专题：智能网联汽车

# 车联网中基于stacking集成学习的攻击检测模型

徐会彬，方龙，张莎
（湖州师范学院信息工程学院，浙江 湖州 313000）

**摘　要：**由于无线网络的开放性，车联网容易受到网络攻击，如拒绝服务、模糊和欺骗攻击。为此，提出融合随机森林（random forest，RF）和梯度提升决策树（gradient boosting decision tree，GBDT）的堆叠（stacking）的入侵检测（RG-IDS）模型。首先，RG-IDS模型利用自适应合成采样（adaptive synthetic sampling，ADASYN）算法对不平衡类别的数据样本进行近邻采样，进而生成更多同类别的近似样本，形成相对平衡的样本数据。其次，RG-IDS模型利用GBDT评估特征的重要性，并选择具有重要特征的样本数据，建立轻量级分类器。最后，RG-IDS采用基于$k$折交叉验证的堆叠方法，降低过拟合的概率。将RF、GBDT和LightGBM分类器作为基学习器。采用数据集CICIDS 2017和NSL-KDD对RG-IDS模型进行实验测试。实验结果表明，RG-IDS模型可实现较高的F1值。

**关键词：**车联网；入侵检测；自适应合成采样；梯度提升决策树；堆叠
**中图分类号：**TP183
**文献标志码：**A
**doi**: 10.11959/j.issn.1000−0801.2024257

# Attack detection model based on stacking ensemble learning for Internet of vehicles

XU Huibin, FANG Long, ZHANG Sha
School of Information Engineering, Huzhou University, Huzhou 313000, China

**Abstract:** Due to openness of wireless communication, Internet of vehicles (IoV) is vulnerable to many cyber-attacks such as denial of service, spoofing and fuzzy attacks. Therefore, random forest (RF) and gradient boosting decision tree-based stacking intrusion detection (RF-IDS) model was proposed. Firstly, the adaptive synthetic sampling (ADASYN) algorithm was adopted to generate more similar samples through the nearest neighbor sampling strategy in order to balance the training samples of different categories, and form a relatively symmetric dataset. Secondly, GBDT was used to evaluate the importance of features and select sample data with important features to build a light-weight classifier. Finally, the *k*-fold cross-validation stacking method was used to reduce the probability of overfitting. RF, GBDT and LightGBM classifiers serve were used as base-learner. The RG-IDS model was tested by CICIDS

收稿日期：2024−10−18；修回日期：2024−11−08
通信作者：徐会彬，02623@zjhu.edu.cn
基金项目：湖州市自然科学基金资助项目（No.2024YZ08）
**Foundation Item:** The Huzhou Natural Science Foundation (No.2024YZ08)

2017 and NSL-KDD datasets. The experimental results demonstrate that RG-IDS model can achieve a higher F1-score.

**Key words:** Internet of vehicles, intrusion detection, ADASYN, GBDT, stacking

## 0　引言

随着物联网（Internet of things，IoT）和车联网（Internet of vehicles，IoV）技术的快速发展，自动驾驶汽车和互联汽车等网络控制汽车开始逐步取代传统汽车[1]。车联网系统主要由车内网络（intra-vehicle network，IVN）和外部网络组成。在IVN中，控制器区域网络（controller area network，CAN）总线是实现电子控制单元之间通信以完成各种功能的核心基础设施[2-3]。外部网络则利用V2X（vehicle-to-everything）技术实现智能汽车与其他实体（如路边单元、基础设施和智能设备）之间的连接。

随着网络攻击面的不断扩大，车联网系统面临越来越多的安全威胁。此外，由于CAN数据包的长度较短，在处理CAN数据包时并未构建认证或加密机制，在缺乏基本安全机制的情况下，攻击者能够轻易地将恶意消息插入IVN，并执行各种攻击，如拒绝服务（denial of service，DoS）攻击[4]、模糊攻击和欺骗攻击[5]。由于联网汽车需要连接外部网络，这进一步加剧了车辆遭受外网攻击的风险。

入侵检测系统（intrusion detection system，IDS）已成为保护IoV系统和智能汽车免受网络攻击的有效技术[6]。为了保护IVN，IDS可以部署在CAN总线的顶部以识别恶意CAN消息[7]。IDS也可以集成到网关中，以检测来自外部网络的恶意数据包。限于篇幅，本文仅讨论车载外部网络攻击的检测问题。

随着机器学习（machine learning，ML）技术的不断发展，基于ML的IDS受到广泛关注和研究[8-9]。研究者通过分析网络流量数据，利用ML的学习分类能力可以构建基于分类器的IDS方案，进而检测各种网络攻击[10-11]。文献[12]针对车联网系统，提出了一种带有梯度惩罚的Wasserstein生成对抗网络（Wasserstein generative adversarial network with gradient penalty，WGAN-GP）和残差网络（residual network，ResNet）的车联网入侵检测方法。该方法利用对抗量化变分自编码器处理数据的不平衡问题，并通过ResNet和改进的分段残差神经网络对输入的样本数据进行联合学习，进而预测攻击类型。尽管该方法在车联网数据集CICIDS 2017上的F1值达到了99.86%，但其复杂度较高。文献[13]提出了一种基于Transformer和自适应模糊神经网络推理系统的混合车联网入侵检测方法。该方法采用自适应合成采样（adaptive synthetic sampling，ADASYN）算法对数据进行增强，以处理数据样本的不平衡问题。同时，通过Transformer的自注意力机制来增强特征表达能力。该检测模型在数据集CICDIS 2017上的F1值达到了98.31%。

尽管现有的关于车联网的IDS的检测框架已展现出较好的性能，但仍具有较大的性能改进空间。由于数据规模的不断增大和数据维度的不断提升，单个分类器对样本的预测能力变得有限。为此，研究人员提出了不同的集成方法。文献[14]提出了一种基于集成学习和改进的粒子群优化-遗传算法（particle swarm optimization-genetic algorithm，PSO-GA）特征选择的入侵检测方法。该方法通过粒子群优化与遗传混合算法来提取特征，选择重要性较高的特征进行模型训练。该模型在数据集CICIDS 2017上的检测精确率为95%。文献[15]提出了一种基于极端梯度提升（extreme

gradient boosting，XGBoost）的集成检测方法。该方法在数据集KDD Cup 99上的准确率达到99.95%。文献[16]提出了一种基于树的ML模型的stacking集成框架，通过该框架实现了车联网的入侵检测。性能分析表明，该框架在CICIDS数据集上具有良好的检测性能。文献[17]将深度学习模型引入IDS中，并提出了基于深度神经网络（deep neural network，DNN）、长短期记忆和深度信念网络的检测方法。

相比单个分类器，基于集成学习的检测模型具有更优的检测性能。为此，针对IoV网络，本文提出了融合随机森林（random forest，RF）和梯度提升决策树（gradient boosting decision tree，GBDT）的stacking集成学习的入侵检测（RG-IDS）模型。该模型采用两层stacking方法，将RF、GBDT和LightGBM（light gradient boosting machine）作为第一层基学习器，再将LightGBM作为第二层元学习器。本文采用ADASYN方法增强数据，以处理样本不平衡问题。同时，为了降低模型的复杂度，本文利用GBDT评估特征的重要性，选择重要特征的数据样本训练模型。在数据集NSL-KDD和CICIDS 2017上分析RG-IDS模型的性能，仿真结果表明，提出的RG-IDS模型能够准确地检测攻击，其在数据集NSL-KDD和CICIDS 2017上的F1值均可达到99%以上。

## 1　随机森林和提升决策树概述

决策树（decision tree，DT）是最基本的ML算法之一，利用树结构并结合分治技术进行决策。树中的每个内部节点表示一个特征属性上的测试，每个分支表示一个测试输出，每个叶节点则代表一种类别。随机森林是一种利用多棵树对样本进行训练并预测的分类器，它采用bagging思想，将若干个弱分类器的分类结果进行投票选择，从而组成一个强分类器。

GBDT是一种迭代的决策树算法，它先构造一组弱的学习树，再对多个决策树的决策结果进行融合，最终将融合后的决策作为最终的决策结果。

LightGBM是一个由多个DT构建的快速而强大的集成ML模型。相对于其他ML方法，LightGBM的主要优势在于它能够有效地处理大规模和高维数据。基于梯度的单边采样和互斥特征捆绑是LightGBM的两个核心策略。基于梯度的单边采样只保留大梯度的数据样本，并随机丢弃小梯度样本，以加速模型训练并减少内存消耗。互斥特征捆绑是一种特征处理方法，其将互斥特征重组，构成单个特征，进而减少特征尺寸，并提高模型训练效率。通过基于梯度的单边采样和互斥特征捆绑，LightGBM算法能够在不丢失重要信息的前提下极大限度地减少数据量，降低算法的复杂度。

## 2　基于stacking集成学习的车载网络入侵检测模型

### 2.1　模型框架

RG-IDS模型的总体框架如图1所示，该框架主要由数据预处理、特征选择和堆叠（stacking）模块组成。

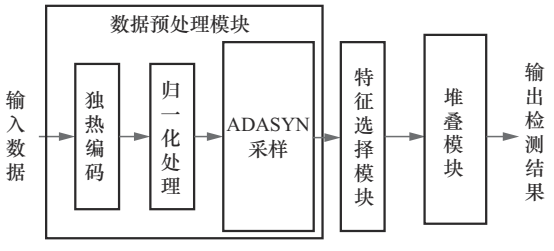


图1　RG-IDS模型的总体框架

数据预处理模块是对数据进行独热编码、归一化处理和ADASYN采样。特征选择模块则是采用GBDT评估特征的重要性，再选择重要特征，利用所选特征的数据对模型进行训练，降低训练和预测的复杂度。堆叠模块采用RF、GBDT

和LightGBM这3个基学习器，并将LightGBM作为元学习器。

### 2.2　数据预处理

通过独热（one-hot）编码，本文将离散特征取值映射到欧氏空间，使离散特征的某个取值对应欧氏空间中的某个点，这样不仅可以方便地计算特征之间的距离，还有利于后续的归一化处理。

考虑数据集中常包含多种不同类型的特征，如协议类型、源端口、目的端口等。为便于处理和分析这些特征，需要对这些数据进行归一化处理，将数据的取值映射至[0,1]。为此，RG-IDS模型采用最大−最小值方法对样本数据进行归一化处理。

$$x_n=\frac{x-x_{\min}}{x_{\max}-x_{\min}} \tag{1}$$

其中，$x$表示原始数据，$x_{\min}$和$x_{\max}$分别表示原始数据所在列的最小值和最大值，$x_n$表示对$x$归一化后的数据。

在现实生活中，网络在大部分时间处于正常状态，攻击标签的实例往往较少。因此，网络数据常存在类别不平衡的问题。为此，本文采用了ADASYN算法[18]，该算法是基于合成少数过采样技术（synthetic minority over-sampling technique，SMOTE）算法的改进算法。ADASYN算法能自动确定少数类样本需要合成的新样本数量，为少数类样本合成更多的新样本，从而补偿数据集的偏态分布。这种自适应机制能够根据数据的实际情况动态地调整样本合成策略，进而提高模型的泛化能力。

具体而言，ADASYN算法能够在低密度特征空间中生成多个实例，而在高密度特征空间中生成少数实例，进而为少数较难学习的类别样本合成更多的训练数据，从而尽可能降低类别不平衡分布所产生的负面影响。执行ADASYN算法的伪代码如下所示。

**输入：** $S=\{(x_i,y_i),i=1,2,\cdots,N\}$；参数$N^-$，$N^+$；采样率$\gamma$；近邻数$K$。

**步骤1**　依据训练集$S$，生成多数类样本训练集$S^-$和少数类样本训练集$S^+$

**步骤2**　**For** $i\leftarrow 1$ to $N^+$ **do**

**步骤3**　　对于样本$x_i\in S^+$,找到样本$x_i$的$K$个近邻,并计算这$K$个近邻中多数类样本数$M_i$

**步骤4**　　依据式（2）计算$\beta_i$；

**步骤5**　　依据式（3）计算$g_i$；

**步骤6**　**End For**

**步骤7**　**For** $i=1$ to $N^+$ **do**

**步骤8**　　**For** $i=1$ to $g_i$ **do**

**步骤9**　　调用SMOTE算法为样本$x_i$生成新样本$x_i^{\text{new}}$

**步骤10**　　$S^{\text{New}}=S^{\text{New}}\cup x_i^{\text{new}}$

**步骤11**　　**End For**

**步骤12**　**End For**

**输出：** $S^+=S^{\text{New}}\cup S^+$；

令$S=\{(x_i,y_i),i=1,2,\cdots,N\}$表示训练集。$N$个样本中有些样本数较多，有些样本数较少。用$N^-$和$N^+$分别表示多数类样本数和少数类样本数，且$N^-+N^+=N$。ADASYN算法有$K$和$\gamma$两个主参数，前者表示近邻数，后者表示采样率。

首先，从训练集$S$中取出所有样本。针对$N^-$和$N^+$组成多数类训练样本集$S^-$和少数类训练样本集$S^+$，并设置一个存储新生成的样本集$S^{\text{New}}$。最初，$S^{\text{New}}$为空集。

然后，计算每个少数类样本的主样本频次，如算法的步骤2～步骤7所示。用$x_i$表示$S^+$中第$i$个样本，即$x_i\in S^+$。从训练集$S$中找到$x_i$的$K$个近邻，并统计这$K$个近邻中多数类样本数的个数$M_i$，再利用式（2）计算样本$x_i$的比例$\beta_i$：

$$\beta_i=\frac{M_i}{Z\times K} \tag{2}$$

其中，$Z$表示标准化因子，用以保证$\sum\beta_i=1$。

再利用式（3）计算样本$x_i$的主样本频次$g_i$：

$$g_i=\beta_i\times N^+\times\gamma \tag{3}$$

接着，为样本$x_i$生成合成样本，如算法的步骤9～步骤10所示。调用SMOTE算法为样本$x_i$生成新样本$x_i^{\text{new}}$，并将$x_i^{\text{new}}$添加至$S^{\text{New}}$中，即$S^{\text{New}}=S^{\text{New}}\cup x_i^{\text{new}}$。

最后，将$S^{\text{New}}$合并至集$S^+$，即$S^+=S^{\text{New}}\cup S^+$，并输出过采样后的训练集$S^+$。

### 2.3 基于$k$折交叉验证的多模型融合的stacking实现过程

stacking集成学习采用双层堆叠训练模型，第一层的训练模型称为基学习器；第二层称为元学习器。基学习器在训练模型时采用交叉验证方法，以降低模型的过拟合风险，提高模型的准确度。stacking集成学习的框架如图2所示。

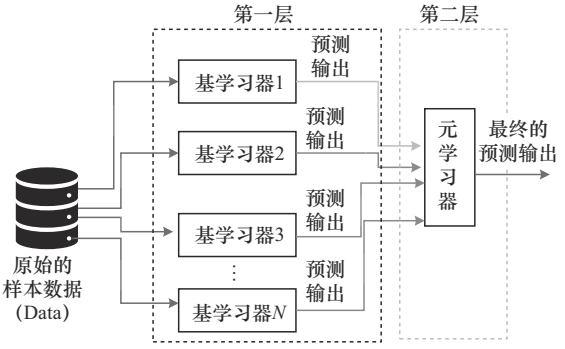


图2 stacking集成学习的框架

stacking集成学习的基本思想：先通过基学习器学习原始数据，然后几个基学习器均对训练数据进行预测，并输出预测结果，再将这几个基学习器所预测的结果作为新的数据特征加入原始样本数据，形成新的样本数据，最后，将新的样本数据作为元学习器的输入，利用元学习器学习样本数据，并对样本数据进行分类预测。

执行stacking集成算法的步骤如下。

第1步，数据拆分。将原始数据Data分割成训练集$\text{Data}_{\text{train}}$和测试集$\text{Data}_{\text{test}}$。为了实现$k$折交叉，将训练集$\text{Data}_{\text{train}}$进一步分成$k$等份：$D_1,D_2,\cdots,D_k$，且$\text{Data}_{\text{train}}=D_1\|D_2\|\cdots\|D_k$，其中“||”表示连接符。

假定$\text{Data}_{\text{train}}$包含$N$个样本数据，上述过程可形式化表示为：

$$\text{Data}_{\text{train}}=\begin{pmatrix}d_1\\d_2\\\vdots\\d_N\end{pmatrix}=D_1+D_2+\cdots+D_k \tag{4}$$

其中，$D_1,D_2,\cdots,D_k$的定义如下所示。

$$\begin{cases}D_1=\left(d_1\,d_2\cdots d_\sigma\right)^{\text{T}}\\D_2=\left(d_{\sigma+1}\,d_{\sigma+2}\cdots d_{\sigma+\sigma}\right)^{\text{T}}\\\vdots\\D_{k-1}=\left(d_{(k-2)\sigma+1}\,d_{(k-2)\sigma+2}\cdots d_{(k-2)\sigma+\sigma}\right)^{\text{T}}\\D_k=\left(d_{(k-1)\sigma+1}\,d_{(k-1)\sigma+2}\cdots d_{(k-1)\sigma+\sigma}\right)^{\text{T}}\end{cases} \tag{5}$$

其中，$\sigma=N/k$。

第2步，基学习器训练模型。对于任意一个基学习器需要学习$k$等份数据（$D_1,D_2,\cdots,D_k$）$k$次。在每次的学习过程中，从$D_1,D_2,\cdots,D_k$中选择一份数据作为模型测试数据，其余的作为模型训练数据。执行了$k$次测试后，将这些测试结果（预测值）拼接成一列，作为训练集$\text{Data}_{\text{train}}$数据的新特征。此外，基学习器在每次训练后，也对测试集$\text{Data}_{\text{test}}$数据进行测试，并记录每次测试结果（预测值），计算$k$次测试所产生预测值的平均值，最后，将此预测的平均值作为测试集$\text{Data}_{\text{test}}$的新特征。

为了更好地理解上述过程，以$k$=3为例分析基学习器1（Model1）如何执行上述过程。基学习器1训练示例如图3所示，先将训练集$\text{Data}_{\text{train}}$拆分成3等份（$D_1,D_2,D_3$），每一份执行3次。第1次选择$D_3$为测试数据，其余的$(D_1,D_2)$为训练数据；第2次选择$D_2$为测试数据，其余的$(D_1,D_3)$为训练数据；第3次选择$D_1$为测试数据，其余的

$(D_2, D_3)$为训练数据。将这 3 次测试所产生的预测值拼接成新特征 1。同时，Model1 在每次训练后也对 $\text{Data}_{\text{test}}$ 进行测试，并取 3 次测试的平均值作为预测值，加入样本数据的新特征，最终形成完整的新特征 1。

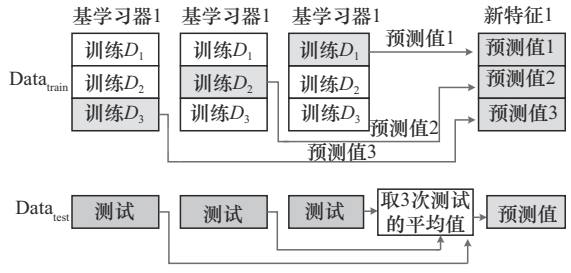


图3　基学习器 1 训练示例（$k$=3）

第 3 步，元学习器训练模型。基学习器执行了第 2 步后，就将新生成的训练样本数据 $\tilde{\text{D}}\text{ata}_{\text{train}}$ 作为元学习器的输入数据，即元学习器利用 $\tilde{\text{D}}\text{ata}_{\text{train}}$ 进行训练。最后，利用经元学习器训练后的模型对 $\text{Data}_{\text{test}}$ 进行预测。

## 2.4　特征选择

利用少量特征训练检测模型有助于建立轻量级分类器，降低过拟合的概率，且轻量级分类器容易在网络平台上实施。

选用 GBDT 估计特征的重要性，使所用特征的重要性之和为 1。为了选择重要特征，依据特征重要性，对所有特征从高到低进行排序，依次将重要性高的特征加入所选特征集 $\psi$ 中，直到特征集 $\psi$ 中所有特征的重要性之和达到 0.9。最后，只利用 $\psi$ 中特征信息对样本数据进行训练。

GBDT 采用基尼指数评估特征的重要性[19]。假定样本数据中有 $M$ 个特征，这些特征构成特征集 $\Theta=\{f_i|1\leqslant i\leqslant M\}$，其中 $f_i$ 表示第 $i$ 个特征。令 $\text{GI}_n(p)$ 表示节点 $n$ 的基尼指数，其定义如下所示。

$$\text{GI}_n(p)=\sum_{k=1}^{K}p_{nk}(1-p_{nk}) \tag{6}$$

其中，$K$ 表示 $K$ 个类别，$p_{nk}$ 表示节点 $n$ 中类别 $k$ 所占的百分比。

用 $I_{i,n}$ 表征特征 $f_i$ 对节点 $n$ 的重要性，其定义如下所示：

$$I_{i,n}=w_n\times\Delta G \tag{7}$$

其中，$w_n$ 表示节点 $n$ 所占用的样本量占总样本量的比例，$\Delta G$ 表示分枝前后的基尼指数的变化值。

若节点 $n$ 位于第 $j$ 棵树中，则特征 $f_i$ 在第 $j$ 棵树的重要性可表示为：

$$TI_{j,i}=\sum\nolimits_{\ell\in\Lambda}I_{i,\ell} \tag{8}$$

其中，集合 $\Lambda$ 表示由特征 $f_i$ 在第 $j$ 棵树中出现的全部节点所构建的集合，$\ell$ 表示集合中的任意一个节点。

若 GBDT 共有 $T$ 棵树，则特征 $f_i$ 的重要性可表示为：

$$\text{AI}_i=\sum_{j=1}^{T}TI_{j,i} \tag{9}$$

RG-IDS 模型利用 GBDT 估计数据集中所有特征的重要性，并将各特征的重要性进行归一化，使得所有特征的重要性之和为 1。令 $\tilde{\text{A}}\text{I}_i$ 表示对 $\text{AI}_i$ 归一化后的特征 $f_i$ 的重要性。依据特征重要性从高到低排序，令 $\hat{\Theta}$ 表示对 $\Theta$ 排序后的特征集。然后，从 $\hat{\Theta}$ 中选择第 1 个特征加入所选择特征集 $C$ 中，然后再选择第 2 个特征加入 $C$ 中，直至特征集 $C$ 中所有特征的重要性之和达到 0.9，就停止加入。换而言之，选择 90% 的重要特征进行模型训练，只丢失 10% 的不重要特征信息。

## 2.5　算法流程

RG-IDS 模型检测攻击的主要流程如图 4 所示，该流程主要包含 3 个部分。（1）数据预处理：对样本数据进行独热编码，并对特征进行归一化处理，进而平衡各个特征的贡献；引用 ADASYN 算法进行过采样，平衡各类样本数。（2）特征选择：利用 GBDT 算法估计特征的重要性，进而选择具有重要特征的数据进行训练，降低训练模型的时间。（3）训练及测试 RG-IDS 模型：通过 stacking 集成学习完成对模型的训练，并对样本进行预测。

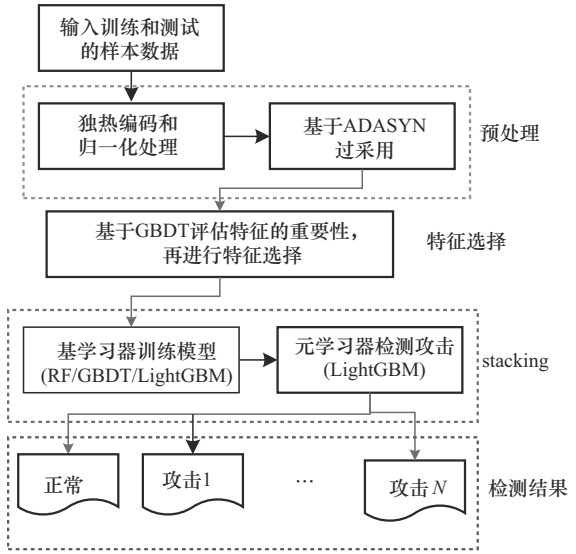


图4　RG-IDS模型检测攻击的主要流程

## 3　实验结果与分析

### 3.1　实验环境

利用Jupyter Notebook软件编写检测程序。运行程序的计算机参数：Intel四核处理器（i5-10210U 1.6 GHz），64位操作系统，Windows 11。

RF、GBDT和LightGBM是基础的分类器，它的默认参数通常是最优的参数，因此这3个分类器均采用默认参数。例如，RF的决策树数量为100；每个节点的最大特征数为输入特征数量的平方根；GBDT的学习率为0.1，最大迭代次数为100，采用对数似然损失函数；LightGBM的迭代次数为100，叶子节点数为31。

### 3.2　数据集概述

为了更好地分析RG-IDS模型的检测性能，选择数据集NSL-KDD和CICIDS 2017进行训练和测试。NSL-KDD是基于著名KDD 99数据集的优化版，其广泛应用于入侵检测领域。

NSL-KDD数据集由训练集（KDD Train+）和测试集（KDD Test+）两部分组成，NSL-KDD数据集分布情况见表1。每条样本由41个特征和1个标签组成。标签由正常流量（Normal）、DoS、用户到根（user to root，U2R）、远程到本地（remote to local，R2L）和探测（Probing）5类攻击组成。本文将KDD Train+和KDD Test+合并，然后从中取80%作为训练数据集，剩余的20%作为测试数据集。

表1　NSL-KDD数据集分布情况

| 数据集 | 类型 | 样本编码 | 样本数量 | 所占比例 |
|---|---|---|---|---|
| KDD Train+ | Normal | 4 | 67 343 | 53.46% |
| | DoS | 0 | 45 927 | 36.46% |
| | Probing | 1 | 11 656 | 9.25% |
| | R2L | 2 | 995 | 0.79% |
| | U2R | 3 | 52 | 0.04% |
| KDD Test+ | Normal | 4 | 9 711 | 43.07% |
| | DoS | 0 | 7 548 | 33.08% |
| | Probing | 1 | 2 421 | 10.74% |
| | R2L | 2 | 2 754 | 12.22% |
| | U2R | 3 | 200 | 0.89% |

CICIDS 2017数据集是入侵检测的专用数据，包含了多个良性和最新的常见网络攻击，广泛用于分析车载网络的入侵检测算法的性能。考虑原CICIDS 2017的样本数据较大，从中随机抽取了56 661条样本数据进行训练和测试。CICIDS 2017数据集分布情况见表2，表2列出了样本数据在良性（Benign）、DoS攻击、端口扫描（Portscan）攻击、暴力（Bruteforce）攻击、Web攻击（Webattack）、Bot攻击和渗入威胁（Infiltration）攻击7类样本上的数量及分布情况。

表2　CICIDS 2017数据集分布情况

| 类型 | 样本编码 | 样本数量 | 所占比例 |
|---|---|---|---|
| Benign | 0 | 22 731 | 40.12% |
| DoS | 3 | 19 035 | 33.59% |
| Portscan | 5 | 7 946 | 14.02% |
| Bruteforce | 2 | 2 767 | 4.88% |
| Webattack | 6 | 2 180 | 3.85% |
| Bot | 1 | 1 966 | 3.47% |
| Infiltration | 4 | 36 | 0.06% |

### 3.3　评价指标

选用准确率（Accuracy）、精确率（Precision）、召回率（Recall）和F1值作为评价指标。

将正常流量（Normal）和良性（Benign）作为正例，其他的攻击全部视为反例。TP 表示真实值为正例，即正确地预测为正例的样本数目；TN 表示真实值为反例，即正确地预测为反例的样本数目；FP 表示真实值为反例，即被错误地预测为正例的样本数目；FN 表示真实值为正例，即被错误地预测为反例的样本数目[20]。

Accuracy 表示分类正确的样本数量与总样本数量的比值，其定义如式（10）所示。不失一般性，准确率越高，表明算法的分类能力越好。

$$\text{Accuracy}=\frac{\text{TP}+\text{TN}}{\text{TP}+\text{FP}+\text{TN}+\text{FN}}\times 100\% \quad (10)$$

与 Accuracy 不同，Precision 表示预测为正例的样本中，真实值为正例的比例，其定义如下所示。

$$\text{Precision}=\frac{\text{TP}}{\text{TP}+\text{FP}}\times 100\% \quad (11)$$

召回率（Recall）反映了检测算法对正例的分类能力，其等于正确预测为正例的样本数占总的正例样本数的比值，其定义如下所示。

$$\text{Recall}=\frac{\text{TP}}{\text{TP}+\text{FN}}\times 100\% \quad (12)$$

由于 Precision 和 Recall 具有一定的片面性，F1 值对它们进行了融合，其定义如式（13）所示。通常 F1 值越大，模型分类的综合性能越好。

$$\text{F1}=2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}\times 100\% \quad (13)$$

### 3.4　交叉验证中参数 $k$ 的分析及选取

本节分析交叉验证中参数 $k$ 对 RG-IDS 模型的预测性能的影响。选择数据集 CICIDS 2017 进行训练和预测。选择 $k=3,5,8,10$ 进行实验。$k$ 值对 RG-IDS 模型的性能影响（CICIDS 2017）见表 3，表 3 列出了 4 次实验对 7 类攻击的准确率、精确率、召回率和 F1 值。由表 3 可知，在 $k=3,5,8,10$ 的 4 次实验中，精确率、召回率和 F1 值均在 99.75% 以上，且 4 次实验的数据相差不大。表 3 中加粗的值表示所在行的最大值。考虑 $k$ 值的增加对复杂度的提升（$k$ 值越大，计算负担越重），在后续实验中，选取 $k$=5 进行实验。

表 3　$k$ 值对 RG-IDS 模型的性能影响（CICIDS 2017）

| 性能指标 | $k$=3 | $k$=5 | $k$=8 | $k$=10 |
|---|---|---|---|---|
| 准确率 | 98.870 5% | **99.002 9%** | 98.720 5% | 98.720 5% |
| 精确率 | 99.761 9% | 99.753 3% | **99.779 5%** | 99.770 7% |
| 召回率 | 99.761 7% | 99.752 9% | **99.779 4%** | 99.770 5% |
| F1 值 | 99.760 2% | 99.751 5% | **99.779 0%** | 99.769 0% |

### 3.5　基于数据集 CICIDS 2017 的模型性能分析

首先，分析 RG-IDS 模型在数据集 CICIDS 2017 上的测试性能。取数据集 CICIDS 2017 中 80% 的数据进行训练，剩余的 20% 进行测试。CICIDS 2017 的训练数据和增强后的数据分布情况见表 4。

表 4　CICIDS 2017 的训练数据和增强后的数据分布情况

| 类型 | 原训练数据 | 经 ADASYN 增强后的训练数据 |
|---|---|---|
| Benign | 18 184 | 18 184 |
| DoS | 15 228 | 18 201 |
| Portscan | 6 357 | 18 202 |
| Bruteforce | 2 213 | 18 176 |
| Webattack | 1 744 | 18 190 |
| Bot | 1 573 | 18 168 |
| Infiltration | 29 | 18 184 |

由表 4 可知，原训练数据在 7 类样本中的数量并不相同，且相差较大，存在显著的分布不平衡问题。例如，Benign 类样本的训练数据达到 18 184 个，而 Bot 类和 Infiltration 类样本的训练数据只有 1 573 和 29 个。经 ADASYN 算法增强后，7 类样本的分布达到平衡。RG-IDS 模型就利用这些增强后的数据训练模型。

RG-IDS 模型对 7 类样本的预测性能（CICIDS

2017）见表5。由表5可知，除了对Bot和Infiltration两类样本，RG-IDS模型对其他所有样本的F1值均达到100%。这说明RG-IDS模型能够有效地对攻击进行检测和分类。

表5　RG-IDS模型对7类样本的预测性能（CICIDS 2017）

| 样本编码 | 精确率 | 召回率 | F1值 |
|---|---|---|---|
| 0 | 100% | 100% | 100% |
| 3 | 100% | 100% | 100% |
| 5 | 100% | 100% | 100% |
| 2 | 100% | 100% | 100% |
| 6 | 100% | 100% | 100% |
| 1 | 98% | 99% | 98% |
| 4 | 100% | 71% | 83% |

测试Bot和Infiltration两类样本的F1值分别只达到98%和83%，原因在于它们的原训练数据样本较少。尽管通过ADASYN算法对它们的训练样本进行增强（生成新训练样本数），但是生成的训练样本数据的质量不及原始的训练样本数据。

考虑RG-IDS采用了stacking集成技术，其中将RF、GBDT和LightGBM作为第一层的基学习器，将LightGBM作为第二层的元学习器，3个基学习器和RG-IDS模型对同一份测试数据的F1值（CICIDS 2017）见表6。

表6　3个基学习器和RG-IDS模型对同一份测试数据的F1值（CICIDS 2017）

| 模型 | 7类样本的F1值 | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| RF | 99% | 93% | 100% | 100% | 77% | 100% | 96% |
| GBDT | 98% | 94% | 99% | 99% | 83% | 100% | 94% |
| LightGBM | 100% | 98% | 100% | 100% | 77% | 100% | 100% |
| RG-IDS | 100% | 98% | 100% | 100% | 83% | 100% | 100% |

由表6可知，RG-IDS模型的预测性能优于3个基学习器的预测性能。例如，在对样本类型1的预测中，RG-IDS模型的F1值达到98%，而RF和GBDT的F1值分别为93%和94%。这说明通过集成技术可以提升单个分类器的预测性能。

此外，由于攻击类型4的原训练数据样本较少，RG-IDS模型对其预测的性能不佳，F1值低至83%，原因在于攻击类型4（Infiltration样本）数量极少，只有29条。尽管通过ADASYN算法对其进行了增强，但是增强后的样本质量不及初始的样本质量，利用这些增强后的样本训练模型，导致模型的性能不高，最终导致RG-IDS模型对攻击类型4的F1值较低。

### 3.6　基于数据集NSL-KDD的模型性能分析

本节分析RG-IDS模型在数据集NSL-KDD上的测试性能。取数据集NSL-KDD中80%的数据进行训练，剩余的20%进行测试。NSL-KDD的训练数据和增强后的数据分布情况见表7。

由表7可知，数据集NSL-KDD中5类样本数分布存在不均衡问题。其中，U2R攻击的样本数只有162条，而DoS攻击的样本数达到43 478条。经ADASYN算法增强后，5类样本的分布达到平衡。RG-IDS模型就利用这些增强后的数据训练模型。

表7　NSL-KDD的训练数据和增强后的数据分布情况

| 类型 | 原训练数据 | 经ADASYN增强后的训练数据 |
|---|---|---|
| Normal | 61 642 | 61 642 |
| DoS | 43 478 | 61 793 |
| Probing | 11 302 | 61 528 |
| R2L | 2 228 | 61 660 |
| U2R | 162 | 61 662 |

RG-IDS模型对数据集NSL-KDD中5类样本的预测性能见表8。由表8可知，对于Normal和DoS样本的预测性能最好，F1值达到100%，而对于U2R攻击的预测性能最差，F1值只有85%，原因在于U2R的原训练样本数最少，尽管经

ADASYN算法进行了增强，但是增强后的样本质量不及原训练样本的质量。

表8　RG-IDS模型对NSL-KDD中5类样本的预测性能

| 样本编码 | 精确率 | 召回率 | F1值 |
|---|---|---|---|
| 0 | 100% | 100% | 100% |
| 1 | 99% | 99% | 99% |
| 2 | 95% | 93% | 94% |
| 3 | 97% | 75% | 85% |
| 4 | 100% | 100% | 100% |

3个基学习器和RG-IDS模型对NLS-KDD中5类样本的F1值见表9。由表9可知，RG-IDS模型的预测性能优于3个基学习器的预测性能。例如，在对类型2的预测中，RG-IDS模型的F1值达到94%，而RF、GBDT和LightGBM的F1值分别为93%、63%和94%。这说明通过集成技术可以提升分类器的预测性能。此外，由于类型3和类型2样本的原训练数据样本较少（只有162条和2 228条），RG-IDS模型对其的预测性能不佳，F1值低至85%和94%。

表9　3个基学习器和RG-IDS模型对NSL-KDD中5类样本的F1值

| 模型 | 5类样本的F1值 | | | | |
|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 |
| RF | 100% | 99% | 93% | 82% | 100% |
| GBDT | 99% | 96% | 63% | 38% | 97% |
| LightGBM | 100% | 99% | 92% | 82% | 99% |
| RG-IDS | 100% | 99% | 94% | 85% | 100% |

### 3.7　特征选择对RG-IDS模型性能的影响

首先，针对数据集CICIDS 2017进行了基于特征选择的操作（如第2.4节描述），然后再分别对未选择特征和选择特征的样本数据进行了RG-IDS模型训练和预测，特征选择对RG-IDS模型性能的影响（CICIDS 2017）见表10。由表10可知，利用GBDT选择特征，并没有降低对数据集CICIDS 2017中7类样本的预测性能，只有编码为4的类样本的精确率、召回率和F1值有所下降。类似地，也对数据集NSL-KDD进行了同样的分析，特征选择对RG-IDS模型性能的影响（NSL-KDD）见表11。由表11可知，经特征选择后对分类预测的影响非常有限，只有编码为2和3的类样本的F1值下降。

表10　特征选择对RG-IDS模型性能的影响（CICIDS 2017）

| 样本编码 | 未选择特征 | | | 选择特征 | | |
|---|---|---|---|---|---|---|
| | 精确率 | 召回率 | F1值 | 精确率 | 召回率 | F1值 |
| 0 | 100% | 100% | 100% | 100% | 100% | 100% |
| 3 | 100% | 100% | 100% | 100% | 100% | 100% |
| 5 | 100% | 100% | 100% | 100% | 100% | 100% |
| 2 | 100% | 100% | 100% | 100% | 100% | 100% |
| 6 | 100% | 100% | 100% | 100% | 100% | 100% |
| 1 | 98% | 99% | 98% | 99% | 99% | 99% |
| 4 | 100% | 71% | 83% | 80% | 57% | 67% |

表11　特征选择对RG-IDS模型性能的影响（NSL-KDD）

| 样本编码 | 未选择特征 | | | 选择特征 | | |
|---|---|---|---|---|---|---|
| | 精确率 | 召回率 | F1值 | 精确率 | 召回率 | F1值 |
| 0 | 100% | 100% | 100% | 100% | 100% | 100% |
| 1 | 99% | 99% | 99% | 99% | 99% | 99% |
| 2 | 95% | 93% | 94% | 93% | 94% | 93% |
| 3 | 97% | 75% | 85% | 86% | 80% | 83% |
| 4 | 100% | 100% | 100% | 100% | 0.99 | 100% |

特征选择对数据集CICIDS 2017和NSL-KDD的运算时间影响如图5所示。

由图5可知，无论在哪个数据集上，选择特征后，都可有效地降低运算时间。例如，在数据集CICIDS 2017上，若未选择特征，执行RG-IDS模型需要2 269.513 5 s，而选择特征再执行RG-IDS模型只需要931.308 2 s，有效缩短了运算时间。结合表10和表11可知，通过选择特征

可略微降低模型的分类性能，但有效地缩短了运算时间。

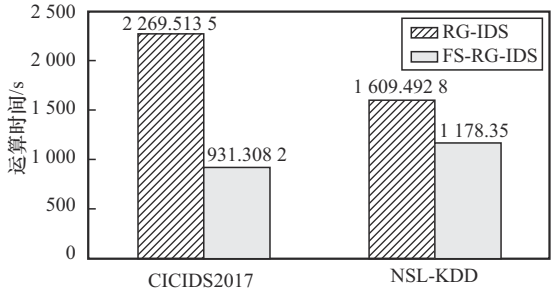


图5　特征选择对数据集CICIDS 2017和NSL-KDD的运算时间影响

### 3.8　与深度学习模型的比较

为了更好地评估RG-IDS模型的性能，选择卷积神经网络（convolutional neural network，CNN）、时序卷积网络（temporal convolutional network，TCN）和门控循环单元（gated recurrent unit，GRU）作为基准模型，同RG-IDS模型进行性能对比分析，分析它们的平均准确率性能。此外，为了分析ADASYN算法在提升检测性能的优势，选择采用SMOTE算法作为一个基准，并标记为RG-SM。RG-SM模型与RG-IDS模型的不同之处在于：RG-IDS模型采用ADASYN算法选择特征，而RG-SM模型采用SMOTE算法选择特征。

不同模型在数据集CICIDS 2017和NSL-KDD上的平均准确率如图6所示。由图6可知，相比于GRU、CNN和TCN这3个深度学习模型，提出的RG-IDS模型提升了平均准确率。例如，在数据集CICIDS 2017上，RG-IDS模型的平均准确率达到98.4%，而GRU、CNN和TCN模型的平均准确率分别只有85.3%、89.1%和94.3%。此外，相比于RG-SM模型，RG-IDS模型均提高了在数据集CICIDS 2017和NSL-KDD上的平均准确率。这证实了ADASYN算法在选择特征方面的优势，通过选择信息表征能力更强的特征，可有效地提升检测精度。

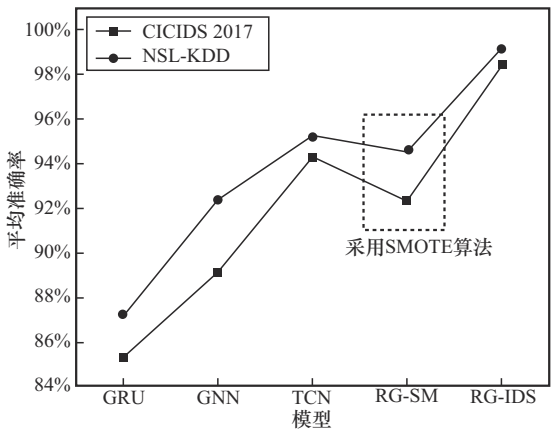


图6　不同模型在数据集CICIDS 2017和NSL-KDD上的平均准确率

## 4　结束语

本文针对车联网的入侵检测问题，提出了融合RF和GBDT的stacking的检测模型。该模型利用ADASYN算法对样本数据进行增强，形成相对平衡的数据，并通过GBDT评估特征重要性，丢弃无关特征，进而减少运算量，提升训练模型的速度。同时，采用$k$折交叉验证的stacking方法，降低过拟合的概率。利用数据集CICIDS 2017和NSL-KDD测试验证RG-IDS模型的性能。测试结果表明，提出的RG-IDS模型在两个数据集中的绝大多数样本上的F1值达到99%以上。此外，通过选择特征可有效降低计算量，提升了模型的检测效率。

## 参考文献：

[1] 江荣旺, 魏爽, 龙草芳, 等. 基于增量学习的车联网恶意位置攻击检测研究[J]. 信息安全研究, 2024, 10(3): 268-276.
JIANG R W, WEI S, LONG C F, et al. Research on location attack detection of VANET based on incremental learning[J]. Journal of Information Security Research, 2024, 10(3): 268-276.

[2] 于天琪, 胡剑凌, 金炯, 等. 基于移动边缘计算的车载CAN网络入侵检测方法[J]. 计算机科学, 2021, 48(1): 34-39.
YU T Q, HU J L, JIN J, et al. Mobile edge computing based in-vehicle CAN network intrusion detection method[J]. Computer

Science, 2021, 48(1): 34-39.

[3] 银鹰, 周志洪, 姚立红. 基于LSTM的CAN入侵检测模型研究[J]. 信息网络安全, 2022, 22(12): 57-66.
YIN Y, ZHOU Z H, YAO L H. Research on LSTM-based CAN intrusion detection model[J]. Netinfo Security, 2022, 22(12): 57-66.

[4] 宋秀兰, 李洋阳, 何德峰. 外部干扰和随机DoS攻击下的网联车安全$H_\infty$队列控制[J]. 自动化学报, 2024, 50(2): 348-355.
SONG X L, LI Y Y, HE D F. Secure $H_\infty$ platooning control for connected vehicles subject to external disturbance and random DoS attacks[J]. Acta Automatica Sinica, 2024, 50(2): 348-355.

[5] 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.
ZHOU M. Research on attack detection method based on behavior analysis in vehicle networking[D]. Wuhan: Huazhong University of Science and Technology, 2022.

[6] 周建华, 侯英哲, 吕臣臣, 等. 智能网联汽车安全防护技术研究综述[J]. 武汉大学学报(理学版), 2023, 69(5): 617-635.
ZHOU J H, HOU Y Z, LYU C C, et al. A review on security and defense technologies for intelligent connected vehicle[J]. Journal of Wuhan University (Natural Science Edition), 2023, 69(5): 617-635.

[7] 赵丽, 孙敏. 基于CAN的现代车辆入侵检测[J]. 计算机应用与软件, 2024, 41(2): 328-332.
ZHAO L, SUN M. Modern vehicle intrusion detection based on can[J]. Computer Applications and Software, 2024, 41(2): 328-332.

[8] 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.

[9] 冯光升, 蒋舜鹏, 胡先浪, 等. 面向物联网的入侵检测技术研究新进展[J]. 信息网络安全, 2024, 24(2): 167-178.
FENG G S, JIANG S P, HU X L, et al. New research progress on intrusion detection techniques for the Internet of Things[J]. Netinfo Security, 2024, 24(2): 167-178.

[10] DORIGUZZI-CORIN R, MILLAR S, SCOTT-HAYWARD S, et al. Lucid: a practical, lightweight deep learning solution for DDoS attack detection[J]. IEEE Transactions on Network and Service Management, 2020, 17(2): 876-889.

[11] CIL A E, YILDIZ K, BULDU A. Detection of DDoS attacks with feed forward based deep neural network model[J]. Expert Systems with Applications, 2021, 169: 114520.

[12] 魏明军, 李凤, 刘亚志, 等. 基于改进WGAN-GP和ResNet的车联网入侵检测方法[J]. 郑州大学学报(工学版), 2024, 45(4): 30-37.
WEI M J, LI F, LIU Y Z, et al. An intrusion detection method for Internet of vehicles based on improved WGAN-GP and ResNet[J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(4): 30-37.

[13] 方介泼, 陶重犇. 应对零日攻击的混合车联网入侵检测系统[J]. 计算机应用, 2024, 44(9): 2763-2769.
FANG J P, TAO C B. Hybrid internet of vehicles intrusion detection system for zero-day attacks[J]. Journal of Computer Applications, 2024, 44(9): 2763-2769.

[14] 王军, 司昌馥, 王凯鹏, 等. 基于集成学习和改进的PSO-GA算法特征选择的入侵检测方法[J]. 吉林大学学报(工学版), 2024: 1-9.
WANG J, SI C F, WANG K P, et al. Intrusion detection method based on ensemble learning and improved PSO-GA feature selection[J]. Journal of Jilin University (Engineering and Technology Edition), 2024: 1-9.

[15] BHATI B S, CHUGH G, AL-TURJMAN F, et al. An improved ensemble based intrusion detection technique usingXGBoost[J]. Transactions on Emerging Telecommunications Technologies, 2021, 32(6): 4076-4090.

[16] YANG L, MOUBAYED A, HAMIEH I, et al. Tree-based intelligent intrusion detection system in Internet of vehicles[C]//Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM). Piscataway: IEEE Press, 2019: 1-6.

[17] ELMASRY W, AKBULUT A, ZAIM A H. Evolving deep learning architectures for network intrusion detection using a double PSO metaheuristic[J]. Computer Networks, 2020, 168: 107042.

[18] 陈虹, 姜朝议, 金海波, 等. 融合改进堆叠编码器和多层BiLSTM的入侵检测模型[J]. 计算机工程与应用, 2024: 1-12.
CHEN H, JIANG C Y, JIN H B, et al. Fusion of improved stacked encoder and multi-layer BiLSTM for intrusion detection model[J]. Computer Engineering and Applications, 2024: 1-12.

[19] 周杰英, 贺鹏飞, 邱荣发, 等. 融合随机森林和梯度提升树的入侵检测研究[J]. 软件学报, 2021, 32(10): 3254-3265.
ZHOU J Y, HE P F, QIU R F, et al. Research on intrusion detection based on random forest and gradient boosting tree[J]. Jour-

nal of Software, 2021, 32(10): 3254-3265.

[20] 王秀玉, 吴晓鸽, 冯永晋. 融合过-欠采样与GAN的网络入侵检测方法[J]. 小型微型计算机系统, 2024: 1-8.
WANG X Y, WU X L, FENG Y J. Network intrusion detection method combining over-undersampling with GAN[J]. Journal of Chinese Computer Systems, 2024: 1-8.

[作者简介]

**徐会彬**（1982–），男，博士，湖州师范学院信息工程学院讲师、硕士生导师，主要研究方向为VANET安全、路由技术。

**方龙**（2000–），男，湖州师范学院信息工程学院硕士生，主要研究方向为车联网安全及入侵检测。

**张莎**（2002–），女，湖州师范学院信息工程学院在读，主要研究方向为电子信息、计算机应用技术。