端到端的基于深度学习的网络入侵检测方法＊

王金华1

(1. 四川大学网络空间安全学院，四川省，成都市，610207)

**摘要**：当前网络入侵检测多使用人工特征，而人工特征对于出现的新型攻击需要重新设计，并且手工特征需要在总览全局流量情况下才能计算得出，针对该问题提出一种深度学习算法，该算法从网络流量数据中提取会话作为样本，将样本视为不同的数据形式同时送入两个神经网络，1：会话样本中一系列有顺序的数据包视为一维序列送入GRU中提取特征；2：整个会话的流量视为二维图像送入CNN中提取特征。最后合并两部分特征作为总特征送入神经网络中检测会话是否是异常流量。本文算法同时关注了会话中数据包的时序信息，以及会话的整体空间信息，无需人工设计特征，不需要全局流量参与计算。实验证明，使用CICIDS-2018网络流量，文章方法在入侵检测上，识别僵尸网络异常流量的准确率达99.99%，f1为99.99%，DDoS异常流量的识别准确率达99.99%，f1为99.99%，效果好于人工设计的特征。

**关键词**：入侵检测；深度学习；流量分割；流量可视化;DDoS检测

**中图分类号：作者补充**　　**文献标识码：作者补充** 　　**文章编号：**1009-8054(2019)12-0000-00（5号宋体）

**End-to-end NETWORK intrusion detection method based on deep learning**

#### WANG Jinhua1

#### School of Cyberspace Security, Sichuan University, Chengdu Sichuan 610207, China;）

**Abstract:** Currently, network intrusion detection mostly uses artificial features, and artificial features need to be redesigned for new attacks that appear, and manual features need to be calculated under the condition of an overview of global traffic. To solve this problem, this paper proposes a deep learning algorithm, which starts from Extract the session as a sample from the network traffic data, treat the sample as different data forms and send it to two neural networks at the same time. 1: A series of sequential data packets in the session sample are sent to the GRU as a one-dimensional sequence to extract features; 2: The traffic of the entire conversation is regarded as a two-dimensional image and sent to the CNN to extract features. Finally, the two features are combined as the total feature and sent to the neural network to detect whether the conversation is abnormal traffic. The algorithm in this paper also pays attention to the timing information of the data packets in the session, and the overall spatial information of the session, without the need for manual design features, and no need for global traffic to participate in the calculation. Experiments have proved that using CICIDS-2018 network traffic, the article method in intrusion detection, the accuracy of identifying abnormal botnet traffic is 99.99%, f1 is 99.99%, the recognition accuracy of DDoS abnormal traffic is 99.99%, and f1 is 99.99%. The effect is better than the artificially designed feature.

**Keywords:** Intrusion detection; deep learning; traffic segmentation; traffic visualization; DDoS detection

## **0　引　言**

网络入侵检测的目标是识别那些通过网络企图跨越被保护系统安全边界的行为，因此任何会造成被保护系统安全性受损的入侵行为都应被识别。异常入侵检测对于网络安全有着重要意义的问题，通过对正常用户行为的建模它可以识别出异常的流量，从而检测新型的攻击。然而，研究者首先面临的是网络上收集的网络流量数据，对于这些数据，一种处理方法是根据专家知识和统计学提取人工特征，形成特征集，对这些人工特征数据做检测，将识别某条记录是异常流量还是正常流量。

KDD-99数据集[1]是为了1999 KDD杯挑战赛而提出的入侵检测数据集，该数据集对网络流量进行预处理，提取了几大类人工特征：基本特征，内容特征，基于主机的流量特征，基于时间的流量特征，共计41个特征，共4种大类攻击类型，总计包含400万条记录，保存为txt文件。在使用KDD-99数据集验证检测方法存在着诸多的不足，如攻击类型定义模糊，冗余的记录过多等问题。针对KDD-99数据集的缺点，Mahbod Tavallaee等在该数据集的基础上，删除了冗余项，规划了训练集和测试集的比例，提出了NSL-KDD数据集[2]，此数据集对于41维特征没有做出改变，主要优点是类别分配平衡，是用于入侵检测方法比较的基准数据集。

上述两种数据集并没有提供网络流量数据文件，而在2015年公开的UNSW-NB15[3]和2018年公开的CICIDS-2018[4]数据集则提供了网络流量数据pcap文件和已经预处理提取好特征的csv文件。Nour Moustafa等对UNSW-NB15数据集做了评估分析[5]，认为其包含新型的攻击方式，更具复杂性，可以取代KDD-99数据集成为新的基准数据集。上述3种数据集（KDD-99和NSL-KDD算一种）人工提取的特征中有着高度的相似性，均包含数据流的持续时间，目的端口，协议，数据包发送速率，统计TCP数据流的设置PSH和URG次数，数据包大小，以及一些其他的流统计特征：流数据包数目，到达时间偏差，其他基于时间的（如最近100个连接中具有相同源地址的个数）统计特征。不同的是不同数据集有一些独特设计的人工特征。值得注意的是，这些特征往往需要整体的流量情况做支撑，才能计算出统计学特征。

对于深度学习方法，如CNN和RNN等，这些方法其强大的地方在于可从复杂的原始数据如图像，文本等数据中自动提取特征，从而用这些人类无法确切解释，但是十分有效的特征进行一些AI任务如对象检测或者语音识别，机器翻译[6]等。将人工特征送入CNN或RNN中进行分类，只是将其做为了一个强大的分类器，没有利用到深度学习方法其强大的表示学习能力。尽管目前研究者所提出的方法在入侵检测分类上得到了良好的效果，但是对于应对各种新型的攻击类型，人工提取的特征是否还能胜任检测工作，以及对于检测方法的实时性要求，也需要我们的检测方法所依赖的特征可以快速的获取，而不依赖在整个流量都得到之后再进行计算，因此研究者开始考虑直接使用网络流量的方法。

## **相关工作**

随着深度学习的兴起，使用深度学习方法替代传统机器学习算法做入侵检测的趋势逐渐流行，但也正如在引言中所述，这些方法很多是基于已有的人工特征的分类算法，很少直接针对网络流量数据进行设计,或者忽视了网络流量数据某一方面的特征。本章会简要介绍其他研究者在入侵检测领域所作的工作。

Hu Wenjie等人[7]研究了鲁棒支持向量机(Robust-SVM),传统SVM,K最近邻（KNN）在KDD-99数据集上的性能表现，实验表明使用鲁棒SVM在准确率和误报率上表现更好。Dong BO 和Wang Xue等人[8]则对支持向量机(SVM),决策树C4.5,贝叶斯(Naive Bayes)和使用了深度玻尔兹曼机(RMB)的SVM-RBM进行了比较，实验表明结合了RBM技术的SVM在kdd-99数据集中具有最好的性能。Ravi V 等人[9]则综合评估了卷积神经网络CNN，循环神经网络RNN(LSTM,GRU)，以及结合使用CNN和RNN的深度神经网络，他们将这些在 KDD-99 上表现良好的 DNN 模型应用于其他数据集，例如 NSL-KDD、UNSW-NB15、Kyoto、WSN-DS 和 CICIDS-2017，对其进行了全面而完整的分析，实验结果表明使用深度学习作为分类器，学习人工特征所含信息做出决策的能力要比传统机器学习更好。此外，Xu Congyuan等人 [10]详细对比了采用长短期记忆网络（LSTM）和门控循环单元[11]（GRU）的循环神经网络在入侵检测数据集NSL-KDD和KDD 99上的性能表现，实验表明采用双向GRU能够达到更好的性能，当然这只是在特定领域上，Junyoung Chung等人[12]研究表明了在复调音乐建模和语言信号建模两种任务上，这两种网络性能相当。

深度学习方法在人工特征集上的任务上表现出其优越的性能，却不能忽略其更强大的能力，从原始低级数据习得高级抽象（自动特征）即表示学习的能力。研究者Wang Zhanyi等人[13]将网络流量数据表示成1000长的一维序列，采用自堆叠的动编码器（SAE）做网络协议识别,取得了良好的效果。Wang Wei等人[14]第一次尝试将网络流量转化为图像，使用卷积神经网络进行恶意软件流量分类，在他们自建的数据集USTC-TFC2016上，取得了平均准确率99.41%的好成绩。他们还对如何表示网络流量取得最佳性能做了对比实验，实验结果表明采取会话即双向流和所有数据包层的流量表示类型会取得最好的结果。他们还使用一维卷积神经网络来鉴别加密流量和未加密流量[15]。也有研究者采用网络流量数据做入侵检测的流量识别，Peng Yujie等人[16]作网络流量作为图像，送入卷积神经网络中检测该流量是否是安卓恶意软件流量。此外，基于循环神经网络，Benjamin J. Radford等人[17]流量数据视为机器之间对话的语言，使用LSTM进行序列建模，并提出了一个简单的基于频率的模型，采用ROC曲线下面积Area Under Curve(AUC)为评估指标,实验在IDS-2017上进行，取得了不错的效果。

本文从网络流量中提取会话样本，将会话中数据包含有的字节作为一个一维长序列送入门控循环单元GRU中，类似与自然语言处理中一句话中一个个单词，循环神经网络提取数据包的时序相关特征；而整个流中各个数据包被合并为一个固定尺寸的二维图像，被送入一个卷积神经网络提取空间特征，最后结合这两种特征做入侵检测，这样做相比先经过卷积然后循环神经网络的线性方法，这种方法是双线并行的，先提取空间特征可能由于卷积层和池化层的存在送入RNN后损失一部分信息，本文方法则避免了这种可能性，而且关注了整个流的空间特征。算法直接从原始字节流量学习深层抽象特征，完成入侵检测任务，是一个端对端的入侵检测算法，一端是数据，一端是任务，无需人工特征提取，并且具有实时检测的优点。

## **算法**

本文提出的算法重点在于如何处理网络流量数据生成可用于深度学习的流量样本，以及构建神经网络模型从这些流量样本中提取出良好的特征用于检测流量是正常流量还是异常流量，整体流程如图1所示，最后得到一个可用于入侵检测的模型。图1还显示了本文算法的一个重要优点是无需人工设计特征，无需专家知识。

图1算法流程

**2.1　流量分割**

在计算机网络中，定义（源IP，目的IP，源端口，目的端口，协议）五元组相同的一系列数据包同属于一个流，源IP和目的IP相反，源端口和目的端口相反，协议相同的另一个流是它的反向流，前向流和反向流就构成了一个双向流即会话，以下统称会话。在这样的定义下，我们的网络数据基本上分为两大类，TCP会话和UDP会话，当然，根据应用层协议的不同在这两类上会衍生出很多小类，如网络流中最多的HTTP会话。对于TCP会话，从第一个Syn报文标志着会话开始，而Fin包之后ACK包表示会话的结束。对于UDP会话，定义在最初始的数据包之后600秒（已经能够包含绝大多数数据包）内是同一个会话，超时过后即使五元组相同也划分为一个新会话，同样对于TCP会话也有超时的设定。流量分割的流程图如图2所示。

值得注意的是，在收到TCP-Fin数据包之后，根据TCP四次握手的规则，还需要再收到一个ACK包才算是结束了整个TCP会话，五元数组只会记录该会话第一个数据包的接收时间，用于与该五元组相同的数据包加入该会话时判断是否超时。



图2流量分割流程图



图3会话样本特征提取示意图

**2.2 模型构建和特征提取**

在流量分割后，获得了以张量形式存储的会话样本x, x是一个二维张量n\*m，其中n表示该会话中数据包的个数，而m是最大字节数，对于字节数超出m的数据包只能截取m个字节，少于m则空缺处填0。如图3所示，对于每个会话样本张量n\*m,按顺序取出n个m长的一维序列送入GRU中，GRU将在这n个序列中共享权重，记忆相关信息，不需要理会这个n长序列的具体位置情况，最后我们在经过n个时间步后得到一个1\*f的向量；将会话样本n\*m组成一个固定大小的w\*w的二维矩阵，类似于只有一个颜色通道的灰度图像，如果n\*m大于w\*w则丢弃样本**x**的多余数据，小于则在剩余位置填0，将其送入卷积神经网络CNN中，最后也会得到一个1\*f的向量。最后将这两个特征向量在第一维拼接起来，得到一个1\*2f的特征向量，至此，完成了特征提取。

在图3所示上半CNN卷积网络中，依顺序分别是3\*3的卷积层，2\*2的池化层，3\*3的卷积层，2\*2的池化层，2\*2的卷积层，最后是一个全连接层。共计3个卷积层，2个池化层，一个全连接层，最后输出1\*128的向量。下半部分为一个循环神经网络的展开图，采用的是长短期记忆网络LSTM的一个较为轻量的网络变种:门控循环单元GRU。一个流的所有数据包按顺序依次作为一个m长向量送入GRU,采用最后输出的output作为特征向量，是1\*128的向量。最后，将这两个向量连接，就得到一个1\*256的总特征向量，就是我们所期望的值。

**2.3**分类

不同于人工设计特征，神经网络模型所得到的特征向量，需要模型先训练，然后在测试集上达到不错的准确度,才认为网络提取的特征有较大可能具备高级抽象表示能力的。而且神经网络模型从特征提取到分类是一个整体，它从输入网络流量数据到网络判别输入数据的类别是端到端的.将前面特征提取网络输出的1\*256向量作为输入，经过两层的全连接层，最后输出1\*Categories(类别数)张量，使用Softmax函数概率化每个类别的置信度，取最高置信度所在下标的类别即为神经网络输出的类别。两个全连接层的网络结构为：1个（256，64）的全连接层，一个（64，2）的全连接层，最后使用Softmax输出置信度。

## **3　实验证明**

实验目的：本文提出的基于深度学习建立的端到端的入侵检测算法和基于人工规则特征的入侵检测算法在入侵检测性能上的对比，证明算法有效性。本文实验环境为：window10-21H1, python-3.8.10, pytorch-1.9.0, scikit-learn-0.24.2, cpu：i5-10400，gpu：1060。

**3.1**处理数据

为了获得算法对当前网络流行攻击的检测效果，本文选择了CICIDS-2018数据集[5],该数据集是加拿大通信安全机构和网络安全研究所的合作项目成果，涵盖了如：暴力攻击，Heartbleed 漏洞攻击，DOS攻击，DDOS攻击，Web攻击（跨站脚本攻击XSS,SQL注入），僵尸网络，渗透攻击共7种攻击场景。该数据集虽然不是真实场景的攻击流量数据集，但是该数据集通过机器学习技术统计分析正常用户的网络行为，从而模拟正常用户流量，使用上述攻击场景模拟异常流量，使用配置文件的方式生成了一个较为全面，一定程度真实可靠的异常检测流量数据集。

CICIDS-2018数据集针对捕获的原始流量，使用CICFlowMeter-V3的工具，从统计分析的角度，按照人类专家规定的特征提取了80维特征，本文算法需要与该人工特征做对比。本文选取了Wednesday-21-02-2018\_Traffic，Friday-02-03-2018\_Trafficr两天的网络流量数据，其包含正常流量，bot，DDoS-LOIC-UDP, DDoS -HOIC三种攻击流量，对该网络流量进行流量分割，其异常流量分布如表1。

表1 CICIDS-2018人工提取特征和本文分割流量攻击类别分布

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Friday -csv** | **Wednesday -csv** | **Friday -网络流量** | **Wednesday-网络流量** |
| **Benign** | 762384 | 360833 | 32790 | 50090 |
| **Bot** | 286191 | 不包含 | 67510 | 不包含 |
| **LOIC-UDP** | 不包含 | 1730 | 不包含 | 200 |
| **HOIC** | 不包含 | 686012 | 不包含 | 49710 |
| **总计** | 1048975 | 1048575 | 100000 | 100000 |

CICIDS-2018提供了基于主机的收集的众多流量pcap文件，对于这些文件，我们需要筛选数据，选取攻击流量攻击目标主机的pcap文件分割得到攻击会话样本，选取其他正常通信主机得到正常会话样本，提取数据后需要对样本进行标记，详情需要参考CICIDS-2018的具体实施攻击场景，按照攻击时间，端口，协议，ip等信息匹配样本，进行标签分配，该工作较为繁琐，需要格外仔细。

CICIDS-2018提供的已经提取好特征的数据集为csv文件，它的每一条记录就是一个会话（双向流），所有的特征是基于双向流的统计特征得出的，如流持续时间，正向包数，反向包数，正向字节数，反向字节数等，两天共计2097150条数据。本文基于2.1所提出的流量分割算法切割网络流量，对于Friday-02-03-2018含有僵尸网络的流量进行分割，提取出双向流会话，共计100001个会话，会话分布如图表1第4列所示，对Wednesday-21-02-2018分割了共计100001个会话，会话分布如表1列5所示。

提取的会话数据以张量的形式存储，一个会话最后的张量形式为n\*256，n是数据包的个数，根据每个会话的情况动态变化，256个字是每个数据包最大接受的字节数，在对网络流量数据进行处理的时候，发现数据包的平均字节数213（Bot），74(DDoS -LOIC-UDP)，179(DDoS -HOIC),最后选取了256字节的数据长度存储数据包，超出256就截取前256个字节，少于256，空余位置填0。

下图4是各种异常会话的可视化图像，其中每一张图片都是一个会话的整体图像，尺寸为28\*28。

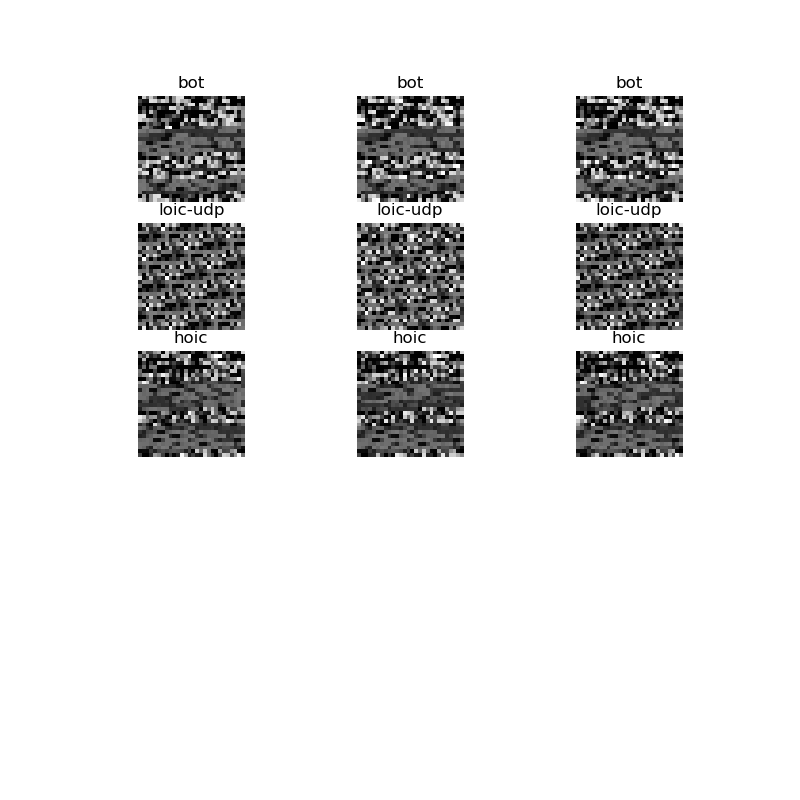


图4　异常会话可视化图

对于数据包数n，实际网络中n可能极大，而在本文分割出的DDoS -LOIC-UDP异常会话中，按照既定规则该会话的数据包数高达十几万个，同时这也是DDoS -LOIC-UDP（分布式拒绝服务攻击-低轨道离子加农炮-UDP洪泛攻击）的特点，该攻击使用多台计算机针对目标服务器开放的端口，短时间内发送了大量伪造源IP地址的UDP数据包，消耗服务器的资源，网络充满无用数据，使其瘫痪不能正常提供服务。通常来说UDP会话在正向流和反向流中数据包数相等，而在该攻击中只会在发送方一侧出现大量数据。当DDoS攻击者已经发送了大量数据包的时候，攻击的目的已经基本完成时，服务器根据网络状况和自己负载状态等条件，很容易就可以判断自己受到了DDoS攻击，然而此时检测到攻击已经过于迟钝了，对于DDoS攻击最好的状态是能够提前发现攻击。本文采用最大数据包数为32，对于大于32的会话会只截取一部分，对于有大量数据包的会话，可以根据局部特征就识别该类攻击，在攻击初见迹象的时候就可以识别它，并对其做出警示。如图4所示，LOIC-UDP攻击可视化完整图像基本就是下半部分的重复。

在实际训练中，数据包并没有冗余造成对计算资源浪费的现象，因为对于实际为空的数据包，其实并不参与实际计算，该种形式是为了样本统一格式，方便处理。

**3.2**训练

3.1中得到的会话总计100,001个，将其打乱，按照7：3的比例，分出70000个会话作为训练集，30001个数据作为测试集。损失函数是交叉熵损失函数CrossEntropyLoss，用于衡量模型分布和经验分布的之间的距离，距离越小，模型就越接近经验分布。优化算法是基于梯度下降的后向梯度传播算法Adam,训练集会被循环使用10个epoch，可以看到训练集在大概第8个epoch的时候就开始收敛。

学习率参数设置，在使用Adam优化算法时，本文算法设置初始的学习率为0.001，在训练中，开始时学习率可以设置的适当大些，以让模型更快的接近最优点，之后应该将学习率减少，以避免跳过最优点，因此本文使用余弦曲率衰减计划，它会按照余弦曲线：初始缓慢减少学习率，之后快速减小学习率，余弦曲线的周期可以指定，本文指定其为10，与训练集的循环次数一致，最低学习率选择默认的0。

**3.3**结果与分析

对于模型性能的指标本文采用混淆矩阵的方式进行计算，下表2为二分类混淆矩阵的示意，可以基于将某一类视为正类，其他类为负类的方式将混淆矩阵扩展到多分类。

表2 混淆矩阵

|  |  |  |
| --- | --- | --- |
|  | 预测为真 | 预测为假 |
| 实际为真 | TP | FN |
| 实际为假 | FP | TN |

准确率(Accuarcy)：模型预测正确的数目占所有样本的比例，如公式（1）所示。在各个类别平衡时，该指标能够表明模型性能，但是在极端类别情况，如正类10个，负类90个，模型只需要输出全负类，就可以达到90%的准确率，然而其对正类的预测能力为0。

 （1）

精确率（precision）：模型预测为类别C的样本中其实际类别为C的样本所占比例，如公式（2）。

 （2）

召回率（recall）:实际类别为C的样本中模型预测为C的样本所占比例，如公式（3）。在上面阐述的类别不平衡分类中，就可以计算出正类的召回率为0/10=0,由此可以判断模型的预测能力完全无效。

 （3）

F1值（f1-score）：精度 (p) 和召回率 (r) 的调和平均值，如公式（4），它综合考虑了precision, recall的值，F1值越高，模型性能越好。

 （4）

本文针对3.1中所提取的网络会话做实验，对Friday提取的100000个会话的含有僵尸网络Bot的数据集进行实验，其结果如图5所示，其accuracy:99.9967%，precision：100%，recall：99.9950%,F1:99.9975%。

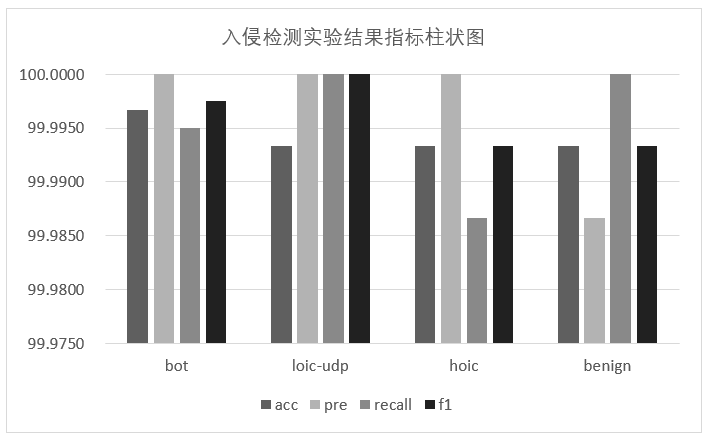


图5　 异常流量实验结果柱状图

对Wednesday提取的含有：正常会话（Benign）, DDoS -LOIC-UDP，DDOS-HOIC两种攻击会话的总计100000个会话进行实验，其结果如图5所示。实验对总体样本识别的准确率为99.99%，对正常流量的识别，对LOIC-UDP和对DDoS -HOIC的识别在各项指标上成绩也为99.99%，错误识别的样本数极少，在对LOIC-UDP的检测攻击中，即使由于最大数据包数的限制，导致算法无法取得该攻击高达十几万数据包整体作为数据输入神经网络中，但是对该攻击的局部特征识别非常成功。

同时，本文也对CIC-IDS-2018提供的基于专家知识和统计学的人工特征数据集（只包含Friday-02-03-2018.csv, Wednesday-21-02-2018.csv这两个文件）在几种机器学习:决策树，随机森林，AdaBoost集成学习，贝叶斯(Naive Bayes)，支持向量机SVM,KNN最近邻的指标表现做了实验，其结果如表3所示，其中除准确率这一列外，每一列下3个百分比分别是对应Bot，DDoS -LOIC-UDP，DDOS-HOIC的指标表现。

从表3结果分析，本文算法对Bot僵尸网络的acc为99.99%，f1为99.99%，对DDOS的两种攻击LOIC-UDP和HOIC的各项指标均高达99.99%%,而人工特征数据集中对Bot的检测效果最好的是决策树算法，其acc为99.98%，f1为99.92%，对LOIC-UDP和HOIC的acc最高为使用决策树算法的100%和99.99%，可以看到本文算法提取的特征检测攻击效果持平人工特征，在Bot检测中略有超出。需要指出的是，这些指标较高的原因有可能是攻击类别较少，该类攻击特征判定条件较为简单的缘故。

表3 CIC-IDS-2018人工特征数据集在传统机器学习上的表现

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型方法 | 准确率 | 精确率 | 召回率 | F1 |
| Decision Tree | 99.98% | [99.91%,  100%,  99.99%] | [99.94%,  100%, 99.99% ] | [99.92%,  100% ,  99.98% ] |
| Random Forest | 99.96% | [99.86%,  100%,  99.97% ] | [99.93%,  100%,  99.99% ] | [99.90%,  100%,  99.98% ] |
| Neural Net | 99.87% | [99.66%, 98.54%, 99.85%] | [99.82%, 99.85%, 99.95%] | [99.74%,  99.19%,  99.90% ] |
| Naive Bayes | 70.15% | [3.06%, 99.70%, 54.22% ] | [0.50%,  100%,  99.98%] | [0.86%,  99.85%,  70.31% ] |
| RBF-SVM | 99.22% | [94.81%,  100%,  99.94%] | [99.97%, 99.41%, 99.98% ] | [97.32%,  99.70%,  99.96%] |
| KNN | 99.94% | [99.74%, 99.70%, 99.97%] | [99.93%, 99.85%, 99.99%] | [99.83%,  99.77%,  99.98% ] |

此外，本文在研究初期，采用提取单向流的流量分割方式，在对Friday-02-03-2018-Trafficr提取的包含Bot僵尸网络的单向流样本数据进行实验，对Bot的准确率为99.98%，f1为99.98%，性能表现不如会话，最后调整策略，采取了提取会话的样本分割方式。

**4 结语**

传统的机器学习入侵检测算法多依赖于人工设计特征，本文提出基于深度学习的异常入侵检测算法，采用端到端的架构，自动学习会话中潜在的高级抽象特征，无需人工设计特征，具有实时检测的优点。实验证明本文算法是有效的，并且效果好于人工设计的特征，此外本文还详细研究了网络流量切割的方法，遵循一般会话的定义来提取双向流，给出了详细的会话分割流程。

未来展望：对抗算法的检测，由于算法提取的流量表示为数字向量，从可视化图像中可以看到，每一种攻击都有较为明显的图形特征，如果攻击者将自己的攻击流量的数据包大小重新设计，添加一些无用数据，打乱其原本的图形特征，那么基于将网络流量视为图像的检测算法是否还可以检测出来；此外，结合人工规则和深度学习的入侵检测算法可以弥补两者的缺点，达成更高的准确率，降低误报率，也是一个研究方向；近几年，图神经网络开始兴起，采用图的结构来表示数据可以承载更加复杂的数据形式，如将主机和会话表示为结点，边是流与主机之间的联系[18]，可以成为检测异常入侵的另一种方向。

**参考文献：**

[1] The UCI KDD Archive，KDD Cup 1999 Data[EB/OL]. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html，1999-10-28/2021-09-22.

[2] Tavallaee M, Bagheri E, LU Wei, et al. A detailed analysis of the KDD CUP 99 data set[C]∥IEEE. 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications. New York：IEEE,2009:1–6.

[3] Moustafa N, Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set) [C]∥IEEE. Military Communications and Information Systems Conference (MilCIS).New York：IEEE,2015:1-6.

[4] Sharafaldin I, Habibi Lashkari A,Ghorbani Ali A, Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization[C]∥SciTePress. Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP).Portugal: SciTePress,2018,108-116.

[5] Moustafa N, Slay J. The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 dataset and the comparison with the KDD99 dataset[J]. Information Security Journal, 2016,25(1-3): 18-31.

[6] Goodfellow Ian, Bengio Yoshua, Courvile Aaron. DEEP LEARNING[M]. 赵申剑，黎彧君，符天凡，李凯译，北京：人民邮电出版社，2017:269.

[7] HU Wenjie, LIAO Yihua, Vemuri VR. Robust anomaly detection using support vector machines[C]∥IEEE. Proceedings of the 2003 International Conference on Machine Learning and Cybernetics. New York：IEEE,2003:282-289.

[8] Dong Bo, WANG Xue. Comparison deep learning method to traditional methods using for network intrusion detection[C]∥IEEE. 2016 8th IEEE International Conference on Communication Software and Networks. New York：IEEE ,2016:581-585.

[9] Ravi V, KP Soman, Poornachandran P. A Comparative Analysis of Deep Learning Approaches for Network Intrusion Detection Systems (N-IDSs): Deep Learning for N-IDSs[J]. International Journal of Digital Crime and Forensics (IJDCF), 2019,11(3):65-89.

[10] XU Congyuan, SHEN Jizhong, DU Xin, ZHANG Fan. An Intrusion Detection System Using a Deep Neural Network With Gated Recurrent Units[J].IEEE Access,2018,6, 48697-48707.

[11] CHO K, Merrienboer B, et al. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches[EB/OL]. https://arxiv.org/pdf/1409.1259v2.pdf,2014/2021.

[12] Chung J, Gulcehre C, Cho KH, Bengio Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling[EB/OL]. https://arxiv.org/pdf/1412.3555.pdf,2014/2021.

[13] WANG Zhanyi. The Applications of Deep Learning on Traffic Identification[EB/OL]. https://www.blackhat.com/docs/us-15/materials/us-15-Wang-The-Applications-Of-Deep-Learning-On-Traffic-Identification-wp.pdf ,2015/2019.

[14] WANG Wei, ZHU Ming, ZENG Xuewen, et al. Malware traffic classification using convolutional neural network for representation learning[C]//IEEE. 2017 International Conference on Information Networking (ICOIN).New York: IEEE,2017:712–717.

[15] WANG Wei, ZHU Ming, WANG Jinlin, ZENG Xuewen, YANG Zhongzhen, End-to-end encrypted traffic classification with one-dimensional convolution neural networks[C]//IEEE. 2017 IEEE International Conference on Intelligence and Security Informatics (ISI).New York:IEEE,2017: 43-48.

[16] PENG Yujie, NIU Weina, ZHANG Xiaosong, Zhou Jie,et al, End-To-End Android Malware Classification Based on Pure Traffic Images[C]//IEEE. 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP).New York:IEEE,2020: 240-245.

[17] Radford B J, Richardson B D. Sequence Aggregation Rules for Anomaly Detection in Computer Network Traffic[EB/OL].https://arxiv.org/pdf/1805.03735v1.pdf,2018/2021

[18] David Pujol-Perich, Jose Su ´ arez-Varela, Albert Cabellos-Aparicio, Pere Barlet-Ros. Unveiling the potential of Graph Neural Networks for robust Intrusion Detection [EB/OL]. https://arxiv.org/pdf/2107.14756.pdf，1999-10-28/2021-09-22.

**作者简介：**



**第一作者照片**

王金华（1996年9月25），男，硕士，四川大学网络空间安全学院研究生在读，主要研究方向为深度学习，入侵检测;

邮箱：859182300@qq.com，电话：19938451778，qq：859182300，微信:19938451778