文章编号:1671-4598(2025)12-0051-07 DOI:10.16526/j.cnki.11-4762/tp.2025.12.007 中图分类号:TP393 文献标识码:A

基于自适应模块和多生成器的车联网入侵检测方法

王 荔,何立明,李 茹
(长安大学 信息工程学院,西安 710064)

摘要:入侵检测是车联网安全关键的防护措施,但仍面对两大局限:一是在实际中网络攻击是不断变化的,导致训练好的模型难以检测最新的未知攻击,二是网络流量中的良性数据与攻击数据是不平衡的;鉴于此提出了ATGCB入侵检测模型,引入自适应模块和多生成器对抗网络;自适应模块是以无监督的方式检测数据漂移的聚类模型,经过聚类识别异常分布样本,触发生成模块生成对应伪数据,进而增量更新分类器提高泛化性;多生成器对抗网络采用高维特征提取器与并行多个生成器同时生成多类型样本重叠低的伪数据,扩充少数类进而平衡数据集;经过在数据集CICIDS-2017和CSE-CICIDS-2018进行实验的结果表明,方法在平衡数据后模型命中率达到97.56%,在面对漂移数据后命中率达到91.12%,由此证明方法在车联网入侵检测中更加适用。

关键词:漂移数据;多生成器对抗网络;入侵检测;平衡数据集;注意力机制

Intrusion Detection Method for Internet of Vehicles Based on Adaptive Module and Multi-Generator

WANG Li, HE Liming, LI Ru
(School of Information Engineering, Chang'an University, Xi'an 710064, China)

Abstract: Intrusion detection is a critical security measure for the Internet of Vehicles (IoV). However, it still faces two major limitations: first, network attacks are constantly evolving in practice, making it difficult for pre-trained models to detect the latest unknown attacks; second, there is an imbalance between benign data and attack data in network traffic. To address these issues, this paper proposes an intrusion detection model named ATGCB (Adaptive and Tmg-GAN-based Clustering for Intrusion Detection), which incorporates an adaptive module and a multi-generator adversarial network. The adaptive module is a clustering model that detects data drift in an unsupervised manner. It identifies abnormally distributed samples through clustering, triggers the generation module to produce corresponding pseudo-data, and then incrementally updates the classifier to improve generalization performance. The multi-generator adversarial network adopts a high-dimensional feature extractor and multiple parallel generators to simultaneously generate multiple types of pseudo-data with low sample overlap, thereby expanding minority classes and balancing the dataset. Experimental results on the CICIDS-2017 and CSE-CICIDS-2018 datasets show that the proposed method achieves a model hit rate of 97.56% after data balancing and 91.12% when dealing with drifted data, which demonstrates that this method is more applicable to intrusion detection in the IoV.

Keywords: drifting data; multi-generator adversarial network; intrusion detection; balance the dataset; attention mechanism

收稿日期:2025-06-20; **修回日期**:2025-07-21。
基金项目:国家自然科学基金项目(51308058)。
作者简介:王 荔(2000-),男,硕士。
通讯作者:何立明(1978-),男,博士,副教授。
引用格式:王 荔,何立明,李 茹.基于自适应模块和多生成器的车联网入侵检测方法[J].计算机测量与控制,2025,33(12):51-57,66.

0 引言

随着车联网应用的深入拓展，海量实时通信数据不断涌现，对外通信接口数量激增，这极大地扩大了车辆系统的攻击面[1]。与传统网络相比，车联网架构更为复杂且异构，其开放的通信环境，加上车载设备本身存在的安全缺陷，使得系统面临多维度的安全威胁[2]。相关研究表明，车联网攻击面比传统 IoT 系统扩大了 37%。作为网络安全防护的关键防线，车联网入侵检测系统必须突破传统方法的设计局限。

伴随人工智能技术的进步，机器学习方法在入侵检测领域的应用愈发普遍。通过对大量历史数据的学习，机器学习模型能够自动探寻并识别复杂的攻击模式，逐步提升检测的准确性与效率。传统的机器学习技术在入侵检测领域中获得了广泛的应用，其中主要涵盖随机森林、支持向量机、自适应增强、决策树以及 k-最近邻[3]等方法。此类方法在早期应对少量攻击有很好性能。

深度学习在车联网入侵检测领域的研究愈发广泛，备受关注。集成学习方法将 5 种基于树结构的监督分类器堆叠集成用于检测已知 DoS 攻击[4]，CNN-BiSRU[5]，CNES 模型[6]，VGG19 模型，Xception 模型，Inception 模型，3 个模型集成 CS-IDS 模型[7]，LSTM 模型[8]，Graph Convolutional Network[9]，COA-GWO[10]，FPA- COA-ANN 混合模型结合了 FPA 和 COA 的优化能力和 ANN 的预测能力[11]，H-KD IDS[12]以上深度学习模型在已知训练数据集上准确度有很大的提升，但其都需要海量的标记数据来进行训练，且没有应对新型攻击的方式。

近年来，强化学习和迁移学习为入侵检测的发展带来了巨大的机遇，CNN-DT[13]，KFRNN[14]，DRL[15]，CNN-ATTENTION[16]，NFIoT-GATE-DTL IDS[17]，ResNet[18]以上入侵检测的方法对未知攻击的检测有一定的提升，减少了大规模标记数据的依赖。但在面对现实情况下的检测还是存在对未知攻击检测不足的问题。

对于数据集平衡性问题的处理。对抗生成网络是一种深度生成模型[19]，通过对抗学习的思想来拟合数据分布，其中包括两个结构生成器和判别器，生成器用于生成类似于真实的分布的伪样本，而采样器用于确定输入样本是否是真实的[20]。该模型通过交替对抗训练进行训练，自生成对抗网络（GAN）被提出以来最终生成器可以生成近似拟合真实的样本的伪样本[21]。便在诸多领域展现出强大的应用潜力，被广泛地应用于图像生成、语音合成以及文本生成等方面。

上述方法极大提升了入侵检测技术，但随着科技进步，网络环境短期内呈现静态数据分布，与现实世界场景的动态特性不符[22]。流量攻击特征分布会发生变化，与旧数据特征分布不一致，此即数据漂移现象，如图 1 所示，数据漂移[23]是机器学习模型部署后面临的核心挑战之一，指模型在实际应用中接收到的实时数据的统计特性（如特征分布、数据模式或结构）逐渐偏离训练阶段所使用的数据分布，导致模型性能随时间持续下降的现象。具体到车辆入侵检测系统，机器学习的应用面临适应数据漂移、泛化模型两大挑战[24]。此外，收集真实攻击样本难度极高，致使攻击数据远少于良性数据，产生数据集不平衡问题。而现有的处理数据不平衡方法，面对少量攻击样本时，均存在种种缺陷。不平衡的样本数据会大大影响模型的性能。从而研究提出了新的车联网入侵检测方法，它由自适应、多生成器对抗网络和检测 3 个模块组成，创造性地解决了当前车联网入侵检测面对的两大问题：一是动态适应性，通过在线学习和生成对抗机制应对新型攻击；二是数据增强，高质量合成数据缓解标注不足问题。为实时性和泛化性的问题提出了新的方法，为解决入侵检测问题带来新突破。

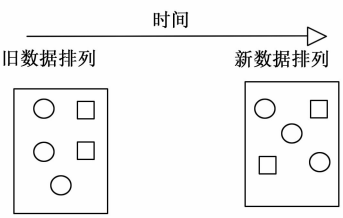

图 1 数据漂移

1 入侵检测算法

研究提出的 ATGCB（Adaptive and Tmg-GAN-based Clustering for Intrusion Detection）入侵检测算法如图 2 所示，主要步骤如下：

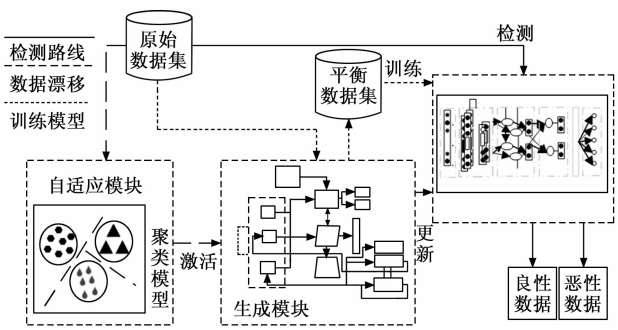

图 2 总体框架

1）检测模型的训练，如图中直线路径，将原始数据经过数据预处理，通过多生成器对抗网络生成少数类攻击数据达到平衡数据的效果，最后将平衡的数据输入检测模型训练。

2）对抗漂移数据，如图中虚线所示，将数据输入自适应模块，通过聚类模型将不同的特征聚类到一块，

当发现未知的攻击，多生成器对抗网络将被激活以生成数据来训练新分类器。

3）入侵检测，如图中省略线所示，将原始数据通过数据预处理，然后将数据输入检测模型分辨出良性与恶性数据。

1.1　自适应模块

研究针对无监督场景下的数据漂移问题，采用了一种基于自适应滑动窗口与聚类算法的检测机制。该机制通过维护一个动态更新的数据窗口，持续存储并刷新网络中的最新流量数据。系统每隔固定时间间隔，自动调用聚类模型对当前窗口内的数据进行无监督评估，以此监测数据分布的变化。如图 3 所示，聚类模型通过计算样本间的特征相关性（如欧氏距离），将数据划分为 K 个互斥类别，自动揭示其内在分布结构。研究选用基于分区的 K-Means 算法实现该过程：首先随机初始化 K 个聚类中心（质心），随后迭代执行两步优化——将每个样本分配至距离最近的质心所属簇，再基于簇内样本均值重新计算质心位置。此过程持续优化目标函数：最小化所有样本到其所属簇质心的平方距离总和（即最小化类内方差），直至质心位置收敛或达到最大迭代次数。聚类算法的核心优势在于完全无需标签即可感知局部分布模式的变化。通过持续监控聚类结构的动态特征（如质心偏移量、新生簇涌现频率或离群样本比例），系统能够敏锐识别数据漂移现象。一旦检测到显著分布偏移（如质心移动超过阈值或出现全新簇类别），自适应模块立即向生成模块和检测模块发送激活信号：生成模块基于新数据分布重构特征空间或合成适配样本，检测模块则动态调整分类器参数或触发增量学习。这种“监测－响应”闭环机制有效保障了入侵检测系统对网络环境持续演化的适应将聚类模型表示为 $f_c: X \to [0,1]^k$ 其中 C 表示聚类中心，X 表示流量，$Y=\{1,2,\cdots,k\}$ 表示伪标签．用 $L(X_i)=\|X_i-C(Y_{X_i})\|^2$ 表示样本 X_i 与其分配的聚类的中心 $C(Y_{X_i})$ 之间的距离。算法中涉及的两个评价指标可以定义如下：

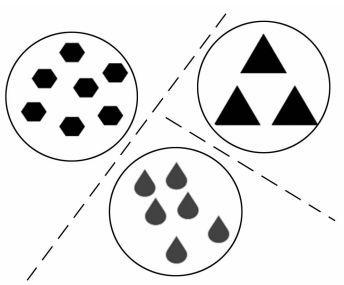

图 3　聚类模型的分类

1）置信区间 $[0 \sim C]$ 定义了所有样本的聚类结果被认为是可靠的范围；超出此区间的样本被认为是离群值。

$$C = E_{X_j \sim D_{old}}[L(X_i)] + \alpha \cdot S_{X_j \sim D_{old}}[L(X_j)] \quad (1)$$

其中：$E_{X_j \sim D_{Old}}[L(X_i)]$ 和 $S_{X_j \sim D_{old}}[L(X_j)]$ 表示从旧分布 D_{old} 中的样本 X_j 到聚类中心 $C(X_j)$，$C = E_{X_j \sim D_{old}}[L(X_i)] + \alpha \cdot S_{X_j \sim D_{old}}[L(X_j)]\alpha$ 表示超参数。

2）离群值比率（OR）定义为新分布 D_{new} 中离群值数量与旧分布 D_{old} 中离群值数量的加权比率：

$$OR = \frac{\frac{1}{d^-}\sum_k^{d^-}\{L(X_K) \notin [0,CI]\}}{\frac{1}{d}\sum_j^{d}\{L(X_j) \notin [0,CI]\}} \quad (2)$$

其中：d^- 表示新分布 D_{new} 中的样本数量，d 表示旧分布 d_{old} 中的样本数量，并且当且仅当 $\{L(X_j) \notin [0, CI]\}$ 条件为真时 $\{L(X_j) \notin [0,CI]\}=1$。因此，当 OR 超过阈值 T 时，定义在新数据分布 D_{new} 中已经发生数据漂移，并且该漂移已经发生在多个局部分布中。在检测到数据漂移后，生成模块将被激活以生成旧数据来训练新分类器。K 和 α 是该模块中的两个关键超参数，它们的值直接影响其性能。

在聚类算法中，若将 K 值设为真实标签数量 L，其模型拟合效果会逊色于将 K 值设为 L 的数倍。我们的目标是使每个聚类都符合高斯分布，即聚类后每个聚类内的样本均遵循高斯分布。基于此，存在一个合理假设：具有相同标签的数据在特征层面会呈现多种分布，而增加聚类数量能够捕捉到这种多分布现象。不过，为 K 选择过大的值并非最优方案。当 K 值过大时，每个聚类可能仅包含少量样本，这会导致模型出现过拟合问题，同时降低泛化能力。

α 的值直接影响旧数据分布中离群值的数量，进而影响 OR 和该模块的检测性能。我们将通过数学推导过程探索离群值数量、OR 和 α 之间的关系。

首先，假设第 i 个聚类的数据分布如下：其中 μ_i 和 σ_i 分别表示分布 D_i 的平均值和标准差：

$$D_i(x) = \frac{1}{\sqrt{2\pi}\sigma_i}\exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right) \quad (3)$$

其次：我们假设当 $x > N_i$ 时，所有样本都是离群值。离群值的数量可以通过 $D_i(x)$ 来估计：

$$\text{Outliers} = \sum_{i=1}^{|C|} 2d_i \int_{Ni}^{\infty} D_i(x)\mathrm{d}x \quad (4)$$

$|c|$ 表示簇的数量，假设 $t=(x-\mu_i)/\sigma_i$，则 t 服从标准正态分布：

$$\int_{N_i}^{\infty} D_i(x)\mathrm{d}x = \int_{(N_i-\sigma_i)/\sigma_i}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right) = 1-\Phi\left(\frac{N_i-\mu_i}{\sigma_i}\right) \quad (5)$$

$$\Phi\left(\frac{N_i-\mu_i}{\sigma_i}\right) = \int_{-\infty}^{(N_i-\mu_i)/\sigma_i} \frac{1}{\sqrt{2\pi}}\exp(-\frac{t^2}{2})\mathrm{d}t \quad (6)$$

根据高斯误差函数，可以表示为：

$$\Phi\left(\frac{N_i-\mu_i}{\sigma_i}\right) \approx \frac{1}{2}\left[1+erf\left(\frac{N_i-\mu_i}{\sqrt{2}\sigma_i}\right)\right] \quad (7)$$

$$erf\left(\frac{N_i-\mu_i}{\sigma_i}\right)=\frac{2}{\sqrt{\pi}}\int_0^{(N_i-\mu_i)/\sigma_i}e^{-t^2}\,\mathrm{d}t \tag{8}$$

其中：$erf(x)$ 表示高斯误差函数。其可用级数近似表示为：

$$erf(x)=1-\frac{e^{-x^2}}{n}\left[1-\frac{1}{2x^2}+\frac{1}{4x^2}+o\left(\frac{1}{x^6}\right)\right] \tag{9}$$

设 $N_i=C_i+CI$，C_i 是第 i 个簇的中心。根据上述式子可得离群值的数量可以表示为：

$$Outliers(\alpha)\approx\sum_{i=1}^{c}d_i\frac{e^{-(A\alpha+B)^2}}{(A\alpha+B)\sqrt{\pi}} \tag{10}$$

我们假设在 D_i 的置信区间之外存在一个新的数据分布 D_j。因此，随着 α 从 0 开始增加，置信区间逐渐扩大，导致 D_i 和 D_j 从非重叠状态过渡到重叠状态，最终完全包含。考虑到非重叠状态，D_j 中的所有样本都是离群值，OR_1 可以表示为：

$$OR_1(\alpha)=\frac{\frac{1}{\mathrm{d}},d'}{\frac{1}{d}Outliers(\alpha)}=\frac{d}{Outliers(\alpha)} \tag{11}$$

考虑到重叠或包含状态，D_i 几乎没有异常值，我们将其设置为 1。与此同时，D_j 中的离群值数量也将逐渐减少。同样，它的下降趋势应该类似于离群值（α），我们将其表示为 $IDs(\alpha)$。因此，OR_2 可以表示为：

$$OR_2(\alpha)=\frac{\frac{1}{d},ID_S(\alpha)}{\frac{1}{d}\times 1}=\frac{d}{d},ID_s(\alpha) \tag{12}$$

由上述公式可知，随着 α 的增大，离群值呈现指数衰减趋势。异常率（OR）起初会随 α 的增长而上升，但当 α 过度增大时，异常率也会随之呈指数衰减。值得注意的是，α 在异常值拐点附近的取值可能是最优的，此时能够有效地检测数据漂移。

1.2 多生成器对抗网络

如图 4 所示，多生成器由 4 个基本子结构组成：生成器、判别器、分类器和特征提取器。生成器用于生成伪样本数据；判别器用于判断样本是真实的还是伪造的；分类器用于判断样本类别；特征提取器用于在分类层之前提取高维特征。且使用一种多生成器结构，它可以同时处理多种类型的数据，并结合特征提取器以生成类别重叠较少的样本。模块，它既可以用作伪样本分类器，也可以用作噪声过滤器。此外，使用了一个高维特征提取器，用于计算每种类型的数据与原始数据以及其他类型数据之间的余弦相似度，以便将计算得到的余弦值用于更新相应的生成器参数，如此可以提升生成样本的质量，减少不同类别样本之间的重叠区域，提高分类的边界清晰度。

过采样通过在少数类样本间插值生成新样本，本质是对已有样本的“近似复制”，易导致生成样本多样性

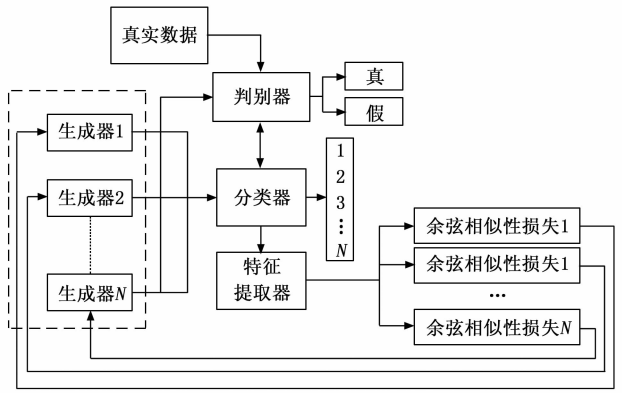

图 4 多生成器框架

不足，甚至引入冗余特征，欠采样通过减少多数类样本数量平衡数据，会直接丢失多数类中可能包含的关键信息，破坏数据原有分布。多生成器对抗网络通过捕捉多模态分布、生成高真实性与多样性样本，在平衡数据、保留特征、提升模型泛化能力等方面，显著优于简单复制或删除样本的过采样和欠采样方法，尤其适用于复杂数据的不平衡问题处理。接下来主要介绍多生成器对抗网络模型的训练，我们假设由生成器生成的样本是 $x_k=G_k(z)$ 对应的原始样本是 x_k。在通过进行特征提取之后，两个样本的高维特征是 $F[G_k(z)]$ 和 $F(x_k)$。然后生成的样本与原始样本之间的余弦相似度可以由以下等式获得：

$$\uparrow O_k[F(\overline{X}_k),F(x_k)]=\left|\frac{F(\overline{X}_k)F(x_k)}{\|F(\overline{X}_k)\|\|F(x_k)\|}\right|,\quad k\in\{1,\cdots,N\} \tag{13}$$

类似地，生成的样本 $\overline{X}_k=G_k(z)$ 与其他类型的生成的样本 $\overline{X}_j=G_j(z),j=\{1,2,\cdots,N\}$ 且 $j\neq k$ 之间的余弦相似度可以通过以下等式获得：

$$\downarrow O_k[F(\overline{X}_k),F(\overline{X}_j)]=\frac{1}{N-1}\sum_j\left|\frac{F(\overline{X}_k)F(\overline{X}_j)}{\|F(\overline{X}_k)\|\|F(\overline{X}_j)\|}\right|,\quad j\in\{1,\cdots,N\}\text{and}(j\neq K) \tag{14}$$

因为希望生成的样本 $\overline{X}_k$ 的分布更接近原始样本 x_k，所以余弦相似度 $O_k[F(\overline{X}_k),F(x_k)]$ 越大越好。相反，希望生成样本 $\overline{X}_k$ 的分布不会与其他类型生成样本 $\overline{X}_j$ 的分布重合，因此余弦相似度 $O_k[F(\overline{X}_k),F(\overline{X}_j)]$ 越小越好。基于此，最终用于更新生成器 G_k 的余弦相似性损失表示如下：

$$O_k=O_k[F(\overline{X}_k),F(\overline{X}_j)]-O_k[F(\overline{X}_k),F(x_k)],\quad\{k,j\}\in\{1,\cdots,N\}\text{and}(j\neq k) \tag{15}$$

1.3 检测模型

研究采用的检测模型如图 5 所示。因为网络入侵中拥有很强的时空性，所以首先考虑使用 CNN（卷积神经网络）提取数据空间特征，对于时间特征。起初用 LSTM（长短时记忆网络）提取时间特征，不过它只能单向遍历序列。于是采用 BiLSTM（双向长短时网络）

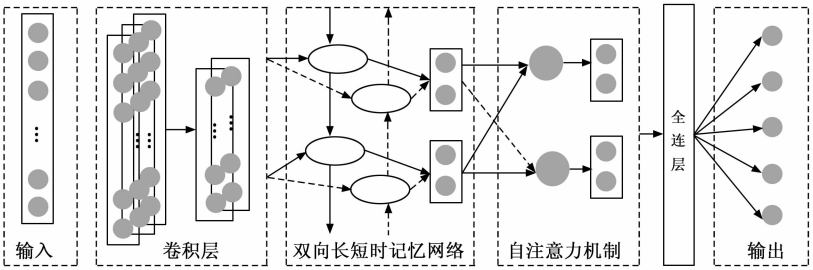

图 5　检测模型框架

结构，能够捕获长距离上下文依赖。注意力机制可选择性关注对分类影响大的特征，赋予高权重，提升模型分类精度。该机制能够动态地、有选择性地为 CNN-BiLSTM 输出的特征分配不同的权重，自动聚焦于对当前入侵检测判断最为关键的特征片段，抑制噪声或不重要的信息。这种结合了 CNN 空间特征提取、BiLSTM 双向时序建模以及注意力机制特征加权聚焦的设计，共同兼顾了网络入侵数据的局部性、复杂时序动态性和关键特征重要性，显著提升了模型在入侵检测任务中的分类精度和整体性能表现。序性和重要性，在入侵检测任务中表现优异。

2　实验结果与分析

2.1　数据集与预处理

择了两个公开可用且广泛使用的数据集 CICIDS-2017 和 CSE-CICIDS-2018 来进行我们的实验。这两个数据集都使用 CICFlowMeter 特征提取工具从原始数据包中提取与流相关的特征。从表 1 和表 2 中可以发现

表 1　CICIDS-2017 数据集

种类	数量/条
Benign	13 484 000
Brute Force	423 000
Web Attack	1500
DOS/DDOS	1 200 000
Botnet	286 000
Infiltration	1 600
ortScan	400 000
APT	4 800

表 2　CSE-CICIDS-2018 数据集

种类	数量/条
Benign	2 273 097
Brute Force	13 911
Heartbleed	11
Web Attack	1 610
Infiltration	30
DOS/DDOS	252 661
PortScan	158 930
Botnet	1 966

CICIDS-2017 和 CSE-CICIDS-2018 数据集中类不平衡的问题，这会使检测模型难以检测到少数类，然而攻击数据就存在于少数类，所以数据平衡问题的重要性油然而生。

由于原始入侵数据集不能直接用作模型的输入，因此需要对数据进行预处理。数据预处理主要包括以下两个部分。数据的数字化。采用一热编码技术将离散的非数值特征转化为数值特征，解决了模型只能传输数值数据的问题。数据标准化。采用 Min-Max 归一化方法将数据归一化到［0，1］区间，提高了模型的收敛速度。x_1 是归一化数据，x 是当前数据，$x_{\min}$ 是最小数据值，$x_{\max}$ 是最大数据值：

$$x_1 = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{16}$$

2.2　实验指标与设置

研究所采用的评价指标涵盖准确率（Acc）、召回率（又称检测率，DR）以及假阳性率（FPR），公式分别为 1.8，1.9，1.10。其中，准确率主要用于评估模型对正常数据与攻击数据进行准确分类的能力。召回率，着重反映模型对攻击类数据的检测能力。假阳性率则用于评价模型的误分类能力：

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

$$DR = \frac{TP}{TP + FN} \tag{18}$$

$$FPR = \frac{FP}{FP + TN} \tag{19}$$

实验中，操作系统为 Ubuntu 22.04.3，CPU 为 Intel Xeon Platinum 8375 C，GPU 为 NVIDIA GeForce GTX1080TI（11 G 显存），编程语言为 Python 3.10。将 CICIDS-2017 数据集作为模型的训练集，其中 70% 为训练集，30% 为测试集，将 CSE-CICIDS-2018 数据集作为模型验证数据漂移的测试数据集。Adam 被选为模型的优化器。epoch 设置为 50，初始学习率 $r=0.001$，正则化超参数 $T=0.000\ 5$。最大训练轮次为 1 000。使用提前停止训练的方法来调整学习率，权重衰减和训练周期。

2.3　检测模型的评估

为了验证 CNN-BiLSTM-Attention 模型在网络入侵检测中的分类性能，采用分层 K-Fold 交叉验证方法，并在 CICIDS-2017 进行了分类实验。K 值范围设置为 2～10。多分类实验结果如表 3 所示：在 CICIDS-2017 数据集上，最佳性能出现在 $K=10$ 时。

选择的模型在 $K=10$ 时几乎均达到了最佳结果。这是因为随着 K 值的增加，不同攻击类别的训练样本数量也相应增加，提升了模型的分类性能。

表 3 在 CICIDS-2017 数据集上测试

K	准确率/%↑	召回率/%↑	假阳率/%↓
2	81.55	93.72	8.03
4	85.49	96.51	4.72
6	87.32	97.41	2.65
8	88.36	**98.15**	2.26
10	**88.95**	97.86	**1.32**

2.4 使用数据平衡后的入侵检测模型评估

为系统验证本文采用的数据平衡方法的优势，选取当前网络安全领域 4 种数据平衡策略方法构建的对比实验体系。1D-CNN-BILSTM 是面向数据不平衡的二阶段网络入侵检测方法。SRFCNN-BILSTM 是融合改进采样技术和 BiLSTM 的入侵检测方法。设计的一种 FBS-RE 混合采样算法。ADASYN-WGAN 采用 ADASYN 算法生成少数类样本利用 WGAN 算法生成符合原始数据集分布规律的少数类样本，构建平衡数据集。PSSNS-RF 提出了系统化数据预处理与混合采样相结合的网络入侵检测算法。如表 4 所示，基于统一评估框架的实验结果表明，采用本方法优化后的检测模型在关键性能指标上展现出显著优势[25]。实验结果表明，传统数据平衡方法在应对网络入侵检测的复杂场景时均存在明显局限：过采样技术易引入噪声干扰，欠采样策略导致关键信息丢失，而混合架构的异构设计加剧了数据分布不平衡的影响。研究提出的方法 B-CNN-Bilstm-Attention（平衡后的检测模型）通过创新的动态平衡机制，在保持模型轻量化的同时，显著提升了小样本攻击类别的识别能力。这一进展验证了所提方法在解决数据不平衡问题具有重要意义。

表 4 各种平衡方法后的检测模型在 CICIDS-2017 数据集

模型	准确率/%↑	召回率/%↑	假阳率/%↓
1D-CNN-BILSTM[26]	96.26	98.66	0.64
SRFCNN-BILSTM[27]	95.54	98.73	0.73
ADASYN-WGAN[28]	95.23	97.46	0.56
PSSNS-RF[29]	96.33	**99.73**	0.45
B-CNN-Bilstm-Attention	**97.56**	99.69	**0.22**

2.5 自适应模块的评估

异常值与 k 的关系，在范围［9，100］内选取不同的 K 值开展实验。每次实验中，都对旧数据分布以及 OR（这里假设 OR 是某个与实验相关的特定指标）中的离群值进行记录。起初，随着 K 值增大，离群值数量逐渐减少，且减少速度逐渐变缓，直至稳定在某个恒定值附近。从离群值的变化情况可知，选择大于真实标签数量的 K 值，对减少离群值更为有利，这意味着聚类模型的收敛效果更佳。不过，K 值过大并非最优选择，可能会增加模型过拟合的风险。另外，随着 K 值增加，OR 也逐步上升，这有利于更精准地识别数据漂移现象。同样，K 值过大并不会带来更好的结果。在［40，50］这个范围内选取 K 值似乎最为适宜，而非直接选用真实标签的数量（本实验中是 1 个良性类别和 8 个攻击类别）。异常值与 α 关系我们进行了 100 次实验，α 值设定在（0，10）这个区间。每次实验都记录旧数据中的离群值分布以及 OR。异常值、OR 与 α 的关系通过坐标图进行说明，如图 6 所示（这些实验使用相同的 K 值）。我们观察到，随着 α 值增加，离群值数量呈指数下降趋势，直至趋近于 0。对于 OR 而言，一开始它会随着 α 值上升而增加，但当离群值变化到达拐点（在图 9 中大约是 $\alpha=7.0$ 的位置）时，OR 的变化趋势从上升转变为下降。这表明，当 α 值接近 7 时，OR 达到最大值，最佳反映出数据漂移的出现。

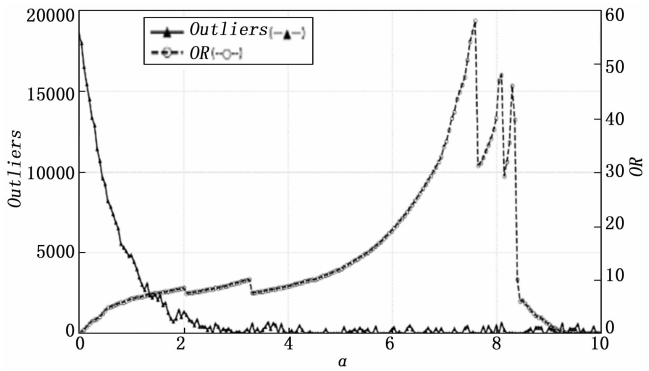

图 6 异常值与 α 关系

2.6 消融实验

研究通过对比实验数据验证了模型改进的有效性。如表 3、表 4 所示，经平衡数据集训练的入侵检测模型在核心性能指标上实现显著提升，其中精准率提高 10%，召回率和 F_1 值也有一定的增长。然而实验发现，尽管优化后的模型在训练数据上达到 90%＋的指标水平，但面对未训练数据集时检测准确率骤降 15%，这暴露出传统检测系统在动态网络环境中的泛化缺陷。深入分析表明，攻击模式的持续演变导致模型出现特征偏移问题。值得注意的是，集成自适应模块的模型在新场景下仅出现 7%的指标波动，这种对比结果验证了提出的双重改进方案（数据平衡处理＋自适应机制）在提升模型鲁棒性方面的协同效应：前者有效缓解了类别不均衡导致的过拟合问题，后者通过动态特征校准机制增强了模型对新型攻击的适应能力。

2.7 总体评估

为了验证提出的方法在应对数据漂移问题上的显著优势，研究通过系统性实验对比分析了传统模型与新方法在未知攻击场景下的泛化能力。如表 4 所示，在标准训练集上各模型均能取得 90%以上的检测准确率、F_1 值和召回率，这表明现有方法在已知攻击模式下的性能

差异并不显著。然而，当我们将训练好的模型迁移至完全未参与训练的新数据集时（如表 5），数据漂移带来的挑战开始凸显：传统模型的三大核心指标平均下降幅度高达 15%，这直接暴露出现有检测系统在动态网络环境中的脆弱性。这种性能断崖式下跌的根本原因在于网络攻击模式具有持续演变的特性，攻击者会不断开发新的漏洞利用手段，导致训练数据与实时流量特征分布产生显著偏移，即典型的“概念漂移”现象。

表 5　在 CSE-CICIDS-2018 数据集各模型的训练结果

模型	准确率/% ↑	召回率/% ↑	假阳率/% ↓
1D-CNN-BILSTM[26]	76.33	82.13	2.66
SRFCNN-BILSTM[27]	79.22	83.66	3.12
ADASYN-WGAN[28]	86.2	87.12	1.33
PSSNS-RF[29]	84.22	86.55	2.13
ATGCB	**91.12**	**92.63**	**1.56**

值得关注的是，研究提出的自适应检测框架在相同测试条件下展现出更强的鲁棒性，其综合指标仅出现的有限降幅，仍然保持巨大优势。这种优势源于方法中创新的动态特征提取机制和对抗性学习策略，通过构建可扩展的特征空间，系统能自动识别新型攻击流量中的潜在模式，这些实证数据不仅验证了数据漂移问题的严峻性，更凸显了构建自适应网络安全体系的重要性，在网络攻击日益激烈的现实环境中，仅依靠历史数据进行静态建模已无法满足防护需求，必须建立具有持续学习能力的检测机制。

3　结束语

研究指出车联网入侵检测面临数据漂移，数据集不平衡的问题。针对问题，提出 ATGCB 入侵检测模型，采用聚类算法和自适应滑动窗口以无监督方式检测未知攻击，多个生成器网络同时生成少数类。以此提高了泛化性与稳定性。实验证明，提出的方法能有效检测车联网攻击数据，特别是在攻击数据变化大的情况下有独特的优势。网络攻击数据特征有一定的相似性，对物联网和网络安全领域的泛化性问题也有很高的借鉴意义，为持续学习提供了可行的技术路径，也为下一代智能安全系统的发展奠定了重要基础。不过方法在未来研究中仍有提升空间，虽可通过无监督学习识别未标记新数据中的数据漂移训练新模型，但训练新模型仍需标记新数据，需领域专家参与。聚类和分类模型能在此过程中提供指导，相信未来这些问题将得到更好的解决。

参考文献：

[1] LI Y L，LI F，JIA Q S. The research of random forest intrusion detection model based on optimization in internet of vehicles [J]. Journal of Physics：Conference Series，2021，1757 (1)：012149.

[2] 周　漫. 车联网中基于行为分析的攻击检测方法研究 [D]. 武汉：华中科技大学，2022.

[3] 孙浩然. 面向不平衡样本的车载网入侵检测系统设计与实现 [D]. 哈尔滨：哈尔滨工业大学，2022.

[4] 郭健忠，王　灿，谢　斌，等. 面向车联网 DoS 攻击的混合入侵检测系统 [J]. 计算机系统应用，2025，34 (3)：85－93.

[5] 曹　磊，温　蜜，何　蔚. 基于深度学习的车联网的路网监测系统的 DoS 和 DDoS 攻击的入侵检测方法 [J]. 计算机应用与软件，2025，42 (1)：303－311.

[6] 张　锐. 面向车联网的基于卷积神经网络的入侵检测模型 [J]. 电信科学，2024，40 (12)：51－62.

[7] 刘　沛，刘昌华，林俏伶. 基于优化特征堆叠与集成学习的车联网入侵检测模型 [J]. 计算机工程与科学，2024，46 (12)：2186－2195.

[8] VIJAYALAKSHMI S，BOSE S，LOGESWARI G，et al. Smart parking：intelligent intrusion detection system in VANET enabled car parking system [J]. Automatika，2025，66 (2)：281－299.

[9] BASHEERL P R. A deep learning framework for intrusion detection system in smart grids using graph convolutional network [J]. Engineering Research Express，2025，7 (1)：015257.

[10] KUMARI D，PRANAV P，SINHA A，et al. A hybrid cheetah and grey wolf optimization algorithm for network intrusion detection [J]. Engineering Research Express，2025，7 (1)：015256.

[11] KUMARI D，PRANAV P，SINHA A，et al. A novel approach to intrusion detection system using hybrid flower pollination and cheetah optimization algorithm [J]. Scientific Reports，2025，15 (1)：13071－13076.

[12] 徐煊翔，杜彦辉，陈李舟，等. 一种混合知识蒸馏的轻量化 CAN 总线入侵检测方法 [J/OL]. 武汉大学学报 (工学版)，1－16[2025-12-04]. https://link.cnki.net/urlid/42.1675.T.20241021.1052.002.

[13] QIU L，XU Z，LIN L，et al. Design and optimization of hybrid CNN-DT model-based network intrusion detection algorithm using deep reinforcement learning [J]. Mathematics，2025，13 (9)：1459－1466.

[14] PASUMPONTHEVAR K M，JEYARAJ R P. Kalman reinforcement learning-based provably secured smart grid false data intrusion detection and resilience enhancement [J]. Electrical Engineering，2024，107 (3)：1－19.

[15] NAJAFLI S，HAGHIGHAT T A，KARASFI B. A novel reinforcement learning-based hybrid intrusion detection system on fog-to-cloud computing [J]. The Journal of Supercomputing，2024：1－23.

（下转第 66 页）