支撑。

2) 为有效解决生成数据协议语义合规性问题，提出了一种用于强化学习的协议感知混合奖励函数，通过将判别器反馈与协议规则校验相结合，实

现对生成策略的精准引导与优化，以生成统计逼真且协议合规的5G网络异常流量。

3) 基于5G-NIDD数据集和实验数据集的仿真结果表明，所提方法能够在保证生成数据在统计分布相似性、特征间内在关联与真实数据高度一致的同时，有效缓解异常检测模型的训练数据匮乏问题。

1　系统架构

在5G通信网络架构不断演进的背景下，网络功能模块呈现出高度分离化与虚拟化的趋势，尤其是用户面与控制面的彻底分离，使得用户面成为网络攻击的重点目标。用户面流量面临如分布式拒绝服务攻击（distributed denial of service, DDoS）、漏洞利用攻击、中间人攻击等多种威胁，这些攻击不仅影响网络服务的可用性，还可能造成敏感数据泄露和业务中断等严重后果[25]。相较于传统互联网，5G用户面的协议体系更复杂、数据处理更具时效性，这进一步增加了异常数据的获取与生成难度。

为应对上述挑战，本文提出了一种5G通信网络异常数据生成模型系统架构，如图1所示。该方案在用户平面功能（user plane function, UPF）网元处进行数据采集与生成。左侧描绘了5G网络环境中，攻击者可能从外部互联网或内部接入网发起多种攻击，这些攻击流量最终汇入5G核心网。中部展示了数据流经核心网，在UPF网元处部署数据采集点，抓取原始数据包。右侧是本文模型的核心部分，采集到的真实数据被送入异常网络流数据生成器。该生成器学习真实异常样本的分布特性，生成符合真实分布的、语义合规的新数据样本。这些生成的数据可以有效扩充稀缺的异常流量数据集，最终支撑上层的威胁检测与异常识别和追踪系统。

系统流程如下：首先，在UPF网元处实时采集符合数据包捕获（packet capture, PCAP）文件格式的数据包，并将其传输至远程服务器进行集中存储与后续处理；随后，在服务器端，系统对原始数据包进行深度解析，提取包括五元组、时间特征、流量统计特征等在内的多维度特征信息，从而构建结构化的原始网络流量数据集；最终，基于已标注异常类型的结构化特征数据集，对本文模型进行训练。当实际应用场景中出现数据不平衡或异常样本稀缺的情况时，系统可按需调用训练好的生成模型，执行目标异常类型的流量样本生成任务。

在模型设计方面，本文基于GAN框架，提出一种引入强化学习机制的变体模型——RL-WGAN，用于生成贴近真实分布的5G用户面异常数据。首先，通过大规模真实流量的采集与预处理，构建高质量的训练样本，为生成器与判别器的对抗训练提供充分数据支持。为提高生成样本的多样性与真实性，模型在生成器中引入强化学习机制，通过奖励函数引导生成器在生成过程中强化对潜在异常特征的学习，特别是在协议行为层面上对异常样本特征的建模能力。

强化学习模块中的奖励函数设计包括两个部分：一是协议层奖励，用于评估生成流量在协议结

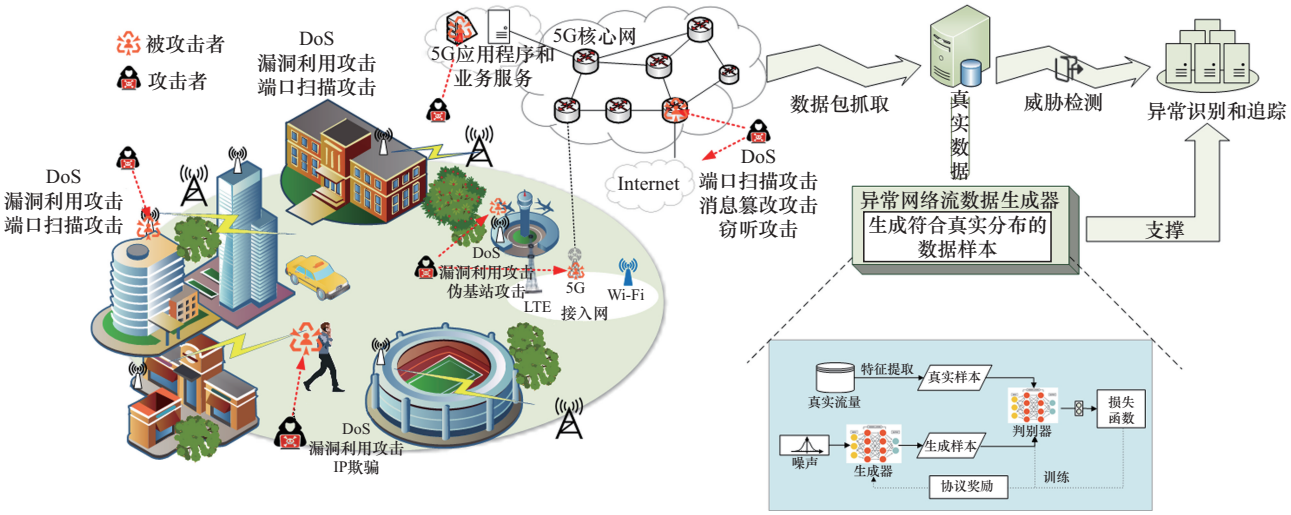


图1　5G通信网络异常数据生成模型系统架构

构和行为特征上的合法性，确保生成样本在语义层面与真实异常流量保持一致；二是对抗性奖励，由判别器输出的反馈信号组成，用于衡量生成器在生成真实异常流量方面的能力。这些奖励信号共同引导生成器策略的更新，使其能够在保持协议一致性的同时生成更加多样、符合攻击行为特征的流量数据。

在判别器设计中，本文引入Wasserstein距离作为损失函数，用于衡量生成样本与真实样本在分布层面的差异。相比传统判别损失，Wasserstein距离在处理高维稀疏特征时表现出更强的鲁棒性。WGAN采用Wasserstein距离取代詹森–香农（Jensen-Shannon, JS）散度，其关键优势在于，即使在分布不重叠的情况下，Wasserstein距离仍能为生成器提供平滑且有意义的梯度。这一特性从根本上保证了训练过程的稳定性，避免了梯度消失问题。同时，由于损失函数能够可靠地反映生成样本的质量，它能有效引导生成器学习真实数据的完整分布，从而显著缓解模式坍塌问题，确保了生成数据的多样性与保真度。

2　异常数据生成方法

本文提出的5G异常数据生成流程如图2所示，其核心环节涵盖数据预处理、模型训练与生成，以及闭环优化。

首先，在数据预处理阶段，系统对原始5G用户面流量包进行多步处理：1)通过通用分组无线服务隧道协议用户面（general packet radio service tunnelling protocol user plane, GTP-U）隧道解封装技术剥离外层协议，精确提取内层IP报文；2)实施基于网络流的特征工程，将离散的数据包聚合成流单元并计算多维度特征；3)采用均值–方差法等统计学方法进行离群值检测与处理；4)最后对特征数据进行最小–最大标准化，消除量纲差异。

经过预处理后，高质量的结构化数据被划分为训练集与测试集。模型训练与生成阶段以训练集为输入，通过对抗训练机制优化生成模型。同时为确保生成质量，闭环优化阶段对生成样本进行严格评估。该环节将生成样本与测试集进行多维度对比分析，重点考察其统计分布相似性与特征结构一致性。评估结果以反馈信号的形式指导模型参数的迭代调优，从而形成一个自适应的优化闭环，持续提升生成数据的真实性与可用性。

为应对网络异常流量的多样性挑战，本文提出一种如图3所示的RL-WGAN并行生成架构。该架构的核心在于为每种异常类型训练一个专用的RL-WGAN模型，并将其集成至统一的模型库中。当需要合成特定攻击流量时，系统仅需调用相应的预训练模型。每个模型作为一个独立的“生成隧道”，能够精准复现特定攻击的流量特征与数据分布。通过该生成策略，模型能够合成高质量的异构流量。这些样本用于扩充原始数据集，进而有效增强入侵检测系统的训练效果与泛化能力。

2.1　5G通信网用户面流量预处理

鉴于机器学习模型对数据的尺度与范围高度敏感，对5G通信网用户面流量开展数据预处理至关重要。该环节的核心在于将原始数据转化为适应模型训练的结构化格式，旨在显著提升模型的性能与泛化能力。

流量特征提取是预处理流程的关键环节。传统逐包扫描方法在处理原始PCAP数据时面临显著效

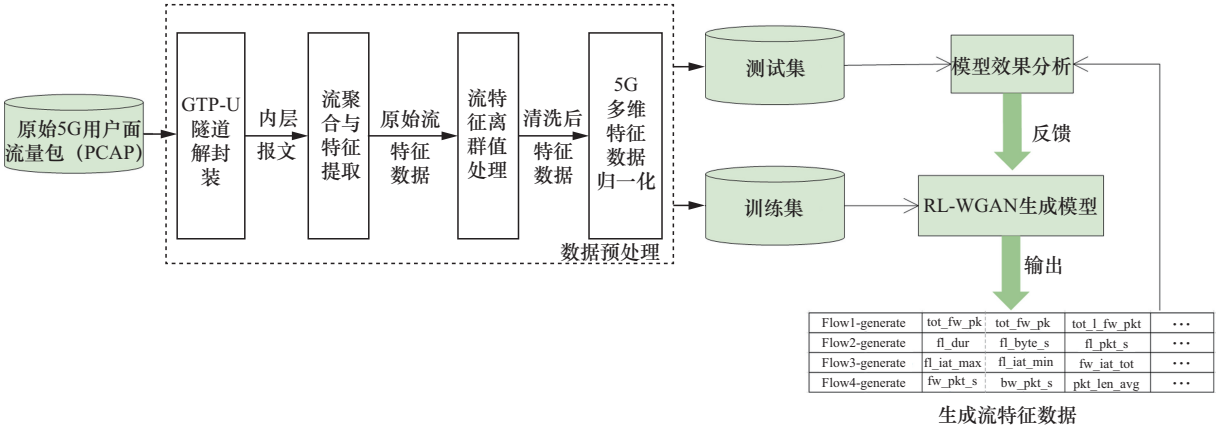


图2　5G异常数据生成流程

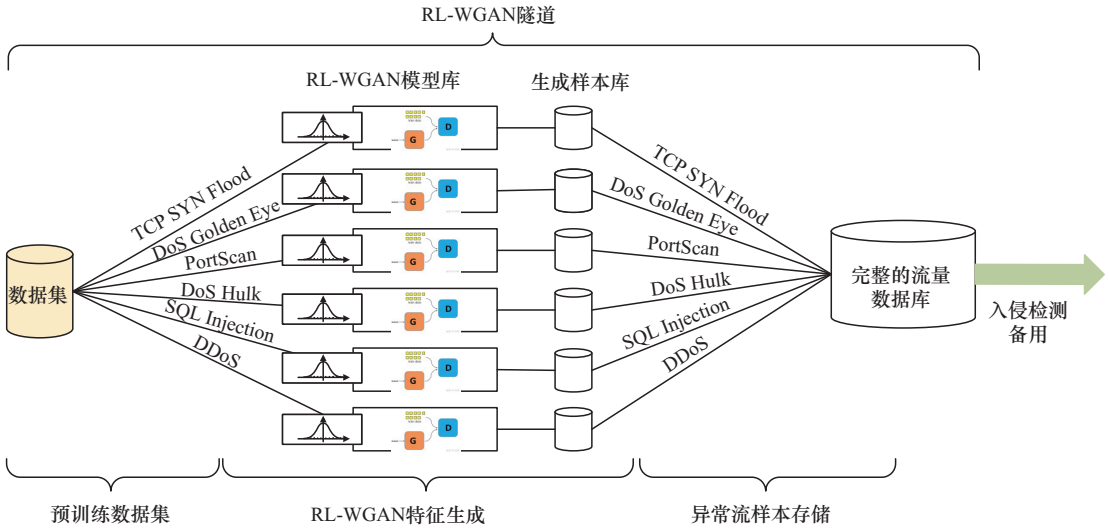


图3　RL-WGAN并行生成架构

率瓶颈。此类数据通常混杂多种协议和报文格式，缺乏统一规范，使逐个包解析开销巨大且难以有效识别复杂攻击行为。为解决此双重困境，本文采用基于网络流的分析方法，通过聚合相同五元组（源/目的IP地址、端口、协议）的数据包形成逻辑流单元。该转换通过结构化聚合显著降低了网络流数据总量，提升了分析处理效率。

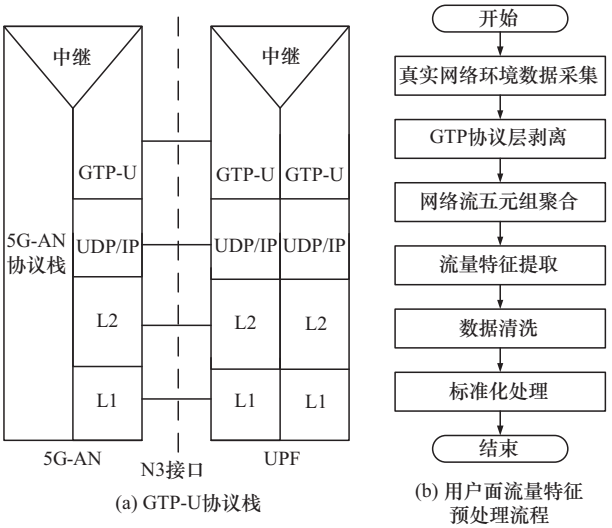
然而，在5G网络特有的用户面架构下，传统的流分析方法面临失效的风险。其核心挑战在于用户面数据包均需经由通用分组无线服务（general packet radio service, GPRS）隧道协议GTP-U封装。此封装机制将用户原始数据包的真实五元组信息隐藏于外层GTP-U协议头之下，致使传统分析方法难以直接提取有效特征。因此，依赖五元组进行流聚合的传统方法无法准确解析和区分不同用户的真实网络行为。

为解决上述由GTP-U封装引发的技术问题，本文设计并实现了一套GTP-U隧道解封装方案。图4展示了该方案的核心流程，包括协议识别、头部解析与内层报文提取等关键步骤。其具体操作步骤如下。

1) GTP-U报文识别：系统对捕获的PCAP数据进行初步筛选，依据外层用户数据报协议（user datagram protocol, UDP）协议头的目标端口号识别GTP-U隧道报文。

2) 协议头解析与载荷定位：对识别的报文执行逐层协议头解析，精确定位GTP-U头部起始位置。

3) 内层IP报文提取：基于解析所得的协议头长度与偏移量信息，程序从原始字节流中完整剥离外层隧道封装，提取内层IP报文原始数据，并重构为标准格式的IP数据包。



(a) GTP-U协议栈　　(b) 用户面流量特征预处理流程

图4　基于GTP-U协议的5G用户面流量处理方法

为确保本文模型的通用性与可行性，本文设计的GTP-U隧道解封装方法严格遵循3GPP TS29.281标准规范。该标准定义了5G核心网用户面中GTP-U协议的统一行为，是所有设备商在实现UPF网元时必须遵守的基准。因此，本文模型在理论上具备跨设备商的通用性。无论是独立组网还是非独立组网架构，其用户面数据均通过N3接口上的GTP-U隧道传输，故本文模型同样适用于这两种主流组网模式。尽管标准协议提供了一致性保障，但实际部署中不排除存在厂商特定的实现细节差异，对这些

边缘情况的适配将作为未来工作的优化方向。

在成功提取内层IP报文后，系统采用基于真实五元组的标准流聚合技术，将离散数据包重组为表征用户行为的网络流基元并进一步提取所需的网络流量特征与用户行为特征。

在特征提取完成后，为保证后续模型训练的稳定性与数据质量，需对数据进行清洗以识别并处理离群值。5G网络流量呈现出极大的动态范围和复杂性：一方面，eMBB等应用可产生极高的合法流量；另一方面，各类网络攻击，无论是产生瞬时巨大流量的DDoS，还是维持长连接、低速率的“慢速”攻击，都会在数据集中形成异常值。这些异常值可能导致模型在训练过程中偏离正常流量分布，从而影响其泛化能力与异常检测准确性。

为此，本文采用均值–方差法对数据集中每一个数值型特征F_j进行离群值检测。首先，计算该特征F_j的均值$\mu(F_j)$和标准差$\sigma(F_j)$

$$\mu(F_j)=\frac{1}{n}\sum_{i=1}^{n}f_{ij},\quad \sigma(F_j)=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_{ij}-\mu(F_j))^2} \tag{1}$$

其中，F_j代表数据集中第j个特征的集合，f_{ij}代表该特征的第i个特征值，n为样本总数。

针对5G网络流量呈现的极大动态范围和复杂性，本文采用3σ准则进行离群值处理。5G网络中，eMBB等合法应用与DDoS等网络攻击均可能产生极端流量值。在此背景下，3σ准则的适用性体现在其作为一种高效务实的启发式方法，能够有效地识别并处理这些极端离群点，其核心目标是保障后续RL-WGAN模型训练的稳定性和收敛速度。通过此方法约束数据范围，可以为生成器和判别器提供一个更优化的学习空间，从而与负责学习复杂真实分布的RL-WGAN模型形成了有效互补。根据3σ准则，任何落在$[\mu(F_j)-3\sigma(F_j),\mu(F_j)+3\sigma(F_j)]$之外的特征值$f_{ij}$都将被视为离群值并予以处理。

为消除不同特征之间的量纲差异，确保模型训练的稳定与高效，在剔除离群值后，必须对数据进行标准化处理。为此，本文采用最小–最大标准化法，将每个特征F_j内的所有特征值f_{ij}线性映射到[0,1]内，其转换公式为

$$f'_{ij}=\frac{f_{ij}-\min(F_j)}{\max(F_j)-\min(F_j)} \tag{2}$$

其中，f_{ij}是特征F_j的原始特征值，f'_{ij}是其标准化后的值，$\min(F_j)$和$\max(F_j)$分别是特征F_j在整个数据集中的最小值和最大值。

2.2　生成对抗网络中的强化学习方法设计

强化学习作为指导生成器G_θ生成网络流量并更新迭代的核心方法，贯穿于整体模型之中。在强化学习框架下，智能体根据当前状态选择最优动作并执行，执行后环境会根据动作反馈奖励，并返回智能体的下一个状态，循环迭代直至收敛，从而实现预期目标。在本文提出的基于强化学习的数据生成方法中，生成器作为智能体，学习最优策略以生成网络流量序列$\boldsymbol{x}_{1:T}=(x_1,\cdots,x_t,\cdots,x_T)$，其中$x_t$表示在时隙$t$生成的特征值，$T$为完整流量序列的特征维度。智能体在环境中的状态表示为$\boldsymbol{s}_t=(x_1,\cdots,x_{t-1})$，智能体$G_\theta$需要根据当前环境状态$\boldsymbol{s}_t$选择下一步动作$a_t$，即决定生成的下一个流量特征值$x_t$。执行动作$a_t$之后，智能体$G_\theta$接收奖励，并基于新的状态$\boldsymbol{s}_{t+1}=(x_1,\cdots,x_t)$继续生成。

本文将整个流量序列的生成任务解构为一个自回归的马尔可夫决策过程。

环境：正在被逐步构建的流量序列本身。

状态：在时间步t，状态$\boldsymbol{s}_t=(x_1,\cdots,x_{t-1})$，即为已经生成的部分序列。它为模型决策下一步动作提供了完整的历史上下文。

动作：生成器在状态$\boldsymbol{s}_t$下所采取的动作a_t，就是选择并生成下一个具体的特征值x_t。

状态转移：当动作a_t被执行后，生成的特征值x_t会被追加到当前序列的末尾，从而使环境进入下一个状态$\boldsymbol{s}_{t+1}=(\boldsymbol{s}_t,x_t)=(x_1,\cdots,x_t)$。

采用此建模方式的动机在于，一个真实有效的网络流量样本，其内部各特征字段并不是相互独立的，而是存在紧密的逻辑关联。通过将每一步的生成都视为一个依赖于历史上下文的决策，强化学习智能体能够学习到这种内在结构，从而确保最终生成的序列在整体上连贯、真实且符合协议规范。

在强化学习中，奖励R通常由环境反馈至智能体。相较于传统生成对抗网络关注生成样本与真实样本的整体一致性，本文考虑到网络流量数据的协议特性及标志位字段等关键特征，因此将奖励R规定为两个组成部分：判别奖励$Q_{D_w}^{G_\theta}$和协议奖励R_p。

在基于强化学习的生成对抗训练中，生成器与

判别器持续博弈，并通过最大化判别器输出的奖励信号不断优化生成策略。判别奖励由 Wasserstein 距离导出，生成分布越接近真实分布，该距离越小，表明生成效果越优。由于判别器仅评估完整序列，针对未生成完整样本的情形，引入蒙特卡罗采样补全序列，以此评估当前策略的潜在收益。最终，拼接序列用于计算判别奖励，具体公式为

$$Q_{D_w}^{G_\theta}(s=\boldsymbol{x}_{1:t-1},a=x_t)=\begin{cases}\dfrac{1}{N}\displaystyle\sum_{n=1}^{N}D_w(\boldsymbol{x}_{1:T}^n),\boldsymbol{x}_{1:T}^n\in \\ \left[\boldsymbol{x}_{1:t-1},\mathrm{MCSampling}\left(\boldsymbol{x}_{t,T},t\right)\right], & t<T \\ D_w(\boldsymbol{x}_{1:T}), & t=T\end{cases} \tag{3}$$

其中，$\mathrm{MCSampling}\left(\boldsymbol{x}_{t,T},t\right)$ 表示从 t 位置进行蒙特卡罗采样得到的长度为 $T-t$ 的序列，$\boldsymbol{x}_{1:T}^n$ 表示生成序列 $\boldsymbol{x}_{1:t-1}$ 与蒙特卡罗采样序列拼接后的临时序列。

在流量生成任务中，不同的协议类型通常对应不同的威胁模式，确保协议的规范性对于构建合理的样本至关重要。为此，协议奖励 R_p 依据以下 4 个维度进行计算：协议特征字段、协议对应的端口字段、协议标志字段及服务与协议的映射关系。

$$R_p=C_1\cdot C_2\cdot C_3\cdot C_4 \tag{4}$$

其中，C_1 表示协议特征字段的合规性，即该字段是否符合协议规范并满足预定义的格式要求。若协议特征字段 F_{pr} 处于协议号特征范围[0,255]内，则 C_1 记为 1，否则记为 0，表示协议奖励无效。

$$C_1=\begin{cases}1, & F_{\mathrm{pr}}\in F_{\mathrm{protocol}} \\ 0, & F_{\mathrm{pr}}\notin F_{\mathrm{protocol}}\end{cases} \tag{5}$$

其中，F_{pr} 表示协议特征字段，F_{protocol} 表示协议特征范围集合。

C_2 用于验证端口号是否正确映射到相应的应用层服务，其计算式为

$$C_2=\begin{cases}1, & F_{\mathrm{po}}\in F_{\mathrm{port}} \\ 0, & F_{\mathrm{po}}\notin F_{\mathrm{port}}\end{cases} \tag{6}$$

其中，F_{po} 表示端口号特征，F_{port} 表示端口号特征的有效范围集合。

C_3 用于检查协议标志字段的合规性。传输控制协议（transmission control protocol, TCP）流量包含确认标志（acknowledgment, ACK）、重置标志（reset, RST）、同步标志（synchronize, SYN）及结束标志（finish, FIN）等特定标志位，而 UDP 流量无此类标志。因此，在一致性校验中，若协议字段为 TCP，且至少一个标志位存在，则 $C_3=1$；若协议字段为 UDP，且所有标志位均不存在，同样判定 $C_3=1$。其计算式为

$$C_3=\begin{cases}1, & F_{\mathrm{pr}}=F_{\mathrm{pr\_TCP}}\ , \\ \mathrm{Num}_{F_{\mathrm{ACK}}}+\mathrm{Num}_{F_{\mathrm{RST}}}+\mathrm{Num}_{F_{\mathrm{SYN}}}+\mathrm{Num}_{F_{\mathrm{FIN}}}\neq 0 & \\ 1, & F_{\mathrm{pr}}=F_{\mathrm{pr\_UDP}}\ , \\ \mathrm{Num}_{F_{\mathrm{ACK}}}\times\mathrm{Num}_{F_{\mathrm{RST}}}\times\mathrm{Num}_{F_{\mathrm{SYN}}}\times\mathrm{Num}_{F_{\mathrm{FIN}}}=0 & \\ 1, & F_{\mathrm{pr}}\notin F_{\mathrm{pr\_else}} \\ 0, & \text{其他}\end{cases} \tag{7}$$

其中，F_{pr} 表示协议特征，$F_{\mathrm{pr\_UDP}}$ 和 $F_{\mathrm{pr\_TCP}}$ 分别表示协议号特征范围集合，$\mathrm{Num}_{F_{\mathrm{ACK}}}$、$\mathrm{Num}_{F_{\mathrm{RST}}}$、$\mathrm{Num}_{F_{\mathrm{SYN}}}$、$\mathrm{Num}_{F_{\mathrm{FIN}}}$ 分别表示 ACK、RST、SYN、FIN 标志位特征的数量。

C_4 用于验证服务与协议的对应关系。在通信网中，不同的威胁和异常的信息可能通过不同的协议完成传输，因此在生成数据时还需要检查攻击所用的服务特征和传输协议之间的映射是否有效。

$$C_4=\begin{cases}1, & F_{\mathrm{service}}\rightarrow F_{\mathrm{protocol}} \\ 0, & F_{\mathrm{service}}\not\rightarrow F_{\mathrm{protocol}}\end{cases} \tag{8}$$

其中，F_{service} 表示服务特征，F_{protocol} 表示协议集合，$\rightarrow$ 表示映射成功，$\not\rightarrow$ 表示映射失败。

当前的样本通过所有条件的协议核查后可以得到协议奖励 R_p。协议奖励 R_p 将作为奖励 R 的一部分参与到最终的计算。为了加强判别器对生成器的引导性，奖励 R 将判别奖励作为主导成分，协议奖励 R_p 作为辅助成分，引入折扣因子 α，最终计算式为

$$R=Q_{D_w}^{G_\theta}-\alpha\cdot R_P \tag{9}$$

其中，$Q_{D_w}^{G_\theta}$ 表示判别奖励，R_p 表示协议奖励，R 为奖励。奖励 R 将作为生成器更新的依据参与到生成器损失函数的计算中。

2.3 基于 RL 的生成器和判别器设计

本文提出的 RL-WGAN 模型主要由生成器和判别器两个部分组成。生成器采用双层双向门控循环单元（bidirectional gated recurrent unit, Bi-GRU）结构，以强化学习策略进行训练优化；判别器则基于 WGAN 框架构建，并利用深度神经网络（deep neural network, DNN）增强判别性能。

通过引入对抗训练机制和策略梯度优化，模型在提升生成样本质量和多样性的同时，增强了模型的收敛稳定性。

生成器负责将随机噪声或嵌入后的特定输入转换为高质量的流量序列。为提升序列建模能力，本文在生成器中采用嵌入层、双层BiGRU网络和全连接层结构。

在流量序列生成任务中，序列以特征值形式表示，每个特征值对应唯一整数索引。由于神经网络更擅长处理向量化数据，因此需将整数索引转换为词向量。嵌入层负责该转换，具体映射为

$$\boldsymbol{s}_t' = \mathrm{Embedding}\left(\boldsymbol{s}_t\right) \tag{10}$$

其中，$\boldsymbol{s}_t$表示时间步t的输入状态，$\boldsymbol{s}_t'$为其对应的嵌入向量表示。

生成器采用双层BiGRU结构，其核心优势在于该双向信息处理机制能够全面捕获5G网络流量序列的时序上下文依赖。相较于仅能利用历史信息的单向模型，BiGRU通过前向与后向并行处理序列，使模型在生成任意时间步的特征时，能够整合过去与未来的全部上下文信息。因此，模型得以精准地学习并复现攻击流量中复杂的时序模式与协议状态转移的内在逻辑，从而确保了生成样本不仅在统计分布层面逼近真实数据，更在协议行为的逻辑连贯性与语义真实性上达到高度保真。第一层BiGRU计算式为

$$\overrightarrow{\boldsymbol{h}_t^{(1)}} = \mathrm{GRU}\left(\boldsymbol{s}_t', \overrightarrow{\boldsymbol{h}_{t-1}^{(1)}}\right) \tag{11}$$

$$\overleftarrow{\boldsymbol{h}_t^{(1)}} = \mathrm{GRU}\left(\boldsymbol{s}_t', \overleftarrow{\boldsymbol{h}_{t+1}^{(1)}}\right) \tag{12}$$

其中，$\overrightarrow{\boldsymbol{h}_t^{(1)}}$和$\overleftarrow{\boldsymbol{h}_t^{(1)}}$分别表示向量的前向和后向隐藏状态。

第二层BiGRU计算式为

$$\overrightarrow{\boldsymbol{h}_t^{(2)}} = \mathrm{GRU}\left(\overrightarrow{\boldsymbol{h}_t^{(1)}}, \overrightarrow{\boldsymbol{h}_{t-1}^{(2)}}\right) \tag{13}$$

$$\overleftarrow{\boldsymbol{h}_t^{(2)}} = \mathrm{GRU}\left(\overleftarrow{\boldsymbol{h}_t^{(1)}}, \overleftarrow{\boldsymbol{h}_{t+1}^{(2)}}\right) \tag{14}$$

其中，$\overrightarrow{\boldsymbol{h}_t^{(2)}}$和$\overleftarrow{\boldsymbol{h}_t^{(2)}}$分别表示第二层BiGRU的前向和后向隐藏状态；$\overrightarrow{\boldsymbol{h}_{t-1}^{(2)}}$表示该层上一时间步的前向隐藏状态；$\overleftarrow{\boldsymbol{h}_{t+1}^{(2)}}$表示该层的下一时间步的后向隐藏状态。在第二层中，将第一层的BiGRU的输出作为输入，继续进行BiGRU的计算。

每一层BiGRU由两个GRU单元堆叠而成，分别计算前向和后向序列状态。具体而言，每个单元包含更新门、重置门和隐藏状态的计算。对于输入的向量$\boldsymbol{s}_t'$，具体的门控循环单元的计算式为

$$\boldsymbol{z}_t = \delta\left(\boldsymbol{W}_z \cdot \left[\boldsymbol{s}_t', \overrightarrow{\boldsymbol{h}_{t-1}^{(1)}}\right] + \boldsymbol{b}_z\right) \tag{15}$$

$$\boldsymbol{r}_t = \delta\left(\boldsymbol{W}_r \cdot \left[\boldsymbol{s}_t', \overrightarrow{\boldsymbol{h}_{t-1}^{(1)}}\right] + \boldsymbol{b}_r\right) \tag{16}$$

$$\tilde{\boldsymbol{h}}_t = \tanh\left(\boldsymbol{W} \cdot \left[\boldsymbol{s}_t', \boldsymbol{r}_t \odot \overrightarrow{\boldsymbol{h}_{t-1}^{(1)}}\right] + \boldsymbol{b}\right) \tag{17}$$

$$\boldsymbol{h}_t = \left(1 - \boldsymbol{z}_t\right) \odot \overrightarrow{\boldsymbol{h}_{t-1}^{(1)}} + \boldsymbol{z}_t \odot \tilde{\boldsymbol{h}}_t \tag{18}$$

其中，$\boldsymbol{z}_t$表示更新门，$\boldsymbol{r}_t$表示重置门，$\tilde{\boldsymbol{h}}_t$表示更新隐藏层状态，$\boldsymbol{h}_t$表示当前GRU单元的隐藏状态；δ表示sigmoid函数；$\boldsymbol{W}_z$、$\boldsymbol{W}_r$、$\boldsymbol{W}$是权重参数；$\boldsymbol{b}_z$、$\boldsymbol{b}_r$、$\boldsymbol{b}$是偏置参数，$\odot$表示逐元素乘法。

全连接层用于将BiGRU输出映射至特定特征空间，学习复杂的非线性关系，并转换为最终输出分布，从而提升生成数据的质量与多样性。具体而言，首先拼接第一层与第二层BiGRU输出，形成当前隐藏状态$\boldsymbol{h}_t'$。

$$\boldsymbol{h}_t' = \left[\overrightarrow{\boldsymbol{h}_t^{(1)}}; \overleftarrow{\boldsymbol{h}_t^{(2)}}\right] \tag{19}$$

然后，将隐藏层状态输入输出层，计算当前时间步的预测概率。

$$\hat{\boldsymbol{y}}_t = \mathrm{softmax}\left(\boldsymbol{W}_o \boldsymbol{h}_t' + \boldsymbol{b}_o\right) \tag{20}$$

其中，$\boldsymbol{W}_o$为输出层权重矩阵，H为隐藏状态维度，F为输出维度，故$\boldsymbol{W}_o$维度为$F \times 2H$；$\boldsymbol{b}_o$为偏置向量；softmax函数将输出转换为概率分布。通过采样得到当前时间步的动作x_t。

判别器结构包括嵌入层、多层感知机、随机失活层和输出层。嵌入层处理为

$$\boldsymbol{X}' = \mathrm{Embedding}\left(X\right) \tag{21}$$

其中，X为输入流量序列，$\boldsymbol{X}'$为对应的嵌入向量，Embedding为嵌入层映射函数。

多层感知机堆叠多个隐藏层，提取输入的局部与全局特征，提升判别器区分生成样本与真实样本的能力。作为非线性模型，多层感知机（multilayer perceptron, MLP）能更好地拟合复杂数据。在WGAN框架下，流量序列数据复杂性较高，因此设计三层MLP计算嵌入向量$\boldsymbol{X}'$。

$$h^{(1)} = \mathrm{ReLU}\left(\boldsymbol{W}_{\mathrm{in}} \cdot \boldsymbol{X}' + \boldsymbol{b}_{\mathrm{in}}\right) \tag{22}$$

$$h^{(2)} = \mathrm{ReLU}\left(\boldsymbol{W}_{\mathrm{hid1}} \cdot h^{(1)} + \boldsymbol{b}_{\mathrm{hid1}}\right) \tag{23}$$

$$h^{(3)} = \mathrm{ReLU}\left(\boldsymbol{W}_{\mathrm{hid2}} \cdot h^{(2)} + \boldsymbol{b}_{\mathrm{hid2}}\right) \tag{24}$$

其中，$\boldsymbol{W}_{\mathrm{in}}$、$\boldsymbol{W}_{\mathrm{hid1}}$、$\boldsymbol{W}_{\mathrm{hid2}}$是各层的权重矩阵，$\boldsymbol{b}_{\mathrm{in}}$、$\boldsymbol{b}_{\mathrm{hid1}}$、$\boldsymbol{b}_{\mathrm{hid2}}$是各层的偏置向量。

为了提高模型的泛化能力，判别器引入Dropout机制，以缓解因参数过多导致的过拟合问题。在训练过程中，Dropout通过随机丢弃隐藏层神经元，以一定概率p使部分单元在前向传播与反向传播中失效，从而减少神经元间的共适应性，提高模型的鲁棒性。Dropout操作可表示为

$$h_{\mathrm{drop}} = h^{(3)} \odot \mathrm{mask},\ \mathrm{mask} \sim \mathrm{Bernoulli}(p) \tag{25}$$

$$\mathrm{Output} = \boldsymbol{W}_{\mathrm{out}} \cdot h_{\mathrm{drop}} + \boldsymbol{b}_{\mathrm{out}} \tag{26}$$

其中，$h^{(3)}$为MLP最后一层的输出，mask为与$h^{(3)}$形状相同的二元掩码向量，按Bernouli分布生成，每个元素以概率p取0，概率$1-p$取1；$\boldsymbol{W}_{\mathrm{out}}$和$\boldsymbol{b}_{\mathrm{out}}$分别为输出层的权重矩阵和偏置向量；Output为判别器输出，以衡量输入样本与真实数据分布的差异。

此外，强化学习策略在训练中通过奖励机制引导生成器学习更加逼真的样本结构。生成器接收判别器输出作为奖励信号，结合策略梯度方法更新生成策略，在对抗博弈中不断提升性能。在RL-WGAN中，生成器和判别器在迭代过程中的对抗目标为

$$\min_{G_\theta} \max_{Dw} L\left(G_\theta, D_w\right) = E_{X \sim p_{\mathrm{data}}}[D_w(x)] - E_{X \sim G_\theta}[D_w(x)] \tag{27}$$

其中，x是输入，G_θ表示参数为θ的生成器，D_w表示参数为w的判别器。

在判别器优化阶段，固定生成器G_θ的参数。判别器D_w通过最小化Wasserstein距离损失函数以更新其参数w。参数更新梯度$\nabla_w L(G_\theta, D_w)$具体为

$$\nabla_w L(G_\theta, D_w) = \nabla_w\left(E_{x \sim P_{\mathrm{data}}}[D_w(x)] - E_{x \sim G_\theta}[D_w(x)]\right) \tag{28}$$

在生成器优化阶段，固定判别器D_w的参数。生成器G_θ被视为强化学习中的智能体，其优化目标是最大化策略$\pi_\theta(x)$下的期望累积奖励$E\left[R_{\mathrm{total}}\right]$。为此，采用策略梯度方法更新生成器的参数$\theta$。其损失函数定义为$L_G = -E[R_{\mathrm{total}}]$，参数更新的梯度$\nabla_\theta L_G$可表示为

$$\nabla_\theta L_G = -E_{x \sim G_\theta}[R_{\mathrm{total}} \nabla_\theta \ln \pi_\theta(x)] \tag{29}$$

其中，R_{total}是结合了判别器反馈和协议约束的总奖励，$\pi_\theta(x)$是生成器G_θ生成序列x的策略。

基于上述强化学习方法及生成器与判别器模型，构建RL-WGAN奖励计算函数与训练算法。奖励计算函数结合协议奖励判定与判别器反馈，最终通过奖励折扣因子计算序列的综合奖励值。具体算法流程如算法1所示。

算法1 RL-WGAN训练与更新

输入 初始化参数的生成器G_θ、判别器D_w，奖励折扣因子α，专家样本数据集X

输出 更新参数后的G'_θ和D'_w

1) 计算判别奖励$Q_{D_w}^{G_\theta} = D_w\left(x_g\right)$
2) 计算协议奖励$R_P = C_1 \cdot C_2 \cdot C_3 \cdot C_4$
3) 计算总奖励$R = Q_{D_w}^{G_\theta} - \alpha \cdot R_P$
4) 更新参数G_θ、D_w
5) for episodes = 0, 1, ⋯, E do
6) for each discriminator_iteration do
7) 获取真实数据样本X_r;
8) 获取生成数据样本X_g;
9) 计算判别损失$L_W\left(D_w\left(X_r, X_g\right)\right)$;
10) 更新判别器参数D_w;
11) end for
12) for each generator_iteration do
13) for t = 0, 1, ⋯, T do
14) 获取当前状态s_t;
15) 生成网络流序列时间步t的动作$x_t = G_\theta\left(s_t\right)$;
16) 时间步序列拼接(x_g=x_1,x_2,⋯,x_t,MCSampling($x_{t,T}$,t));
17) 计算奖励$R = \mathrm{Reward}\left(x_g\right)$;
18) 更新状态$s_{t+1} = \left(x_1, x_2, \cdots, x_t\right)$;
19) end for
20) 计算回合总奖励R_{total};
21) 计算生成损失$L_G\left(R_{\mathrm{total}}\right)$;
22) 更新生成器参数G'_θ;
23) end for
24) end for

3　性能分析

3.1　参数设定

生成器参数配置如下：嵌入层维度为64，双层BiGRU隐藏层维度为128，全连接层维度为64，学习率为0.001，批量大小为64，优化器选用随机梯度下降（stochastic gradient descent, SGD）。判别器参数设置如下：嵌入层维度为64，多层感知机隐藏层维度为128，Dropout率为0.2，学习率为0.000 1，批量大小为64，优化器采用Adam。强化学习参数按每训练周期1次的频率更新。

为评估关键超参数对模型性能的影响并验证其配置的鲁棒性，本文进行了系统的敏感性分析。

首先，本文对奖励函数中的折扣因子α进行了敏感性分析，如表1所示，以确定其最优取值。折扣因子α用于平衡判别奖励与协议奖励。通过网格搜索方法，在[0.1, 0.9]内以0.2为步长选取了不同的α值，并使用标签一致性比率（label consistency ratio, LCR）作为综合评价指标。实验结果表明：当α值过小时，模型对协议规则的约束不足，导致生成样本的合规性下降；当α值过大时，模型过度关注协议规则，抑制了对真实数据复杂分布的学习，导致分布保真度降低。综合来看，当$\alpha = 0.5$时，模型在协议合规性与分布保真度之间达到最佳平衡，LCR最高。因此，本文后续实验均采用$\alpha = 0.5$作为默认配置。

表1　　加权因子α敏感性分析

α值	LCR（5G-NIDD DoS）	LCR（实验数据集DDoS）
0.1	0.852	0.861
0.3	0.883	0.889
0.5	**0.895**	**0.902**
0.7	0.874	0.885
0.9	0.841	0.853

其次，本文对其他关键超参数进行了分析，如图5和图6所示。实验结果表明：在对抗训练框架下，生成器与判别器的最优学习率分别为0.001和0.000 1，该差异为采用非对称学习率策略提供了经验证据。此外，对BiGRU隐藏层维度的测试验证了维度为128时模型性能达到最优，此配置在模型复杂度与过拟合风险之间取得了理想的权衡。因此，该系列分析为本文所选超参数组合的合理性与有效性提供了充分的实验支撑。

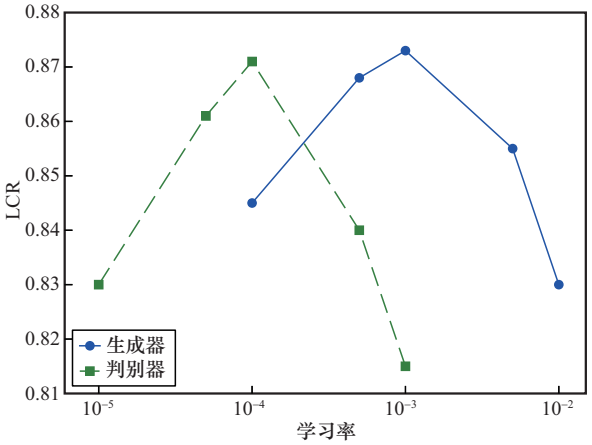


图5　生成器与判别器学习率敏感性

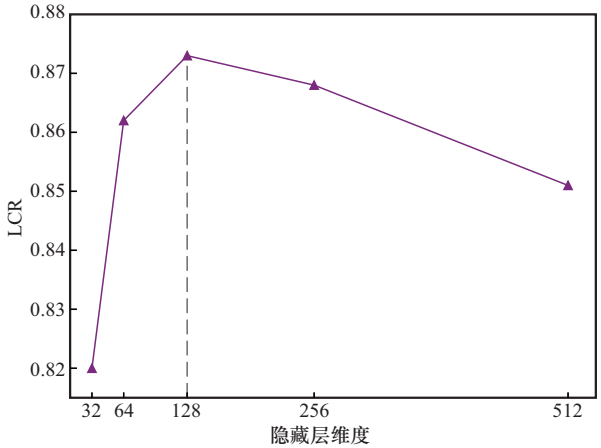


图6　BiGRU隐藏层维度敏感性

3.2　数据来源

本实验采用以下两类数据源。

1) 5G-NIDD数据集：该数据集采集自芬兰奥卢大学真实5G测试环境，包含DoS攻击及端口扫描两类攻击样本。

2) 实验数据集：通过XPRO Replay专业仪表模拟5G核心网环境，生成跨站脚本（cross-site scripting, XSS）攻击、代码注入攻击、规避性用户数据报协议（Evasive UDP）攻击、DDoS、结构化查询语言（structured query language, SQL）攻击、暴力破解6种常见攻击流量。

相较于5G-NIDD数据集，实验数据集涵盖了更广泛的网络威胁与更复杂的流量行为。选择这6种攻击类型是为了构建一个多层次、递进式的技术挑战基准，用以从不同维度系统性地评估RL-WGAN模型的生成能力。每个类别的攻击都对模型提出了独特的要求，具体如下。

DDoS 被用作检验模型复现流量统计分布能力的基准。此类攻击的核心在于其宏观的统计特征，如极高的包速率和连接数，这构成了对生成模型基础能力的考验，旨在验证其学习和再现极端数据分布的能力。

SQL、代码注入及 XSS 攻击则用于深度检验模型生成复杂载荷内容的能力。这 3 种攻击的恶意特征主要嵌入在数据包载荷的微观字节层面，而流量元数据可能与正常业务无异。这严格考验了模型在确保协议结构合法的前提下，精准生成包含特定恶意字符串的高维稀疏载荷数据的细节捕捉与生成能力。

暴力破解与 Evasive UDP 攻击被用于检验模型生成复杂行为与隐蔽模式的能力。暴力破解的异常体现在连续的时间行为上，考验着模型对时序依赖关系的建模能力。同时，Evasion UDP 攻击通过操纵协议字段来规避检测，这直接对模型中强化学习与协议奖励的机制进行了验证，旨在评估其引导生成器产出在协议层面具有欺骗性的复杂流量的有效性。

3.3 相似性分析

在生成模型性能评估中，传统机器学习评价指标因无法有效表征生成样本的分布特性而存在局限性。本文通过概率分布相似性量化方法，验证生成数据与真实数据在统计空间中的逼近程度。

在计算数据相似度时，首先将生成数据与真实数据视为两个概率分布，并采用核密度估计（kernel density estimation, KDE）方法量化其相似性。

使用式(30)来计算核密度估计值。

$$\hat{f}(x)=\frac{1}{nh}\sum_{i=1}^{n}K\left(\frac{x-X_i}{h}\right) \tag{30}$$

其中，$\hat{f}(x)$ 表示数据集在点 x 处的核密度估计，n 为数据集的样本数，X_i 为样本点，$K(\cdot)$ 为核函数，h 为数据集的带宽参数。高斯核函数的带宽可按式(31)进行调整。

$$h=\left(\frac{4\hat{\sigma}^5}{3n}\right)^{\frac{1}{5}}\approx 1.06\hat{\sigma}n^{-\frac{1}{5}} \tag{31}$$

本文使用随机森林对特征进行评分，选取了 22 个重要特征作为生成器输出的数据特征序列。这些特征包括：流持续时间（Flow Duration）、前向到达间隔最大值（Fwd IAT Max）、后向到达间隔最大值（Bwd IAT Max）、空闲时间标准差（Idle Std）、流字节速率（Flow Byte/s）、活跃时间最大值（Active Max）、包长度方差（Pkt Len Var）、前向头长度（Fwd Header Len）、前向最小段大小（Fwd Seg Size Min）、后向头长度（Bwd Header Len）、后向包总长度（TotLen Bwd Pkt）、流包速率（Flow Pkt/s）、目的 IP 地址（Destination IP）、源 IP 地址（Source IP）、目的端口（Destination Port）、源端口（Source Port）、协议（Protocol）、标签（Label），以及结束标志计数（FIN Count）、同步标志计数（SYN Count）、重置标志计数（RST Count）和确认标志计数（ACK Count）。为验证模型的优势，本节设计了与本文模型具有相同网络结构和参数的时间序列生成对抗网络（time-series generative adversarial network, TimeGAN）[26] 及带梯度惩罚的 Wasserstein 生成对抗网络（Wasserstein GAN with gradient penalty, WGAN-GP）[27] 模型进行对比实验。通过式(30)与式(31)对各模型生成的关键特征与真实数据进行相似性对比，得到相似性分布对比结果如图 7 和图 8 所示。

由图 7 可知，TimeGAN 在拟合时间动态特征方面展现出明显优势。例如在 Idle Std 和 Flow Duration 上，其生成分布与真实数据高度一致，充分说明其在学习连续时间规律上的能力。然而，对于 Destination IP 和 Fwd Seg Size Min 等非连续动态特征，TimeGAN 的表现则明显不足，无法准确复现其复杂结构。相比之下，WGAN-GP 作为改进型基准模型，在 Fwd Seg Size Min 等特征上实现了合理拟合，但其生成结果往往过于平滑，尤其在 Source Port 等多模态分布中未能捕捉到关键的峰谷特征。

与上述基准模型形成鲜明对比的是，RL-WGAN 在几乎所有被测特征上均表现最优。以 Source Port 和 Destination Port 为例，RL-WGAN 生成的分布曲线能够精准捕捉真实流量的复杂波动，明显优于 WGAN-GP 的平滑趋势和 WGAN 的偏差拟合。在 Destination IP 和 ACK Count 这类复杂或稀疏特征的拟合中，RL-WGAN 同样展现出稳定而卓越的效果，既保持了对整体分布趋势的把握，又能准确复现细节波动。更重要的是，在 TimeGAN 表现突出的 Idle Std 与 Flow Duration 等时间动态特征上，RL-WGAN 依旧展现出最高保真度。这一结果充分证明，通过将分布学习能力与协议感知的强化学习策略相结合，RL-WGAN 实现了全面而精准的泛化能力，能够忠实刻画 5G 网络流量的复杂内在规律。

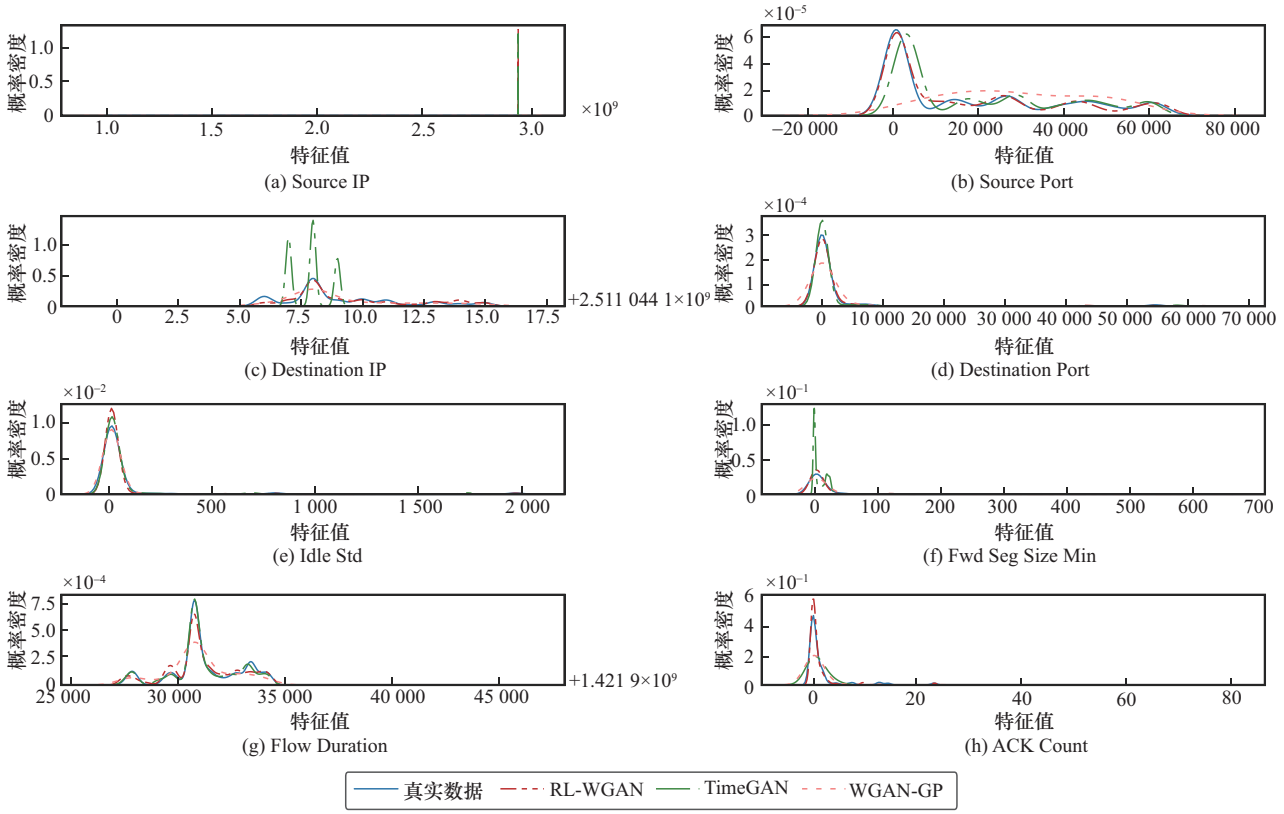


图7　生成数据与真实数据相似性对比（5G-NIDD数据集）

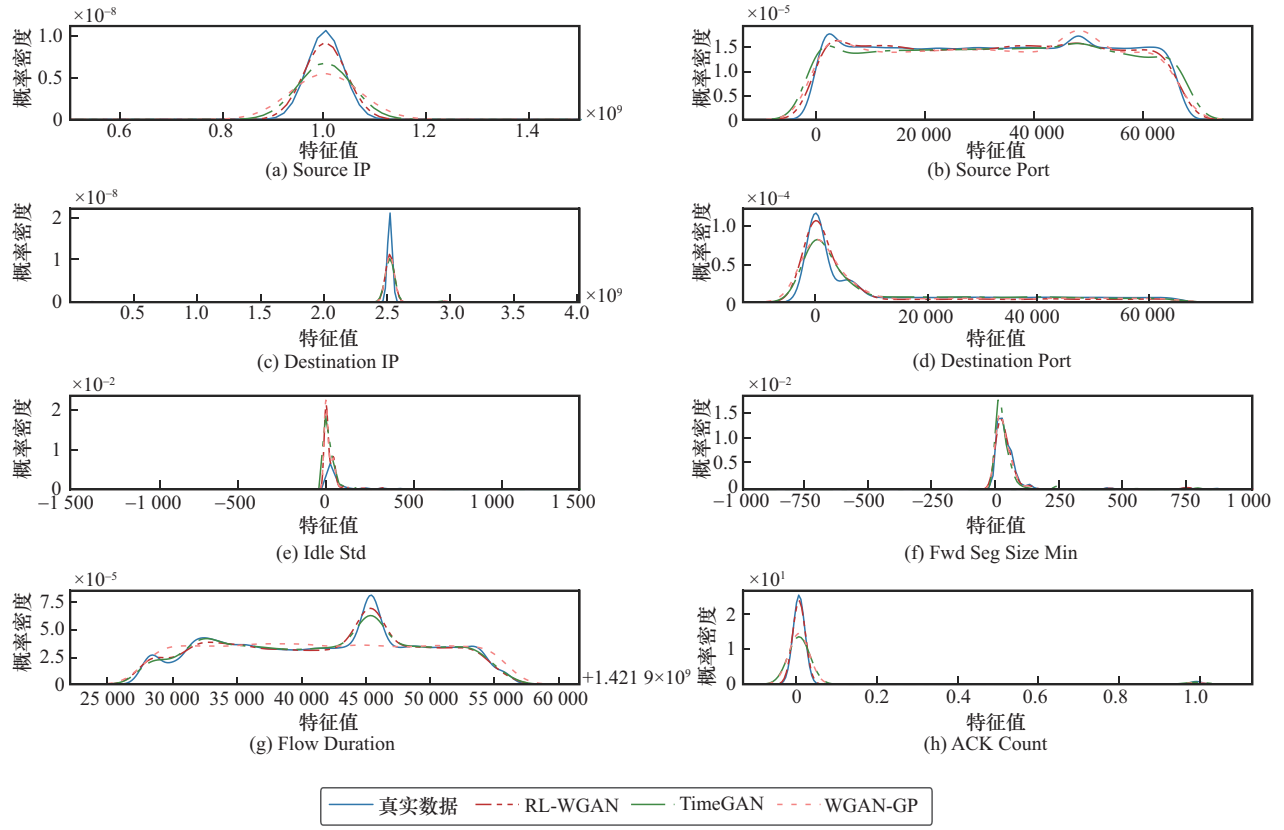


图8　生成数据与真实数据相似性对比（实验数据集）

图 8 表明，在更具挑战性的实验数据集上，由于其流量特征无法完全对应理想的攻击场景，所有模型的拟合精度均有所下降。然而，RL-WGAN 在这种非理想条件下，其生成结果仍然最大程度地贴合了真实数据的整体趋势，且实验结果明确优于所有对比模型，显示出其在复杂场景中的性能优势。

数据分布箱线图结果如图 9 和图 10 所示。从图 9 和图 10 可以清晰地看到，RL-WGAN 在精准捕捉真实数据核心统计特征上的全面优势，其生成数据在中位数、分布范围及四分位距等关键指标上均与真实数据高度一致，显示出极高的数据保真度与统计一致性。相比之下，TimeGAN 与 WGAN-GP 在多项统计特性上则暴露出各自的显著缺陷。TimeGAN 的主要问题在于其生成样本的多样性严重不足，这在 Destination Port 等特征上表现为被极度压缩的分布范围。WGAN-GP 则倾向于生成高度同质化的样本，导致其统计分布完全收窄，无法形成有效的箱体。此外，两种模型均伴随着大量不切实际的离群点和噪声样本的生成。

上述现象反映出，在缺乏针对性引导时，不同的生成模型架构会因其内在偏置而陷入困境，难以在维持生成样本多样性的同时避免统计形态的畸变。相较之下，RL-WGAN 通过引入基于强化学习的奖励机制，有效优化了特征空间的探索过程，实现了统计特性精准复现与分布结构鲁棒性的统一。

为量化评估生成分布与真实分布的差异，本文采用 Wasserstein 距离作为概率分布的相似性度量指标如表 2 和表 3 所示。该方法通过计算两分布间的最优传输代价，综合考虑概率分布的空间结构特征，避免了传统度量方法对分布形态的先验假设依赖。Wasserstein 距离越小，表明真实分布与生成分布越相似。实验结果显示生成样本与真实样本的 Wasserstein 距离稳定收敛于 0~0.8，验证了生成模型在复杂分布建模中的有效性。该结果从数学优化角度证实，模型能够以可接受的偏差阈值生成符合真实数据分布规律的样本。

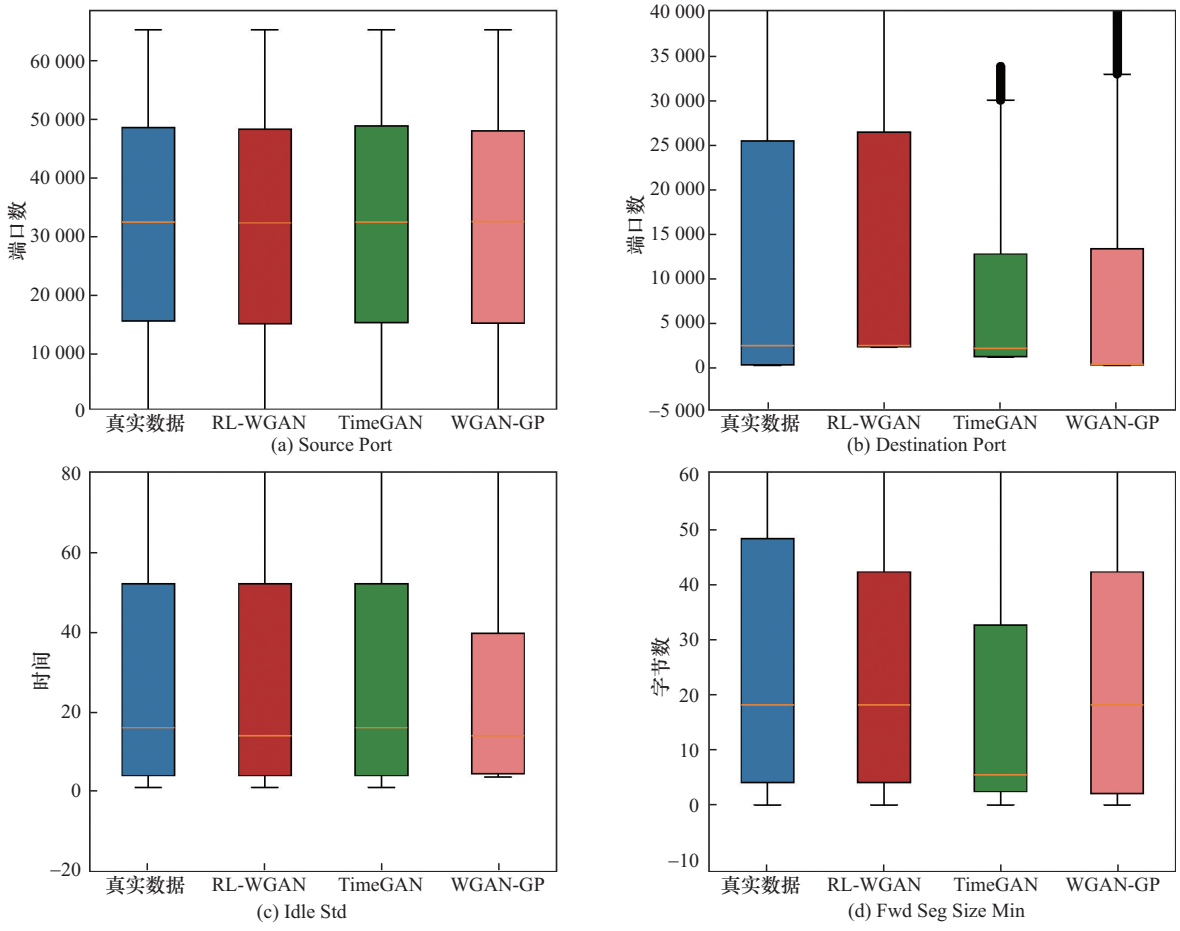


图 9　数据分布箱线图（5G-NIDD 数据集）

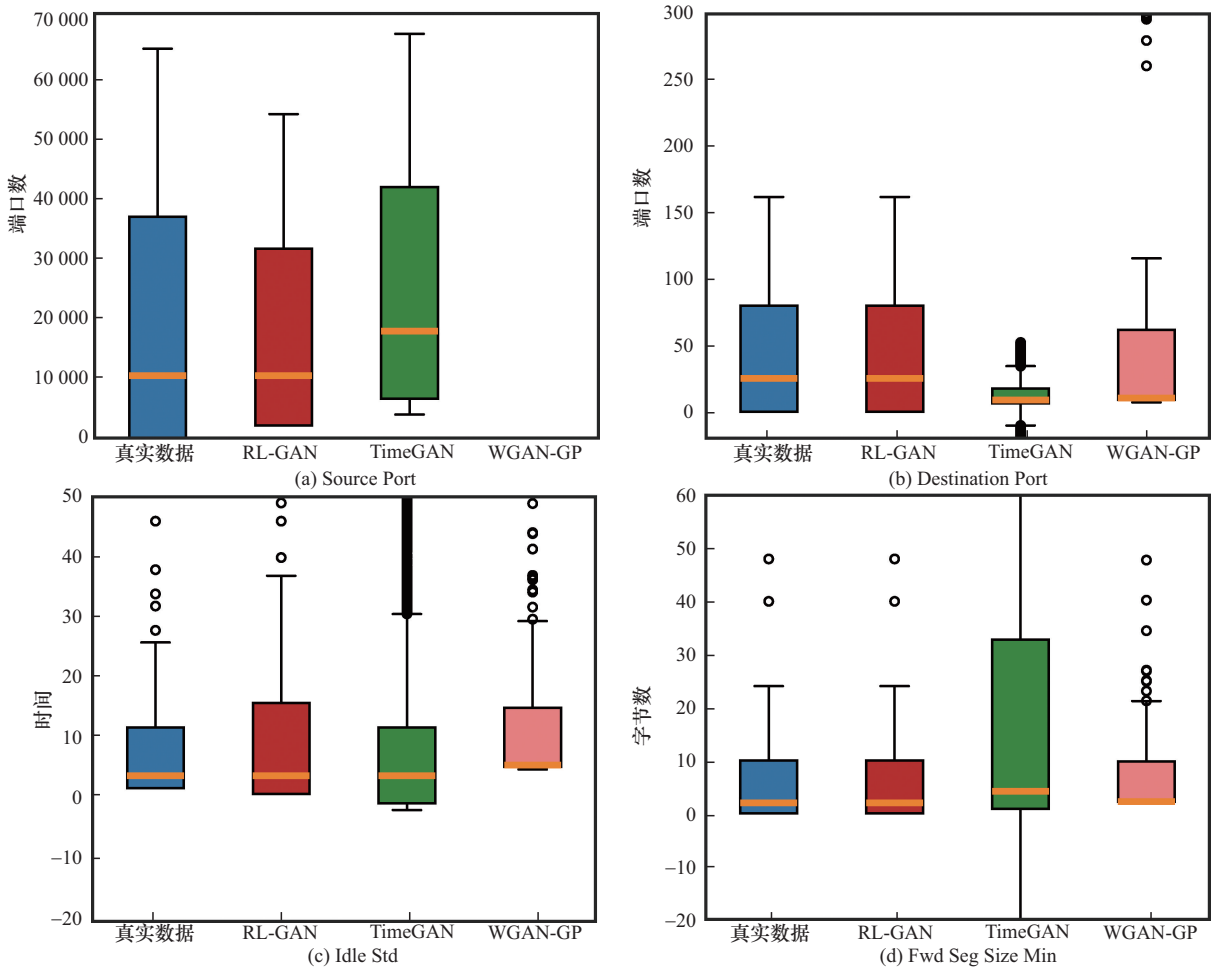


图 10　数据分布箱线图（实验数据集）

表 2　生成数据与真实数据的 Wasserstein 距离（5G-NIDD 数据集）

特征字段	距离	特征字段	距离
Flow Duration	0.176 9	Flow Byte/s	0.067 1
Fwd IAT Max	0.194 3	Active Max	0.378 5
Bwd IAT Max	0.159 6	Pkt Len Var	0.504 2
Idle Std	0.364 8	Fwd Header Len	0.120 3
Fwd Seg Size Min	0.064 9	Bwd Header Len	0.138 6
TotLen Bwd Pkt	0.217 8	Flow Pkt/s	0.349 5

表 3　生成数据与真实数据的 Wasserstein 距离（实验数据集）

特征字段	距离	特征字段	距离
Flow Duration	0.165 4	Flow Byte/s	0.721 9
Fwd IAT Max	0.129 5	Active Max	0.378 1
Bwd IAT Max	0.386 7	Pkt Len Var	0.647 2
Idle Std	0.659 1	Fwd Header Len	0.138 3
Fwd Seg Size Min	0.399 8	Bwd Header Len	0.375 7
TotLen Bwd Pkt	0.137 5	Flow Pkt/s	0.218 5

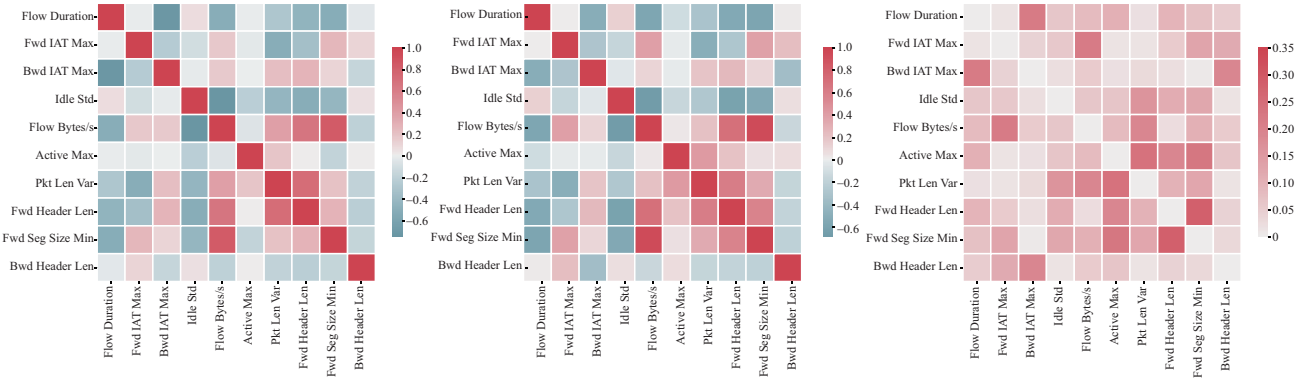
3.4　生成数据关系结构分析

除了验证数据样本分布的相似性，对于生成的样本而言，样本特征之间结构关系也需要与真实样本一致。通过皮尔逊（Pearson）相关系数的计算可以衡量变量之间的关系，并将关系强度量化表达。

$$r=\frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2}} \tag{32}$$

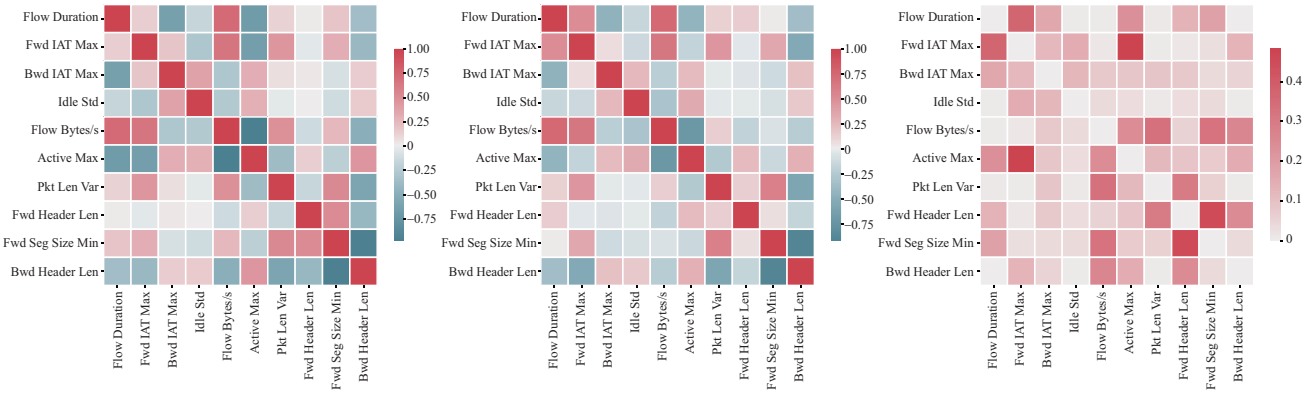
其中，x_i、y_i分别表示样本变量x、y的第i个值，$\bar{x}$、$\bar{y}$分别表示样本变量x、y的均值。

本文基于相关系数公式的量化分析结果，引入热力图对真实样本和生成样本的特征关联结构进行可视化对比，如图 11 和图 12 所示。实验结果表明，生成数据与真实数据在相关系数矩阵的空间分布模式上具有显著相似性。具体而言，对图 11(a)、图 11(b)与图 12(a)、图 12(b)的对比分析显示，两类数据在



(a) 真实样本特征关系　(b) 生成样本特征关系　(c) 关系结构误差

图 11　真实样本和生成样本特征关系结构热力图（5G-NIDD 数据集）



(a) 真实样本特征关系　(b) 生成样本特征关系　(c) 关系结构误差

图 12　真实样本和生成样本特征关系结构热力图（实验数据集）

特征间的相关性强度与方向性方面高度一致。为进一步量化该相似性，研究计算了相关系数的绝对误差，并通过热力图展示于图 11(c)与图 12(c)。误差分布分析表明，特征关联差异主要集中在 0~0.35，验证了生成模型在保持特征空间结构完整性方面的有效性。

进一步分析表明，热力系数分布图在整体结构和局部特征关联上均表现出高度一致性。多种特征在真实与生成数据中展现出相似的关联模式，进一步增强了两者在特征结构上的相似性认知。综上，热力图比较结果表明生成模型能够有效复现真实数据的特征关联结构。

3.5　模型效果分析

为评估生成数据与真实数据的分布一致性，本文构建基于随机森林的监督分类框架：首先使用原始测试集训练分类器实现类别特征空间的最大可分性；然后将生成样本输入分类器获取预测标签，通过计算预测与生成标签的 LCR 验证生成质量。

为进一步验证本文所采用的双层 BiGRU 生成器架构的优越性，基于此评估框架，将其与另外两种主流的序列生成架构——长短期记忆（long short-term memory, LSTM）网络和 Transformer 进行了横向对比。所有模型均在同等条件下训练，详细的 LCR 对比结果及评估方法的有效性验证如表 4 所示。由表 4 可以看出，在 5G-NIDD 与实验数据集场景下，分类器测试集准确率分别达到 94.1% 和 93.7%，证实了基于该分类器计算 LCR 这一评估策略的可靠性。

核心架构对比结果清晰地表明，采用 BiGRU 的 RL-WGAN 在所有攻击类型上均取得了最高的 LCR，显著优于采用 LSTM 和 Transformer 的变体。这主要得益于 BiGRU 的双向信息处理机制，它能够更全面地捕捉网络流量序列中复杂的时序上下文依赖关系，从而更精准地学习和复现攻击流量的内在逻辑。相比之下，单向的 LSTM 在处理需要前后文关联的复杂模式时能力受限，而 Trans-

former架构在处理此类高度结构化的流量特征序列时，并未展现出超越循环神经网络的性能优势。

表4　生成模型性能对比及评估有效性验证

数据集类型	攻击类型	BiGRU	LSTM	Transformer	分类器准确率
5G-NIDD数据集	DoS	0.895	0.881	0.875	94.1%
	端口扫描	0.851	0.839	0.842	
实验数据集	DDoS	0.902	0.889	0.884	93.7%
	暴力破解	0.881	0.873	0.865	
	Evasive UDP攻击	0.855	0.842	0.837	
	SQL注入	0.830	0.815	0.821	
	XSS	0.824	0.809	0.811	
	代码注入	0.819	0.803	0.807	

此外，从纵向来看，模型对不同攻击类型的复现能力存在差异。对于DDoS和暴力破解这类在流量统计维度上模式明确的攻击，模型易于捕捉和复现，故LCR最高；对于SQL、XSS等行为特征深藏于数据载荷、与正常业务流量边界模糊的攻击，生成挑战更大，LCR也相对更低。尽管存在这些细微差异，但本文模型在两个数据集的各类攻击上均展现出稳定且优异的生成性能，证明其具备良好的实际应用潜力。

本文目标是通过生成式数据增强解决流量样本的类别不均衡问题。如图13和图14所示，经RL-WGAN隧道方案处理后的数据集完成类别均衡化重构，各异常类型样本比例达到统计均衡状态，验证了生成模型在少样本学习场景下的工程适用性。

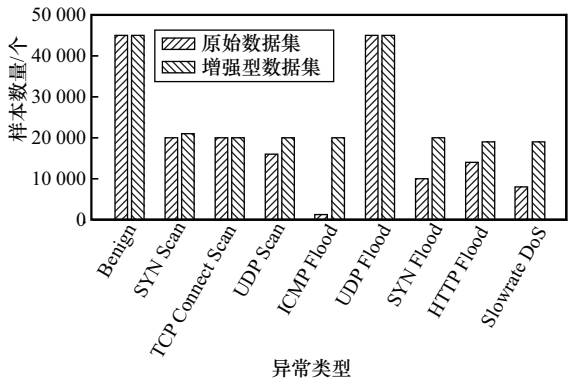


图13　生成样本数据集统计(5G-NIDD数据集)

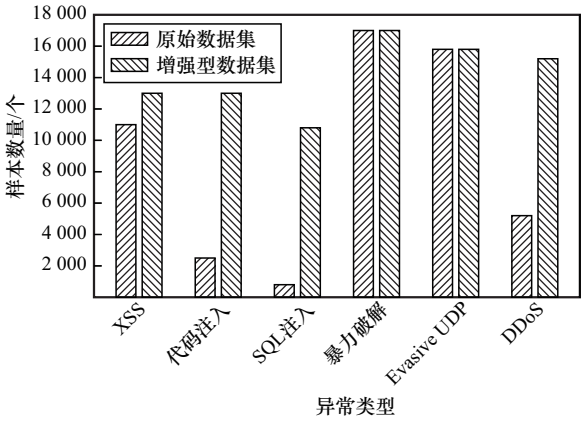


图14　生成样本数据集统计(实验数据集)

3.6　模型复杂度与部署可行性分析

RL-WGAN深度融合了生成对抗网络与强化学习，其训练阶段属于计算密集型任务。本文的全部实验在一台配置了Intel® Xeon® Silver 4210R CPU、256 GB内存及NVIDIA GeForce RTX 4090 GPU的服务器上完成。针对单一攻击类型，模型在5G-NIDD数据集上的完整训练时长约为7 h；在规模更大、流量行为更复杂的实验数据集上，训练时长则增至约11 h。

尽管训练成本较高，但这与模型的应用定位和部署模式相符。RL-WGAN的核心价值是作为一种数据增强工具，旨在解决5G NIDS因异常样本稀缺而面临的训练困境。为此，模型采用“离线训练，按需生成”的策略部署。在离线训练阶段，于后端服务器上利用真实流量对模型进行充分训练，以构建一个覆盖多种攻击类型的预训练模型库。此过程独立于实时网络业务，因此其计算与时间开销可控且合理。在按需生成阶段，当需要训练或更新NIDS时，可调用相应的预训练模型，高效地批量生成特定类型、高保真的异常流量样本。这些样本可直接用于扩充与均衡训练数据集，进而提升NIDS的检测性能与泛化能力。

4　结束语

本文提出了一种基于RL-WGAN的异常数据生成方法。针对5G通信网络用户面流量的特殊性，设计了适用于GTP-U协议的数据预处理机制。结合强化学习，构建了RL-WGAN模型，并设计了基于BiGRU的生成器和改进的DNN判别器。通过判别奖励和协议奖励的协同作用，优化了生成器的学习过程。本文模型能够有效学习流量序列的内在结

构，生成与真实数据分布高度相似的样本，从而有效扩充了当前紧缺的5G异常数据样本空间。

然而，本文模型仍存在一些待改进之处。首先，RL-WGAN模型的训练过程涉及复杂的对抗博弈和策略优化，计算开销较大，尤其是在处理大规模、高维度流量特征时对计算资源要求较高。其次，协议奖励函数的设计依赖于预定义的专家规则，这虽然保证了生成数据的合规性，但在面对利用未知协议漏洞的新型攻击时，其有效性可能受限。

因此，未来的研究将致力于探索更轻量化、可扩展的模型架构，并研究自适应的奖励机制以提升模型效率和对未知威胁的建模能力。

参考文献：

[1] Chen X, Chen Y F, Feng W, et al. Real-time DDoS defense in 5G-enabled IoT: a multidomain collaboration perspective[J]. IEEE Internet of Things Journal, 2023, 10(5): 4490-4505.

[2] 周光海, 宁兆龙, 陈志奎, 等. 基于核偏最小二乘法的物联网无线传感网络故障分析与研究[J]. 通信学报, 2017, 38(S2): 94-98.
Zhou G H, Ning Z L, Chen Z K, et al. Fault analysis and research of wireless sensor network based on kernel partial least squares[J]. Journal on Communications, 2017, 38(S2): 94-98.

[3] Cui Z H, Zhao Y R, Cao Y, et al. Malicious code detection under 5G HetNets based on a multi-objective RBM model[J]. IEEE Network, 2021, 35(2): 82-87.

[4] Chettri L, Bera R. A comprehensive survey on Internet of things (IoT) toward 5G wireless systems[J]. IEEE Internet of Things Journal, 2020, 7(1): 16-32.

[5] Li S X, Song L Y, Wu X Y, et al. Multi-class imbalance classification based on data distribution and adaptive weights[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(10): 5265-5279.

[6] Zhou L, Deng X F, Ning Z L, et al. When generative AI meets semantic communication: optimizing radio map construction and distribution in future mobile networks[J]. IEEE Network, 2025, 39(3): 47-55.

[7] Zhou L, Deng X F, Wang Z, et al. Semantic information extraction and multi-agent communication optimization based on generative pre-trained transformer[J]. IEEE Transactions on Cognitive Communications and Networking, 2025, 11(2): 725-737.

[8] Dablain D, Krawczyk B, Chawla N V. DeepSMOTE: fusing deep learning and SMOTE for imbalanced data[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(9): 6390-6404.

[9] Zhou Q Y, Feng Z Y, Gu Q Q, et al. Context-aware mixup for domain adaptive semantic segmentation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(2): 804-817.

[10] Li C, Yao W, Wang H D, et al. Bayesian evolutionary optimization for crafting high-quality adversarial examples with limited query budget[J]. Applied Soft Computing, 2023, 142: 110370.

[11] 王小洁, 刘子依, 唐守泽, 等. 面向支付通道网络的多优先级资源调度算法[J]. 通信学报, 2025, 46(2): 83-96.
Wang X J, Liu Z Y, Tang S Z, et al. Multi-priority resource scheduling algorithm for payment channel networks[J]. Journal on Communications, 2025, 46(2): 83-96.

[12] 朱晓荣, 张佩佩. 基于GAN的异构无线网络故障检测与诊断算法[J]. 通信学报, 2020, 41(8): 110-119.
Zhu X R, Zhang P P. Fault detection and diagnosis method for heterogeneous wireless network based on GAN[J]. Journal on Communications, 2020, 41(8): 110-119.

[13] Zhang G H, Sikdar B. Synthetic time-series data generation for smart grids using 3D autoencoder GAN[J]. IEEE Transactions on Industrial Informatics, 2025, 21(7): 5047-5058.

[14] Liang X Y, Wang Z H, Wang H. Synthetic data generation for residential load patterns via recurrent GAN and ensemble method[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 2535412.

[15] Tan S S, Zhong X X, Tian Z Y, et al. Sneaking through security: mutating live network traffic to evade learning-based NIDS[J]. IEEE Transactions on Network and Service Management, 2022, 19(3): 2295-2308.

[16] Prakash C D, Karam L J. It GAN do better: GAN-based detection of objects on images with varying quality[J]. IEEE Transactions on Image Processing, 2021, 30: 9220-9230.

[17] She R, Fan P Y. From MIM-based GAN to anomaly detection: event probability influence on generative adversarial networks[J]. IEEE Internet of Things Journal, 2022, 9(19): 18589-18606.

[18] Cui T Y, Gou G P, Xiong G, et al. 6GAN: IPv6 multi-pattern target generation via generative adversarial nets with reinforcement learning[C]// Proceedings of the IEEE INFOCOM 2021 - IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2021: 1-10.

[19] 宁兆龙, 张凯源, 王小洁, 等. 基于多智能体元强化学习的车联网协同服务缓存和计算卸载[J]. 通信学报, 2021, 42(6): 118-130.
Ning Z L, Zhang K Y, Wang X J, et al. Cooperative service caching and peer offloading in Internet of vehicles based on multi-agent meta-reinforcement learning[J]. Journal on Communications, 2021, 42(6): 118-130.

[20] Shi J C, Zhou G Q, Bao S D, et al. MultiselfGAN: a self-guiding neural architecture search method for generative adversarial networks with multicontrollers[J]. IEEE Transactions on Cognitive and Developmental Systems, 2023, 15(2): 544-554.

[21] Hui S D, Wang H D, Li T, et al. Large-scale urban cellular traffic generation via knowledge-enhanced GANs with multi-periodic patterns[C]// Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2023: 4195-4206.

[22] Kasgari A T Z, Saad W, Mozaffari M, et al. Experienced deep reinforcement learning with generative adversarial networks (GANs) for model-free ultra reliable low latency communication[J]. IEEE Transactions on Communications, 2021, 69(2): 884-899.

[23] 刘涛涛, 付钰, 王坤, 等. 基于VAE-CWGAN和特征统计重要性融合

的网络入侵检测方法[J]. 通信学报, 2024, 45(2): 54-67.
Liu T T, Fu Y, Wang K, et al. Network intrusion detection method based on VAE-CWGAN and fusion of statistical importance of feature[J]. Journal on Communications, 2024, 45(2): 54-67.

[24] 段雪源, 付钰, 王坤. 基于VAE-WGAN的多维时间序列异常检测方法[J]. 通信学报, 2022, 43(3): 1-13.
Duan X Y, Fu Y, Wang K. Multi-dimensional time series anomaly detection method based on VAE-WGAN[J]. Journal on Communications, 2022, 43(3): 1-13.

[25] Wang X J, Li T F, Xiong X R, et al. Federation chain for data privacy protection in industrial Internet of Things: the perspective from 5G core networks[J]. IEEE Internet of Things Journal, 2025, 12(19): 39260-39271.

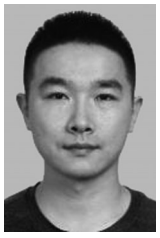
[26] Zhang Y F, Zhou Z H, Liu J W, et al. Data augmentation for improving heating load prediction of heating substation based on TimeGAN[J]. Energy, 2022, 260: 124919.

[27] Zhu G Y, Zhou K, Lu L, et al. Partial discharge data augmentation based on improved Wasserstein generative adversarial network with gradient penalty[J]. IEEE Transactions on Industrial Informatics, 2023, 19(5): 6565-6575.

[作者简介]

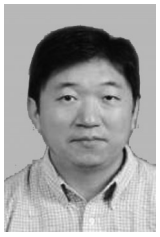
宁兆龙（1986– ），男，辽宁沈阳人，博士，重庆邮电大学教授，主要研究方向为边缘智能、应急通信网络、网络资源优化。

邹道远（2001– ），男，江苏徐州人，重庆邮电大学硕士生，主要研究方向为5G通信安全。



周力（1988– ），男，湖北汉川人，博士，国防科技大学副研究员，主要研究方向为智能通信网络、无人集群组网、语义通信。

欧阳瑞崎（1998– ），男，四川绵阳人，重庆邮电大学硕士生，主要研究方向为5G通信安全。



熊炫睿（1976– ），男，四川德阳人，博士，重庆邮电大学副教授，主要研究方向为人工智能应用、网络入侵检测、车联网网络安全、5G/6G网络安全。