基于 Stacking 的 DDoS 攻击检测方法

付国庆 李俭兵 高雨薇
(重庆邮电大学通信与信息工程学院 重庆 400065)
(重庆邮电大学通信新技术应用研究中心 重庆 400065)
(重庆信息设计有限公司 重庆 401121)

摘 要 近年来 DDoS 攻击检测多采用机器学习的方法,Stacking 便是其一,现阶段 Stacking 初级学习器的配置方法多为固定搭配,但由于 DDoS 攻击的复杂性和动态性,静态的配置策略显得灵活性较差。对此提出 QGA-Stacking 算法,即利用量子遗传算法(QGA)动态地选取 Stacking 中评价指标最高的一组学习器组合,从而提高检测模型的准确性和灵活性;提出一组最佳特征集来节省计算成本。经过实验对比,充分证明了 QGA-Stacking 算法相较于其他 3 种主流算法,其检测性能更加显著,最佳特征集的选取也较为合理。

关键词 网络空间安全 DDoS 攻击检测 集成学习 Stacking 量子遗传算法

中图分类号 TP393.08 **文献标志码** A DOI:10.3969/j.issn.1000-386x.2024.03.049

DDOS ATTACK DETECTION METHOD BASED ON STACKING

Fu Guoqing Li Jianbing Gao Yuwei
(*School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*)
(*Research Center of New Telecommunication Technology Applications, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*)
(*Chongqing Information Technology Designing Co., Ltd., Chongqing 401121, China*)

Abstract In recent years, DDoS attack detection has mostly adopted machine learning methods, and Stacking is one of them. The current stacking base-learner configuration method is mostly fixed collocation. Due to the complexity and dynamics of DDoS attacks, static configuration strategy is obviously less flexible. In this regard, the QGA-Stacking algorithm is proposed, which uses quantum genetic algorithm (QGA) to dynamically select a group of learner combinations with the highest evaluation index in Stacking, thereby improving the accuracy and flexibility of the detection model. At the same time, a set of optimal feature sets was proposed to save computational cost. Through experimental comparison, it is fully proved that the QGA-Stacking algorithm has more significant detection performance than the other three mainstream algorithms, and the selection of the best feature set is more reasonable.

Keywords Cyberspace security DDoS attack detection Ensemble learning Stacking Quantum genetic algorithm (QGA)

0 引 言

DDoS 攻击是当前网络空间安全的主要威胁之一,它衍生于传统的 DoS 攻击之中,因其易实现、技术成本低、追踪困难、破坏力强等特点,受到了众多攻击者的追捧,致使现如今的 DDoS 攻击事件与日俱增。

据绿盟科技 2019 年统计数据:全年 DDoS 攻击次数为 16.74 万次,攻击总流量为 43.68 万 TB,与 2018 年同期相比,攻击次数增加了 30.2%[1]。由此可以看出当前 DDoS 攻击横行网络,安全防御形势依旧十分严峻。在当前大数据、云计算、物联网高速发展的大环

收稿日期:2020-11-23。重庆市教育委员会科学技术研究项目(KJQN202000647)。**付国庆**,硕士生,主研领域:机器学习,网络安全。**李俭兵**,高工。**高雨薇**,硕士生。

境下,各式各样的设备和虚拟机数量暴增的同时,也衍生出了诸多问题,如廉价的虚拟机安全隐患多,用户的安全意识薄弱等,这就有可能为攻击者组建僵尸网络提供机会。因此,在当前时代解决 DDoS 攻击的问题将会是一项任重而道远的研究课题。

1 研究现状

随着人工智能的发展,传统机器学习、神经网络、深度学习等算法被人们应用于 DDoS 的攻击检测中。文献[2]通过改进 KNN 算法来提高检测效率,但存在计算量大且速度较慢等问题;文献[3]中结合长、短期记忆(LSTM)和贝叶斯方法进行 DDoS 检测,LSTM 模块以高置信度输出识别 DDoS 攻击的部分,然后使用贝叶斯方法判断置信度较低的输出,从而提高其准确性;文献[4]中分别使用遗传算法(GA)和人工神经网络(ANN)进行特征选择和攻击检测,但参数较多,训练时间过长;文献[5]结合了功率谱密度(PSD)熵函数和 SVM,从而在正常流量中检测出攻击流量。尽管这些单一的分类方法取得了飞速发展,但在实际操作中仍会遇到这些方法不能有效解决的问题。于是,Hansen 等[6]提出了新的机器学习方法—集成学习,以实现 DDoS 攻击的精准检测。文献[7]采用 GBDT 算法进行 DDoS 检测,结果表明,其检测精度高于其他算法,测试时间也短于其他算法;文献[8]提出了一种利用模糊逻辑选择正确的分类器的集成框架,以达到检测 DDoS 攻击的目的;文献[9]利用 Voting 机制组合了 DT-C4.5、KNN、NN、SVM 等分类器,并利用具有简化特征的数据集(NSL-KDD),最终得到了一个效果上佳的检测结果。

长期以来,DDoS 攻击检测一直是网络空间安全领域的研究热点,然而,DDoS 的攻击和检测一直是一个相生相克的过程,随着检测和防御技术的不断升级、更新,攻击者们也在不断挖掘新的漏洞,发起新的攻击。目前,DDoS 攻击的手段越来越复杂多变、越来越先进,攻击目标也越来越明确,其所产生的危害不容小觑。因此,快速提高 DDoS 攻击检测的准确性成了当务之急。

本文的研究内容主要有两点:(1) 为了提升 DDoS 攻击检测的准确率,提出一种基于集成学习的 QGA-Stacking 算法,以检测模型输出的准确率作为评价指标,然后通过 QGA 动态地选取评价指标最高的一组分类器组合,作为初级学习器,来提高检测模型的准确性和灵活性;(2) 本文使用数据全、攻击场景新、攻击种类多的 CICDDoS2019 数据集作为实验数据,并对其进行分析处理,提出特征数量少的最佳特征集,做到在尽量不损失检测性能的前提下节省计算成本,降低计算复杂度。

2 DDoS 攻击检测方法

2.1 集成学习

集成学习是通过在数据上构建多个模型,并将所有模型的建模结果进行集成,最终得到的一个综合成果,以此来获取比单个模型更好的表现效果。集成又分为同质和异质,同质是指集成中只包含同种类型的个体学习器,如 Bagging[10] 系列算法(随机森林,RF)和 Boosting 系列算法(AdaBoost[11])。异质是指集成中包含不同类型的个体学习器,如 Stacking 系列算法。本文便是基于异质的 Stacking 算法进行 DDoS 攻击检测。

严格来说,Stacking[12] 并不是一种算法,而是一种集成策略,它可以聚合多种不同的模型,来追求比单一模型更好的表现效果,著名机器学习竞赛网站 Kaggle 也经常存在 Stacking 的身影。这种集成方法一般分为两个阶段:第一阶段是训练出多个模型并得到各自的结果;第二阶段再用前一阶段的输出结果训练次级模型,最后由第二阶段的训练模型输出最终的预测结果。Stacking 的算法伪代码如下:

输入:训练集 D;初级学习算法$\varsigma_1,\varsigma_2,\cdots,\varsigma_T$;次级学习器算法$\varsigma$。

输出:$H(x)=h'(h_1(x),h_2(x),\cdots,h_T(x))$。 //预测结果

```
for t = 1,2,…,T do                    //依次训练基学习算法
  h_t = ς_t(D);                       //产生基学习器
end for
D' = ∅;                               //初始化次级训练集
for i = 1,2,…,m do
for t = 1,2,…,T do
  z_it = h_t(x_i);                    //次级训练集的特征样本
end for
D' = D' ∪ ((z_i1,z_i2,…,z_iT),y_i);   //产生次级训练集
end for
h' = ς(D');                           //在次级训练集上产生元学习器
```

2.2 量子遗传算法

量子遗传算法[13] 属于智能优化算法的一种,它在 GA 的基础上引入了量子位和量子门的概念,该算法的提出推动了量子计算和智能优化算法的融合发展。

QGA 相较于传统的 GA 的不同之处在于,其染色体上的基因采用概率幅的方式表示,不再是非 0 即 1 的两种状态,而是存在三种不同的状态,分别是 $|0\rangle$

态、|1〉态、|0〉和|1〉的叠加态,在同等长度下量子位的编码方式比传统遗传算法能承载更多的信息量。一个量子位可定义为$\begin{bmatrix}\alpha\\ \beta\end{bmatrix}$,则 m 个量子位可以定义为:

$$q=\left[\begin{array}{c|c|c|c}\alpha_1 & \alpha_2 & \cdots & \alpha_m\\ \beta_1 & \beta_2 & \cdots & \beta_m\end{array}\right] \tag{1}$$

式中:α、β 均是复数,分别表示|0〉态和|1〉态的概率幅,且满足 $|\alpha_i|^2+|\beta_i|^2=1, i=1,2,\cdots,m$。这种描述的优点是可以表达任意量子叠加态,所以相较于传统 GA 拥有更好的种群多样性和收敛性。

QGA 是一种概率搜索算法,拥有一个量子种群 $Q(t)=\{q_1^t,q_2^t,L,q_n^t\}$。其中:$n$ 表示种群的规模;t 表示遗传的代数;q_j^t 表示一条量子染色体。其定义如式(2)所示。

$$q_j^{\ t}=\left[\begin{array}{c|c|c|c}\alpha_1^t & \alpha_2^t & \cdots & \alpha_m^t\\ \beta_1^t & \beta_2^t & \cdots & \beta_m^t\end{array}\right] \tag{2}$$

式中:m 是量子位数,即染色体的长度,$j=1,2,\cdots,n$。QGA 流程如图 1 所示。

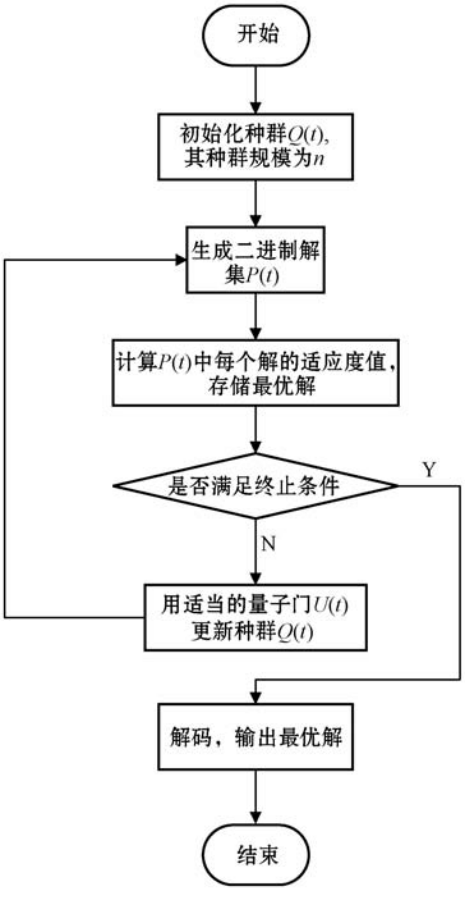

图 1　量子遗传算法流程

由于在量子编码作用下的染色体不再是单一的纯态,所以 QGA 没有使用传统的选择、交叉和突变等操作,而是采用量子旋转门来实现对种群 $Q(t)$ 的更新,其更新过程可用式(3)进行描述。

$$\begin{bmatrix}\tilde{\alpha}_i\\ \tilde{\beta}_i\end{bmatrix}=\begin{bmatrix}\cos\theta_i & -\sin\theta_i\\ \sin\theta_i & \cos\theta_i\end{bmatrix}\begin{bmatrix}\alpha_i\\ \beta_i\end{bmatrix} \tag{3}$$

式中:(α_i,β_i) 是第个量子位;θ_i 为旋转角。

2.3　QGA-Stacking 算法

集成算法的缺点是随着初级学习器的增加,其学习速率和存储空间也是在急剧增加,为了解决这类问题,选择集成[14]的思想被提出,旨在使用尽可能少的学习器来达到更好的学习性能;选择集成按照选择的方法可以分为静态选择法和动态选择法[15],本文便是使用智能启发算法中的 QGA 来实现集成的动态选择。智能启发算法常用于参数调节、特征优化、集成选择、原型优化、加权投票集成、核函数学习[16]。

对于 Stacking,其学习器的搭配问题一直是学术界研究的热点,如有 n 个初级学习器和 m 个次级学习器可供选择,其组合方案总数多达 $m\cdot 2^n-1$ 种,若通过遍历的方法选取最优的组合方案会使计算成本急剧增加。鉴于现阶段的 Stacking 学习器的配置方法多为固定搭配,但由于 DDoS 攻击的复杂性和动态性,固定搭配显得灵活性较差,其主要原因有两个:(1) 固定搭配的方法过于局限,经常会随着样本数量和样本特征的不同,其最终的预测结果也不尽相同;(2) 效果出众的 Stacking 配置一般由两层模型组合决定的,固定的配置方法多是基于次级学习器的选择,从而忽略了初级学习器的选择。

对此,本文利用 QGA 来解决 Stacking 的学习器配置问题,它可以动态地选择出准确率最高的一组分类器组合,作为初级学习器;初级学习器的配置取决于评价指标的选取,评价指标的选择多种多样,如准确率、精确率、召回率等。精确率代表对正样本结果中的预测准确程度,若单独使用,有一定的局限性;召回率是在实际为正的样本中被预测为正样本的概率,也具有一定的局限性,如多增加样本便可使其值增高。准确率则代表整体的预测准确程度,包括正样本和负样本,并且其计算简单,时间复杂度也较低,虽说样本不平衡的问题会使其受到影响,但可以通过多种方法(如上下采样)使数据集的正负样本保持平衡,加之本文主要也是为了提升 DDoS 攻击的准确率,所以模型选用准确率作为其评价指标。

为增强 Stacking 模型的泛化性能,次级学习器的选择也是其研究热点之一。已有研究表明,对于初级学习器输出的类概率分布,次级学习器选用 MLR 算法,在诸多数据集中的表现效果均较好[17],所以本文也采用 MLR 算法作为本文模型的次级学习器。QGA-

Stacking 流程如图 2 所示。

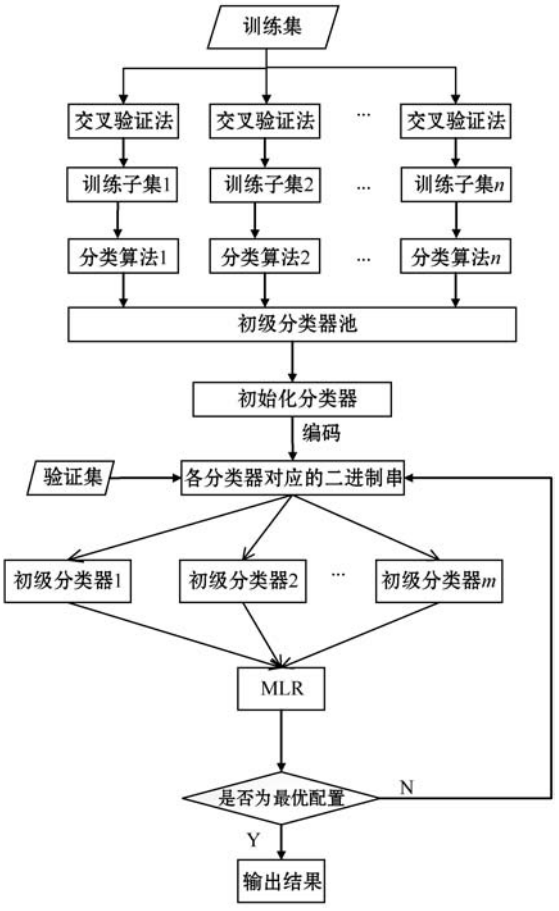

图 2 QGA-Stacking 算法流程

在初级分类器的选择问题中，首先将训练集 D 通过交叉验证法训练 n 个分类算法 $\varsigma_i(i=1,2,\cdots,n)$ 得到 n 个分类器 $B_i(i=1,2,\cdots,n)$，将其构成算法池 $B_P=\{B_1,B_2,\cdots,B_n\}$。

初始化分类器池，将所有分类器的概率幅都初始化为 $1/\sqrt{2}$。这样做的目的是使遗传代为 0 的情况下，分类器能以相同的概率存在于线性的叠加态之中，即：

$$|\psi_{q_j^0}\rangle = \sum_{k=1}^{2^m} \frac{1}{\sqrt{2^m}} |s_k\rangle \tag{4}$$

式中：s_k 由二进制串 $(x_1,x_2,\cdots,x_m)$ 描述的第 k 状态，$x_i=0,1,i=1,2,\cdots,m$。通过观察分类器池的状态来生成二进制解集 $P(t)=(x_1^t,x_2^t,\cdots,x_n^t)$，每个解 $x_j^t(j=1,2,\cdots,n)$ 是一个长度为 m 的二进制串，其值由相应量子位的概率 $|\alpha_i^t|^2$ 或 $|\beta_i^t|^2(i=1,2,\cdots,m)$ 决定。

Stacking 配置可以表示为 $M\cdot B_C$，其中：M 代表次级分类器 MLR；B_C 为算法池中选出的初级分类器的组合。则 Stacking 分类器的最终组合配置可由式(5)表示。

$$I=\{M\cdot B_C \mid B_C\subseteq B_P\} \tag{5}$$

通过验证集，I 中包含的每一种配置 s 都对应一个评价指标 $E(s)$，其评价指标以模型的准确率为准。Stacking 配置选择的目标就是要从 I 中寻找评估值最优的配置 H，由式(6)表示。

$$H=\arg\max_{s\in I}\{E(s)\} \tag{6}$$

s 对应个体的二进串为 P_s，其适应度值由式(7)定义。

$$Fitness(P_s)=E(s) \tag{7}$$

计算每个解的适应度值，选取当前状态下的配置方案，与存储的配置方案进行比较，选取最优配置方案将其更新，最后使用量子门来对种群进行更新。经过不断循环，直至获得全局最优配置，最后再解码输出结果。

3 实 验

3.1 数据集

基于机器学习的检测方法严重依赖于一个好的数据集，但目前大多数 DDoS 攻击的数据集存在许多缺点和问题，例如流量数据不完整、匿名数据和过时的攻击场景等，这使得研究人员难以找到全面而有效的数据集来测试、评估其提出的检测和防御模型，针对目前存在的主要缺点和局限性，本文实验采用 CICDDoS2019 数据集[18]，该数据集涵盖各种 DDoS 攻击技术和方案，包含 NTP、TFTP、DNS、LDAP、NetBIOS、SNMP、MSSQL、SSDP、SYN Flood、UDP Flood 和 UDP-Lag 等 11 种攻击方式，其攻击类别的样本数量分布如图 3 所示。

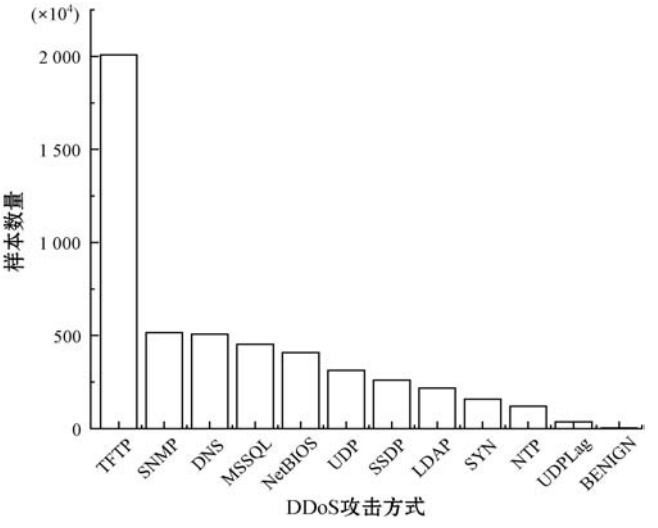

图 3 CICDDoS2019 数据类分布

(1) CICDDoS2019 数据集处理。获取数据后一般要对其进行预处理和特征选择。数据预处理是去除噪声，让数据集更适应模型的需求，主要包括数据清洗、数据分类和无量纲化等。先将 11 种不同攻击方式的数据整合后进行数据清洗，因存有缺失值的样本很少，所以将缺失样本删除；对时间戳特征进行切分和再重组，对源 IP 地址和目的 IP 地址特征等进行有效值提取；然后进行数据分类与编码，标签分为 0 和 1，0 代表

所有的攻击样本,1 代表正常流量的样本;无量纲化分为归一化和标准化,本文选择数据标准化方法对数据进行缩放,其公式如下:

$$x_{\text{std}}^{i}=\frac{x^{i}-\mu_{x}}{\sigma_{x}} \tag{8}$$

式中:μ_x和 σ_x分别为一组特征数据中的均值和标准差;x^i是特征中的第 i 组数据。

特征选择的目的是降低计算成本、提升模型上限,主要方法有过滤法(Filter)、嵌入法(Embedded)和包装法(Wrapper)。过滤法计算速度快,但比较粗糙,包装法和嵌入法更精确,但计算量大、运行时间长。在特征数目相同时,包装法与嵌入法的效果能够匹敌,但包装法计算速度更快。由此本文选用过滤法结合包装法的方法,首先用过滤法对数据特征进行粗略筛选,再使用包装法对特征进行精确选择。

过滤法分为方差过滤、卡方过滤、F 检验和互信息法。本文首先使用方差过滤,因为特征方差越小,代表特征对于样本的区分程度越低,所以优先消除方差为 0 的特征,然后使用 threshold 为中位数时进行卡方过滤,此时数据集的特征数量已从 80 降至 34,经测试,过滤后的数据表现效果并未降低,则在此基础上再进行卡方过滤。卡方过滤能依照卡方统计量由高到低为特征排名,并能选出前 k 个得分最高的特征,所以 k 值的选取尤为重要。卡方检验会返回卡方值和 p 值两个统计量,其中卡方值很难界定有效范围,所以一般会根据 p 值选择 k,即 p 值小于 0.05 的特征就是和标签相关的特征。经过卡方过滤特征数量降至 27 个。

包装法的目标函数选用递归特征消除法(Recursive Feature Elimination,RFE),它是一种贪婪的优化算法,旨在找到性能最佳的特征子集。它反复创建模型,并在每次迭代时保留最佳特征或剔除最差特征,下次迭代时,它会使用上一次建模中没有被选中的特征来构建模型,直到所有特征耗尽后,根据保留或剔除特征的顺序来排名,最终选择出最佳子集。通过绘制学习曲线得到最佳阈值为 12,所以选取排名前 12 的特征作为最优特征,分别为:Packet Length Max, Subflow Bwd Bytes, ACK Flag Count, Fwd Packet Length Max, Flow IAT Max, Flow IAT Mean, Fwd IAT Max, Fwd IAT Mean, min_seg_size_forward, Flow Duration, Protocol, Init_Win_bytes_forward

(2) 构造实验所需数据集。本实验构造两大类数据集,将原有数据集经过上述数据预处理过程后得到的数据集定义为 Set1 系,在 Set1 系的基础上利用上述特征选择方法可得出一组最优特征集合,将其定义为 Set2 系。Set1 系和 Set2 系又各分为样本数量为 5 万条、10 万条、20 万条和 50 万条的 8 个数据集,如表 1 所示。正常样本数量仅为 56 863 条,对于正常样本数量不足所造成的不平衡现象,采用 SMOTE 算法的上采样使得正常样本和异常样本的数据量保持一致。

表 1　实验数据集分类

Set1 系	特征数量	样本数量	Set2 系	特征数量	样本数量
Set11	80	5 万	Set21	12	5 万
Set12	80	10 万	Set22	12	10 万
Set13	80	20 万	Set23	12	20 万
Set14	80	50 万	Set24	12	50 万

构建的两大系数据集,可以对比出不同算法在不同的特征数量、样本数量中的表现差异;且通过 Set1 系和 Set2 系在不同样本数量、不同算法中的对比,也能验证最优特征集合的选择是否合理。

3.2　实验设计

(1) 实验环境。实验平台为 CPU Intel i7-8750H,内存 16 GB,Python 3.6.5。

(2) 实验方案。为验证 QGA-Stacking 算法的表现效果,在表 2 的 8 个数据集上,将其与目前主流的 RF、SVM 和 XGBoost 算法进行对比。

实验中本文仅训练 6 种具有代表性的单一分类算法作为 QGA-Stacking 的候选学习器,详见表 2。

表 2　实验算法分类

学习器	算法名称	算法描述
初级学习器	ID3	基于信息增益的决策树算法
	CART	基于 Gini 系数的决策树算法
	BernoulliNB	伯努利分布的朴素贝叶斯算法
	KNN	基于欧氏距离的 K 近邻算法
	MLP	基于深度学习的神经网络算法
	LR	不能拟合非线性的逻辑回归算法
次级学习器	MLR	可以拟合非线性的逻辑回归算法

(3) 评估标准。混淆矩阵是评判模型结果的指标,如表 3 所示。

表 3　混淆矩阵

真实情况	预测结果	
	正例	反例
正例	T_P(真正例)	F_N(假反例)
反例	F_P(假正例)	T_N(真反例)

本文的评价标准包括准确率(A_{cc})和误报率(F_{PR}),公式如下:

$$A_{cc}=\frac{T_P+T_N}{T_P+T_N+F_P+F_N} \tag{9}$$

$$F_{PR}=\frac{F_P}{F_P+T_N} \tag{10}$$

3.3 结果分析

将 QGA-Stacking 和其他三种算法(RF、SVM 和 XGBoost)在样本数量为 50 万的数据集 Set14 和 Set24 上进行对比,详情可见表 4。

表 4 4 种算法的攻击检测评价指标比较

数据集(Set)	分类算法	准确率/%	误报率
Set14	RF	95.98	0.036 4
	SVM	95.23	0.030 8
	XGBoost	96.33	0.021 2
	QGA-Stacking	**99.01**	**0.009 3**
Set24	RF	95.02	0.039 7
	SVM	94.62	0.034 9
	XGBoost	96.04	0.028 5
	QGA-Stacking	**98.43**	**0.010 5**

由表 5 不难得出,QGA-Stacking 算法的准确率均优于其他三种算法,且误报率也最低。

在特征个数相同、样本数量不同的情况下,图 4 为拥有 80 个特征的 Set1 系数据集在样本数量分别是 5 万条、10 万条、20 万条和 50 万条时 4 种算法的准确率对比。可以看出,随着样本数量的增加各算法的准确率也随之增加,其中 QGA-Stacking 算法的准确率最高,且随着样本数量的逐渐增加,其准确率的提升幅度也最明显。

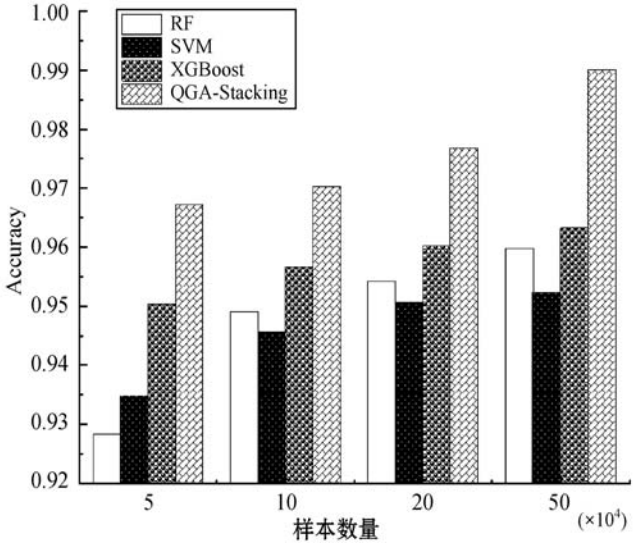

图 4 Set1 系数据集的准确率对比

在样本数量相同、特征个数不同的情况下,图 5 为 50 万条数据时,4 种算法分别在 80 个特征和 12 个特征上的准确率对比。可以看出,Set2(12 个特征)的准确率稍低于 Set1(80 个特征),就 QGA-Stacking 算法来看,虽然其准确率下降了 0.58%,但却大大节省了计算成本,由此可以证明最优特征集合的选择是合理的。

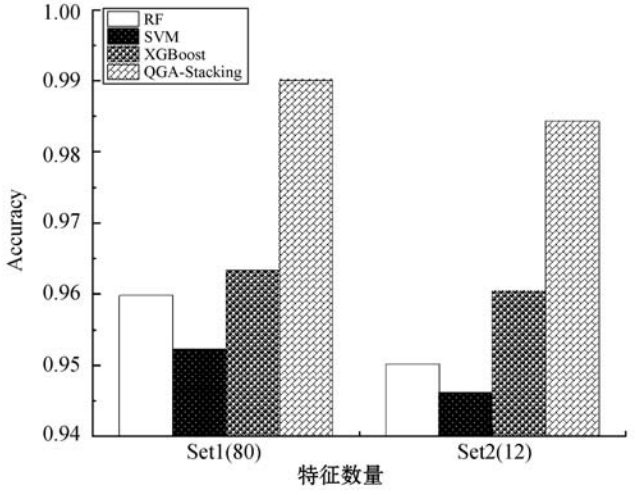

图 5 不同特征数的准确率对比

图 6 是在不同特征数量的数据集中 4 种算法的误报率对比情况。显而易见的是,随着样本数量的逐渐增加,4 种分类算法的误报率都随之减少,对比之下不难发现,QGA-Stacking 算法与其他 3 种算法相比误报率是最低的。

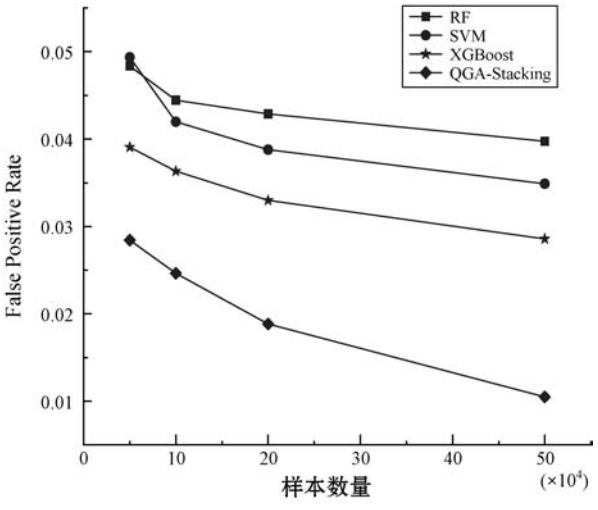

图 6 Set2 系数据集的误报率对比

经过上述各项指标之间的对比,可以充分地证明本文在数据集的选择和处理上是合理的,QGA-Stacking 算法与其他 3 种分类算法相比是有明显优势的,可以在 DDoS 攻击检测中发挥出显著的效果。

4 结 语

本文提出一种基于智能启发算法和选择集成的 QGA-Stacking 算法,旨在提升 DDoS 攻击检测的性能。该算法以检测模型输出的准确率作为最终的评价指

标,然后通过 QGA 动态地选取准确率最高的一组分类器组合,作为初级学习器,来提高检测模型的准确性和灵活性;次级学习器采用 MLR 算法,从而构建 DDoS 攻击检测模型。数据集选用 CICDDoS2019,经过分析处理提出一组最优特征集,并和原有的所有特征分别构造了 Set2 系和 Set1 系数据集,在这两类数据集上分别验证了 4 种分类算法的准确率。实验结果表明,各算法在 Set2 系中的各项指标只是稍低于 Set1 系,且本文所提出的 QGA-Stacking 算法在各项评价指标中均优于 RF、SVM 和 XGBoost 等三种分类算法。这充分证明了本文算法的检测性能出众,最佳特征集的选取也较为合理。下一步的研究方向将致力于在当前大数据背景下,解决大规模网络中海量数据的 DDoS 攻击检测性能和检测的实时性问题。

参 考 文 献

[1] 绿盟科技. 2019-DDoS 攻击态势报告[R]. 绿盟科技,2019.

[2] 肖甫,马俊青,黄洵松,等. SDN 环境下基于 KNN 的 DDoS 攻击检测方法[J]. 南京邮电大学学报(自然科学版),2015,35(1):84 – 88.

[3] Li Y, Lu Y. LSTM-BA: DDoS detection approach combining LSTM and Bayes[C]//2019 7th International Conference on Advanced Cloud and Big Data,2019:180 – 185.

[4] Barati M, Abdullah A, Udzir N, et al. Distributed denial of service detection using hybrid machine learning technique [C]//2014 International Symposium on Biometrics and Security Technologies,2014:268 – 273.

[5] Zhang N, Jaafar F, Malik Y. Low-rate DoS attack detection using PSD based entropy and machine learning[C]//2019 6th IEEE International Conference on Cyber Security and Cloud Computing/2019 5th IEEE International Conference on Edge Computing and Scalable Cloud,2019:59 – 62.

[6] Hansen L, Salamon P. Neural network ensembles[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990,12(10):993 – 1001.

[7] Zhang J, Liang Q, Jiang R, et al. A feature analysis based identifying scheme using GBDT for DDoS with multiple attack vectors[J]. Applied Sciences,2019,9(21):4633.

[8] Alsirhani A, Sampalli S, Bodorik P. DDoS detection system: Using a set of classification algorithms controlled by Fuzzy logic system in Apache spark[J]. IEEE Transactions on Network and Service Management,2019,16(3):936 – 949.

[9] Das S, Mahfouz A, Venugopal D, et al. DDoS intrusion detection through machine learning ensemble[C]//2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion,2019:471 – 477.

[10] Breiman L. Bagging predicators[J]. Machine Learning, 1996,24:123 – 140.

[11] Freund Y, Schapire R. A decision-theoretic generalization of on-line learning and an application to boosting[J]. Journal of Computer and System Sciences,1997,55(1):119 – 139.

[12] Wolpert D. Stacked generalization[J]. Neural Networks, 1992,5(2):241 – 259.

[13] Han K, Kim J. Genetic quantum algorithm and its application to combinatorial optimization problem[C]//2000 Congress on Evolutionary Computation,2000:1354 – 1360.

[14] Zhou Z, Wu J, Tang W. Ensembling neural networks: Many could be better than all[J]. Artificial Intelligence,2002, 137(1/2):239 – 263.

[15] 张春霞,张讲社. 选择性集成学习算法综述[J]. 计算机学报,2011,34(8):1399 – 1410.

[16] 沈焱萍,郑康峰,伍淳华,等. 智能启发算法在机器学习中的应用研究综述[J]. 通信学报,2019,40(12):124 – 137.

[17] Ting K, Witten I. Issues in stacked generalization[J]. Journal of Artificial Intelligence Research,1999,10(1):271 – 289.

[18] Sharafaldin I, Lashkari A, Hakak S. Developing realistic distributed denial of service(DDoS) attack dataset and taxonomy[C]//2019 International Carnahan Conference on Security Technology,2019:1 – 8.

(上接第 289 页)

[27] 陈聿,田博今,彭云竹,等. 联合手肘法和期望最大化的高斯混合聚类电力系统客户分群算法[J]. 计算机应用,2020,40(11):3217 – 3223.

[28] Rand W M. Objective criteria for the evaluation of clustering methods[J]. Journal of the American Statistical Association, 1971,66(336):846 – 850.

[29] Brandmaier A M. PDC: An R package for complexity-based clustering of time series[J]. Journal of statal software,2015, 67(5):1 – 23.

[30] Petitjean F, Ketterlin A, Gançarski P. A global averaging method for dynamic time warping with applications to clustering[J]. Pattern Recognition,2011,44(3):678 – 693.

[31] Ferreira L N, Zhao L. Time series clustering via community detection in networks[J]. Information Sciences,2016,326: 227 – 242.

[32] Singhal A, Seborg D E. Clustering multivariate time-series data[J]. Journal of Chemometrics,2010,19(8):427 – 438.

[33] Rodriguez J, Alonso C J, Maestro J A. Support vector machines of interval-based features for time series classification [J]. Knowledge-Based Systems, 2005, 18(4 – 5): 171 – 178.