研究与开发

基于时间卷积网络的无监督入侵检测模型

廖金菊[1]，丁嘉伟[1]，冯光辉[2]
（1. 郑州工业应用技术学院信息工程学院，河南 郑州 451150；
2. 广州大学计算机科学与网络工程学院，广东 广州 510006）

摘　要：现有的多数入侵检测模型通过长短期记忆（long short-term memory，LSTM）网络评估数据之间的时间依赖性。然而，LSTM网络处理序列数据增加了训练模型的计算复杂度和存储成本。为此，提出了基于多头注意力机制和时间卷积网络的无监督入侵检测模型（unsupervised intrusion detection model based on multi-head attention mechanism or temporal convolutional network，UDMT）。UDMT不依赖于LSTM网络，它利用时间卷积网络和多头注意力机制构建生成对抗网络的生成器和决策器，实现计算的并行化，进而降低复杂度。同时，UDMT不依赖于标签的攻击数据，它具有检测已知攻击和未知攻击的能力。此外，UDMT采用不同的隐藏层模式，配置灵活，以满足不同的检测率和检测时延的要求。相比于两个同类的检测模型，提出的UDMT能获取更高的检测率和更低的检测时延。

关键词：入侵检测模型；长短期记忆网络；生成对抗网络；多头注意力机制；时间卷积网络
中图分类号：TN393
文献标志码：A
doi: 10.11959/j.issn.1000−0801.2025001

Unsupervised intrusion detection model based on temporal convolutional network

LIAO Jinju[1], DING Jiawei[1], FENG Guanghui[2]
1. School of Information Engineering, Zhengzhou University of Industrial Technology, Zhengzhou 451150, China
2. School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou 510006, China

Abstract: Most existing intrusion detection models rely on long short-term memory (LSTM) networks to consider time-dependencies among data. However, LSTM's sequential data processing significantly increases computational complexity and memory consumption during training. Therefore, unsupervised intrusion detection model based on multi-head attention mechanism and temporal convolutional network (UDMT) was proposed. UDMT didn't rely on LSTM networks. Instead, it used temporal convolutional network and multi-head attention mechanism in the generative adversarial network generator and discriminator networks to enable more computation parallelization, and re-

收稿日期：2024−07−06；**修回日期：**2024−11−20
基金项目：教育部产学合作协同育人项目（No.220602236285739）；广东省自然科学基金面上项目（No.2022A1515011386）
Foundation Items: The Industry-University Cooperative Education Project of Ministry of Education (No.220602236285739), The Natural Science Foundation of Guangdong Province (No.2022A1515011386)

duced computational complexity. Moreover, UDMT was capable of detecting both known and zero-day attacks without relying on labeled attack data. In addition, UDMT can adopt different privacy layer modes, and the configuration was flexible to meet the requirements of different detection rates and detection delays. Experiment results show that the proposed UDMT has higher detection rate and lower detection latency than two state-of-the-art intrusion detection models.

Key words: intrusion detection model, long short-term memory network, generative adversarial network, multi-head attention mechanism, temporal convolutional network

0　引言

物联网（Internet of things，IoT）是信息通信技术领域的重要研究方向和发展趋势，其技术的发展不仅提升了生产效率[1-2]，还便捷了人们的日常生活。目前，物联网已在诸多领域得到广泛应用，如智慧农业、智能家居、智慧工厂等。

然而，无线通信的广播特性，导致IoT系统易遭受网络攻击，这些攻击可能会损害网络的完整性，使节点出现感测错误数据或者被恶意控制。尽管传统互联网中已有多种解决方案，但是因物联网中物理层的限制以及物联网内部节点的异构性，这些解决方案并不适用于物联网。例如，不同的接入技术、应用和需求所带来的异构性，增加了物联网被攻击的可能性[3]。此外，节点的电池容量有限和计算能力不强，进一步阻碍了大多数基于密码和身份验证的安全机制在物联网中的部署。

为了克服这些不足，近年来发展出入侵检测系统（intrusion detection system，IDS）这一动态监控、预防和抵御入侵行为的安全机制[4-5]。相比于其他安全机制，基于异常的IDS机制通过测量输入数据与正常数据间的偏差来检测网络攻击。尽管机器学习算法的最新进展激发了新的IDS机制[6]，但仍存在一系列亟待解决的问题。

文献[7]融合了基于贝叶斯优化的高斯过程和决策树分类算法，用于检测物联网中的僵尸网络攻击。类似地，文献[8]利用遗传算法优化随机森林模型，再利用优化的随机森林模型检测僵尸网络攻击。然而，这些检测方法均属于监督学习方法，无法检测未知攻击，并且这些方法依赖于已标签的数据训练模型，而未知攻击往往对网络损害极大，且不容易被发现。

由于缺乏带标签的数据以及不断涌现的新攻击手段，IDS必须具备检测未知攻击的能力，增强防御零日攻击的能力。即使对于已知攻击的检测，获取带标签的数据也是一项艰巨的挑战，收集这些数据相当耗时，有时甚至不可能实现。因此，设计无监督的IDS迫在眉睫。

文献[9]提出了基于生成对抗网络（generative adversarial network，GAN）的异常检测方法，该方法通过结合GAN的生成器和重构损失来检测系统中的异常行为。然而，该方法需要通过为评估中的每个数据模式解决优化问题来检测攻击，这增加了检测时间，导致其不适合时延受限的应用程序。此外，文献[10]提出了低时延的无监督IDS，该系统利用自编码器重构数据，并利用长短期记忆（long short-term memory，LSTM）网络构建GAN的生成器和决策器。

上述工作主要依赖LSTM网络捕获数据间的依赖关系，所提出的检测模型过分依赖LSTM。然而，近期的研究表明，LSTM网络存在一些局限性，这使其作为序列建模任务的标准架构的地位受到质疑[11]。例如，LSTM网络需要按顺序处理数据，这不仅增加了计算复杂度，而且在训练过程中还需要大量内存来存储多个门单元的处理数

据。尽管文献[12]提出了无LSTM的基于GAN的IDS，该系统只采用全连接和常规的神经网络，但其检测性能低于文献[10]的检测性能。

为此，本文采用时间卷积网络（temporal convolutional network，TCN）和多头注意力机制代替LSTM，提出基于多头注意力机制和时间卷积网络的无监督入侵检测模型（unsupervised intrusion detection model based on multi-head attention mechanism or temporal convolutional network，UDMT）。UDMT利用TCN和多头注意力机制构建GAN的生成器和决策器网络。由于TCN和多头注意力机制便于实现计算并行化，其性能优于LSTM[13]。此外，UDMT可灵活配置隐藏层，通过优化配置，能够在检测率和检测时间之间达到平衡，进而满足不同的应用需求。

1 系统模型

1.1 基于GAN的入侵检测系统

GAN是用于训练生成器和决策器神经网络的强大框架。若只使用正常数据进行训练，则生成器可隐式地对系统进行建模，并学习如何生成类似于正常数据的合成数据。生成器学习从潜在Z空间中的分布$P(z)$中提取的随机变量z，映射到与正常数据相似的数据模式，以使得生成器的输出尽可能与正常数据模式相似。

决策器则学习如何辨别输入样本数据是正常数据还是由生成器产生的合成数据。因此，决策器的输出实质上就是判断输入样本x是正常数据还是合成数据。换而言之，不论是已知攻击还是未知攻击，决策器都通过估计与正常数据间的偏差来判断输入样本的类型。

生成器和决策器的神经网络以对抗方式进行训练，致使生成器输出的数据能骗过决策器，尽可能地提高决策器进行错误辨别的概率，而决策器则尽可能地降低此概率。

本文提出的UDMT采用WGAN（wassertein GAN）框架，WGAN框架如图1所示。D和G分别表示决策器和生成器。生成器最大化决策器误判的损失$G_{\text{Loss}}=D(G(z))$，而决策器最小化误判的损失$D_{\text{Loss}}=D(G(z))-D(x)$，其中$x$表示输入样本。相比于原始的GAN框架，WGAN框架更易训练，且不受梯度消失问题的影响。

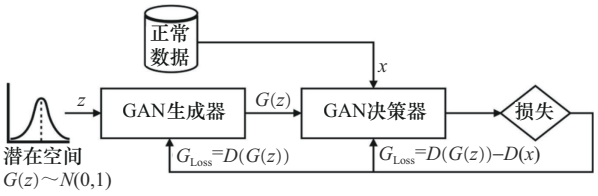


图1 WGAN框架

现存的基于GAN的IDS严重依赖于LSTM网络评估数据之间的时间依赖性。然而，LSTM网络中处理序列数据的计算复杂度高，且内存消耗大。因此，本文通过时间卷积网络和多头注意力机制替代LSTM网络。

1.2 时间卷积网络

时间卷积网络是针对序列预测任务而改进的卷积网络架构，是一种基于卷积神经网络结构的方法。由于TCN可并行计算，已被广泛应用于语音识别、动作检测、时间序列分类等领域。

TCN将输入序列映射到相同长度的输出序列，并使用因果卷积，即只利用历史信息进行卷积。因此，时刻t的输出仅由t时刻之前的信息卷积而成。此外，由于序列建模任务可能需要更多的历史信息，TCN还需要使用扩张因果卷积来使架构能更深入地研究历史信息。

1.2.1 因果卷积

因果卷积通过掩膜方式去除网络中的非必要连接，它只通过前向连接，使网络满足时间上的前后依赖约束，是一个严格的时间约束模型。因果卷积结构如图2所示，其利用掩膜方式逐层去除部分连接，保留那些从前往后的连接[14]。图2中虚线部分表示去除的连接，而实线部分表示保留的连接。

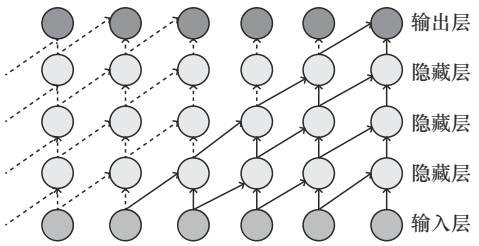


图2　因果卷积结构

然而，因果卷积存在覆盖历史信息不足的问题。以图2为例，最后的输出结果只能基于第一层输入中的5个神经元计算，但更前面的输入信息未被利用，导致信息丢失。尽管可通过增加网络层数捕捉更多的历史信息，但是增加网络层数必然会加大训练难度。

1.2.2　扩张因果卷积

为了克服因果卷积的不足，本文采用扩张因果卷积。扩张因果卷积通过扩大神经网络对历史信息的感受野，消除因果卷积在训练过程中的维数问题。具体而言，对于一个输入序列 $x=(x_0,x_1,\cdots,x_t,\cdots,x_s)\in\mathbb{R}^n$ 和一个卷积核 f，它在序列元素 s 上的扩张因果卷积为：

$$F(s)=\left(x*_d f\right)(s)=\sum_{i=0}^{k-1} f(i)\cdot x_{s-d\cdot i} \tag{1}$$

其中，f 是一个卷积核，且 $f:\{0,\cdots,k-1\}\to\mathbb{R}$，$k$ 为卷积核大小，d 为扩张因子，表示在卷积过程中利用过往信息的深度，$*_d$ 表示扩张卷积操作符。

1.2.3　残差模块

为了防止在训练深层神经网络中出现过拟合和退化问题，在TCN中引入了残差模块，并通过残差模块替换传统的网络结构，进而解决网络退化问题。残差模块先合并历史信息和转换后的信息，再以此信息作为输入。残差模块的激励函数如下。

$$O(x)=\Phi(x+T(x)) \tag{2}$$

其中，$O(x)$ 表示残差模块，Φ 表示激活函数，$T(x)$ 表示转换之后的信息。

TCN能从数据中提取时间依赖关系，并且已被证明，在建模序列方面，TCN具有比LSTM网络更明显的优势。因此，本文采用TCN。TCN结构如图3所示，它由单一扩张因果卷积和ReLU函数组成。此外，为了避免过度拟合，TCN引入了一个归一化层和用于正则化的Dropout层。图3中虚线为一个基本元，根据应用需求，可将基本元重复 N_{TCN} 次。例如，$N_{\mathrm{TCN}}=1$，则TCN结构中只有1个基本元，若 $N_{\mathrm{TCN}}=2$，则TCN结构中有2个基本元。

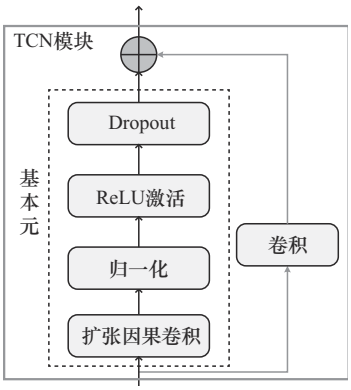


图3　TCN结构

1.3　多头注意力机制

自注意力（self-attention）机制允许模型在处理序列数据时，考虑序列中的所有位置，进而捕捉序列内的长距离依赖关系。实施自注意力机制主要有4个步骤。

（1）输入表示。模型先接收一个输入序列，并将此序列转换成输入矩阵。

（2）评估注意力分数。模型通过计算输入矩阵中前后元素间的关系，估算注意力分数。计算分数的过程中涉及查询矩阵 $\boldsymbol{Q}$、键矩阵 $\boldsymbol{K}$ 和值矩阵 $\boldsymbol{V}$ 这3个矩阵。

（3）加权求和，得到加权后的上下文矩阵。利用Softmax函数对上一步所计算的分数进行归一化处理。

（4）模型输出。将上下文矩阵作为模型输出。

令 $\boldsymbol{C}$ 为上下文矩阵，其定义如下：

$$\boldsymbol{C}=\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\text{Softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V} \quad (3)$$

其中，d_k表示数据维度，当输入为x时，则$\boldsymbol{Q}=x\boldsymbol{W}^{\boldsymbol{Q}}$、$\boldsymbol{K}=x\boldsymbol{W}^{\boldsymbol{K}}$和$\boldsymbol{V}=x\boldsymbol{W}^{\boldsymbol{V}}$，其中$\boldsymbol{W}^{\boldsymbol{Q}}$、$\boldsymbol{W}^{\boldsymbol{K}}$和$\boldsymbol{W}^{\boldsymbol{V}}$分别表示矩阵$\boldsymbol{Q}$、$\boldsymbol{K}$和$\boldsymbol{V}$的权重矩阵。

多头注意力（multi-head attention，MHA）机制是一种多头注意力模型，其将注意力机制扩展到多个头，从而增强模型对于不同特征的关注度。多头注意力机制的网络结构模型如图4所示。先通过x与$\boldsymbol{W}^{\boldsymbol{Q}}$、$\boldsymbol{W}^{\boldsymbol{K}}$和$\boldsymbol{W}^{\boldsymbol{V}}$相乘，得到矩阵$\boldsymbol{Q}$、$\boldsymbol{K}$和$\boldsymbol{V}$，再将矩阵$\boldsymbol{Q}$、$\boldsymbol{K}$和$\boldsymbol{V}$分别转换成$h$个矩阵：

$$\langle \boldsymbol{Q}_1,\boldsymbol{Q}_2,\cdots,\boldsymbol{Q}_h;\boldsymbol{K}_1,\boldsymbol{K}_2,\cdots,\boldsymbol{K}_h;\boldsymbol{V}_1,\boldsymbol{V}_2,\cdots,\boldsymbol{V}_h\rangle \quad (4)$$

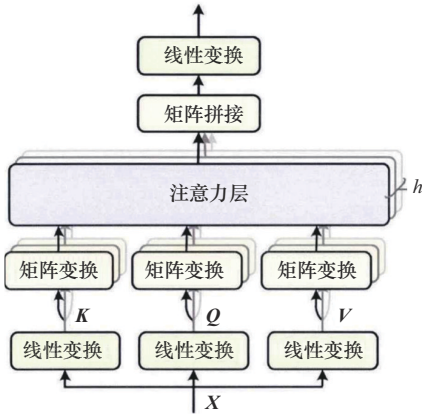


图4　多头注意力机制的网络结构模型

这h组中的每一组$\boldsymbol{Q}_i$、$\boldsymbol{K}_i$和$\boldsymbol{V}_i$都对应一个注意力层，其中$i=1,2,\cdots,h$，经注意力机制处理后，将处理后的结果进行拼接，再将拼接后的结果作为线性变换的输入，最终得到MHA的输出：

$$\text{MHA}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\boldsymbol{W}^0\text{Concat}(\text{head}_1,\cdots,\text{head}_h) \quad (5)$$

其中，$\boldsymbol{W}^0$是权重矩阵，Concat(·)表示拼接操作，head_i表示第i个头的输出，$i=1,2,\cdots,h$。

$$\begin{aligned}\text{head}_i&=\text{Attention}\left(\boldsymbol{Q}\boldsymbol{W}_i^{\boldsymbol{Q}},\boldsymbol{K}\boldsymbol{W}_i^{\boldsymbol{K}},\boldsymbol{V}\boldsymbol{W}_i^{\boldsymbol{V}}\right)\\&=\text{Softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{W}_i^{\boldsymbol{Q}}\left(\boldsymbol{K}\boldsymbol{W}_i^{\boldsymbol{K}}\right)^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V}\boldsymbol{W}_i^{\boldsymbol{V}}\end{aligned} \quad (6)$$

与TCN结构类似，多头注意力机制可以提取数据之间的依赖关系，并且在几个序列建模任务中的表现优于LSTM网络。相比于LSTM网络，TCN结构可有效地提取特征，有助于训练出更准确的模型。

本文所采用的多头注意力模块如图5所示，其由一个MHA模块、归一化层和Dropout层组成。归一化层和Dropout层用于防止过拟合。与图3类似，多头注意力模块中也引用了基本元，根据应用需求，可将基本元重复N_{MHA}次。例如，$N_{\text{MHA}}=1$，则结构中只有1个基本元；若$N_{\text{MHA}}=2$，则自注意力模块中有2个基本元。

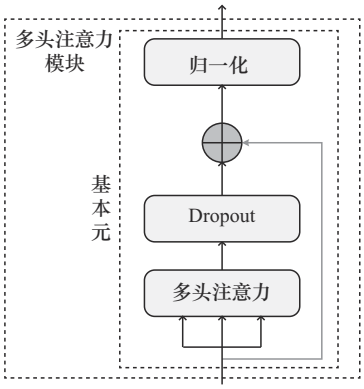


图5　多头注意力模块

1.4　UDMT的结构

为了提高UDMT的扩展性，UDMT采用灵活的结构。GAN中的生成器和决策器的网络架构如图6所示。

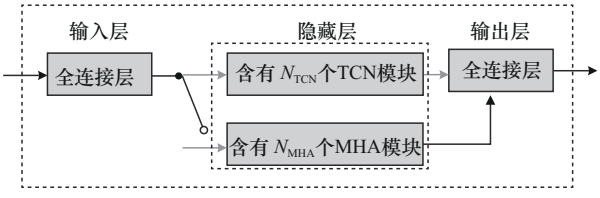


图6　GAN中的生成器和决策器的网络架构

输入层和输出层均为全连接层。隐藏层内采用灵活的网络结构，以TCN模块或者MHA模块为隐藏层。此外，若以TCN或者MHA模块作为隐藏层时，它们的基本元的单元数可灵活匹配。例如，若$N_{\text{TCN}}=2$，则表示采用了两个基本元的

TCN 结构。在后续的实验中，将分析 N_{TCN} 或 N_{MHA} 对 UDMT 的性能。

1.5　执行 UDMT 的流程

UDMT 的主要流程如图 7 所示，其主要由数据预处理、模型训练和模型测试这 3 部分组成。

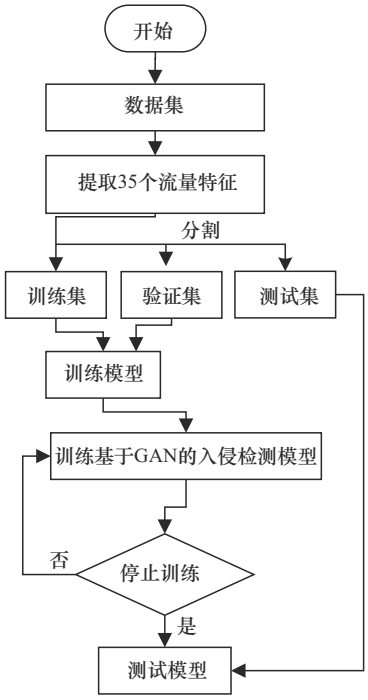


图 7　UDMT 的主要流程

2　系统分析与仿真

2.1　基于 Optuna 的超参数设定

多数基于 GAN 的入侵检测模型以 LSTM 构建生成器和决策器。本文提出的 UDMT 则是将 TCN 或 MHA 模块作为隐藏层代替 LSTM。为此，以基于 LSTM 为隐藏层的 GAN 作为一个基准（简写 LSTM），便于分析用 TCN 或 MHA 代替 LSTM 模型的性能。

为了选取最优的超参数，本文采用 Optuna 框架。Optuna 是一个用于超参数优化的开源框架，广泛应用于机器学习领域。它提供了一种高效的方式设置超参数，并自动地选择最佳的超参数组合。UDMT 所涉及的参数有学习率、优化器、批量大小、隐藏层维数、卷积核大小、注意力头数以及 N_{TCN} 或 N_{MHA} 值。通过 Optuna 优化超参数，优化后的超参数取值见表 1。

表 1　优化后的超参数取值

参数	UDMT-LSTM	UDMT-TCN(1)	UDMT-TCN(2)	UDMT-MHA(1)
最大的 Epoch 次数	50	50	50	50
优化器	Adam	Adam	Adam	Adam
Dropout 率	0.25	0.25	0.25	0.20
决策器学习率	0.002 84	0.004 85	0.007 53	0.011 56
生成器学习率	0.000 04	0.000 15	0.000 001	0.000 039
TCN 核尺寸	—	2	2	—
注意力头数	—	—	—	40

2.2　数据集概述

本文利用 CICDDoS 2019 数据集验证 UDMT 的检测性能。CICDoS 2019 数据集包含最新的分布式拒绝服务（distributed denial of service，DDoS）攻击，与真实世界的数据相似。这些 DDoS 攻击包括 Syn、UDP、UDPLag、MSSQL、NetBIOS、LDAP 和 Portmap 等多种攻击类型。此外，本文使用文献[15]所定义的 35 个相关的网络流特征，如流持续时间、前向和后向的数据包总数，以及识别网络流的源 IP 地址、目的 IP 地址、源端口、目的端口和协议这 5 个特征。

为了训练和评估 UDMT，先构建训练集、验证集和测试集。从 1 个训练日中的正常网络流量中随机抽取 80% 构建训练集，剩余的 20% 和一部分 DDoS 攻击样本构建验证集，通过此验证集数据优化模型的超参数，训练集、验证集和测试集见表 2。训练集只包含正常流量，而验证集既包含正常流量样本，也包含 DDoS 攻击样本。

表 2　训练集、验证集和测试集

数据集	正常样本数	DDoS 攻击样本数
训练集	45 408	0
验证集	11 342	68 052
测试集	50 000	50 000

从1个测试日中抽取50 000个正常流量样本和50 000个恶意流量样本构建测试集。此外，虽然测试集的恶意网络流来自各种类型的DDoS攻击，但UDMT依赖于一个二进制分类器来区分正常和恶意的网络流，而不是按攻击类型分类。

训练集、验证集和测试集所包含的DDoS攻击类型和样本数见表3。尽管验证集包含了Syn、UDP、UDPLag、MSSQL、NetBIOS、LDAP和Portmap，但只是用于训练模型的超参数。此外，为了评估模型防御零日攻击的能力，验证集中不包含Portmap攻击样本，但测试集中包含此攻击样本，即将Portmap攻击作为未知攻击。

表3 训练集、验证集和测试集所包含的DDoS攻击类型和样本数

DDoS攻击类型	训练集	验证集	测试集
Syn	0	11 342	8 021
UDP	0	11 342	8 021
UDPLag	0	11 342	1 873
MSSQL	0	11 342	8 021
NetBIOS	0	11 342	8 021
LDAP	0	11 342	8 021
Portmap	0	0	8 022

2.3 检测结果分析

本节将分析UDMT检测攻击的性能。为了更好地分析UDMT的性能，选择两个同类模型作为基准：（1）文献[10]提出的FID-GAN模型。FID-GAN模型采用LSTM作为GAN生成器和决策器，通过LSTM捕获数据间的时间关系。（2）文献[12]提出的ALAD模型。ALAD模型没有采用LSTM网络，仅利用全连接层和规则的神经网络作为GAN生成器和决策器。

2.3.1 模型的ROC曲线

本文引用接收者操作特征（receiver operating characteristic，ROC）曲线评估UDMT的性能，并计算ROC的曲线下方面积（area under the curve，AUC）。

UDMT的ROC曲线如图8所示，横坐标为假阳率，纵坐标为真阳率。真阳率表示实际为正样本的样本中被正确预测为正样本的比例，而假阳率表示负样本的样本中被错误预测为正样本的比例。图8考虑了4种情况。（1）利用LSTM作为隐藏层，标记为UDMT-LSTM。（2）一个基本元的TCN（$N_{TCN}=1$）作为隐藏层，标记为UDMT-TCN(l)。（3）两个基本元的TCN（$N_{TCN}=2$）作为隐藏层，标记为UDMT-TCN(2)。（4）一个基本元的MHA（$N_{MHA}=1$）作为隐藏层，标记为UDMT-MHA(l)。

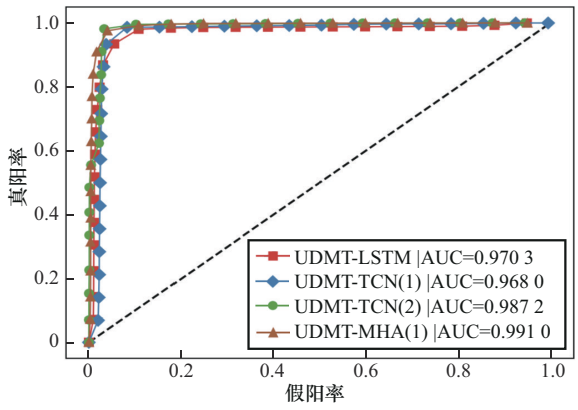
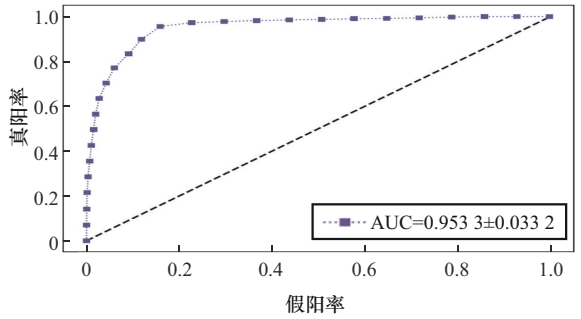


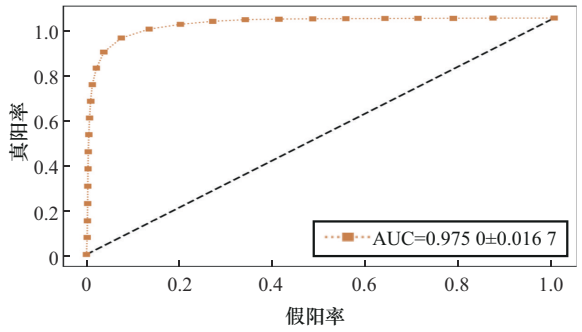
图8 UDMT的ROC曲线

由图8可知，所考虑的上述4种情况下UDMT均可获取接近于1的AUC，这说明利用所提出的UDMT可同时确保低的误报率和漏报率，减少误报率可维持网络运行，而减少漏报率可确保系统安全。此外，相比于另外3种情况，MHA（$N_{MHA}=1$）作为隐藏层可获取更高的AUC。

ALAD模型和FID-GAN模型的ROC曲线如图9所示。



（a）ALAD模型



（b）FID-GAN模型

图9　ALAD模型和FID-GAN模型的ROC曲线

由图9可知，ALAD模型和FID-GAN模型的AUC分别为0.953 3±0.033 2和0.975 0±0.0167，低于UDMT的AUC。UDMT下$N_{TCN}=1$、$N_{TCN}=2$和$N_{MHA}=1$这3种情况的AUC分别为0.968 0±0.021 7、0.987 2±0.012 4和0.991 0±0.010 2，原因在于ALAD模型仅用全连接和规则化的卷积层，并没有考虑数据间的依赖性。

2.3.2　检测率

UDMT、ALAD和FID-GAN模型的检测性能见表4，检测性能包括准确率、精确率、召回率和F1值。

表4　UDMT、ALAD和FID-GAN模型的检测性能

检测模型	准确率	精确率	召回率	F1值
UDMT-I LSTM	0.940 5	0.941 1	0.940 8	0. 941 0
UDMT -TCN(1)	0.958 6	0.958 8	0.957 8	0.958 0
UDMT-TCN(2)	0.970 7	0.970 5	0.971 0	0.970 7
UDMT-MHA(I)	0.968 1	0.968 2	0.969 0	0.968 4
FID- GAN	0.920 8	0.920 3	0.920 1	0.921 0
ALAD	0.886 1	0.886 0	0.886 9	0.887 0

由表4可知，无论是采用TCN（$N_{TCN}=1$或$N_{TCN}=2$）作为隐藏层，还是采用MHA作为隐藏层，UDMT的检测性能均优于FID-GAN和ALAD模型。考虑F1值融合了准确率和召回率，以F1值为例分析各模型的性能，其中，UDMT-TCN(2)模型的检测性能最优，其F1值达到0.970 7，而FID-GAN和ALAD模型的F1值分别为0.920 3和0.886 0。此外，利用TCN或者MHA模块代替LSTM，可提升UDMT的检测性能。

UDMT、ALAD和FID-GAN模型对各类攻击的检测性能见表5。由于UDMT没有考虑Portmap攻击样本（将Portmap攻击作为未知攻击），在分析模型的检测性能时，重点关注各模型对Portmap攻击样本的检测性能。由表5可知，UDMT-TCN(2)模型检测Portmap攻击的检测率高于其他模型，这说明UDMT对未知攻击的检测性能优于同类检测模型，检测率达到0.996 0。

2.4　复杂度及检测时延分析

为了评估UDMT检测攻击所需要的时间，实验记录了UDMT检测攻击的平均检测时间。

表5　UDMT、ALAD和FID-GAN模型对各类攻击的检测性能

样本类型	UDMT-LSTM	UDMT-TCN(1)	UDMT-TCN(2)	UDMT-MHA(1)	FID-GAN	ALAD
Syn	0.729 0	0.972 8	0.824 2	0.889 7	0.605 2	0.572 6
UDP	0.999 4	0.994 6	1.000 0	1.000 0	0.951 6	0.819 3
UDPLag	0.921 5	0.661 0	0.985 6	0.977 0	0.865 5	0.883 1
MSSQL	0.999 9	0.986 8	1.000 0	1.000 0	0.989 8	0.966 2
NetBIOS	0.947 9	0.935 5	0.999 0	0.943 9	0.992 5	0.973 6
LDAP	0.999 6	0.999 4	1.000 0	0.999 8	0.999 8	0.999 9

UDMT、ALAD和FID-GAN模型的检测时延如图10所示。

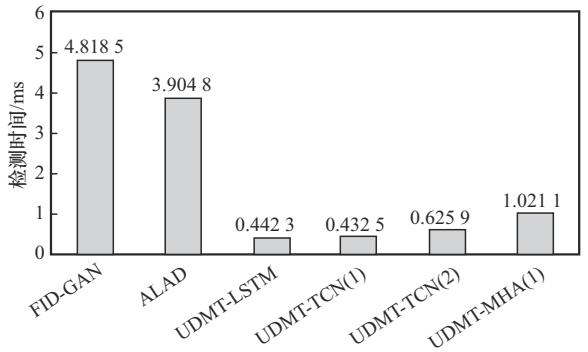


图10 UDMT、ALAD和FID-GAN模型的检测时延

由图10可知，UDMT-TCN(l)模型的检测时延最低，低至0.432 5 ms，而FID-GAN模型的检测时延最高，达到4.818 5 ms。FID-GAN模型采用了更复杂的GAN结构，并通过自编码的神经网络计算重构误差，因此训练一个自编码的神经网络需要消耗更多时间。

结合图8，再进一步观察图10可知，选择不同的隐藏层，可平衡检测性能与检测时延。例如，在所提出的UDMT中，UDMT - MHA(l)模型具有最长的检测时延，但其却有最大的AUC面积（检测性能最优），而UDMT - TCN(l)模型具有最低的检测时延，但其AUC面积最小（检测性能最差）。因此，可依据应用需求，确定是以追求最低的检测时延为目标，还是以追求最高的检测性能为目标，从而选择不同的模型配置。

最后，实验分析了采用不同隐藏层的UDMT复杂度，见表6。表6给出了UDMT所迭代的Epoch次数、每次Epoch所消耗的平均训练时间（简写为单次训练时间）、总的收敛时间、GAN生成器和决策器的参数。由于迭代停止条件不同，每个模型所迭代的Epoch次数不同。所迭代的Epoch次数越小，意味着收敛速度越快。其中，UDMT - LSTM模型收敛速度最快，所需要迭代的Epoch次数最少。此外，UDMT - TCN(l)模型具有最少的参数，每次Epoch所消耗的平均训练时间也最短。

表6 采用不同隐藏层的UDMT复杂度

参数	UDMT-LSTM	UDMT-TCN(1)	UDMT-TCN(2)	UDMT-MHA(1)
Epoch次数	33	50	45	42
单次训练时间	T2.03	10.07	13.64	1942
收敛时间	397.00	503.29	613.80	815.80
决策器的参数	4234	3 101	4 916	5 921
生成器的参数	3 618	2 330	8 180	5 520

3 结束语

本文提出了一个无监督的基于GAN的入侵检测模型UDMT。该模型利用TCN和多头注意力机制构建GAN的生成器和决策器网络，降低训练模型的复杂度，同时采用灵活配置，实现不同的隐藏层模式，进而在检测率和检测时延间达成平衡，以满足不同的应用需求。性能分析表明，相比于同类的检测模型，本文提出的UDMT具有更大的AUC值。此外，UDMT还具有更低的检测时延，进而使UDMT更适合于时延受限的应用。

参考文献：

[1] 罗国宇, 汪学舜, 戴锦友. 物联网入侵检测的随机特征图神经网络模型[J]. 计算机工程与应用, 2024, 60(21): 264-273.
LUO G Y, WANG X S, DAI J Y. Random feature graph neural network for intrusion detection in Internet of things[J]. Computer Engineering and Applications, 2024 , 60(21): 264-273.

[2] 李聪宇, 赵利辉, 安洋. 基于图神经网络的物联网入侵检测研究[J]. 中北大学学报(自然科学版), 2024, 45(2): 194-204.
LI C Y, ZHAO L H, AN Y. Research on intrusion detection of Internet of things based on graph neural network[J]. Journal of North University of China (Natural Science Edition), 2024, 45(2): 194-204.

[3] 冯绮航. 考虑属性加密的物联网隐私数据跨域安全共享模型[J]. 现代电子技术, 2023, 46(1): 91-95.
FENG Q H. Internet of things privacy data cross domain security sharing model considering attribute encryption[J]. Modern Electronics Technique, 2023, 46(1): 91-95.

[4] 项睿涵, 潘巨龙, 李玲艺, 等. 一种物联网入侵检测和成员推理攻击研究[J]. 传感技术学报, 2024, 37(2): 317-325.

XIANG R H, PAN J L, LI L Y, et al. A new study of an IoT intrusion detection and membership inference attack[J]. Chinese Journal of Sensors and Actuators, 2024, 37(2): 317-325.

[5] 吴昊, 郝佳佳, 卢云龙. 物联网场景下基于蜜场的分布式网络入侵检测系统研究[J]. 通信学报, 2024, 45(1): 106-118.
WU H, HAO J J, LU Y L. Research on distributed network intrusion detection system for IoT based on honeyfarm[J]. Journal on Communications, 2024, 45(1): 106-118.

[6] SHONE N, NGOC T N, PHAI V D, et al. A deep learning approach to network intrusion detection[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2018, 2(1): 41-50.

[7] INJADAT M, MOUBAYED A, SHAMI A. Detecting botnet attacks in IoT environments: an optimized machine learning approach[C]//Proceedings of the 2020 32nd International Conference on Microelectronics (ICM). Piscataway: IEEE Press, 2020: 1-4.

[8] MOUBAYED A, INJADAT M, SHAMI A. Optimized random forest model for botnet detection based on DNS queries[C]//Proceedings of the 2020 32nd International Conference on Microelectronics (ICM). Piscataway: IEEE Press, 2020: 1-4.

[9] ZHANG Z J, LI W Z, DING W X, et al. STAD-GAN: unsupervised anomaly detection on multivariate time series with self-training generative adversarial networks[J]. ACM Transactions on Knowledge Discovery from Data, 2023, 17(5): 1-18.

[10] FREITAS DE ARAUJO-FILHO P, KADDOUM G, CAMPELO D R, et al. Intrusion detection for cyber–physical systems using generative adversarial networks in fog environment[J]. IEEE Internet of Things Journal, 2021, 8(8): 6247-6256.

[11] HUANG S H, LIU Y, FUNG C, et al. HitAnomaly: hierarchical transformers for anomaly detection in system log[J]. IEEE Transactions on Network and Service Management, 2020, 17(4): 2064-2076.

[12] ZENATI H, ROMAIN M, FOO C S, et al. Adversarially learned anomaly detection[C]//Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM). Piscataway: IEEE Press, 2018: 727-736.

[13] LI Y D, ZHANG L, LV Z, et al. Detecting anomalies in intelligent vehicle charging and station power supply systems with multi-head attention models[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(1): 555-564.

[14] 张海涛, 李文娟, 李雪峰, 等. 基于变分模态分解和时间注意力机制TCN网络的光伏发电功率预测[J]. 电测与仪表, 2024: 1-8.
ZHANG H T, LI W J, LI X F, et al. Photovoltaic power forecasting based on TPA-TCN model and variational modal decomposition[J]. Electrical Measurement & Instrumentation, 2024: 1-8.

[15] JIA Y Z, ZHONG F T, ALRAWAIS A, et al. FlowGuard: an intelligent edge defense mechanism against IoT DDoS attacks[J]. IEEE Internet of Things Journal, 2020, 7(10): 9552-9562.

[作者简介]

廖金菊（1980– ），女，郑州工业应用技术学院信息工程学院副教授，主要研究方向为隐私保护、机器学习。

丁嘉伟（1986– ），男，郑州工业应用技术学院信息工程学院副教授，主要研究方向为人工智能、管理信息系统。

冯光辉（1982– ），男，广州大学计算机科学与网络工程学院博士生，主要研究方向为隐私保护、联邦学习。