第53卷第5期
2021年10月

南京航空航天大学学报
Journal of Nanjing University of Aeronautics & Astronautics

Vol. 53 No. 5
Oct. 2021

DOI:10.16356/j.1005-2615.2021.05.010

基于三支决策和数据增广的入侵检测算法

张师鹏，李永忠

（江苏科技大学计算机学院，镇江 212100）

摘要：针对传统的入侵检测方法在未知攻击上表现不佳、且没有考虑信息不足的情况对于决策的影响的问题，本文提出了一种基于三支决策和数据增广的入侵检测算法CGAN-3WD。算法利用条件生成对抗网络来满足三支决策理论对数据信息的需求。首先基于三支决策理论对网络行为做出决策，将网络行为划分至正域、负域以及边界域中；之后基于条件生成对抗神经网络来完成数据增广，生成新的样本数据，从而为分类器提供更多的信息以支撑分类器将边界域转化为正域或者负域。NSL-KDD数据集被用于本文的实验中，实验证明，本文提出的算法CGAN-3WD在对入侵行为的检测上要优于对比的方法，能够有效地检测出入侵行为。

关键词：入侵检测；条件生成对抗网络；三支决策

中图分类号：TP309　　**文献标志码：**A　　**文章编号：**1005-2615(2021)05-0735-08

Intrusion Detection Algorithm Based on Three-Way Decisions and Data Augmentation

ZHANG Shipeng, *LI Yongzhong*

(School of Computer Science, Jiangsu University of Science and Technology. Zhenjiang 212100, China)

Abstract: Since traditional intrusion detection methods perform poorly on unknown attacks and do not consider the impact of insufficient information on decision-making, an intrusion detection algorithm based on three three-way decisions and data augmentation called CGAN-3WD is proposed. The conditional generative confrontation nets are used to meet the data requirements of the three-way decisions. First, the three-way decisions theory is used to make decisions about network behavior, and it can catagorize network behavior into the positive domain, the negative domain or the boundary domain. Second, new samples are generated by the conditional generative adversarial nets, and the new samples can provide more information for the classifier to put the boundary domain into the positive or the negative domain. Third, the NSL-KDD dataset is used in the experiments. Experiments have proved that the CGAN-3WD model has indeed achieved good performance in intrusion detection, and it can effectively detect intrusions.

Key words: intrusion detection; conditional generative adversarial networks; three-way decisions

随着网络技术的不断发展，尤其是5G等技术的成功应用，网络越来越多地改变着人们的生活。在以网络构建的互联网中，每天都在传输用户的大量隐私，因此对互联网的攻击每天都在发生，如何确保网络的安全，已经成为学术界以及企业界都在关注的研究方向。检测入侵是确保网络安全的一个重要的步骤，入侵检测系统是对入侵行为进行检测的重要安全机制之一，它可以监测活动的网络流量并识别出可疑或者异常的网络行为[1]，因此对于入侵检测的研究一直都是一个网络信息安全领域

基金项目：国家自然科学基金(61471182)资助项目；江苏省研究生科研与实践创新计划(KYCX20_3163)资助项目。
收稿日期：2020-10-25；**修订日期：**2020-11-07
通信作者：李永忠，男，教授，E-mail：lunwenyong20@163.com。

引用格式：张师鹏，李永忠. 基于三支决策和数据增广的入侵检测算法[J]. 南京航空航天大学学报，2021，53(5)：735-742. ZHANG Shipeng, LI Yongzhong. Intrusion detection algorithm based on three-way decisions and data augmentation [J]. Journal of Nanjing University of Aeronautics & Astronautics，2021，53(5)：735-742.

的一个重要的研究方向。

随着人工智能等技术的发展,研究人员逐渐发现了机器学习等算法在入侵检测领域的应用前景,基于机器学习算法的模型可以通过经验的累积自动提升性能,符合入侵检测系统针对外部入侵行为通过自我学习进行入侵检测的原则[2-3],因此机器学习算法越来越多地被应用到入侵检测的研究中。例如,Zhou等[4]基于改进的曲线下面积自适应增强算法构建了入侵检测模型,该模型可以更有效地检测网络入侵;Alzubi等[5]提出了一种改进的特征选择算法并应用于入侵检测模型,提升了入侵检测系统的性能;Tao等[6]根据遗传算法和支持向量机的特点,提出了基于遗传算法和支持向量机的入侵检测算法,加快了算法的收敛速度,提升了检出率,降低了误报率;Zhao等[7]针对基于神经网络的入侵检测模型训练时间长,易陷入局部最优等不足提出了一种基于深度信念网络和概率神经网络的入侵检测方法,利用深度信念网络获取原始数据的低维表示,并利用概率神经网络进行分类。

研究人员关于机器学习算法在入侵检测领域的应用取得了一定的成果,也为未来的研究开辟了道路。然而目前已有的基于机器学习(包括深度学习)的入侵检测模型的分类方法都是基于传统的二支决策。所谓二支决策,即对于每一个待分类的样本数据,无论分类器所学习到的信息是否足够,都会立即对该样本做出一个确定的决策,例如对于一个网络行为数据,分类器会将其立即打上入侵或者正常的标签。这种决策方式忽略了信息不足带来的影响,但是如果分类器所获取的信息不足,对于某些样本数据分类器可能并不能立即做出一个合理的决策,盲目进行决策会导致出错。本文提出了一种基于三支决策的入侵检测算法,当信息不足时,根据三支决策理论,可以对要分类的样本数据延迟决策,从而降低信息不足的情况下盲目决策所产生的负面影响。而对于采取延迟决策的样本数据,若要对其做出一个合理的决策,则需要获取新的信息。导致信息不足的最主要的原因在于训练数据的不充足,因此在决策的过程中利用条件生成对抗神经网络(Conditional generative adversarial nets, CGANs)[8]生成新的样本数据用于训练,会为分类器模型提供更多的新的信息,从而对延迟决策的样本做出最终的决策。

1 相关理论

1.1 条件生成对抗网络

CGAN是在生成式对抗网络(Generative adversarial nets, GANs)[9]的基础上发展而来的。GAN在生成样本数据的时候对生成器几乎没有任何的约束,因此生成过程过于自由。而CGAN改善了GAN在生成数据中过于自由的问题,可以使生成过程按照既定的方向前进,即利用CGAN可以生成指定标签的数据,可以把CGAN看作是一种对GAN加了条件约束的变种GAN。CGAN的理论推导和GAN相似,区别在于CGAN中的生成器G以及判别器D都被加上一个隐含的标签y,y是一种约束,通过y可以生成指定标签(类型)的样本。目标函数的公式为

$$\min_{\phi}\max_{\theta}L(D,G)=E_{x\sim p_r(x)}[\log D(x|y)]+E_{z\sim p(z)}[\log(1-D(G(z|y)))] \tag{1}$$

式中:ϕ为生成器网络G的参数;θ为判别器网络D的参数。对于参数的更新一般都是通过梯度下降算法。

式(1)中,右边第1项是判别器网络需要优化的目标函数,第2项为生成器网络需要优化的目标函数。两个目标函数一般为交叉熵损失函数,式(1)是通过将生成网络的目标函数以及判别网络的目标函数进行合并而得到的,具体的推导步骤可参见文献[8,9]。

1.2 三支决策

三支决策[10]是近些年来在学术界得到关注的一种新的理论方法。三支决策是传统的二支决策理论的一个推广,三支决策理论在接受(正)以及拒绝(负)之外增加了一个延迟决策的选项,通常将被采取接受决策的数据所组成的区域称为正域(POS),将被采取拒绝决策的数据组成的区域称为负域(NEG),而被延迟决策的数据则被称为边界域(BND)。

对于一个二分类问题,真实的分类标签可以表示为P(正),N(负),用一个状态集$\Omega=\{X,\neg X\}$来表示。三支决策的决策集可以表示为$D=\{D_{\mathrm{P}},D_{\mathrm{B}},D_{\mathrm{N}}\}$,分别表示接受决策,边界决策以及拒绝决策。所有决策的代价损失函数如表1所示。记$\lambda_{\mathrm{PP}},\lambda_{\mathrm{BP}},\lambda_{\mathrm{NP}}$分别表示当前数据属于$X$的时候,采取行动$D_{\mathrm{P}},D_{\mathrm{B}}$以及$D_{\mathrm{N}}$时的损失,$\lambda_{\mathrm{PN}},\lambda_{\mathrm{BN}},\lambda_{\mathrm{NN}}$分别表示当前数据不属于$X$的时候,采取行动$D_{\mathrm{P}},D_{\mathrm{B}}$以及$D_{\mathrm{N}}$时的损失。

假设$0\leqslant\lambda_{\mathrm{PP}}\leqslant\lambda_{\mathrm{BP}}<\lambda_{\mathrm{NP}},0\leqslant\lambda_{\mathrm{NN}}\leqslant\lambda_{\mathrm{BN}}<\lambda_{\mathrm{PN}}$,根据文献[11]的推演证明,可以得到如下2个相关阈值的计算公式

表 1　三支决策的代价函数

Table 1　Cost function of three-way decisions

决策	P	N
D_P	λ_{PP}	λ_{PN}
D_B	λ_{BP}	λ_{BN}
D_N	λ_{NP}	λ_{NN}

$$\alpha=\frac{(\lambda_{PN}-\lambda_{BN})}{(\lambda_{PN}-\lambda_{BN})+(\lambda_{BP}-\lambda_{PP})} \tag{2}$$

$$\beta=\frac{(\lambda_{BN}-\lambda_{NN})}{(\lambda_{BN}-\lambda_{NN})+(\lambda_{NP}-\lambda_{PP})} \tag{3}$$

式中 $0\leqslant\beta\leqslant\alpha\leqslant1$。可以得到如下 3 条应用到入侵检测领域的规则：

(1) 如果 $P(X|[x])>\alpha$，则该网络行为被归为正类，即该网络行为是入侵行为；

(2) 如果 $P(X|[x])<\beta$，则该网络行为被归为负类，即该网络行为是正常行为；

(3) 如果 $\beta\leqslant P(X|[x])\leqslant\alpha$，则表示当前信息下，无法对该行为采取任何决策，则该行为需要被划分到边界域以等待进一步的处理。

$[x]$表示样本在属性集下的等价类，$P(X|[x])$表示将等价类$[x]$分为X的概率，在入侵检测的领域则表示为一个网络行为属于入侵行为的概率。

2　CGAN-3WD入侵检测算法

本文利用深度生成模型CGAN以及三支决策理论，提出了一种基于数据增广和三支决策的入侵检测算法CGAN-3WD。

算法模型主要由两部分组成：(1)利用CGAN生成样本数据的过程；(2)基于三支决策理论进行分类的过程。

2.1　基于三支决策理论的分类

假设原始的训练集为 T_r，测试集为 T_e，分类器为 f，则 f 的目的是将 T_e 中的每个样本数据都做出尽可能准确的决策，假设 f 对 T_e 做出的最终决策集合为 Y，则 $Y=\mathrm{POS}\cup\mathrm{NEG}$，其中 POS 为正域，NEG 为负域，若 Y 不是最终的决策集合，则 $Y=\mathrm{POS}\cup\mathrm{NEG}\cup\mathrm{BND}$，其中 BND 为边界域，即被采取延迟决策的样本数据的集合，而在最终的决策之前，$\mathrm{BND}=\mathrm{POS}\cup\mathrm{NEG}\cup\mathrm{BND}$，BND 为在 BND 基础上得到的新的边界域。

算法 1 为基于三支决策和数据增广的入侵检测算法的步骤。

算法 1　基于三支决策和数据增广的入侵检测算法

输入：训练集 $T_r=\left\{(x^{(i)},y^{(i)})\right\}_{i=1}^{m}$，测试集 $T_e=\left\{(x^{(j)})\right\}_{j=1}^{n}$，生成对抗神经网络 CGAN，阈值 α,β，初始分类器 f，正域 $\mathrm{POS}=\varnothing$，边界域 $\mathrm{BND}=\varnothing$，负域 $\mathrm{NEG}=\varnothing$，控制阈值下降速率参数 ρ.

输出：POS∪NEG(最终分类结果)

1：While T_e不为空

1.1：根据训练集 T_r 训练分类器 f；

1.2：由模型 f 得到的测试集中的每个数据属于正类的概率 $P=f(T_e)$；

1.3：For $p,t_e\in P,T_e$

1.3.1：If $p>\alpha$：$\mathrm{POS}=\mathrm{POS}\cup t_e$；

1.3.2：Else if $p<\beta$：$\mathrm{NEG}=\mathrm{NEG}\cup t_e$；

1.3.3：Else：$\mathrm{BND}=\mathrm{BND}\cup t_e$；

End if

End for

1.4：$T_e=\mathrm{BND}$；

1.5：对阈值 α,β 做出调整：$\alpha=a-\rho\times0.01$，$\beta=\beta-\rho\times\dfrac{\alpha-0.5}{0.5-\beta}\times0.01$；

1.6：If $\alpha\leqslant0.5$ or $\beta\geqslant0.5$：基于传统的二支决策对剩余边界域中的样本数据进行强制分类并跳出迭代；

1.7：训练CGAN模型 $\mathrm{CGAN}(T_r)$；

1.8：生成新的样本集：$T_r'=\mathrm{CGAN}(\varepsilon)$，$\varepsilon$ 为服从标准正态分布的噪声数据；

1.9：从 T_r 中选取部分样本数据，记为 $\bar{T}_r$，$T_r=\bar{T}_r\cup T_r'$；

End while

2：输出 POS∪NEG。

在算法流程的初始阶段，训练集代表原始的训练集，首先通过原始的训练集对分类器模型 f 进行训练，由于在实验的过程中要计算出每个网络行为数据属于正域的概率，因此对于分类器的选择以软分类模型为主，本文选择多层感知机模型作为基分类器。通过基分类器 f，会得到每个网络行为数据属于正域的概率 p，将 p 与阈值 α,β 进行比较，如算法 1 中的 1.3 步骤的表述，可以将所有的测试集中的样本数据划分到正域、负域以及边界域中。而后，令边界域中的数据组成新的测试集，并对阈值 α,β 调整，减小 α，增大 β，如算法 1 中步骤 1.5 所示。假设第 i 次迭代产生的边界域为 BND1，第 $i+1$ 次迭代产生的边界域为 BND2，则 BND2 中的样本数据相对于 BND1 中的样本数据更难划分；若阈值一直保持不变，则可能会出现部分数据最终难以被划

归到任何一个确定的域中。因此,适当地对阈值做出调整令样本数据被划归到正域以及负域中的条件逐渐变宽,不仅可以保证最终所有的样本数据都被划归到正域以及负域中,而且可以降低迭代所产生的时间代价。在调整阈值的过程中,以0.01作为一个调整单位,并利用参数ρ控制调整的速度,为了能够让α,β同时靠近0.5,在调整的过程中,以α的调整为基准,为β的调整参数乘以$(\alpha-0.5)/(0.5-\beta)$,从而控制α,β可以同时靠近0.5。如果$\alpha\leqslant 0.5$或者$\beta\geqslant 0.5$,则会根据普通的二分类模型的标准,将样本数据强制归为正域或者负域,并结束程序的运行。如果上述情况没有出现,则边界域的存在会触发CGAN模型,当前的训练集被用于训练CGAN模型的数据集,经过训练后,CGAN模型会根据相应标签生成新的样本数据,而后将生成的样本数据与之前选择的样本数据组成新的训练集T_r,用于下一轮对分类器模型的训练。

2.2 基于CGAN的数据增广

在基于三支决策理论进行分类决策的过程中,有一个重要的环节,即利用CGAN完成的数据增广。CGAN在生成数据的过程中可以根据标签生成相应的数据。

算法1中的步骤1.7以及1.8是基于CGAN的数据增广过程,可以扩展为如算法2所示的步骤。

算法2 基于CGAN的数据增广

输入:生成样本数量m,训练数据集T_r,迭代次数e;

输出:生成样本数据X_{gen};

1:随机采样m条噪声数据,并根据需求为每条噪声数据连接相应的标签,记为input;

2:初始化生成器G以及判别器D的相关参数;

3:For i=1 to e:

3.1:得到生成数据:$X_{gen}=G(\text{input})$;

3.2:判别器返回对真实数据的判别结果:$d_{real}=D(T_r)$;

3.3:判别器返回对生成数据的判别结果:$d_{gen}=D(X_{gen})$;

3.4:计算生成器损失l_{gen};

3.4:计算判别器损失l_{dis};

3.6:更新判别器参数;

3.7:更新生成器参数;

End for

4:输出生成样本数据X_{gen}。

CGAN由生成器G以及判别器D组成,在对生成器G进行训练的过程中需要对判别器D的参数进行固定,而对判别器D进行训练的过程中需要对生成器G进行固定。

本文是利用CGAN生成网络行为数据。输入的训练数据集T_r是连接了每条网络行为对应标签的数据集。由于要生成固定数量的网络行为,因此在生成数据的过程中,采用了生成多少数据就采样多少噪声的做法,噪声采样自标准正态分布,采样后为噪声打上相应的标签,由于网络行为数据的标签要么为1,要么为0,因此每条噪声数据的标签要么为1,要么为0。生成的数据被用于处理边界域中的样本,因此生成数据的数量要与边界域中数据的数量有关系,在本文的实验设置中,将生成样本数量m设置为边界域数量的3倍,若边界域中样本数量的3倍小于1 000,则默认m为1 000,而后从原始的训练集中随机选取数量为边界域数量的2倍的数据,这个样本数量的最小值默认设置为500,随机选取的样本与生成样本的样本数据组合成新的训练集用于训练分类器。

在基于CGAN进行样本扩增的过程中,边界域的存在会触发CGAN,对CGAN的使用不止一次,因此在每次使用CGAN时,会将上次使用的CGAN的参数作为本次CGAN的初始化参数,即除第一次外,每一次都只需对CGAN进行微调即可。网络行为本身相似性较大,通过微调CGAN网络即可得到生成的网络行为数据。

3 实验结果与分析

本文所提出的是基于三支决策和数据增强的入侵检测算法,其中基于CGAN实现数据增强,基于三支决策理论进行分类。

3.1 数据集

本文实验所采用的数据集是NSL-KDD数据集[12]。NSL-KDD数据集由41个特征属性和1个类属性组成。KDD数据集包括训练集和测试集两种,总共包含38种攻击,其中训练集包含22种攻击,而测试集中包含训练集中的20种攻击,除此之外还包含在训练集中没有见过的17种攻击类型。因此可以使用测试集测试入侵检测方法在未知攻击上的表现。38种攻击类型可以分为以下4种主要的攻击类型:拒绝服务攻击(Dos)、远程攻击(R2L)、本地用户非法提升权限的攻击(U2R)、网络刺探(Probe)。

3.2 评价指标

本文选择准确率A,误报率F,检出率D,查准

率P_R与F_1分数F_1作为评判指标。评价指标的计算公式如下

$$A=\frac{TP+TN}{TP+FP+TN+FN} \tag{4}$$

$$D=\frac{TP}{TP+FN} \tag{5}$$

$$P_R=\frac{TP}{TP+FP} \tag{6}$$

$$F=\frac{FP}{TN+FP} \tag{7}$$

$$F_1=\frac{2TP}{2TP+FP+FN} \tag{8}$$

式中：TP和TN分别表示攻击记录和正常记录已正确分类；FP代表被误认为是攻击的正常记录；FN代表错误分类为正常记录的攻击记录。

3.3　样本选取

实验中随机选取选择原始训练集中的20%作为训练集。由于攻击类型为U2R的样本数据极少，只有52个，因此在选择数据集的时候，选取所有的攻击类型为U2R的数据。实验数据的分布如表2所示。

表2　数据分布

Table 2　Data distribution

数据集	DOS	Probe	R2L	U2R	Normal
训练集	9 234	2 289	209	52	13 449
测试集	7 458	2 421	2 754	200	9 710

3.4　消融实验

在本文的实验中，为了探究基于CGAN的数据增强以及基于三支决策理论进行分类对实验结果的影响，进行了消融实验。实验中在保证同样使用基于三支决策的分类方法的同时，对比使用了基于CGAN数据增强的入侵检测算法CGAN-3WD与没有使用数据增强的入侵检测算法3WD；在保证同样使用CGAN扩充数据的情况下对比使用了基于三支决策的入侵检测算法与没有使用基于三支决策的入侵检测算法的实验结果，对比算法的分类器分别选用了多层感知机(CGAN-MLP)。

本文所提出的方法CGAN-3WD以及本节所提及的几种对比方法的实验结果如表3所示。从表3中可以发现，3种方式在对入侵行为的检出上达到的效果都非常好，而且检测效果非常接近。

表3　消融实验结果

Table 3　Results for ablation experiment

方法	A/%	D/%	F/%	P_R/%	F_1/%
CGAN-3WD	<u>96.43</u>	<u>96.13</u>	<u>3.17</u>	<u>97.56</u>	<u>96.84</u>
3WD	93.97	94.63	6.88	94.79	94.70
CGAN-MLP	94.88	93.63	3.47	97.27	95.42

3种模型的检出率都达到了93%以上，F_1值都超过了94%，但基于CGAN-3WD方法构建的入侵检测模型在这5个指标上的表现都要更好。

相对于没有使用数据增强的3WD，CGAN-3WD的表现更好，CGAN-3WD模型拥有更低的误报率的同时也拥有更高的检出率。对于没有使用数据增强方法的分类器来说，原始数据所包含的信息非常不充分，尤其是在本文的实验中，训练集的样本和测试集样本的数量基本一致，这也说明了分类器模型不会通过训练集获得充分的信息。对于没有使用数据增强的基于三支决策的分类方法来说，对于边界域的处置，只能通过特征提取模型获取到的多粒度的特征集。虽然这种方式在一定程度上可以解决边界域，但是毕竟只是在原始的训练集上进行特征提取。从数据层面来看，并没有增加训练集的信息，只是对数据集从新的角度做了新的分析。

相比较同样使用了基于CGAN的数据增强但是没有使用基于三支决策理论进行分类的方法，CGAN-3WD在几个评价指标上的表现也要更好。两种入侵检测模型都使用了多层感知机模型作为分类方法，但是利用了三支决策理论进行分类的CGAN-3WD模型取得效果还是要优于基于二支决策的CGAN-MLP。两种入侵检测模型都使用了CGAN进行数据扩增，之所以会在对入侵行为的检测上产生差异，主要是因为CGAN-MLP模型一开始便是利用CGAN生成的数据以及原始训练集对测试集中的样本数据进行分类。因此，对于测试集中的部分原本可以被轻易分类的样本可能会因为生成数据的加入而产生不可控的类似过拟合的现象，即CGAN-MLP模型中的分类器模型因为太早学习到了生成的样本数据中的信息，反而导致了分类器模型对部分样本数据进行了错误的分类。对于一个分类器来说，可能一次只能对要分类器的数据集中的一部分做出合理的决策。基于三支决策理论的分类方法，每一次分类过程中都将原始的待分类数据分为正域、负域以及边界域。对于正域以及负域中的数据，下一次的分类已经和它们没关系，若是因为迭代次数过多产生过拟合的问题，通过这种方法也可以缓解过拟合所产生的影响。

图1为根据3种入侵检测模型的检测结果得到的ROC曲线图，ROC曲线以真阳性率为纵坐标，假阳性率为横坐标绘制的曲线，ROC曲线可以被用于评判分类以及检测的结果的好坏。AUC表示

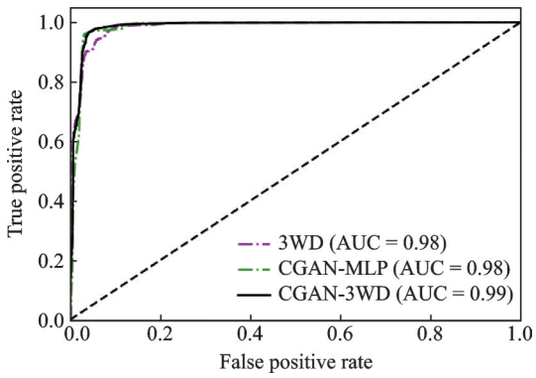

图1 ROC曲线图

Fig.1 ROC curve

ROC曲线下的面积，主要用于衡量模型的泛化性能。从图中可以发现，几种模型得到的ROC曲线围成的面积非常接近，说明3种模型在对入侵行为的检出上是非常接近的，但是CGAN-3WD模型得到的AUC面积要稍微大于另外两种模型，说明基于CGAN完成的数据增广以及基于三支决策理论的分类都对入侵检测产生了积极的影响。

3.5 检测效果对比实验

本节实验所探究的是本文提出的模型CGAN-3WD与其他模型的性能对比。

3.5.1 对比方法

本文对比方法为：文献[13]提出的一种PSO-XGBOOST模型，该模型利用粒子群算法自适应搜索XGBOOST的最优结构；文献[14]提出的一种基于自适应主成分(A-PCA)和增量极限学习机(I-ELM)的方法APCA-IELM，该方法自适应地选择网络流量的相关特征，通过I-ELM算法获得最佳的检测精度；文献[15]提出的一种基于深度学习和半监督学习的入侵检测方法DL-SSL；文献[16]提出的一种基于深度生成模型的半监督入侵检测方法SS-DGM。

3.5.2 在异常行为上的表现

在实验过程中，把所有入侵行为的标记设为1，正常样本的标记设为0。不同模型得到的实验结果如表4所示。

表4 性能对比实验结果

Table 4 Results for performance comparison

方法	A/%	D/%	F/%	P_R/%	F_1/%
CGAN-3WD	96.43	96.13	3.17	97.56	96.84
APCA-IELM	91.31	88.10	4.43	96.34	92.04
DL-SSL	90.46	87.35	5.09	95.84	91.22
PSO-XGBOOST	88.67	89.72	12.70	90.37	90.02
SS-DGM	92.21	86.23	4.85	89.81	88.04

从表4中可以发现，几种入侵检测模型取得的实验效果都不错，准确率都达到了90%以上，检出率也都超过了85%，而最能够反映出一个模型综合性能的F_1值也都超过了90%。实验数据表明，目前已有的入侵检测方法在对入侵行为的检出方面效果都不错。但是在误报率上，有些模型却表现得不尽如意，例如POS-XGBOOST模型的误报率达到了12%，已经是一个非常高的数值。而本文提出的CGAN-3WD模型在几种评价指标上的表现都是最好的，既做到了高检出率又做到了低误报率。

图2为根据几种模型的实验结果绘制的ROC曲线图。从ROC曲线图中也可以发现，根据CGAN-3WD模型的实验结果绘制的ROC曲线的AUC面积达到了0.99，是几种模型中最高的，反映出CGAN-3WD模型的泛化性能要好于其他几种对比模型。

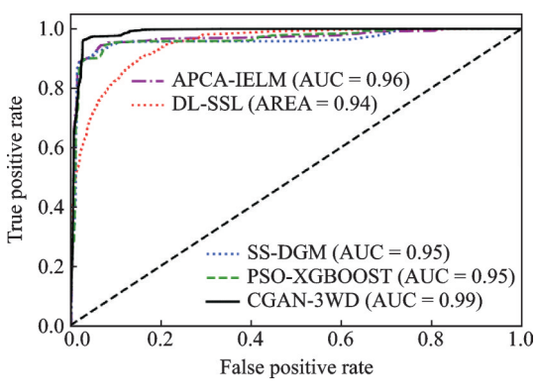

图2 ROC曲线图

Fig.2 ROC curve

3.5.3 在不同行为类型上的表现

本节实验不仅要区分出正常的网络行为和入侵行为，还要区分出入侵行为的具体类型。具体做法：每次选择一种类型的入侵行为，将其标记为1，其余的网络行为标记为0。本节实验中利用SMOTE方法先对样本数据比较少的攻击行为进行过采样，而后用过采样后的数据训练CGAN模型，实验结果如表5所示。

从表5中可以发现，几种入侵检测模型在对有些攻击的检测上表现不错，如DOS，Probe，而对有些攻击行为的检测则相对很差，如U2R，R2L。这主要是因为训练集中攻击类型为R2L以及U2R类型的样本数据非常少，虽然在训练的过程中利用SMOTE方法进行了过采样，但是SMOTE过采样方法本身就存在缺陷[17]。

CGAN-3WD模型和几种对比方法比较，取得

表 5　入侵实验结果对比

Table 5　Results of intrusion behavior comparison

入侵类型	评价指标	CGAN-3WD	DL-SSL	PSO-XG-BOOST	SS-DGM	APCA-IELM
DOS	A	93.15	87.65	90.30	91.34	91.44
	P_R	91.92	81.53	86.12	87.51	87.87
	D	86.78	81.03	84.27	86.13	86.00
	F	3.71	9.08	6.72	6.08	5.87
	F_1	89.35	81.28	85.19	86.81	86.93
Probe	A	95.72	93.73	92.08	95.58	93.92
	P_R	86.33	70.00	64.63	85.75	76.45
	D	71.45	72.78	58.03	70.59	60.41
	F	1.36	3.75	3.82	1.41	2.31
	F_1	78.22	71.36	61.15	77.44	66.94
R2L	A	98.91	89.90	97.61	98.50	91.02
	P_R	77.57	65.93	30.51	49.72	79.10
	D	40.31	35.12	45.21	41.45	36.12
	F	0.10	2.51	1.61	0.63	1.31
	F_1	53.23	45.91	30.54	45.23	49.51
U2R	A	90.89	89.93	90.49	84.91	87.51
	P_R	98.61	97.83	98.41	31.12	47.84
	D	25.78	18.01	22.51	19.48	21.41
	F	0.05	0.06	0.06	6.00	3.21
	F_1	40.88	30.42	36.64	23.91	29.61

了比较好的效果。首先在对DOS以及R2L攻击行为的检测上，几种模型的检测结果都不错，没有出现特别差的情况。但是有些模型对Probe攻击的检测效果并不是特别好，如DL-SSL，PSO-XG-BOOST；在对R2L以及U2R的检测上几种方法效果都不理想，其中也包括CGAN-3WD模型。训练数据集的严重不平衡导致了几种模型表现较弱，但是相比较而言，CGAN-3WD模型对两种攻击的检测效果较好，除了在对R2L的检出率上没有达到最好的结果之外，其他的评价指标在几种对比方法中都是最好的。

4　结　　论

本文提出了一种基于数据增广和三支决策的方法CGAN-3WD，通过条件生成对抗网络生成新的数据用以满足三支决策对于信息的需求，经过实验证明，本文提出的方法与对比方法比较，取得了比较好的结果。

在未来的工作中，可以考虑改进生成模型，使得生成的样本更加地具有多样性。

参考文献：

[1] ZAVRAK S, İSKEFIYELI M. Anomaly-based intrusion detection from network flow features using variational autoencoder [J]. IEEE Access, 2020, 8: 108346-108358.

[2] FANG Weijian, TIAN Xiaoling, WILBUR D. Application of intrusion detection technology in network safety based on machine learning[J]. Safety Science, 2020, 124: 104604.

[3] MAGÁN-CARRIÓN R, URDA D, DÍAZ-CANO I, et al. Towards a reliable comparison and evaluation of network intrusion detection systems based on machine learning approaches[J]. Applied Sciences, 2020, 10(5): 1775.

[4] ZHOU Ying, MAZZUCHI T A, SARKANI S. M-AdaBoost-A based ensemble system for network intrusion detection[J]. Expert Systems with Applications, 2020, 162: 113864.

[5] ALZUBI Q M, ANBAR M, ALQATTAN Z N M, et al. Intrusion detection system based on a modified binary grey wolf optimisation[J]. Neural Computing and Applications, 2020, 32: 6125-6137.

[6] TAO Peiying, SUN Zhe, SUN Zhixin. An improved intrusion detection algorithm based on GA and SVM [J]. IEEE Access, 2018, 6: 13624-13631.

[7] ZHAO Guangzhen, ZHANG Cuixiao, ZHENG Llijuan. Intrusion detection using deep belief network and probabilistic neural network[C]//Proceedings of 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC). [S.l.]: IEEE, 2017: 639-642.

[8] MIRZA M, OSINDERO S. Conditional generative adversarial nets[EB/OL]. (2014-11-6)[2020-10-24]. https://arxiv.org/abs/1411.1784.

[9] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge, USA: MIT Press, 2014: 2672-2680.

[10] YAO Yiyu. Three-way decisions with probabilistic rough sets[J]. Information Sciences, 2010, 180(3): 341-353.

[11] 刘盾，梁德翠. 广义三支决策与狭义三支决策[J]. 计算机科学与探索，2017，11(3)：502-510.
LIU Dun, LIANG Decui. Generalized three-way decisions and special three-way decisions[J]. Journal of Frontiers of Computer Science and Technology,

2017, 11(3): 502-510.

[12] EVER Y K, SEKEROGLU B, DIMILILER K. Classification analysis of intrusion detection on NSL-KDD using machine learning algorithms [C]//Proceedings of International Conference on Mobile Web and Intelligent Information Systems. Cham, Germany: Springer, 2019: 111-122.

[13] JIANG Hui, HE Zheng, YE Gang, et al. Network intrusion detection based on PSO-Xgboost model [J]. IEEE Access, 2020, 8: 58392-58401.

[14] GAO Jianlei, CHAI Senchun, ZHANG Baihai, et al. Research on network intrusion detection based on incremental extreme learning machine and adaptive principal component analysis[J]. Energies, 2019, 12(7): 1223.

[15] LI Yongzhong, ZHANG Shipeng, LI Yi, et al. Research on intrusion detection algorithm based on deep learning and semi-supervised clustering [J]. International Journal of Cyber Research and Education, 2020, 2(2): 38-60.

[16] 曹卫东, 许志香, 王静. 基于深度生成模型的半监督入侵检测算法[J]. 计算机科学, 2019, 46(3): 197-201.
CAO Weidong, XU Zhixiang, WANG Jing. Intrusion detection based on semi-supervised learning with deep generative models [J]. Computer Science, 2019, 46(3): 197-201.

[17] FERNÁNDEZ A, GARCIA S, HERRERA F, et al. SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary [J]. Journal of Artificial Intelligence Research, 2018, 61: 863-905.

(编辑:张蓓)