文章编号： 1673-3193（2024）02-0194-11

基于图神经网络的物联网入侵检测研究

李聪宇，赵利辉，安　洋

（中北大学 软件学院，山西 太原 030051）

摘　要： 针对物联网入侵检测中网络设备的异构性以及设备间的复杂关联性，本文基于图神经网络（Graph Neural Network，GNN）提出一种GraphSAGE-GAT模型，可以有效捕捉物联网设备之间的关联关系，并还原物联网设备之间的通信拓扑，从而达到提升物联网异常检测准确率的目的。首先，基于物联网设备间的网络流数据构建了设备关联关系图，然后利用GraphSAGE（Graph Sample and Aggregate）算法对相邻设备节点进行采样，从而可利用相互关联设备节点信息增强设备节点的嵌入信息表示；再利用图注意力网络（Graph Attention Network，GAT）为提取到的关联设备节点之间的关系自动化地学习到相关性权重，并通过多层聚合函数将关联设备节点的表示进一步融合，得到设备关联图节点的嵌入表示向量，从而进一步增强各设备节点的表示能力。最后，根据融合后的图节点嵌入表示向量实现对设备网络节点样本的良性和攻击分类。基于数据集NF-ToN-IoT-v2和NF-BoN-IoT-v2进行了实验验证，结果表明，本文所提出的模型GraphSAGE-GAT在物联网入侵检测上的准确率分别高达97.25%和98.62%，均优于现有最新的基线检测模型，可进一步保障网络数据的通信安全。

关键词： 物联网；入侵检测；特征选择；GraphSAGE；图注意力网络

中图分类号： TP393　　**文献标识码：** A　　**doi:** 10.3969/j.issn.1673-3193.2024.02.009

引用格式： 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 Congyu，ZHAO Lihui，AN Yang. 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.

Research on Intrusion Detection of Internet of Things Based on Graph Neural Network

LI Congyu，ZHAO Lihui，AN Yang

（School of Software，North University of China，Taiyuan 030051，China）

Abstract： Aiming at the heterogeneity of network devices and the complex correlation among devices in the internet of things intrusion detection，this paper proposed a GraphSAGE-GAT model based on the graph neural network，which could effectively capture the correlation between internet of things devices and reduced the communication topology between internet of things devices，so as to improve the accuracy rate of internet of things anomaly detection. Firstly，the device association graph was constructed based on the network flow data among iot devices，and then the graph sample and aggregate algorithm was used to sample adjacent device nodes，so that the embedded information representation of device nodes could be enhanced by using the information of interconnected device nodes. Then，through the graph attention

收稿日期： 2023-07-09

作者简介： 李聪宇（1999－），男，硕士生，主要从事图神经网络与网络安全方向的研究。

通信作者： 赵利辉（1979－），男，副教授，博士，主要从事网络测试与网络安全方向的研究。E-mail：leehwi@nuc.edu.cn。

network，the correlation weight was automatically learned for the relationship between the extracted associated device nodes，and the representation of the associated device nodes was further fused through the multi-layer aggregation function to obtain the embedded representation vector of the device association graph nodes，so as to further enhance the representation capability of each device node. Finally，the benign and offensive classification of device network node samples was realized according to the fused graph node embedding representation vector. Based on the data sets NF-ToN-IoT-v2 and NF-BoN-IoT-v2，the experiment results show that the accuracy of the proposed model of GraphSAGE-GAT in IoT intrusion detection is as high as 97.25% and 98.62%，respectively，both of which are superior to the latest baseline detection models. Therefore，the model GraphSAGE-GAT proposed in this paper can further guarantee the communication security of network data.

Key words：internet of things；intrusion detection；feature selection；graph sample and aggregate（GraphSAGE）；graph attention network

0 引 言

随着物联网应用范围的不断扩大，安全和隐私问题日益受到关注。物联网设备的异构和动态特性使它们容易受到不同类型的威胁和安全攻击[1]。此外，物联网设备交互信息的高度敏感性增加了设备被欺骗攻击的可能性，这使得物联网在数据安全和数据隐私方面面临严峻挑战。

物联网网络是包含节点和关系的自然图结构。Liu等[2]提出了一种新的基于网络流量图的图神经网络模型（NT-GNN），该模型被用于物联网Android恶意软件的检测，综合考虑了图的节点和边这两个因素，通过在两个数据集CICAndMal2017和AAGM上进行实验，取得了97%的准确率。Nguyen等[3]提出了一种基于文件的物联网僵尸网络检测方法。该方法动态捕捉可执行文件运行期间的行为数据，使用可打印的字符串信息来遍历基于静态分析生成的图形的特征；动态字符串特征被添加到可打印字符串信息图特征中，以消除需要遍历的顶点；它去除了40%的冗余特征，使算法实现时间缩短了30%以上。Caville等[4]提出了一种基于图神经网络的自监督入侵检测模型（Anomal-E）进行异常检测，该方法采用了GraphSAGE网络的归纳模型，并利用了边缘特征和图拓扑结构。在该模型中，作者首先将源节点和目标节点的相邻流进行聚合，然后将其连接起来形成一种流表示；与直接使用流表示不同，该模型学习聚合函数，它能够很轻松地应用于不可见的流。

物联网网络节点具有低能量、低处理能力和较少的内存的特点。Bediya等[5]指出分布式拒绝服务（DDoS）是物联网网络系统中危害最大的一种攻击，并利用区块链技术（BC）设计了一种安全物联网网络的模型（BIoTIDS），该模型将物联网网络中的设备作为区块链节点，在BC上记录这些节点的信息来保证网络数据的完整性和真实性，并通过智能合约来执行策略，监控网络的运行状态，以检测物联网网络入侵的发生和DDoS攻击。Nimbalkar等[6]提出了一种新的IDS特征选择方法，通过使用信息增益（IG）、增益比（GR）和基于过滤器的特征选择技术提供的前50%的特征来检测DoS和DDoS攻击；该方法通过对排名前50%的IG和GR特征所获得的子集进行插入和并运算，过滤掉那些对攻击检测没有贡献的特征，获得最佳特征子集；通过使用JRip分类器在IoT-BoT和KDD99数据集上进行实验，取得了99.992 0%和99.994 3%的准确率。Kumar等[7]提出了物联网环境的新型统一UIDS模型，该模型整合了多个IDS技术，包括基于异常、基于签名和基于主机的检测等，使用4种不同的决策树算法评估UNSW-NB15数据集，并分析了它们在不同类型的攻击（漏洞利用、DDoS、探查攻击和泛型攻击）上的表现。

综上所述，目前大多数基于物联网攻击检测的研究主要集中在模型的性能上，而没有考虑特征权重大小对分类效率影响的问题，在数据集中检验模型的适用性不高；同时采用单一深度学习算法的检测存在特征提取不充分的情况。因此，本文提出一种基于GraphSAGE-GAT模型的物联网攻击检测方法来进行特征提取和模型训练，主要工作包括：1）由于原始GraphSAGE网络中使

用平均聚合函数，各节点之间权重相等，忽略了不同节点的信息差异，所以本文将其修改为基于注意力机制的加权聚合方法；2）将GraphSAGE和GAT进行融合来分析和检测物联网攻击，以提高强特征的影响权重和提高检测性能；3）本文提出的检测方法是归纳式的，可以推广到训练时未见过的数据上，通过对两个著名的物联网入侵检测数据集进行广泛的实验评估，证明了模型可以有效地检测恶意攻击，具有较强鲁棒性和较好的泛化性。

1 相关研究

1.1 物联网入侵检测

目前，身份验证、防火墙、加密认证等常见的物联网防御技术无法保证物联网的绝对安全。为了进一步提升物联网的防御，研究者们引入了入侵检测系统来监控网络中的流量和节点，并判断物联网设备中是否出现未知行为。为了提高入侵检测系统的检测性能，研究者将机器学习技术应用于入侵检测系统上，例如K近邻算法、支持向量机等。将少量低维数据输入到基于机器学习的入侵检测系统中能获得较好的检测性能，然而当数据量以及维度上升到一定程度时，机器学习算法学习数据的能力不足，导致其检测性能得不到保证。在过去的十年里，深度学习技术是人工智能领域中最重要的技术突破之一，深度学习模型在训练过程中避免了人工提取数据特征的过程，从而使其具有优秀的拟合能力，可以适应各种复杂数据的训练。Zhang等[8]将深度学习引入入侵检测系统，并比较了人工神经网络、深度神经网络和循环神经网络在UNSW-NB15数据集上的分类性能。Wu等[9]提出了一种融合了卷积神经网络和长短期记忆网络的分层神经网络模型，具有94.57%的检测准确率。深度学习技术的上述优势促使研究人员将该技术大量地应用于入侵检测系统。

1.2 图神经网络

在入侵检测领域，图神经网络(Graph Neural Network，GNN)是最有前途的深度学习发展方向之一，因为它能够利用物联网领域中遇到的大量数据的固有图形结构，例如社交媒体网络、知识图谱、复杂的文件等。图形格式通过对一组数据进行建模来捕获结构、对象及其关系的信息，与其他数据结构相比，图结构数据通常包含节点或边特征，可以很容易地描述复杂物联网系统中的多种设备交互信息。

图神经网络不仅可以学习特征嵌入，还可以捕获隐藏在图拓扑结构中的空间信息。通过在节点或边之间传递消息来充分利用图结构，使得神经网络能够有效地学习和概括基于图像的数据，并输出低维嵌入向量。

图神经网络的研究与图嵌入或网络嵌入密切相关，网络嵌入旨在通过保留图的网络拓扑结构和节点内容信息，将图中顶点表示为低维向量，以便使用简单的机器学习算法(如支持向量机)进行处理。同时，网络嵌入的深度学习方法也属于图神经网络，包括基于图自动编码器的算法(如DNGR和SDNE)和无监督训练的图卷积神经网络(如GraphSAGE)。图1描述了网络嵌入和图神经网络的区别。

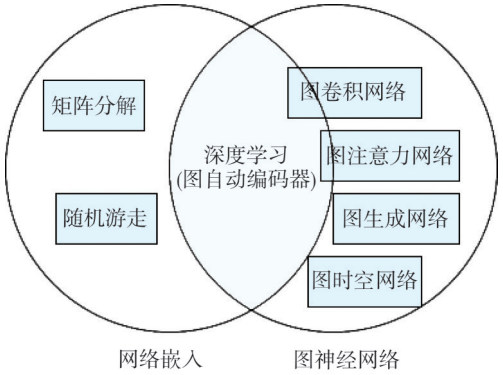

图1 网络嵌入和图神经网络的区别

Fig. 1 Network embedding v. s. graph neural network

1.3 GraphSAGE

GraphSAGE(Graph Sample and Aggregate)算法是著名的GNN之一，由Hamilton等开发[10]。该算法消除了图卷积网络(GCN)一般只能用在直推式学习的局限性。它对传统的GCN进行了两点改进：1）在训练阶段，采样方式将GCN的全图采样优化到部分以节点为中心的邻居抽样，这使得大规模图数据的分布式训练成为可能，并且使得网络可以学习没有见过的节点，这也使得GraphSAGE可以做归纳学习。 2）GraphSAGE研究了若干种邻居聚合的方式，并通过实验和理论分析对比了不同聚合方式的优缺点。GraphSAGE已在图分类、链接预测、推荐系统以及知识图谱等

多个领域取得了成功应用。

传统的网络嵌入算法（基于矩阵分解的算法、基于随机游走的算法）在迭代的过程中需要用到所有节点的信息，学习得到所有节点的向量表示，但是对于新加入的节点需要对所有的节点重新计算，泛化性差。

GraphSAGE学习的结果不再是每个节点的嵌入，而是“聚合函数”。根据已知各个节点的特征和邻居关系，就可以很方便地得到一个新节点的表示。因此，GraphSAGE泛化性好，对新加入的节点可以根据其邻居聚合直接给出其表示学习，而不必对整个网络重新迭代。

GraphSAGE算法分为采样、聚合和更新三个主要阶段，图 2 展示了GraphSAGE的过程。

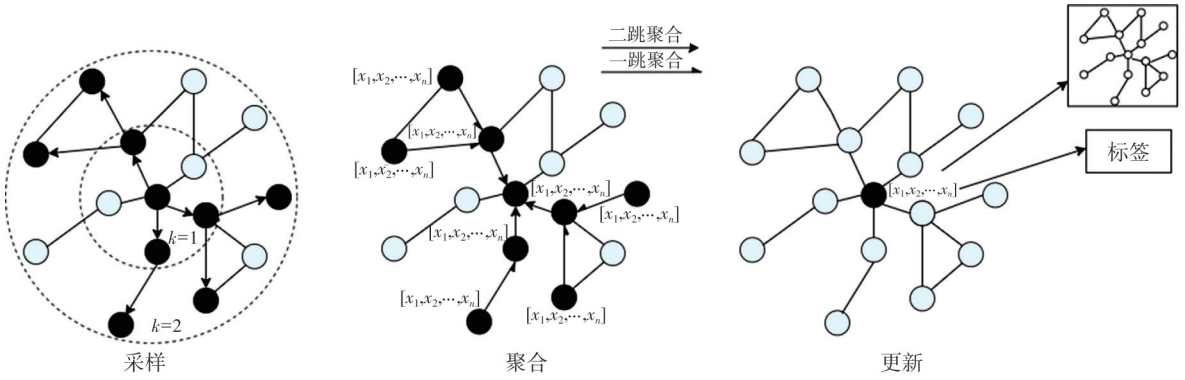

图 2　GraphSAGE算法流程

Fig. 2　Visual illustration of the GraphSAGE sample and aggregate approach

1）采样（Sampling）：对每个节点的邻居进行采样，获取其邻居节点的信息。

2）聚合（Aggregation）：采样后的邻居嵌入向量传到节点上来，并使用一个聚合函数聚合这些邻居信息以更新节点的嵌入向量。

3）更新（Update）：根据更新后的嵌入向量预测节点的标签。

1.4　图注意力网络

图注意力网络[11]中引入了注意力机制，模型中每个节点都有一个向量表示，其余节点的向量表示将用于计算该节点的重要性得分（即注意力权重），这些重要性得分通过激活函数进行归一化，然后加权平均其邻居节点向量表示，以获得聚合后的向量表示。图 3 展示了图注意力层。

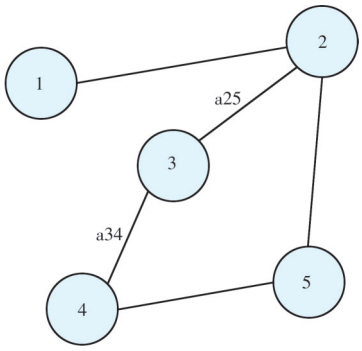

图 3　图注意力层

Fig. 3　Graph attention layer

图 3 中，对于节点 3，它的邻接节点只有节点 2 和节点 4，但不代表这两个节点对节点 3 具有一样的重要性。该重要性可以进行量化，更可以通过网络训练得出。该注意力不满足对称性，即节点 2 对节点 3 的注意力与节点 3 对节点 2 的注意力是不一样的，如果把每条边比作桥，那么该注意力类似桥的宽度。

2　GraphSAGE-GAT 物联网入侵检测模型设计

本文提出的GraphSAGE-GAT模型用以解决现有基于图的物联网入侵检测所遇到的问题，根据设备节点信息及其相关联的设备节点信息的共同依赖性，考虑不同关联设备节点信息对当前目标节点权重的不同，对不同关联设备节点的关系进行加权聚合得到设备关联图的嵌入表示向量，同时，结合了两种现有的图神经网络算法GraphSAGE和GAT的关键优势。

图 4 给出了GraphSAGE-GAT检测算法的研究框架。首先，对原始网络流进行预处理，以生成有代表性的训练和测试图。然后，将训练图输入到GraphSAGE-GAT训练过程中，以训练GraphSAGE-GAT编码器。通过使用GraphSAGE和GAT将当前目标节点与关联设备节点的关系以及节点自身特征结合起来，并自动化学习

出相关性权重，在聚合时对不同关联设备节点的嵌入表示进行加权聚合得到最终的关联图节点的嵌入表示向量。对其中的参数进行调整，并重复该过程以继续训练GraphSAGE-GAT编码器。训练完成后，将训练图嵌入向量输入到常用的异常检测算法(PCA[12]，IF[13]，LOF[13]，DBSCAN[13]，OCSVM[14])，通过对比这5种算法的检测性能，确定最终的模型。

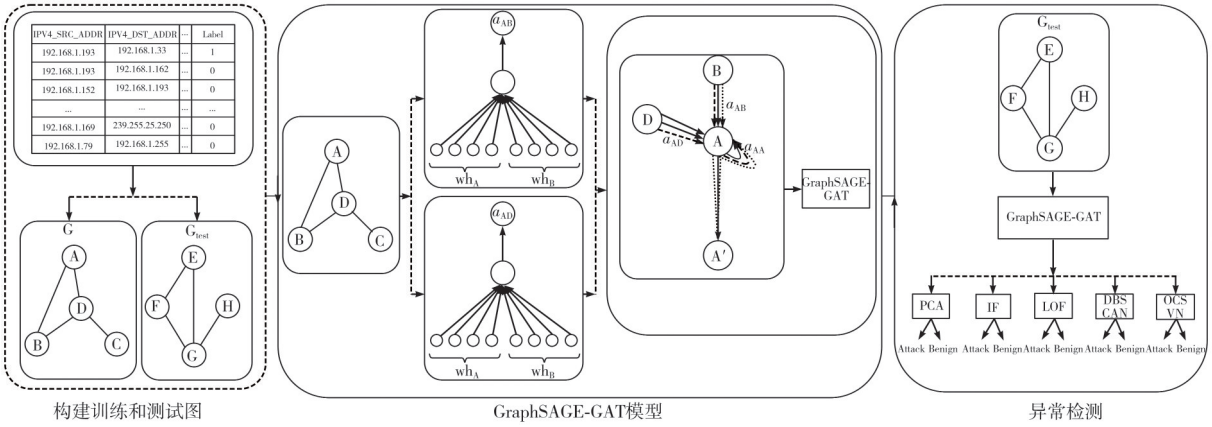

图4 GraphSAGE-GAT检测算法框架

Fig. 4 Framework of GraphSAGE-GAT detection algorithm

2.1 物联网设备关联图的构建

在进行模型训练之前，需要对网络流数据进行结构化处理，从物联网流量数据中构建设备关联关系图。本文所提方法将所有网络流数据示为一张有向图$G(V,E)$，其中，V代表图节点，E代表图边，节点特征表示为x_V。单个物联网主机设备的IP地址被建模为图节点，设备之间的通信被建模为图边。

原始网络流数据通常包含通信的发起方和接收方，同时还记录了该次通信的数据流量大小、通信端口、通信协议、流量持续时间、平均吞吐量、入包数和出包数等。物联网设备之间的通信方向是确定的，且传入字节和传出字节数是不相等的，根据IPV4_SRC_ADDR，IPV4_DST_ADDR可以构建出一条有向边，其权重为BYTES。如果两个节点之间出现了多次通信记录，则边的权重为多次通信的流量之和。

2.2 GraphSAGE-GAT算法模型介绍

GraphSAGE算法对图$G(V,E)$执行图卷积操作，对于GNN的深度，需要设置$k-hop$邻居，这意味着在每次迭代中聚合的是$k-hop$节点邻居的信息。另一方面需要选择可微聚合函数AGG_k，其中$k\in\{1,2,\cdots,K\}$表示邻居信息的聚合。在GraphSAGE中，可以应用各种聚合方法，包括平均聚合、池化聚合或者使用不同类型的神经网络(例如长短期记忆网络)。本文使用了图注意力网络和加权聚合函数。

对于每个设备节点，GraphSAGE模型以迭代方式聚集k跳深度的相邻设备节点。在每次迭代中，对一组关联设备节点进行采样，以降低算法的空间和时间复杂度。同时，利用来自采样的相互关联设备节点信息增强当前设备节点的嵌入信息表示。

在第k层，基于采样邻域$N(v)$，在设备节点v处聚集的设备节点邻域信息$h^k_{N(v)}$，表示为

$$h^k_{N(v)}=AGG_k\left(\{h^{k-1}_u,\forall u\in N_v\}\right),\tag{1}$$

式中：u为相邻设备节点；h^{k-1}_u为上一层的设备节点u处的嵌入，这些相邻的设备节点嵌入表示聚合成节点v在第k层的嵌入表示向量。

聚合过程如图5所示，首先利用聚合函数对每个设备节点的$k-hop$邻域节点特征进行聚合。然后将来自采样邻居的聚合表示与自身的节点表示h^{k-1}_v连接起来。接着应用了模型权重W^k，并将结果传递给非线性激活函数σ(例如ReLU)以获得最终的节点嵌入h^k_v，公式为

$$h^k_v=\sigma\left(W^k\cdot CONCAT\left(h^{k-1}_v,h^k_{N(v)}\right)\right)。\tag{2}$$

最后对所有节点进行归一化，依次传播K层，使用最后一层的嵌入向量作为节点的最终嵌入向量表示。公式为

$$h^k_v\leftarrow\frac{h^k_v}{\|h^k_v\|_2},\ \forall v\in V,\tag{3}$$

$$z_v \leftarrow h_v^K,\ \forall v \in V, \tag{4}$$

式中：z_v表示设备节点v的最终节点表示，代表着最终层K的设备关联图节点的嵌入表示向量。

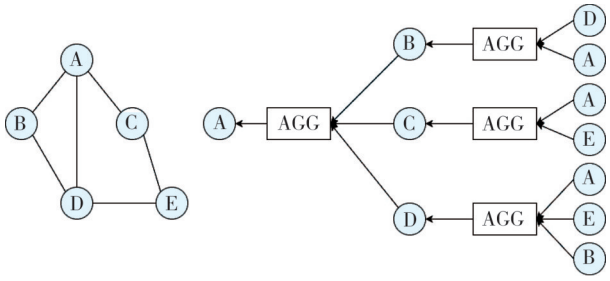

图 5　全领域采样的GraphSAGE架构

Fig. 5　GraphSAGE architecture for full neighborhood sampling

GAT使用一种注意机制来计算图中每个顶点与其相邻顶点的对应权值，并且GAT的训练依赖于相邻节点对，而不是特定的网络结构。假设图中有N个顶点，h_i是第i个顶点的特征表示，维度是F，则

$$\boldsymbol{h} = \{h_1, h_2, \cdots, h_N\},\ \boldsymbol{h}_i \in \boldsymbol{R}^F。 \tag{5}$$

对节点特征向量$\boldsymbol{h}$进行线性变换，可以得到新的特征向量$\boldsymbol{h}'_1$，维度是F'，$\boldsymbol{W}$为线性变换的矩阵，变换公式为

$$\boldsymbol{h}'_i = \boldsymbol{W}\boldsymbol{h}_i,\ \boldsymbol{W} \in \boldsymbol{R}^{F' \times F}, \tag{6}$$

$$\boldsymbol{h}' = \{h'_1, h'_2, \cdots, h'_N\},\ \boldsymbol{h}'_i \in \boldsymbol{R}^{F'}。 \tag{7}$$

通过计算第i个节点相邻节点的注意力得到其节点表示，计算公式为

$$a_{i,j} = \frac{\exp\left(Leaky\,\mathrm{Re}\,LU\left(\boldsymbol{\alpha}^{\mathrm{T}}\left[\boldsymbol{W}\boldsymbol{h}_i \| \boldsymbol{W}\boldsymbol{h}_j\right]\right)\right.}{\sum_{k \in N_i} \exp\left(Leaky\,\mathrm{Re}\,LU\left(\boldsymbol{\alpha}^{\mathrm{T}}\left[\boldsymbol{W}\boldsymbol{h}_i \| \boldsymbol{W}\boldsymbol{h}_j\right]\right)\right.}, \tag{8}$$

式中：$a_{i,j}$表示节点i和节点j之间的注意力系数；权重矩阵$\boldsymbol{W}$是一个$F \times F'$尺寸的矩阵；F表示输入节点特征的维数；F'表示该层输出节点的维数；$\boldsymbol{h}_i$和$\boldsymbol{h}_j$表示节点i和节点j的节点特征；N_i为第i个顶点的相邻顶点的集合；$\|$连接，表示张量的结合。

2.3　异常检测

训练完成后，将生成的设备关联图节点的嵌入表示向量输入到常用的异常检测分类器（PCA、IF、LOF、DBSCAN、OCSVM）执行分类任务。

PCA（Principal Component Analysis）算法是一种常用的降维方法，它也可以利用一些扩展方法改进后用于异常检测。应用传统的PCA算法将正常样本降维到较低维度，提取正常样本数据的相关矩阵。然后将该提取矩阵作为异常值检测模型的标准，通过计算样本与矩阵之间的距离或残差来判断样本是否偏离了正常值。如果样本在一定程度上偏离了正常值，就被认为是异常值。

IF（Isolation Forest）异常检测是一种利用树结构将异常样本从正常样本中分离和识别出来的算法。它通过构建隔离森林，对样本进行划分，并根据样本在树中分割的路径长度来判断样本是否为异常。

LOF（Local Outlier Factor）算法是一种用于检测离群点的算法。它通过计算每个数据点的局部离群因子来评估其在数据集中的异常程度。它的基本思想是将每个数据点与其周围邻居进行比较，衡量其在局部区域内的密度差异。如果一个数据点的邻居密度远远高于该点自身的密度，则该点可能是一个异常点。

DBSCAN（Density-Based Spatial Clustering of Applications with Noise）算法是一种基于密度的聚类算法，用于将数据点分成不同的类别，并且可以识别出异常。它的基本思想是通过寻找具有足够高密度的数据点的连通区域，将其视为一个簇，如果数据点未被分配到任何簇，就被认为是异常。

OCSVM（One-Class Support Vector Machine）算法是一种用于离群点检测的机器学习算法。它是基于支持向量机（SVM）的一类算法，通过学习正常样本的特征来建立一个边界，然后将离边界较远的数据点判定为异常。

3　实验仿真

3.1　数据集

为了评估本文提出的基于GNN的物联网入侵检测模型，在两个物联网入侵检测领域常用的公开数据集（NF-ToN-IoT-v2和NF-BoT-IoT-v2被Sarhan等标记为网络流格式的数据集[15]）上进行了测试，通过对不同分类算法的实验，验证了本文所提算法的有效性。由于数据集的原始表示尺寸较大，数据集被随机且均匀地降采样至10%。在训练中，使用了70%的数据，其余30%的样本用于测试和性能评估。在本文中，将这两个数据集简称为ToN-IoT和BoT-IoT。

NF-ToN-IoT-v2改编自Moustafa等[16-23]提出的物联网遥测数据集ToN-IoT，具有43个特征，数据流总数为16,940,496，其中10,841,027（63.99%）

是攻击样本，6,099,469(36.01%)为正常样本，具有9种攻击类型。数据集的数据分布如表1所示。

表1 ToN-IoT数据集分布

Tab. 1 Distribution of the ToN-IoT dataset

类别	样本	占比/%
Benign	6 099 469	36.005
Backdoor	16 809	0.099
DOS	712 609	4.207
DDOS	2 026 234	11.961
Injection	684 464	4.040
MITM	7 723	0.046
Password	1 153 323	6.808
Ransomware	3 425	0.020
Scanning	3 781 419	22.322
XSS	2 455 020	14.492
合计	16 940 496	100

NF-BoT-IoT-v2是基于BoT-IoT的网络流格式[24-29]，其中有4个攻击类别，数据流总数为37 763 497，其中37 628 460(99.64%)为攻击样本，135 037(0.358%)为正常样本。数据集的数据分布如表2所示。

表2 BoT-IoT数据集分布

Tab. 2 Distribution of the NF-BoT-IoT-v2 dataset

类别	样本	占比/%
Benign	135 037	0.358
Reconnaissance	2 620 999	6.941
DOS	16 673 183	44.152
DDOS	18 331 847	48.544
Theft	2 431	0.006
合计	37 763 497	100

3.2 数据预处理

在训练之前，需要对数据进行清洗工作。图6展示了将每个数据集预处理成训练图和测试图的过程。

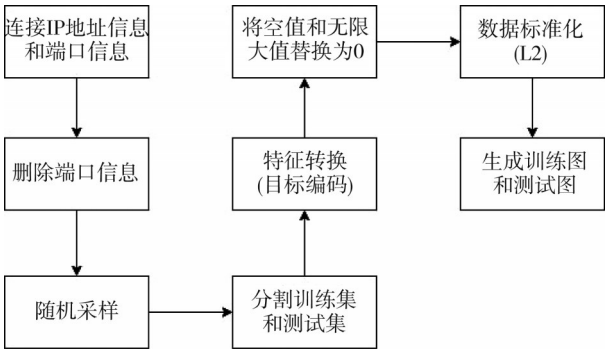

图6 预处理成训练图和测试图的过程

Fig. 6 Pre-processing and graph generetion

对于本研究中的实验，首先将每个流记录中IP地址信和端口信息连接到一起，然后移除源端口和目的端口信息。数据被下采样至其原始大小的10%[30]，在下采样流上，样本被分成训练集和测试集。

使用目标编码技术将数据集中的非数字特征转换为整型数据，进行目标编码的非数字特征包含TCP_FLAGS，L7_PROTO，PROTOCOL，CLIENT_TCP_FLAGS，SERVER_TCP_FLAGS，ICMP_TYPE，ICMP_IPV4_TYPE，DNS_QUERY_ID，DNS_QUERY_TYPE，FTP_COMMAND_RET_CODE，此过程中产生的任何空值或无限值都将替换为值0。

在生成训练图和测试图之前，使用L2归一化方法对训练和测试集进行归一化。L2范数归一化处理操作是对向量$\boldsymbol{X}$的每个维度数据$x_1,x_2,\cdots,x_n$都除以$\|x\|_2$得到一个新向量，计算公式为

$$\boldsymbol{X}_2=\left(\frac{x_1}{\|x\|_2},\frac{x_2}{\|x\|_2},\cdots,\frac{x_n}{\|x\|_2}\right)。\tag{9}$$

最后将训练集和测试集转换成双向图表示，然后将处理后的数据信息作为节点特征向量。

3.3 训练参数设置

GraphSAGE-GAT模型算法需通过设置合适的参数使模型达到可控的拟合效果，本文为使模型的性能达到最佳进行了多轮调参，使最终的实验结果达到较优水平。GraphSAGE-GAT模型参数设置见表3。模型中的图神经网络模块包含一个卷积层和一个注意力层，节点隐藏层维度取128，Dropout层取0.2，损失函数采用BCE，激活函数采用ReLU，优化器采用Adam，学习率取0.01。

表3 GraphSAGE-GAT模型参数设置

Tab. 3 Parameter setings of GraphSAGE-GAT model

参数	参数值
k	1
Attention层数	1
隐藏层大小	128
训练轮次	4 000
Dropout	0.2
损失函数	BCE
激活函数	ReLU
优化器	Adam
学习率	0.001

3.4 评价标准

本文将攻击样本定义为正例样本，使用混淆

矩阵计算得出的4个评价指标来评估机器学习模型，混淆矩阵如表 4 所示。

表 4　混淆矩阵

Tab. 4　Confusion matrix

样本类型	预测为正例样本	预测为负例样本
正例样本	TP(True Positive)	FN(False Negative)
负例样本	FP(False Positive)	TN(True Negative)

通过混淆矩阵，可以计算得出本文的4个评价指标：准确率、精准率、召回率和$F1$，计算公式分别为

$$R_{acc}=\frac{N_{TP}+N_{TN}}{N_{TP}+N_{TN}+N_{FP}+N_{FN}},\tag{10}$$

$$R_{pre}=\frac{N_{TP}}{N_{TP}+N_{FP}},\tag{11}$$

$$R_{rec}=\frac{N_{TP}}{N_{TP}+N_{FN}},\tag{12}$$

$$F1=\frac{2\times R_{pre}\times R_{rec}}{R_{pre}+R_{rec}},\tag{13}$$

式中：R_{acc}为模型的准确率；R_{pre}为模型的精确率；R_{rec}为模型的召回率；$F1$是R_{pre}和R_{rec}两个指标计算得出的值，其范围为[0，1]，1表示模型性能最好，0表示模型性能最差；N_{TP}为正例样本预测为正例样本的数量；N_{TN}为负例样本预测为负例样本的数量；N_{FP}为负例样本预测为正例样本的数量；N_{FN}为正例样本预测为负例样本的数量。

3.5　实验结果分析

考虑GraphSAGE-GAT与入侵检测分类算法相结合的模型性能表现，主要进行了3个实验。实验1将物联网入侵检测领域常用的ToN-IoT数据集的全部特征和使用GraphSAGE-GAT训练生成的嵌入向量分别输入5个分类器，对每个算法进行网格搜索，以确保最佳的参数调整，从而比较不同算法的性能表现。为验证模型的鲁棒性和泛化性，实验2将ToN-IoT数据集替换为BoT-IoT数据集，确保实验环境一致，比较本研究所提出的模型在物联网入侵检测数据中应用的效果。实验3将数据集替换为其他入侵检测数据集，进一步验证本研究所提出的模型的检测性能。

GraphSAGE-GAT是用于生成图嵌入的方法，选择5种经典异常检测算法作为分类方法的对比模型，并在ToN-IoT和BoT-IoT数据集上进行实验，不同分类算法的性能指标对比如表 5 所示。

从表 5 可以看出，本研究提出的模型在两个数据集中具有较高的准确率，精准率和$F1$评分，综合性能表现良好。DBSCAN在两个数据集上都取得了最好的性能，准确率、精确率、召回率和$F1$值都很高。而PCA、IF、LOF和OCSVM的表现则相对较差，性能差异也较大，但仍具有较高的准确率和精确率，也具有一定的分类能力。需要特别注意的是，在BoT-IoT数据集上，IF算法的分类表现明显不如其他 4 种算法，各分类算法的召回率较ToN-IoT数据集也明显降低，分析原因是各数据集样本分布不平衡因素导致的。

表 5　不同分类算法的性能指标对比

Tab. 5　Comparison of performance indicators of different classification algorithms

数据集名称	分类算法	评估指标			
		R_{acc}/%	R_{pre}/%	R_{rec}/%	$F1$
ToN-IoT	PCA	96.95	97.23	97.80	0.9751
	IF	94.76	95.47	96.10	0.9578
	LOF	96.96	97.16	97.53	0.9736
	DBSCAN	97.25	97.41	97.72	0.9756
	OCSVM	95.75	94.94	95.27	0.9510
BoT-IoT	PCA	98.04	98.33	94.24	0.9624
	IF	96.23	95.70	92.64	0.9415
	LOF	97.52	99.62	94.25	0.9686
	DBSCAN	98.62	98.18	95.88	0.9702
	OCSVM	97.67	95.19	94.99	0.9509

图 7 展示了在ToN-IoT数据集中不同分类算法的ROC曲线对比图。

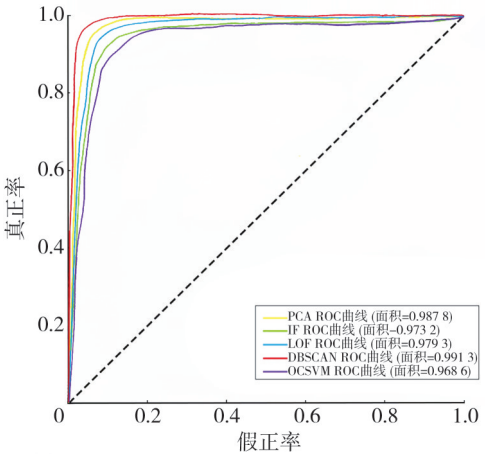

图 7　ToN-IoT数据集中不同算法的ROC曲线对比图

Fig. 7　Comparison of ROC curves of different classification algorithms in the ToN-IoT

二分类问题中，ROC曲线及其下面积AUC值常用于评估分类算法的性能和平衡性。通常情况下，ROC曲线位于$y=x$的上方，AUC取值范围默认为[0.5,1]，数值越大表示算法的预测准确率越高。同时，ROC曲线对正负样本比例不敏感，能够客观反映算法的分类性能。从图 7 可以看出，基于GNN的

物联网入侵检测算法结合DBSCAN聚类算法的模型具有更饱满的ROC曲线，并且靠近左上角位置，该算法曲线完全包住了PCA及其他分类算法的曲线。此外，与其他算法相比，DBSCAN的AUC值比PCA高0.003 5，比IF高0.018 1，比LOF高0.012，比OCSVM高0.022 7，是所有算法中最高的。综合来看，结合DBSCAN聚类算法可以更准确地对数据集中更多的少数样本进行分类，从而实现更好的分类效果。图8展示了在BoT-IoT数据集中根据算法模型测试结果绘制的ROC曲线，每个算法的AUC值都达到了0.94以上，这进一步证明了算法模型的优越性。

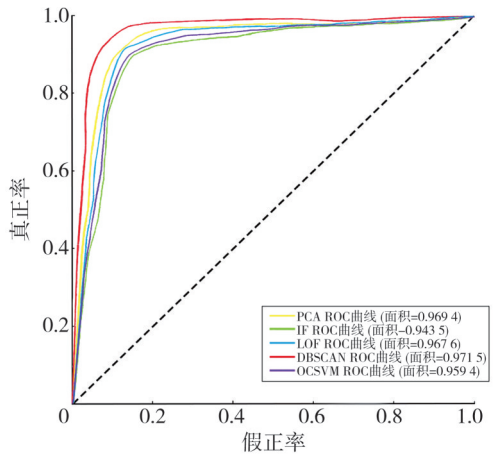

图8 BoT-IoT数据集中不同算法的ROC曲线对比图

Fig. 8 Comparison of ROC curves of different classification algorithms in the BoT-IoT

为进一步验证GraphSAGE-GAT模型的性能，本文对基于同源数据集的现有研究方法进行了仿真，并将模型进行了对比。

Extra Trees Classifier是一种集成学习的分类器算法，它的每棵决策树都是使用原始训练样本构建的，每棵树都采用一个随机特征子集，该子集包含k个特征。每棵决策树都从这个特征子集中选择最佳特征，并根据基尼系数来划分数据。由于特征样本的随机性，导致了多个不相关的决策树被构建出来。最后，收集这些不相关决策树的结果，并将它们聚合起来，得到最终的分类结果。

E-GraphSAGE是GNN的一种扩展形式，它基于GraphSAGE模型，并在节点表示学习过程中引入了边信息，以增强图的表示学习能力。其核心思想是通过聚合邻居节点的特征来更新目标节点的表示。同时，通过采样邻居节点的子集并进行特征聚合，以累积不同领域的信息，从而获取更全面和丰富的节点表示。

IDS-HN是一种基于无监督联邦学习的入侵检测方法，它结合了最先进的异常检测算法(IF和LOF)和深度自编码器，以实现对异构网络的泛化。该方法的关键是通过能量流分类器，在堆叠设置中进行训练。其核心是深度自编码器，由编码器网络、瓶颈段和解码器网络三部分组成。在训练过程中，采用联邦学习方法，利用多个参与方的分布式数据进行模型训练，从而保护数据隐私。

GMM-OCSVM是一种将高斯混合模型(GMM)与单一支持向量机(OCSVM)结合的异常检测方法。它首先使用自编码器来提取正常数据的代表性特征，这可以减轻高维数据对异常检测的影响，并捕捉正常数据的潜在特性。然后使用OCSVM算法对正常数据进行建模，找到一个最小的超平面将正常数据从原点分离开来。将新的数据样本输入到GMM模型中，计算其在GMM中的概率密度值作为异常分数。同时，综合考虑OCSVM的异常分数和GMM的概率密度值，根据综合的异常分数判断样本是否为异常数据。

表6列出了本研究的最佳模型(DBSACN聚类算法)与其他入侵检测方法的4个性能评估指标的对比结果。

表6 本研究方法与其他模型的效果比较

Tab. 6 Comparison of this research method and other models

数据集名称	方法模型	评估指标			
		R_{acc}/%	R_{pre}/%	R_{rec}/%	$F1$
ToN-IoT	本方法模型	97.25	97.41	97.72	0.9756
	Extra Trees Classifier[15]	93.34	94.25	92.68	0.9346
	E-GraphSAGE[31]	96.94	97.32	92.08	0.9463
	IDS-HN[32]	95.42	96.53	94.79	0.9565
BoT-IoT	本方法模型	98.62	98.18	95.88	0.9702
	Extra Trees Classifier[15]	95.64	96.35	92.19	0.9422
	E-GraphSAGE[31]	96.51	98.07	94.36	0.9618
	GMM-OCSVM[33]	95.36	92.21	94.15	0.9394

由表 6 可知，在两个入侵检测数据集中，与其他研究方法相比，本研究方法在精确度与检测率两方面都有较好的表现，并且有更高的$F1$评分。基于GNN物联网入侵检测算法与其他经典学习分类器相结合构建模型时，与DBSCAN聚类算法相结合是最佳的分类模型，因为它在分类准确率、精确率、召回率和$F1$分数方面都取得了最高的准确性，与其他分类方法不同，DBSCAN聚类算法可扩展性高，用法灵活，且易于评估。

为了充分说明本文提出方法的有效性，对其它数据集分别做分类实验，获得的实验结果如表 7 所示。从实验结果可以看出，GraphSAGE-GAT在NSL-KDD数据集、CICIDS2018数据集和UNSW-NB15数据集上均可获得较高的准确率和精确率，进一步验证了GraphSAGE-GAT模型的性能。

表 7　不同数据集的分类对比结果

Tab. 7　Classification and comparison results of different data sets

数据集	R_{acc}/%	R_{pre}/%	R_{rec}/%
NSL-KDD	99.34	99.30	99.35
CICIDS2018	97.96	98.32	98.04
UNSW-NB15	97.62	97.56	97.74

综上所述，基于GNN的物联网入侵检测算法对数据处理后可以有效消除噪声属性且减少了冗余，可以得到低度数据的优良特征子集，并构建物联网设备之间的关联关系。实验结果表明，本文模型获得了更高的准确率，且泛化性能相对较好。

4　结 论

本文提出了一种基于GNN模型的物联网入侵检测模型GraphSAGE-GAT，该模型首先根据物联网设备通信构建网络设备关联关系拓扑结构图，然后根据该关联图获得更优的网络设备关联图的节点嵌入表示向量，最后根据该节点嵌入表示判断该节点是否被攻击。在数据集ToN-IoT和数据集BoT-IoT上的实验结果表明，本文模型的网络入侵检测表现均优于现有模型，获得了更高的准确率、精确度、召回率和$F1$分数。但由于流量数据繁杂，在提取特征项时，可能存在丢失现象。所以，未来将重点关注如何减少数据丢失以及对于GNN同其他深度学习算法复合的研究，以期获得更好的分类效果。

参考文献：

[1] HASSAN W H. Current research on Internet of Things (IoT) security: A survey[J]. Computer networks, 2019, 148: 283-294.

[2] LIU T, LI Z, LONG H, et al. NT-GNN: Network traffic graph for 5G mobile iot android malware detection[J]. Electronics, 2023, 12(4): 789.

[3] NGUYEN T N, NGO Q D, NGUYEN H T, et al. An advanced computing approach for IoT-botnet detection in industrial Internet of Things[J]. IEEE Transactions on Industrial Informatics, 2022, 18(11): 8298-8306.

[4] CAVILLE E, LO W W, LAYEGHY S, et al. Anomal-E: A self-supervised network intrusion detection system based on graph neural networks [J]. Knowledge-Based Systems, 2022, 258: 110030.

[5] BEDIYA A K, KUMAR R. A novel intrusion detection system for internet of things network security[J]. Journal of Information Technology Research, 2021, 14(3): 20-37.

[6] NIMBALKAR P, KSHIRSAGAR D. Feature selection for intrusion detection system in Internet-of-Things (IoT) [J]. ICT Express, 2021, 7(2): 177-181.

[7] KUMAR V, DAS A K, SINHA D. UIDS: A unified intrusion detection system for IoT environment [J]. Evolutionary intelligence, 2021, 14: 47-59.

[8] ZHANG X, ZHENG X, WU D D. Attacking DNN-based intrusion detection models [J]. IFAC-PapersOnLine, 2020, 53(5): 415-419.

[9] WU P, GUO H. LuNET: A deep neural network for network intrusion detection[C]//2019 IEEE Symposium Series On Computational Intelligence (SSCI). IEEE, 2019: 617-624.

[10] HAMILTON W, YING Z, LESKOVEC J. Inductive representation learning on large graphs [J]. Advances in Neural Information Processing Systems, 2017, 30: 1024-1034.

[11] VELICKOVIC P, CUCURULL G, CASANOVA A, et al. Graph attention networks[DB/OL]. (2017-10-30)[2023-07-09]. http://arxiv.org/abs/1710.10903.

[12] SHYU M L, CHEN S C, SARINNAPAKORN K, et al. A novel anomaly detection scheme based on principal component classifier[R]. Coral Gables: University of Miami, 2003.

[13] ARP D, QUIRING E, PENDLEBURY F, et al. Dos and dont's of machine learning in computer security[DB/OL]. (2020-10-19)[2023-07-09]. https://

arxiv.org/abs/2010.09470.
[14] ALAZZAM H, SHARIEH A, SABRI K E. A lightweight intelligent network intrusion detection system using ocsvm and pigeon inspired optimizer[J]. Applied Intelligence, 2022, 52(4): 3527-3544.
[15] SARHAN M, LAYEGHY S, PORTMANN M. Towards a standard feature set for network intrusion detection system datasets[J]. Mobile Networks and Applications, 2022, 27: 1-14.
[16] MOUSTAFA N. A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets[J]. Sustainable Cities and Society, 2021, 72: 102994.
[17] BOOIJ T M, CHISCOP I, MEEUWISSEN E, et al. ToN_IoT: The role of heterogeneity and the need for standardization of features and attack types in IoT network intrusion data sets[J]. IEEE Internet of Things Journal, 2021, 9(1): 485-496.
[18] ALSAEDI A, MOUSTAFA N, TARI Z, et al. TON_IoT telemetry dataset: A new generation dataset of IoT and IIoT for data-driven intrusion detection systems[J]. IEEE Access, 2020, 8: 165130-165150.
[19] MOUSTAFA N, KESHKY M, DEBIEZ E, et al. Federated TON_IoT windows datasets for evaluating AI-based security applications[C]//2020 IEEE 19th international conference on trust, security and privacy in computing and communications (TrustCom). IEEE, 2020: 848-855.
[20] MOUSTAFA N, AHMED M, AHMED S. Data analytics-enabled intrusion detection: Evaluations of ToN_IoT linux datasets[C]//2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom). IEEE, 2020: 727-735.
[21] MOUSTAFA N. New generations of internet of things datasets for cybersecurity applications based machine learning: TON_IoT datasets[C]//Proceedings of the eResearch Australasia Conference, Brisbane, Australia. 2019: 21-25.
[22] MOUSTAFA N. A systemic IoT - fog - cloud architecture for big-data analytics and cyber security systems: A review of fog computing[J]. Secure Edge Computing, 2021: 41-50.
[23] ASHRAF J, KESHK M, MOUSTAFA N, et al. IoTBoT-IDS: A novel statistical learning-enabled botnet detection framework for protecting networks of smart cities[J]. Sustainable Cities and Society, 2021, 72: 103041.
[24] KOROIOTIS N, MOUSTAFA N, SITNIKOVA E, et al. Towards the development of realistic botnet dataset in the internet of things for network forensic analytics: Bot-iot dataset[J]. Future Generation Computer Systems, 2019, 100: 779-796.
[25] KORONIOTIS N, MOUSTAFA N, SITNIKOVA E, et al. Towards developing network forensic mechanism for botnet activities in the IoT based on machine learning techniques[C]//Mobile Networks and Management: 9th International Conference, MONAMI Melbourne, Australia, Springer International Publishing, 2018: 30-44.
[26] KORONIOTIS N, MOUSTAFA N, SITNIKOVA E. A new network forensic framework based on deep learning for internet of things networks: A particle deep framework[J]. Future Generation Computer Systems, 2020, 110: 91-106.
[27] KORONIOTIS N, MOUSTAFA N. Enhancing network forensics with particle swarm and deep learning: The particle deep framework[DB/OL]. (2020-05-02)[2023-07-09]. https://arxiv.org/abs/2005.00722.
[28] KORONIOTIS N, MOUSTAFA, SCHILIRO F, et al. A holistic review of cybersecurity and reliability perspectives in smart airports[J]. IEEE Access, 2020, 8: 209802-209834.
[29] KORONIOTIS N. Designing an effective network forensic framework for the investigation of botnets in the internet of things[D]. New South Wales: UNSW Sydney, 2020.
[30] CHEN P, GUO Y, ZHANG J, et al. A novel preprocessing methodology for dnn-based intrusion detection[C]//2020 IEEE 6th International Conference on Computer and Communications (ICCC). IEEE, 2020: 2059-2064.
[31] LO W W, LAYEGHY S, SARHAN M, et al. E-graphsage: A graph neural network based intrusion detection system for iot [C]//NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium. IEEE, 2022: 1-9.
[32] DE CARVALHO BERTOLI G, JUNIOR L A P, Saotome O, et al. Generalizing intrusion detection for heterogeneous networks: A stacked-unsupervised federated learning approach[J]. Computers & Security, 2023, 127: 103106.
[33] WANG C, SUN Y, LV S, et al. Intrusion detection system based on one-class support vector machine and gaussian mixture model[J]. Electronics, 2023, 12(4): 930.