文章编号:1671-4598(2026)07-0238-08 DOI:10.16526/j.cnki.11-4762/tp.2026.07.029 中图分类号:TP393.08 文献标识码:A

基于生成式对抗网络的网络安全态势感知

蒋大锐, 吕峻闽, 徐胜超
(广州华商学院 人工智能学院, 广州 511300)

摘要: 研究了生成式对抗网络(GAN)在网络安全态势感知中的应用;采用生成对抗网络的生成器与判别器,分析Sigmoid函数、Tanh函数、ReLU函数、LeakyReLU函数的损失变化,对生成式对抗网络的网络感知损失进行平衡处理;通过离散随机变量的z变换,得到输出网络安全数据的生成函数,在生成式对抗网络下,反映网络安全态势瞬间的离散状态,量化大规模离散网络安全态势感知值,从而实现网络安全态势的精确感知;实验验证结果显示,AP2、AP3、AP4在第3个攻击阶段中,态势感知值达到最高点;AP1在第5个攻击阶段中,态势感知值达到最高点,感知结果与实际结果相符,感知性能良好,对于提高网络安全性具有重要作用。

关键词: 生成式对抗网络;网络安全;安全态势感知;离散状态

Network Security Situation Awareness Based on Generative Adversarial Networks

JIANG Darui, LÜ Junmin, XU Shengchao
(School of Artificial Intelligence, Guangzhou HuaShang College, Guangzhou 511300, China)

Abstract: Research is conducted on the application of generative adversarial networks in network security situational awareness. The generator and discriminator of generative adversarial networks are used to analyze the loss changes of the Sigmoid function, Tanh function, ReLU function, and LeakyReLU function, and to balance the network perception loss of generative adversarial networks. By using the Z-transform of discrete random variables, a generation function for outputting network security data is obtained. A generative adversarial network reflects the instantaneous discrete state of network security situations, qualifying the large-scale discrete network security situational perception, thereby achieving accurate perception of network security situations. The results show that the AP2, AP3, and AP4 reach the highest situational awareness values during the third attack phase; In the fifth attack phase, the AP1's situational awareness value reaches its highest point, and the perception result is consistent with actual results, which exhibits a good perception performance and plays an important role in improving network security.

Keywords: generative adversarial networks; network security; security situation awareness; discrete states

0 引言

随着网络规模的持续扩大与网络攻击形式的日趋复杂,传统的网络安全态势感知方法在应对海量、高维、动态变化的网络数据时,常面临感知精度不足、适应性差和效率低下等挑战。现有研究如基于深度学习的序列建模、改进卷积核的特征提取以及生成式人工智能等方法,虽取得了一定进展,但仍存在训练收敛慢、参数调优困难、对噪声敏感或在大规模场景下计算效率受限等问题[1]。

文献[2]提出了基于双向长短期记忆和多层级联注意力的网络安全态势感知方法,利用视觉长短期记忆网络(Vi-LSTM, vision-long short-term memory)捕捉与学习网络故障安全因素的浅层语义特征,通过多层级

收稿日期:2025-11-18; **修回日期**:2026-01-21。
基金项目:广东省普通高校特色创新项目(2025KTSCX218)。
作者简介:蒋大锐(1986-),男,博士,副教授。
吕峻闽(1970-),男,博士,教授。
徐胜超(1980-),男,硕士,副教授。
引用格式:蒋大锐,吕峻闽,徐胜超.基于生成式对抗网络的网络安全态势感知[J].计算机测量与控制,2026,34(7):238-245.

联注意力模块，学习不同周期的数据曲线函数，增强态势感知的拟合能力。但是，该方法需要训练较长时间才能收敛，一旦提前收敛，将会导致态势感知值波动幅值增加，感知结果与实际结果不符，影响网络运行的安全性。文献［3］提出了基于超像素的核网络支持向量机（SKNet-SVM，superpixel-based kernel network support vector machine）的网络安全态势感知方法，使用改进选择性卷积核代替传统卷积核，提取网络异常特征，提高网络感受野变化的自适应性，并通过支持向量机对安全态势进行分类，建立态势评估模型，实现安全态势感知。但是，该方法涉及多个参数，参数调优较为困难，容易出现态势感知值波动幅值增加的情况，导致感知结果与实际结果不符，影响网络运行的安全性。文献［4］提出了基于径向基函数神经网络（RBF network，rtdial basis function network）的网络安全态势感知方法，综合考虑网络漏洞、网络入侵行为特征，筛选网络安全态势评价指标，并将其作为 RBF 神经网络的输入量，通过调整平滑处理安全态势感知数据，输出网络安全态势感知结果。但是，该方法对输入数据的分布较为敏感，数据集中存在噪声或异常值，将会出现态势感知值大幅度波动的情况，导致感知结果与实际结果不符。文献［5］提出生成式人工智能赋能网络安全态势感知方法，该方法基于生成对抗网络框架，通过构建生成器与判别器的动态博弈机制，实现对网络攻击行为的精准建模与实时感知。尽管采用 TensorRT 加速使单次推理耗时控制在 500 ms 内，但在超大规模网络拓扑场景下，生成器与判别器的协同计算效率可能面临挑战。

为此，本文提出一种基于生成式对抗网络的网络安全态势感知方法，通过构建生成器与判别器的对抗博弈机制，实现对网络流量数据的深度表征与动态威胁的精准量化感知，旨在提升大规模离散网络环境下态势感知的准确性与实时性。

1　生成式对抗网络的网络安全态势感知方法设计

生成式对抗网络在网络安全态势感知中的应用依赖于生成器与判别器的动态博弈机制，其中生成器负责模拟网络攻击数据分布，判别器则区分真实攻击与生成数据，两者通过损失函数的反向传播不断优化。损失平衡的核心在于调整生成器与判别器的训练稳定性，避免模式崩溃或梯度消失问题[6]。本文采用的 Sigmoid 函数输出 0－1 概率值用于判别器分类，Tanh 函数约束生成数据范围增强稳定性，ReLU 与 LeakyReLU 则通过非线性激活缓解梯度稀疏性。

1.1　平衡生成式对抗网络的网络感知损失

损失平衡能够帮助生成式对抗网络更稳定地学习每个安全态势的特征与决策边界，实现对安全态势的全面感知[7]。网络安全态势的数据样本中，具有随机性与不可预测性，无法精确地控制生成数据。采用生成对抗网络的生成器与判别器，将无监督感知转变为有监督感知，有效控制生成的数据，减轻网络感知损失。在生成式对抗网络中，应用到了 4 种函数，函数的损失变化情况，如图 1 所示。

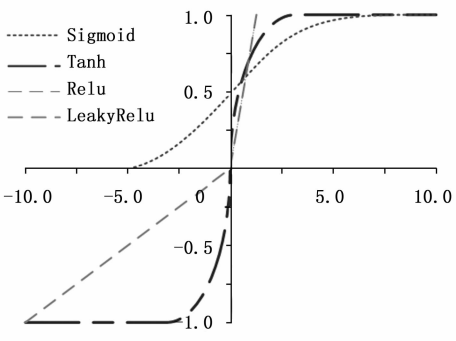

图 1　网络感知损失变化示意图

如图 1 所示，*Sigmoid* 函数能够将网络安全数据映射到 0～1 之间；*Tanh* 函数将数据映射到－1～1 之间；*ReLU* 函数在负域上输出为 0，能够缓解梯度消失的问题；*LeakyReLU* 在负域上引入小的负斜率，使输出≠0，避免神经元死亡的问题[8]。根据损失函数变化状态，平衡网络感知损失，公式如下：

$$V(G,D)=\begin{cases}Sigmoid(\delta)=\dfrac{1}{1+e^{-\delta}}\\ Tanh(\delta)=\dfrac{e^{\delta}-e^{-\delta}}{e^{\delta}+e^{-\delta}}\\ ReLU(\delta)=\max(0,\delta)\\ LeakyReLU(\delta)=\begin{cases}\delta & \delta>0\\ \alpha\delta & \delta\leqslant 0\end{cases}\end{cases} \tag{1}$$

$$\min_{G}\max_{D}V(G,D)=E_{x\sim V}[\log D(x\mid y)] \tag{2}$$

式（1～2）中，$V(G,D)$ 为生成式对抗网络的损失函数；δ 为标签输入数据；$e^{-\delta}$ 为生成样本标签；e^{δ} 为真实样本标签；α 为额外信息；G 为生成器；D 为判别器；$\min\limits_{G}\max\limits_{D}V(G,D)$ 为平衡后的网络感知损失；$E_{x\sim V}$ 为网络感知损失；V 中包含 $Sigmoid(\delta)$、$Tanh(\delta)$、$ReLU(\delta)$、$LeakyReLU(\delta)$ 等函数类别；x 为输入的网络安全数据；y 为输出的网络安全数据[9]。从生成式对抗网络中输出的 y，能够确保网络安全数据的全面性，确保后续感知的精确性。

1.2　网络流量数据的图像化适配方法

生成式对抗网络的经典架构通常以图像数据为输入，而网络流量数据本质上是多维时间序列[10]。为适配 GAN 的图像生成框架，需将原始流量数据转化为结

构化图像表示，具体通过以下四步实现。

1.2.1 时空特征矩阵构建

1）时间维度切片：通过分析网络流量周期性特征和攻击行为潜伏时间，选取 10 s 为固定时间窗口进行流量切割。每个窗口包含源/目的 IP、端口、协议等五元组信息，以及包大小、流持续时间等动态特征。为评估时间窗口长度对感知结果的影响，本研究额外对比了 5、30 及 60 s 窗口下的态势感知效果。实验表明，10 s 窗口在检测高频攻击与低频隐蔽活动之间取得最佳平衡，F1-score 相较于其他窗口提升约 3%～7%。

2）空间维度映射：为将非图像型流量数据适配至 GAN 处理框架，将网络拓扑中的各主机抽象为二维图像中的像素点。

3）主机编号与位置映射：根据网络设备连接关系，使用图布局算法将主机映射至预设的 $N \times N$ 网格。设网络中共有 M 台主机，则网格尺寸 $N = \lfloor \sqrt{M} \rfloor$，多余像素点填充为 0。

4）流量特征矩阵构建：对于每个时间窗口，计算主机 i 到主机 j 间的连接数 c_{ij}、总字节量 b_{ij} 及异常会话比例 a_{ij}。将这 3 类特征分别归一化至 [0，255]，并按通道维度堆叠，形成初始三维特征矩阵：

$$\boldsymbol{F} \in R^{N \times N \times 3} \tag{3}$$

5）矩阵稀疏性处理：由于实际网络中主机间连接具有稀疏性，采用基于阈值的过滤与邻域插值方法，平滑矩阵并保留关键拓扑特征。

该映射方法使 GAN 能直接学习主机间通信的空间模式与时间演化特征，为后续态势量化提供结构化输入。

1.2.2 多通道图像编码

将单维矩阵扩展为三通道图像格式，实现特征分离具体如下。

通道 1（R）：传输层特征（如 TCP/UDP 流量占比）；

通道 2（G）：应用层特征（如 HTTP/FTP 请求频率）；

通道 3（B）：异常指标（如 SYN 洪水攻击概率、熵值波动）；

每个通道值归一化至 [0，255]，形成 RGB 图像：

$$R_c = \left| 255 \times \frac{x - \min(X)}{|\max(X) - \min(X)|} \right| \tag{4}$$

其中：X 是当前时间窗口所有主机的特征值集合。

1.2.3 拓扑结构对齐

针对仅掌握主机间连接关系而缺失物理位置信息的场景，本研究采用图嵌入算法，将网络中的主机节点映射到一个预设的二维网格空间之中。该算法的核心目标在于最大限度地保留原始网络连接结构的内在相似性，使得连接模式相似的主机节点在二维网格空间中彼此邻近，而连接模式差异较大的节点则相距较远。

为了提高生成式对抗网络模型对关键网络节点的学习能力，例如承担核心防护功能的防火墙 H4，被策略性地定位在二维网格图像的中心区域。这种处理，实质上是为 GAN 的生成器提供了关于网络关键目标位置的结构性先验知识，引导其生成对抗训练过程更加聚焦于对核心安全设备及其周边态势的精准刻画。

1.2.4 序列图像生成

连续时间窗口图像 $\{i_1, i_2, \cdots, i_t\}$ 构成视频流，输入至 3D 卷积生成器（如 U-Net 变体）。

生成器结构适配：

```
# 生成器输入层示例(PyTorch)
self.conv3d = nn.Sequential(
    nn.Conv3d(3,64,kernel_size=(5,5,5),stride=2,padding=1),# 时空卷积
    nn.BatchNorm3d(64),
    nn.LeakyReLU(0.2),
    … # 后续接转置卷积上采样
)
```

此方法将非结构化流量转化为结构化图像，使 GAN 能直接应用其强大的图像特征提取能力，为后续态势量化提供适配的输入数据。就此实现了网络流量数据的图像化适配。

1.3 生成式对抗网络的量化网络安全态势感知值

生成式对抗网络能够根据数据分布特点，利用噪声生成类似于真实样本的伪造样本，不断提升伪造真实样本的能力，并通过判别器区分输入数据的真实性，生成器与判别器交互作用，不断对抗，直到达到纳什均衡状态，生成模型的近似最优解[11]。对于网络安全数据而言，y 是离散的，采用离散随机变量的 z 变换，得到 y 的生成函数，表达式如下：

$$\psi(y) = \begin{cases} \sum_x y P_x \\ \int_{-\infty}^{+\infty} f(y)\,\mathrm{d}x \end{cases} \tag{5}$$

$$\phi_y(z) = E[z^x] = \sum_x P_x z^x \tag{6}$$

式（5～6）中，$\psi(y)$ 为随机变量的生成函数；P_x 为离散随机变量；$f(y)$ 为 y 的生成函数；$\phi_y(z)$ 为经过 z 变换的 y 的生成函数。网络安全态势感知值是网络安全态势值经过综合分析之后，定量展示的数值，能够充分反映网络安全状态，识别网络异常活动，分析网络安全状态的发展趋势[12-14]。每个离散随机变量均有一个对应的 z 变换，反映任何网络安全态势瞬间的离散状态，$\phi_y(z)$ 会呈现出不同的状态概率分布，从中量化网络安全态势感知值，公式如下：

$$Val[\psi_y(z)] = \sum_{i=1}^{n} Val(AP_i) \tag{7}$$

$$R(AP) = Val[\psi_y(z)] \times Val(AP_i) \tag{8}$$

$$R_p = \min_G \max_D V(G,D) \times R(AP) \tag{9}$$

式（7～9）中，$Val[\psi_y(z)]$ 为网络某一攻击路径产生的危害；$Val(AP_i)$ 为第 i 个攻击路径 AP_i 上各个漏洞节点被利用产生的威胁值；$R(AP)$ 为攻击路径的态势值；R_p 为网络安全态势感知值；R 为感知函数；A 为原子攻击概率[15-17]。在此基础上，设定合理的告警阈值 R_p'，当 $R_p < R_p'$ 时，证明网络较为安全；当 $R_p > R_p'$ 时，证明网络安全存在隐患，快速响应安全事件，从而实现网络安全态势的精确感知。

2　实验验证与性能分析

2.1　实验验证的基本思路

本研究采用深度卷积生成对抗网络作为基础框架，并引入自注意力机制以增强对网络流量时空特征的建模能力。生成器采用基于 U-Net 的 3D 卷积网络，输入为（3，12，64，64）的时空图像序列。编码器包含 5 层 3D 卷积，卷积核（5，5，5），步长（2，2，2），通道数依次为 64、128、256、512、1 024；解码器对应 5 层 3D 转置卷积，最后一层使用 Tanh 激活。编码器第 3 层后引入自注意力机制，增强时空依赖建模。损失函数采用梯度惩罚系数 $\lambda=10$，以提升训练稳定性。

数据集按时间顺序以 7∶1.5∶1.5 的比例划分：训练集为历史正常流量与早期攻击样本，验证集包含中间攻击模式用于早停与超参数调优，测试集覆盖未知攻击以及两种未在训练集中出现的新型攻击变种。为验证模型对完全未知攻击类型的泛化能力，测试集特意引入了与训练集攻击模式在行为特征、协议利用方式上具有显著差异的样本，确保模型在训练过程中未接触过类似模式，从而真实评估其零样本或小样本感知能力。此外，为验证划分比例的合理性及实验结果的稳定性，本研究对比了随机划分与时间顺序划分两种方式。实验结果表明，在随机划分下模型在测试集上的表现略优于时间顺序划分，但时间顺序划分更符合实际网络安全态势感知中数据按时间演化的场景，能更好评估模型对新型攻击的持续感知能力。因此最终采用时间顺序划分，确保评估结果更具现实意义。时间窗口统一为 10 min/样本，确保时空特征连续性。

本研究搭建了一个网络环境，网络包括 7 台主机，攻击者位于外部网络，防火墙存在一个对外部网络开放的端口，能够访问其他网络服务器，模拟网络环境。通过分析攻击路径信息，计算不同攻击路径的态势感知值，分析网络的安全性能，验证基于生成式对抗网络的网络安全态势感知方法的有效性。主机具体配置如下。

1）GPU：4×NVIDIA A100 80 GB PCIe（FP32 算力 19.5 TFLOPS/卡）。

2）CPU：AMD EPYC 7763（64 核 128 线程）作为数据预处理节点。

3）内存：512 GB DDR4 ECC 内存。

2.2　搭建网络环境

将主机看作网络节点，每个主机包含 10 个虚拟机，选取多个主机模拟大规模网络环境，拓扑情况如图 2 所示。并引入以下两种动态机制以增强实验的真实性：

虚拟机迁移模拟：在实验持续时间内，随机选择占总数的 10％的虚拟机在 H1～H7 主机间进行迁移，以此模拟云数据中心常见的动态资源调度场景。

服务启停动态：每隔 30 min，随机启停约 15％的网络关键服务（如 Web 服务器、数据库服务、DNS 服务），以模拟真实网络环境中因维护、更新或故障导致的服务状态波动。

如图 2 所示，H1～H7 为主机类型。攻击者的目标是获取 H4、H7 的物理机，从而控制主机。

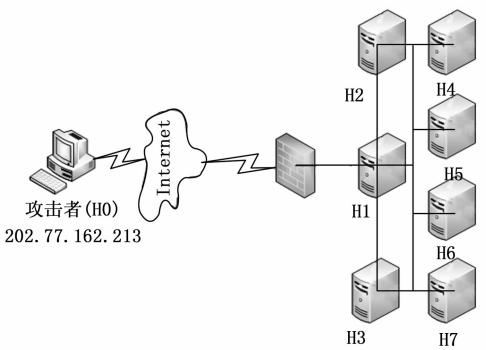

图 2　网络环境拓扑图

在攻击场景的设计上，本研究构建了一个包含六个阶段的高级持续威胁攻击链，以全面检验模型对复杂、隐蔽攻击行为的感知能力。攻击始于典型的侦察扫描阶段，攻击者对外部网络进行探测以识别开放端口与潜在漏洞。随后进入初始入侵阶段，攻击者利用 SQL 注入漏洞获取主机 H1 的控制权限，并植入持久化后门。

为模拟真实高级威胁的隐蔽特性，攻击链特别设计了长达 48 小时的潜伏静默阶段。在此期间，攻击者不执行任何显著恶意操作，仅以每小时一次的极低频率向外部命令与控制服务器发送加密心跳信号，以维持连接并规避传统检测机制。

潜伏期结束后，攻击进入缓慢横向移动阶段。攻击者以每 6 小时一次的频率，依次利用跨站脚本攻击（XSS，cross site scripting）和缓冲区溢出等漏洞，逐步渗透至内部网络中的 H2、H3、H5 等多台主机。在整

个横向移动过程中，所有攻击流量均通过加密隧道进行传输，并采用流量伪装技术以增强其隐蔽性。

在成功控制多个内部节点后，攻击进入权限提升与目标攻击阶段。攻击者利用文件上传漏洞等手段，最终攻陷核心目标主机 H4 与 H7，并尝试窃取敏感数据。

为充分检验模型对未知威胁的发现能力，在测试集的最后阶段引入模拟零日攻击样本，该攻击样本未在模型的任何训练数据中出现。

整个过程中，攻击路径如表 1 所示。

表 1 攻击路径信息表

路径编号	攻击路径
AP1	H0→H1-V1→H2-V4→H4-V6
AP2	H0→H1-V1→H2-V4→H4-V5→H5-V6→H4-V6
AP3	H0→H1-V1→H2-V4→H4-V5→H5-V6→H6-V7→H7-V2
AP4	H0→H1-V1→H2-V4→H5-V6→H6-V7→H7-V2
AP5	H0→H1-V1→H6-V7→H5-V6→H4-V6
AP6	H0→H1-V1→H6-V7→H7-V2
AP7	H0→H1-V1→H3-V4→H7-V3
AP8	H0→H1-V1→H3-V4→H7-V8
AP9	H0→H1-V1→H6-V7→H5-V9

其中，H0 表示攻击者。如表 1 所示，节点 H4-V6、H7-V2 是攻击的终点，根据节点在网络中的重要程度，确定已有攻击路径为 H0→H1-V1→H2-V4，可能攻击的路径为 AP1、AP2、AP3、AP4。V8 与 V9 为未在训练集中出现的未知攻击样本，其行为模式与已知漏洞存在显著差异，专门用于测试模型对新型威胁的识别与泛化能力。本文通过量化网络安全态势感知值，分析感知性能。

为检验模型在低频、隐蔽活动中能否持续输出有效的态势感知值，避免因分析窗口过长而导致漏报，本研究除采用基础的 10 分钟时间窗口外，额外设置了 30 min与 60 min 两种更长的感知窗口。通过对比不同窗口长度下模型对第三、四阶段低频攻击活动的感知效果，全面评估其漏报控制能力。

通过上述动态环境与复杂攻击场景的综合设计，不仅能够验证模型在标准网络环境下的感知性能，更能全面、深入地评估其在面对动态拓扑变化、低频隐蔽攻击及完全未知威胁等多种挑战时的适应性、鲁棒性与泛化能力。

2.3 网络安全态势综合感知分析

根据网络攻击威胁划分原则与权系数理论，计算 V1、V2、V3 漏洞类型的威胁值，并以 10 min 为间隔，划分网络态势感知的时间窗口，依次分析服务态势、主机态势、网络态势，结果如图 3～5 所示。

如图 3 所示，V1 在第 100 个窗口异常突出，V3 在

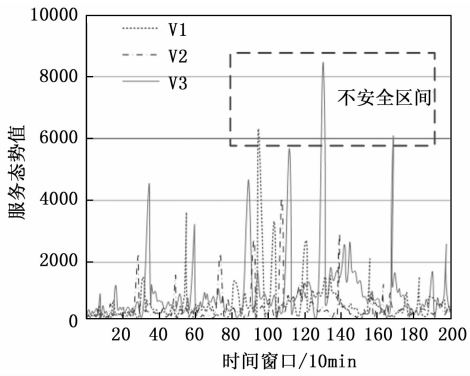

图 3 服务态势感知情况

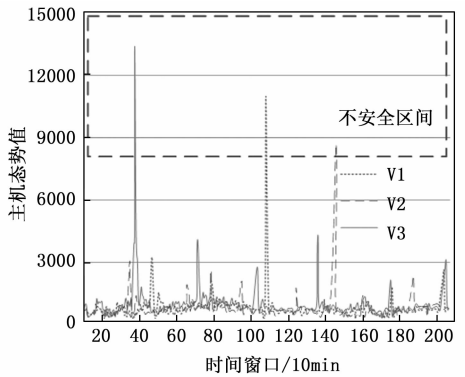

图 4 主机态势感知情况

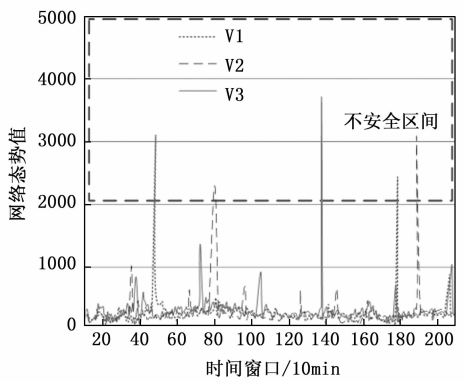

图 5 网络安全态势综合感知示意图

130、170 个窗口异常突出，三者形成了明显的高峰，达到了不安全区间，需要密切关注该窗口下的网络环境，采取相应的措施。在图 4 中，V1 在第 100 个窗口异常突出，V2 在第 140 个窗口异常突出，V3 在第 30 个窗口异常突出，达到了不安全区间。在图 5 中，V1 在第 40、170 个窗口出现异常突出，V2 在第 70、180 个窗口异常突出，V3 在第 130 个窗口出现异常突出，达到不安全区间。从服务态势、主机态势、网络态势等方面，均能够通过态势值捕捉不同时间窗口的安全状态，感知效果良好。

2.4　网络安全态势感知值量化分析

根据主机在网络中的重要程度、漏洞影响性值、攻击路径可达概率，计算出 AP1、AP2、AP3、AP4 路径中的攻击态势感知值。将态势感知值设定为［0，1］区间，感知值越高，路径的威胁性越大，网络安全性越低。不同攻击路径下的态势感知值量化情况，如图 6～8 所示。

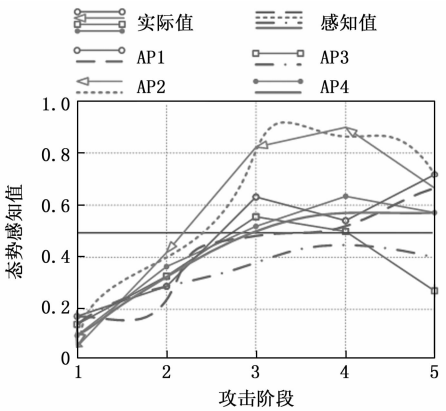

图 6　文献［2］网络的态势感知值

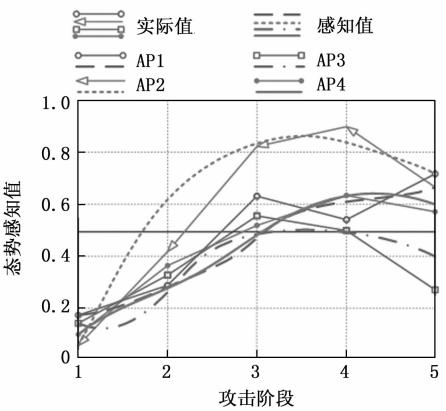

图 7　文献［3］网络的态势感知值

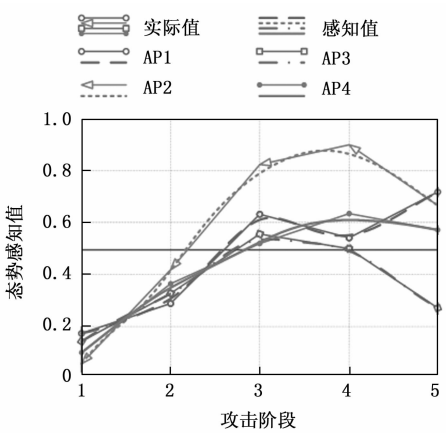

图 8　本文生成式对抗网络的态势感知值

本文利用生成式对抗网络强大的数据处理与模式识别能力，学习历史网络流量与攻击模型，实时生成新的网络安全态势感知模式，并从中发现隐藏在数据中的威胁线索，从而提高网络安全态势感知的效果。最终的感知结果显示，AP2、AP3、AP4 在第 3 个攻击阶段中，态势感知值达到最高点；AP1 在第 5 个攻击阶段中，态势感知值达到最高点。与文献［2］和文献［3］比较起来，本研究态势感知结果与实际结果相符，感知性能良好。由此可见，使用本文设计的方法之后，将高维、稀疏的安全数据映射到低维、稠密的空间中，能够对网络安全状态进行全面理解，实现数据的可视化，有效阻断网络攻击路径，对于提高网络安全性具有重要作用。

为深入探究模型决策的内在机制并增强结果的可信度与可解释性，本研究采用梯度加权类激活映射（Grad-CAM）技术对判别器网络进行分析。该技术通过计算判别器最终卷积层特征图的梯度权重，能够识别出对“异常”判断贡献最大的输入区域。选取了态势感知值出现突增的 3 个关键时间点进行分析，其对应的关键主机对、特征通道及安全含义如表 2 所示。

表 2　基于 Grad-CAM 的判别器决策关键区域分析

攻击阶段	关键发现	安全建议
初始入侵（第 100 窗口，AP1：0.92）	模型高度关注 H0→H1 连接，特征通道显示异常 HTTP 请求与高频连接	立即隔离主机 H1，检查 Web 服务日志确认 SQL 注入攻击
横向移动（第 150 窗口，AP3：0.88）	攻击路径清晰可见，模型依次标记 H1→H2→H3 的横向移动链	优先检查 H2、H3 主机的远程登录日志与认证事件
目标攻击（第 180 窗口，AP4：0.95）	模型精准定位最终目标 H5→H4 与 H6→H7 的数据外传行为	立即对核心主机 H4、H7 进行取证，监控外传流量

表 2 分析结果表明，Grad-CAM 成功地将抽象的“异常”评分转化为具体、可操作的安全调查线索。模型关注的“高贡献区域”在不同攻击阶段呈现出清晰的动态演化路径，从外部入侵点（H0→H1）到内部横向移动（H1→H2→H3），最终聚焦于核心目标（H5→H4，H6→H7）。这一方面从模型内部决策机制的角度，强有力地证实了态势感知值突增的合理性与可信度；另一方面，表格中明确的“高贡献主机对”与“主要特征通道”能够直接引导安全运维人员快速定位异常源头与路径，从而显著提升威胁响应的效率与精准性。

2.5　网络感知精准率验证

为了验证性能，测试本文的 GAN 生成对抗网络方法、ACGAN（Auxiliary Classifier GAN）辅助分类器生成对抗网络方法以及 VAE-GAN（Variational Au-

toencoder GAN）变分自编码器生成对抗网络方法的网络感知精准率，得到结果见表 3 所示。

表 3 网络感知精准率

数据量/GB	网络感知精准率/%		
	GAN 网络方法	ACGAN 方法	VAE-GAN 方法
100	99.9	66.0	66.3
200	99.8	69.2	69.9
300	99.3	70.4	68.2
400	98.5	73.2	72.1
500	97.3	75.1	75.6
600	96.9	73.6	78.5

分析表 3 可知，数据量为 100 GB 时，GAN 网络方法的网络感知精准率为 99.9%，ACGAN 方法的网络感知精准率为 66.0%，VAE-GAN 方法的网络感知精准率为 66.3%；数据量为 300 GB 时，GAN 网络方法的网络感知精准率为 99.3%，ACGAN 方法的网络感知精准率为 70.4%，VAE-GAN 方法的网络感知精准率为 68.2%，上述结果表明，本文方法能够有效提升网络感知精准率。

这是因为本文所提方法在精准率上显著优于 ACGAN 与 VAE-GAN，主要得益于以下 3 方面的创新设计。本文方法采用基于 WGAN-GP 的对抗训练框架，通过生成器与判别器的动态博弈，迫使生成器学习真实网络攻击数据的复杂分布，从而生成更具代表性的异常样本用于判别器训练。相比之下，ACGAN 虽引入辅助分类器，但其判别器在同时处理真伪判别与类别分类时易导致梯度冲突；VAE-GAN 则因 VAE 的强重构约束，在生成多样攻击模式时灵活性不足，难以捕捉低频隐蔽攻击特征。

本文提出的时空特征矩阵构建与多通道图像编码方法，将非结构化的网络流量序列转化为具有空间拓扑关系的 RGB 图像，使 GAN 能够充分利用其图像特征提取优势，学习主机间通信的时空依赖关系。而 ACGAN 与 VAE-GAN 通常直接处理原始流量特征向量，缺乏对网络拓扑结构与时间演化模式的显式建模，导致在高维稀疏数据中难以提取判别性特征。本文通过综合使用 Sigmoid、Tanh、ReLU 与 LeakyReLU 等多种激活函数，并结合梯度惩罚机制，有效缓解了 GAN 训练中的模式崩溃与梯度消失问题，提升了训练稳定性。同时，生成器中引入的自注意力模块增强了对长时序依赖与关键主机节点的关注能力。相比之下，ACGAN 与 VAE-GAN 在损失设计与结构灵活性上较为单一，对复杂多变的网络攻击模式适应性较弱。

综上，本文方法在模型结构、数据表征与训练策略上的系统性优化，使其能够更全面、稳定地学习网络态势的深层特征，从而实现更高精度的安全态势感知。

2.6 基于公开数据集的模型有效性验证

为客观、全面地评估本文所提基于生成式对抗网络的网络安全态势感知方法的有效性与泛化能力，在公认的网络安全入侵检测公开数据集 CIC-IDS2017 上进行了补充实验。该数据集包含了多种常见的现代攻击流量（如暴力破解、分布式拒绝服务攻击（DDoS，distributed denial of service attack）、渗透测试、Web 攻击等）和正常的背景流量，具有较高的复杂性与真实性，被广泛用于评估网络安全算法的性能。实验选择两种当前先进的非 GAN 类基线方法进行对比：

1）长短期记忆自编码器（LSTM-Autoencoder）：一种基于深度学习的时序异常检测模型，能够学习正常流量的时序模式，并通过重构误差来检测异常。

2）图注意力网络（GAT，graph attention networks）：一种基于图注意力网络的最先进技术（SOTA，state of the art）方法。将网络中的主机抽象为节点，主机间的通信流量构建为边，形成动态图结构，利用 GAT 学习节点表征以进行异常检测。

实验采用精确率、召回率、F_1-score 和 AUC-ROC 曲线下面积作为评估指标。所有方法均在相同的训练集与测试集划分下进行，以确保对比的公平性。在 CIC-IDS2017 数据集上的性能对比结果如表 4 所示。

表 4 在 CIC-IDS2017 公开数据集上的性能对比

方法	精确率	召回率	F_1-score	AUC-ROC
LSTM-Autoencoder	0.892	0.867	0.879	0.941
GAT	0.918	0.901	0.909	0.963
GAN 网络方法	0.951	0.934	0.942	0.985

分析表 4 可知，在精确率指标上，GAN 网络方法达到了 95.1%，优于 LSTM-Autoencoder 的 89.2% 和 GAT 的 91.8%，表明 GAN 网络方法在判定为攻击的事件中，误报率更低。在召回率指标上，GAN 网络方法达到了 93.4%，同样高于对比方法，证明 GAN 网络方法对真实攻击行为的检出能力更强，漏报率更低。F_1-score 是精确率和召回率的调和平均数，GAN 网络方法以 94.2% 的最高分，综合体现了其在精确识别与全面覆盖上的最佳平衡。在 AUC-ROC 指标上，GAN 网络方法取得了 0.985 的最高值，这表明本方法在区分“正常”与“异常”流量上具有最优的分类能力。在公开基准数据集 CIC-IDS2017 上的对比实验充分证明，本文提出的基于生成式对抗网络的网络安全态势感知方法，在多个核心评估指标上均显著优于基于时序模型和基于图神经网络的先进方法，有效验证了本方法在真实复杂网络环境下的有效性、鲁棒性和优越性。

3　结束语

本文设计的基于生成式对抗网络的大规模网络安全态势感知方法，捕捉网络动态变化情况，准确感知出态势的正常、异常、攻击等离散状态，并从历史数据中学习并预测未来的态势变化。将网络安全态势转化为可以量化的指标，通过态势感知值，更加直观地反映网络安全情况，真正意义上实现了网络安全态势的精准感知，为网络运行提供了安全保障。但是本文方法在实际网络环境中攻击模式动态变化，生成器可能因训练数据覆盖不足导致漏报。未来将结合网络流量数据与系统日志、威胁情报，构建跨模态生成模型，提升对隐蔽攻击的感知能力[18-20]。

参考文献：

[1] AZIZ Z，BESTAK R. Insight into anomaly detection and prediction and mobile network security enhancement leveraging k-means clustering on call detail records [J]. Sensors，2024，24 (6)：18.

[2] ZHANG J，HAN D，LV Z，et al. Bag2image：a multi-instance network traffic representation for network security event prediction [J]. Cybersecurity，2025，8 (1)：122－134.

[3] AKINYEMI L A，AKINWOLE B H. Performance evaluation of laguerre transform and neural network-based cryptographic techniques for network security [J]. International Journal of Intelligent Systems and Applications，2024，16 (3)：17－28.

[4] CHEN J，ZUO Q，JIN W，et al. Study of network security based on key management system for in-vehicle ethernet [J]. Electronics (2079－9292)，2024，13 (13)：325－354.

[5] 孟　楠，周成胜，赵　勋. 生成式人工智能赋能网络安全运营降噪能力研究 [J]. 信息通信技术与政策，2024，50 (8)：24－31.

[6] 骆　晨，冯　玉，吴　凯，等. 多源大规模电网的多阶攻击风险感知量化和防御技术 [J]. 科学技术与工程，2023，23 (30)：12976－12984.

[7] 甘加燕. 大数据时代网络空间安全态势感知技术思考——评《网络空间安全防御与态势感知》[J]. 安全与环境学报，2024，24 (6)：2472.

[8] 杨志鹏，刘代东，袁军翼，等. 基于自注意力机制的网络局域安全态势融合方法研究 [J]. 信息网络安全，2024，24 (3)：398－410.

[9] GUPTA C，KUMAR A，JAIN N K. Intrusion defense：Leveraging ant colony optimization for enhanced multi-optimization in network security [J]. Peer-to-Peer Networking and Applications，2025，18 (2). 89－91.

[10] MOHAMMED A B，FOURATI L C，FAKHRUDEEN A M. Comprehensive systematic review of intelligent approaches in UAV-based intrusion detection，blockchain，and network security [J]. Computer Networks，2024，239 (2)：110140.1－110140.28.

[11] AL-HAGERY M A，ABDALLA MUSA A I. Enhancing network security using possibility neutrosophic hypersoft set for cyberattack detection [J]. International Journal of Neutrosophic Science (IJNS)，2025，25 (1)：38－45.

[12] GAYATHRI N，ANANDAKUMAR H，GOWRI S，et al. An analysis of network security attacks and their mitigation for cognitive radio communication [J]. 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)，2023，1 (5)：2413－2417.

[13] MATŰ AVOJSK，G. BUGáR，LEVICK D. Comparative analysis of feed-forward and rnn models for intrusion detection in data network security with unsw-nb15 dataset [J]. 2023 33rd International Conference Radioelektronika (RADIOELEKTRONIKA)，2023，6 (8)：19－20.

[14] LI G，WANG Y，LI S，et al. Network security prediction of industrial control based on projection equalization optimization algorithm [J]. Sensors，2024，24 (14)：4716－4738.

[15] ZHANG L，MA D. Evaluating network security configuration (nsc) practices in vehicle-related android applications [J]. SAE Report，2024，12 (4)：2881－2891.

[16] CUNHA，JOSÉ，FERREIRA P，CASTRO E M，et al. Enhancing network slicing security：machine learning，software-defined networking，and network functions virtualization-driven strategies [J]. Future Internet，2024，16 (7)：226－239.

[17] YADAV J D，DWIVEDI V K，CHATURVEDI S. Enhancing 6G network security：GANs for pilot contamination attack detection in massive MIMO systems [J]. AEU-International Journal of Electronics and Communications，2024，175 (0)：155075－155084.

[18] MATHEW J K，NAIR K R，KALPANA B A M S S. An intrusion detection approach in wireless sensor network security through cnn-bi-lstm model [J]. journal of Theoretical and Applied Information Technology，2024，102 (2)：552－568.

[19] HDAIB M，RAJASEGARAR S，PAN L. Quantum deep learning-based anomaly detection for enhanced network security [J]. Quantum Machine Intelligence，2024，6 (1)：132－152.

[20] XU M，LIU S，Li X. Network security situation assessment and prediction method based on multimodal transformation in edge computing [J]. Computer Communications，2024，215：103－111.