网络流量异常检测与分析

调研报告

处理架构组

李侗格 2019

目录

**[摘要](#_Toc20964_WPSOffice_Level1)** **[3](#_Toc20964_WPSOffice_Level1)**

**[第一章 背景](#_Toc1577_WPSOffice_Level1)** **[4](#_Toc1577_WPSOffice_Level1)**

**[第二章 网络流量异常流量概述](#_Toc14339_WPSOffice_Level1)** **[5](#_Toc14339_WPSOffice_Level1)**

**[第三章 网络流量异常检测方法的分类](#_Toc29286_WPSOffice_Level1)** **[8](#_Toc29286_WPSOffice_Level1)**

[3.1基于统计的方法](#_Toc1577_WPSOffice_Level2) [8](#_Toc1577_WPSOffice_Level2)

[3.2基于分类的方法](#_Toc14339_WPSOffice_Level2) [9](#_Toc14339_WPSOffice_Level2)

[3.3基于聚类的方法](#_Toc29286_WPSOffice_Level2) [9](#_Toc29286_WPSOffice_Level2)

[3.4基于信息论的方法](#_Toc21020_WPSOffice_Level2) [10](#_Toc21020_WPSOffice_Level2)

**[第四章 基于机器学习的流量异常检测](#_Toc21020_WPSOffice_Level1)** **[12](#_Toc21020_WPSOffice_Level1)**

[4.1机器学习方法](#_Toc19017_WPSOffice_Level2) [12](#_Toc19017_WPSOffice_Level2)

[4.2基于机器学习的流量异常检测](#_Toc4100_WPSOffice_Level2) [13](#_Toc4100_WPSOffice_Level2)

[4.3 机器学习算法目前存在的问题](#_Toc6013_WPSOffice_Level2) [17](#_Toc6013_WPSOffice_Level2)

**[第五章 总结](#_Toc19017_WPSOffice_Level1)** **[18](#_Toc19017_WPSOffice_Level1)**

**[参考文献](#_Toc4100_WPSOffice_Level1)** **[19](#_Toc4100_WPSOffice_Level1)**

摘要

随着近年来计算机网络的快速发展，网络数据量巨幅增长，网络攻击手段也日趋多样化、复杂化，网络安全问题日益突出。网络流量异常检测网络异常流量检测方法作为一种可靠的网络安全技术，成为了目前为保障网络空间安全的重要研究领域。

本文首先介绍网络流量异常检测技术，总结了网络流量异常类型。之后对流量异常检测方法进行分类，将现有的检测方法分为基于分类的、基于统计的、基于聚类的、基于信息论的等四大类方法，并对每类检测方法进行了详细的描述，着重介绍了机器学习算法在网络流量异常检测中的应用。最后对全文进行总结并探讨分析目前异常检测方法存在的不足。

第一章 背景

网络技术的快速发展以及网络应用的不断普及，潜移默化地影响着人们的生活方式和工作方式。网络已经渗透到社会生活的方方面面，人们对互联网的依赖性日益增加。中国互联网络信息中心(CNNIC)在2019年8月30日发布的第44次《中国互联网络发展状况统计报告》中提到，截至2019年6月，我国网民规模为8.54亿，互联网普及率达到61.2%，且仍处于增长趋势。互联网为网民带来便利的同时，网民面临的风险也日益增加。各种计算机病毒和网络攻击蜂拥而至，给用户带来了不可避免的危害，对我们的信息安全产生了极大的威胁。攻击者借助互联网，利用蠕虫病毒、木马等技术对网络进行攻击，从而获取网络用户的个人隐私信息和私密性较高的资源，这些攻击行为对网民的信息安全影响恶劣。因此，采用有效的网络安全技术和方法对于保障网络安全有着非常重要的作用。

网络流量异常检测作为一种有效的防护手段，能够发现未知攻击行为，可以为网络态势感知提供重要技术支持，近年来受到越来越多的关注。网络流量异常，是指对网络正常使用造成不良影响的网络流量模式，与正常流量差别较大，会引起网络性能下降甚至不可用。引起网络流量异常的原因分为两方面：一是性能原因，指网络结构设计不合理或使用不当造成的异常流量，例如拥塞控制不当、网络设备故障等；二是安全原因，是指网络攻击行为造成的异常流量，例如DDoS攻击、蠕虫病毒等。本文主要研究安全原因造成的网络流量异常。

针对各种复杂的网络异常情况，许多学者在分析研究中可以发现，无论是网络中发生了哪种异常情况都会导致网络流量发生异常变化。网络流量异常检测，是指应用各种异常检测方法分析网络流量并及时发现具有异常行为的流量，对增强网络态势感知能力和维护网络空间安全等发挥着重要作用，具有重要研究意义。

第二章 网络流量异常流量概述

异常(Anomaly)，是指目标数据集中与正常模式的统计特征明显不一致的数据模式，在很多领域的数据分析中都是一个比较常见的现象。造成数据异常的活动多种多样，比如金融数据中的信用卡欺诈，网络数据中的网络攻击等。

按照数据具体表现形式，异常可以分为以下三：

点异常(Pointanomaly)，指某项单独的数据与正常模式差异程度较大，例如，工业数据中，如果某辆车平时的耗油量都是每天五升，但是有一天这个数值变成了五十升，那么这项数据就是点异常。

上下文异常(Contextualanomaly)，指某项数据在特定上下文环境中与正常模式差异较大，例如，金融数据中，信用卡消费记录会在某些节日期间开支变大，例如圣诞节或新年，这属于正常现象，但是如果在非节假日期间某人的信用卡开支突然变大，这就属于上下文异常。

集合异常(Collectiveanomalv)，指一组相似数据共同表现出与整个数据集的正常模式差异较大，例如，医疗数据中，如果某人的心电图出现一段较长时间的连续低值，则很有可能是一种叫做期外收缩的心脏异常现象造成的，这属于集合异常，相反地，单独的一个低值并不属于异常。

引起网络流量异常的原因有操作不当和遭受网络攻击两种。本文主要关注由网络攻击造成的网络流量异常现象。网络攻击，是指任何威胁到计算机网络系统安全正常运行的有特定目的的行为。从形式上分为四类。

1. 拒绝服务(DoS，DenialofService)：在DoS攻击中，系统被大量来自单个连接的请求充斥，直至所有资源被耗尽，然后再也无法处理合法请求。在分布式拒绝服务攻击(DDoS)中攻击者使用分布在互联网上的多个连接且可能是僵尸网络的一部分，这种攻击使得用户无法使用服务或资源。小，中型DoS和DDoS攻击通常被攻击者当作幌子，来隐藏更小但更危险的活动或拿下安全设备如防火墙。
2. 窃听(Probe)：这种攻击的目的是收集目标主机或者网络的信息，本质上是一种侦查活动。对于一个网络的探测攻击可以获知其主机的数量和类型，对于一个主机的探测攻击可以获知其安装的应用程序和类型。探测攻击本身并不会造成实际的损失，但是探测攻击往往是实施进一步网络攻击的第一个步骤，利用收集的目标信息可以实施更加准确和有效的攻击活动。
3. 用户到根(U2R，UsertoRoot)：在这种攻击时，攻击者通常已具有目标系统的本地访问权限，意图利用由普通用户到超级用户/管理员的升级权限漏洞。最常见的U2R攻击是缓冲区溢出(bufferoverflow)攻击，攻击者试图溢出缓冲区以提升其用户权限并执行恶意代码。
4. 远程到本地(R2L，RemotetoLocal)：在此类攻击中，攻击者在目标计算机中没有账户，其尝试获取本地访问。R2L攻击通常于U2R攻击结合在一起。R2L攻击的一个例子是SSHbruteforce。

根据上面对异常分类以及网络攻击分类的介绍，按照文献[2]的划分，如图2．1所示，可以将四类网络攻击对应至三类异常。拒绝服务攻击由于是一系列大量的连接的同时活动，因此属于集合异常。探测属于在某些目的的驱使下收集特定信息，因此属于上下文异常。一般将U2R与R2L划分为点异常。

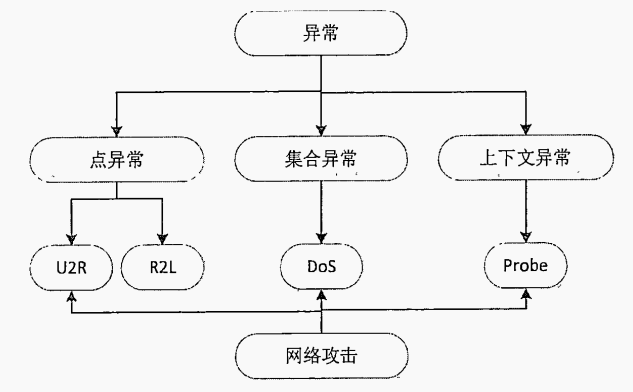


图2.1异常与网络攻击形式的对应关系图

网络流量异常检测，是指应用各种异常检测技术分析网络流量，及时发现网络攻击行为。需要指出的是，和网络流量异常检测非常相近的一个概念是入侵检测，它们之间的区别如下。入侵检测，分为基于主机的入侵检测和基于网络的入侵检测，前者是指通过监视主机日志信息来检测入侵行为，例如对文件系统，注册表的改动信息等，后者是指通过分析进出的网络流量判定入侵行为。基于网络的入侵检测可分为误用检测和异常检测，误用检测是指基于已有攻击提取的流量指纹库进行匹配检测。异常检测是指根据历史数据建立机器学习模型以判定新流量数据是为异常，基于异常的网络入侵检测基本等同于安全相关的网络流量异常检测。

第三章 网络流量异常检测方法的分类

自1987年Denning提出异常检测模型以来，众多国内外的学者针对网络异常检测研究付出了诸多努力，很多具有重要参考价值以及实际应用程度较高的网络异常检测方法被提出。将网络流量异常检测方法具体分为基于分类的、基于统计的、基于聚类的、基于信息论的等四大类方法。下面对四类方法进行简要介绍：

3.1基于统计的方法

最早被提出的异常流量检测方法是基于统计分析的异常检测方法，当前阶段针对其技术的研究已经较为成熟。网络中正常的流量其特征参数参数是稳定的，然而异常流量的特征是不稳定且不相关的。根据以上原理，可以根据正常流量数据的特征来建立统计模型，然后将当前网络流量数据借助统计模型来进行判断，若是当前流量特征不在统计模型的参数范围之内或者是概率很小，则将其看作异常流量。

程光等人提出了一种方法，使用特征维度来鉴别网络中的扫描攻击，这个特征维度就是网络中流量数据的报文长度。算法中认为在正常的网络流量中，流量的数据报文长度相对较为稳定，然而，当网络中发生扫描攻击时，报文长度会发生明显的变化，许多较为短小的报文也会随之产生。除了以上程光等人提出的流量特征之外，还有其他的流量特征可以用来进行统计，如TCP和UDP报文在网络流量中所占的比率，TCP和UDP报文的平均长度等。在网正常的网络流量数据中，这些特征值在都处在一个相对稳定的区间，可以用来构建统计模型。Blazek等人提出的基于流量序列变化值分析的DoS检测方法。

在基于统计分析的异常检测方法中，统计模型在进行正常，异常流量的判别时会自动丰富统计模型中的参数，若是正常流量，则添加进模型并且调整模型中参数的范围；若是异常流量，则能够准确地进行识别。基于统计分析的异常检测方法具有学习性，在不断的检测过程中完善自身，有很大概率能够检测到未知攻击的发生。同时，操作起来较为简单。但是基于统计分析的网络异常检测方法整体环节较为简单，若是网络中的攻击者使用同样的方式来构统计模型，对参数进行修改，就能够将网络异常行为伪装成正常的流量行为。而且基于统计分析的网络异常检测方法由于要比较的特征维度非常多，所以统计模型难以识别出最有价值的统计参数，与此同时，采用此类方法可能导致错报和漏报的情况产生。

3.2基于分类的方法

基于分类的网络流量异常检测方法，是指根据有监督的机器学习的方式，采用基于网络流量分类的方法进行检测。按照检测的粒度，首先确定具体需要的分类数量和类别，例如需要分类五种流量，分别是正常流量，DoS流量，Probe流量，U2R流量，R2L流量，然后使用带标签的网络流量数据训练一个机器学习模型，例如C4.5决策树模型，支持向量机SVM，神经网络模型，之后对新流量进行分类，分类结果即为检测结果。模型通过训练从样本(即标记数据)中进行学习。

Catania C A等人提出了一种对正常流量进行自主标注的方法，以解决支持向量机算法所要求的类分布不平衡的情况。J．Kline等在贝叶斯网络的基础上增加了时间序列，对比传统的单一的贝叶斯网络，这种方法提高了贝叶斯网络的准确性。Manikopoulos等人提出的基于神经网络的入侵检测系统。Radford B J等人证明了递归神经网络能够学习一个模型来表示网络上计算机之间的通信序列，并且可以用来识别离群点网络流量。

此外，基于分类的系统需要定期重新训练以保持高检测率，而由于很难获得定期的标记数据，这是不可行的。即使有标签数据，是否包含所有新攻击也尚不确定。

3.3基于聚类的方法

聚类是一种无监督的检测方法，聚类是将无标记数据集转换为隐藏数据结构的离散集，而不是提供可从其生成的未观察样本的准确特征概率分布。聚类算法在没有任何先验知识的情况下将给定的数据分组(聚类)，以实现高的内部相似性和外部差异性。

聚类检测方法建立在以下三个假设之上。

假设1：只能对正常数据进行聚类，新的数据如果与正常聚类结果差别较大则可判为异常。例如，在基于密度的聚类算法中，噪声数据被看做是异常数据，因此不会为其建立聚类簇。

假设2：如果一个聚类结果同时包含正常数据和异常数据，正常数据通常距离聚类的形心更近，而异常数据通常距离聚类的形心更远，因此可以用一个距离度量方式进行检测。

假设3：在一个包含多个聚类结果的数据集中，小的聚类和稀疏的聚类通常被看做是异常数据，因此，如果一个聚类的尺寸和密度小于一定阈值，则这些数据可判为异常数据。

聚类的方法大致可以分为常规聚类和协同聚类两种。常规聚类：Jianliang M等人采用了K-means算法聚类，分析KDDCup1999数据，实验结果表明K-means算法效率和检测率都很高。协同聚类：这种聚类是指在列和行两个方向上同时应用聚类算法。协同聚类的优势是可以产生更加压缩的数据表示，可以将其看作一种降维技术，计算复杂度也可以大幅降低。例如，Ahmed M提出了一种可用于检测DoS攻击异常流量的检测框架，效率提升非常明显。基于聚类技术系统的主要优势是其在没有任何先验知识的情况下检测未知攻击的能力，消除了对标记数据的需求。其主要缺点是假阳率高。

3.4基于信息论的方法

信息论中的许多概念都可以解释网络流量数据集的特征，例如熵，条件熵，相对熵，信息增益等，因此，利用信息论的方法可以建立相应的网络流量异常检测模型。信息论的基本思路主要基于这样的假设：异常会引起数据的信息内容发生不规则变化。思路如下：若O(D)代表一个给定数据集D的复杂度，信息理论技术的目标是找到D最小的子集I，使O(D)-O(D-I)最大，I中所有的数据实例就为异常。严承华等人采用时问作为单位形成统计元组，在每个统计元组内分别计算各网络属性的信息熵，根据正常网络状态下单位统计元组内流量信息结构的稳定性来判断网络是否发生异常。陈鹏等将信息熵应用于异常检测中，并没计了基于三维可视化的流量监测系统。Tan等针对DoS攻击提出了一种使用多变量相关分析方法来提取流量数据几何相关性特征的检测系统，也可以看做是使用了信息论的相关理论和方法，实际检测效果也比较好。信息熵方法检测异常的速度快，并且对未知的攻击发放有较好的检测效果，具有广泛的应用空间，但是现有的基于信息熵的检测方法抗干扰能力较差，特别是对混合在背景流量中的攻击，容易造成误判和遗漏。

1. 基于机器学习的流量异常检测

传统的通过静态规则匹配的网络异常检测方法在动态、复杂的网络环境中难以检测出未知异常和攻击类型,不能满足网络安全检测的要求。机器学习算法具有自学习的特点，自适应性较强，能够适应复杂多变的网络环境，检测出未知异常，满足实时准确检测的需求。因此，使用机器学习的方法检测网络中海量的流量是一个重要的发展方向。

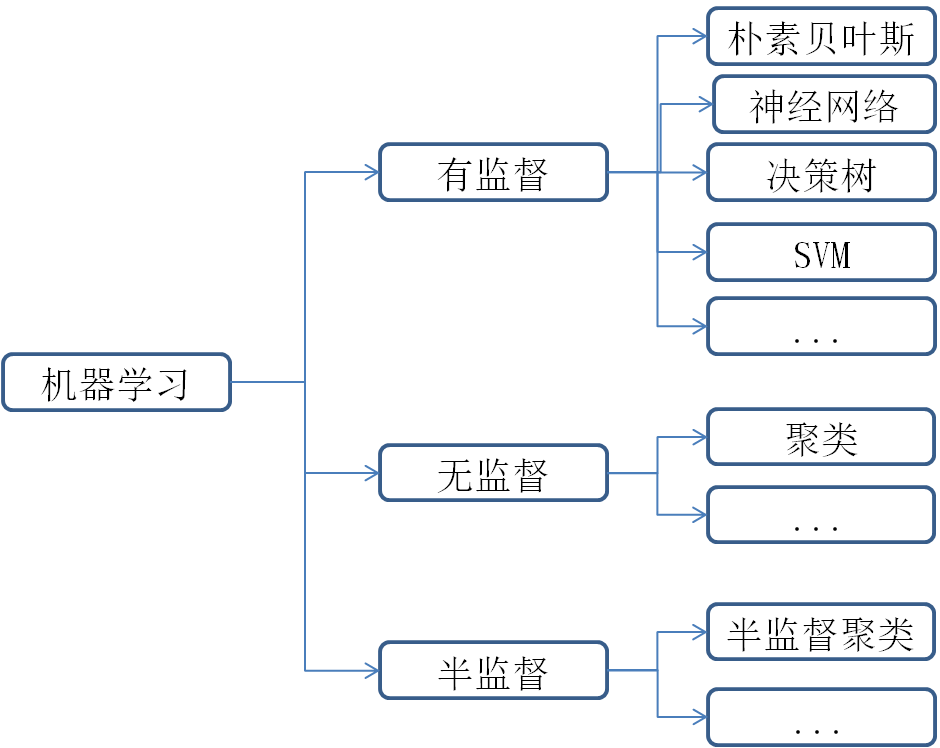
4.1机器学习方法

机器学习方法可分为主要的三类：无监督、半监督和有监督（按照贴标签的程度来划分）。

有监督学习通过将已有的异常样本数据（已知的正常数据和异常数据）去训练得到一个最优的模型，再根据得到的模型去检测其它所有的输入数据所对应的输出结果，由输出结果做一个简单的判断从而实现异常类型的分类。

半监督学习是监督和非监督学习相结合的一种方法，它是将少量标记的样本与大量未标记的样本同时进行训练学习，这样减少了标记大量数据的工作，同时又可以提高检测系统的准确度。

无监督学习的训练模型不需要对异常流量进行标记，可以直接使用现有的流量数据进行建模，省去了手工标记大量数据的麻烦。

图4.1机器学习方法分类

一种机器学习方法，往往由两个阶段组成：训练和测试。具体步骤如下：

实际上，对于大多数机器学习方法，有三个阶段：训练，校验，测试。机器学习方法一般会涉及参数，如：人工神经网络的层数、节点数。在完成训练之后，往往得到一些可用的模型。选定哪个才能在测试集中得到较好的估计误差呢？这就需要有第三个区分的数据集，校验数据集。

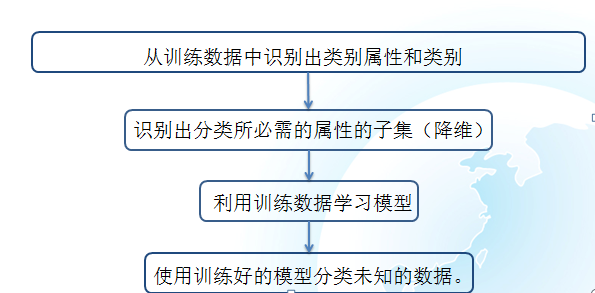


图4.2机器学习流程

在校验数据中上表现最好的那个模型，便是我们需要用的模型。且不应为了提高在测试集上的准确度，再对其进行微调。一旦进行了微调，便不能保证这个模型在其他测试集上的性能了。

4.2基于机器学习的流量异常检测

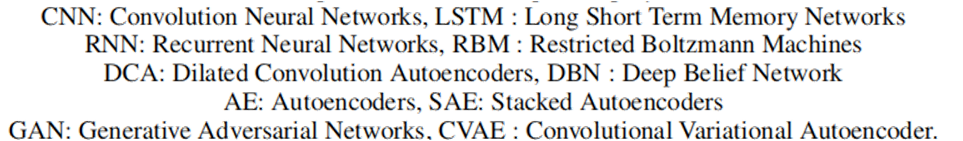
现阶段基于机器学习的流量异常检测成为热点，比较常用的机器学习方法有神经网络，决策树，聚类，SVM等，以下是对这几种算法在流量异常检测上的应用进行简要介绍。

4.2.1神经网络

神经网络，主要指人工神经网络（ArtificialNeural Network，ANN），是机器学习算法中比较接近生物神经网络特性的数学模型。通过模拟人类神经网络的结构和功能，由大量“神经元”构成了一个复杂的神经网络，模拟神经元的刺激和抑制的过程，最终完成复杂运算。

Lippmann和Cunningham提出了一个使用关键词选择和人工神经网络的系统。在telnet会话记录上进行关键字选择，并且统计每个关键字发生的次数。第一个神经网络的输入为关键字统计数据，可以用来判断是否为异常流量。第二个神经网络对被标记为攻击的实例进行分类(即提供攻击名称)。该系统实现了80%的检测，每天大约有1次虚警。

近年来,随着深度学习的快速发展,已经出现了深层神经网络在网络流量异常检测领域的相关应用。例如,Gao提出了一个使用深度信念网络DBN的入侵检测模型,并使用NSL—KDD数据集进行了验证，这项研究使用的输入数据是人工设计的41种流量特征。下表列出了近几年深度学习在入侵检测方面的研究成果。



|  |  |  |
| --- | --- | --- |
| Techniques | Model Architecture | References |
| Generative | DCA,SAE,RBM,DBN,CVAE | [19],[20],[21],[22],[23],[24],[25],[26],[27] |
| Hybrid | GAN | [28],[29],[30],[31],[32],[33],[34],[35] |
| Discriminative | RNN,LSTM,CNN | [19],[36],[37],[38],[39],[40] |

表4.1近几年深度学习在入侵检测方面的研究成果

4.2.2聚类算法

聚类算法是根据相似性度量将目标聚集成簇的方法，聚类后簇内各成员相似，而不同簇之间存在显著差异，是一种无监督学习。目前，聚类算法主要有四种类别：基于划分的聚类算法、层次聚类算法、基于密度的聚类算法和基于网格的聚类算法。

Blowers和Williams使用DBSCAN聚类方法对正常和异常网络数据包进行分组。对KDD数据集进行预处理，以使用相关分析来选择特征。在数据预处理过程中设置攻击比为10%。对于攻击或无攻击检测，报告的性能为98%。

1. Hendry 等人使用称为简单日志文件集群工具(SLCT)的基于密度的聚类方法来创建正常和恶意网络流量的集群。该系统由两个聚类方案组成：第一种用于正常或攻击检测，第二种用于有监督的方式来确定正常流量。该研究采用 KDD数据集进行实验，对于未知攻击有 70%到 80%的检测率，结果可观。

4.2.3决策树

决策树学习采用的是自顶向下的递归方法，其基本思想是以信息熵为度量构造一颗熵值下降最快的树，到叶子节点处，熵值为0。其具有可读性、分类速度快的优点，是一种有监督学习。

EXPOSURE是一个采用大规模被动域名服务(DNS)分析技术来检测与恶意活动有关的域的系统。这个系统由五个部分组成。主要组件：数据收集器、特征属性组件、恶意和良性域收集器、学习模块和分类器。分类器是由WekaJ48决策树程序构建的，它是C4.5算法的一个实现，能够生成剪枝或未修剪的决策树。实验表明，当所有特征组合在一起时，误差达到最小。决策树使用了所有15个特征。

恶意域名来自www.malwaredomains.com， Zeus Block List，Anubis等。初始恶意域列表由3500个域组成。良性域来自Alexa前1000个域列表。

通过对不同周期长度值的实验，确定了该系统的最佳初始训练时间为7天。经过最初的训练，分类器被重新分类。在这最初的训练之后，分类器每天都被重新训练。根据数据集的不同，结果差别很大。总体上，通过十倍交叉验证，恶意域的检测准确率为98.5%，错误接受率为0.9%。

4.2.4 SVM

SVM（support Vector Mac）又称为支持向量机，是一种二分类的模型。当然如果进行修改之后也是可以用于多类别问题的分类。支持向量机可以分为线性核非线性两大类。其主要思想为找到空间中的一个更够将所有数据样本划开的超平面，并且使得本本集中所有数据到这个超平面的距离最短。

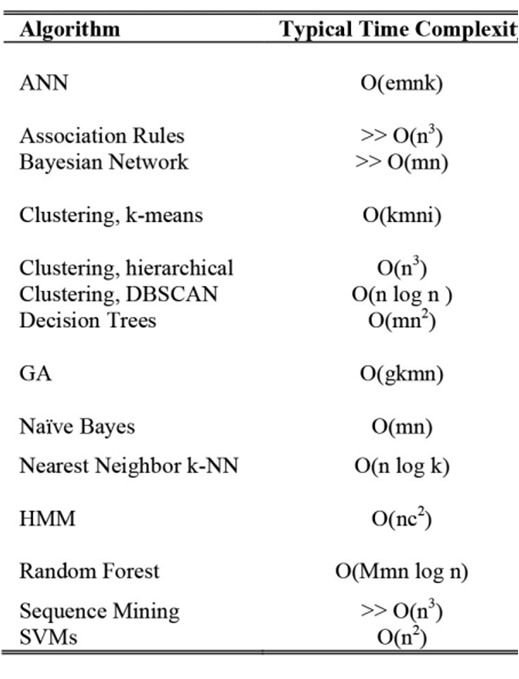
Shon和Moon描述了一种检测网络流量中的新攻击的框架。它采用自组织特征映射(SOFM)、遗传算法和支持向量机相结合的方法。

系统的四个步骤如下：

1. 一种执行SOFM聚类的无监督人工神经网络用于包分析。
2. 使用被动TCP/IP指纹进行数据包过滤。
3. GA选择特征。
4. 数据用增强的支持向量机进行分类，该支持向量机是由一个单类支持向量机和一个有监督的软边缘分类器，前者提供了一类支持向量机的无标记分类能力，后者提供了监督支持向量机的高检测性能。

4.2.5 机器学习算法时间复杂度

总结下图给出了不同机器学习算法的时间复杂度，假设数据有n个实例，每个实例有m个特征。

图4.3 不同机器学习算法的时间复杂

一般来说，O(n)和O(n log n)算法被认为是线性时间，可用于在线方法。O()被认为是大多数实践中可以接受的时间复杂度。O()和更高的算法被认为比较慢的算法，适用于离线方法。

4.3 机器学习算法目前存在的问题

到目前为止，异常流量检测技术取得了很大的进步，从传统的人工化到智能化、经验化到自演进。但是，基于机器学习的异常检测技术依然存在以下问题：

1. 由于机器学习中神经网络算法常常会发生过拟合的情况，所以对于训练时候选取的样本数据需要特别准确。
2. 异常特征选择存在考虑不足的情况，严重影响了异常检测的准确度，需要进一步探讨特征提取的方法。
3. 针对独立的机器学习方法都存在一定的缺陷，混合的机器学习方法只能满足特定的网络环境，缺乏适合多环境的异常检测技术。

第五章 总结

本文首先介绍网络流量异常检测技术，总结了网络流量异常类型。之后对流量异常检测方法进行分类，将现有的检测方法分为基于分类的、基于统计的、基于聚类的、基于信息论的等四大类方法，并对每类检测方法进行了详细的描述，探讨分析了其优点与不足，最后详细介绍了机器学习算法在网络流量异常检测中的应用。

当前异常流量检测已取得了显著的进步和成绩，但在将检测方法应用到实际中时仍然存在以下几个问题：

(1)需要不断更新模型中的特征库。在当前复杂多变的网络环境下，遭受网络攻击时的攻击手段也在不断进步，所以必须不断地更新当前的特征库，重构检测模型，才能提高异常检测的准确率，

(2)实时性要求高。处在一个网络互联的时代中，各种信息密切相关。在异常检测方面，异常流量检测系统需要在发现风险情况时，及时确定攻击的类型，做到实时的处理。

(3)难以对未知攻击做到精确的检测。当前大多是根据异常流量特征码来检测异常，能够有效地识别出已知的网络攻击类型，但是对于当前阶段未知的网络攻击手段难以做到有效的检测和防范。

(4)异常检测方法的检测率不断得到提高，但是针对不同的具体攻击方面有着较高的误报率。

参考文献

1. 中国互联网络信息中心(CNNIC).2019 中国互联网络发展状况统计报告[R].2019
2. 李军军.计算机网络安全问题及其防范措施[J].网络安全技术与应用，2018，(11):1-3.
3. 王伟. 基于深度学习的网络流量分类及异常检测方法研究[D]. 2018.
4. Denning D E . An Intrusion-Detection Model[C]// 1986 IEEE Symposium on Security and Privacy. IEEE, 1986.
5. Ahmed M, Mahmood A N, Hu J. A survey of network anomaly detection techniques.[J]. Journal of Network & Computer Applications, 2016, 60:19-31.
6. 程光, 龚俭, 丁伟. 基于抽样测量的高速网络实时异常检测模型[J]. 软件学报, 2003, 14(3):594-599.
7. Thottan M , Ji C . Anomaly detection in IP networks[J]. IEEE Transactions on Signal Processing, 2003, 51(8):2191-2204.
8. Catania C A , Bromberg F , Carlos García Garino. An autonomous labeling approach to support vector machines algorithms for network traffic anomaly detection[J]. Expert Systems with Applications, 2012, 39(2):1822-1829.
9. Kline J , Nam S , Barford P , et al. Traffic Anomaly Detection at Fine Time Scales with Bayes Nets[C]// International Conference on Internet Monitoring & Protection. IEEE, 2008.
10. Manikopoulos C , Papavassiliou S . Network intrusion and fault detection: a statistical anomaly approach[M]. IEEE Press, 2002.
11. Radford B J , Apolonio L M , Trias A J , et al. Network Traffic Anomaly Detection Using Recurrent Neural Networks[J]. Chinese Journal of Network & Information Security, 2018.
12. Jianliang M , Haikun S , Ling B . The Application on Intrusion Detection Based on K-means Cluster Algorithm[M]. IEEE Computer Society, 2009.
13. Ahmed M, Mahmood A N. Network Traffic Pattern Analysis Using Improved Information Theoretic Co-clustering Based Collective Anomaly Detection[M]// International Conference on Security and Privacy in Communication Networks. 2014.
14. 严承华, 程晋, 樊攀星, et al. 基于信息熵的网络流量信息结构特征研究[J]. 信息网络安全, 2014(3):28-31.
15. 张登银, 廖建飞. 基于相对熵的网络流量异常检测方法[J]. 南京邮电大学学报(自然科学版), 2012, 32(5):26-31.
16. Ahmed M , Mahmood A N , Islam M R . A survey of anomaly detection techniques in financial domain.[J]. Future Generation Computer Systems, 2016, 55(6):278-288.
17. Tan Z , Jamdagni A , He X , et al. A System for Denial-of-Service Attack Detection Based on Multivariate Correlation Analysis[J]. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(2):447-456.
18. Lippmann R P , Cunningham R K . Improving intrusion detection performance using keyword selection and neural networks[J]. Computer Networks, 2000, 34(4):597-603.
19. Yang Y , Jun L , Zhiping C . Network Intrusion Detection through Stacking Dilated Convolutional Autoencoders[J]. Security and Communication Networks, 2017, 2017:1-10.
20. Thing V L L . IEEE 802.11 Network Anomaly Detection and Attack Classification: A Deep Learning Approach[C]// 2017 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2017.
21. Zolotukhin M , Hamalainen T , Kokkonen T , et al. Increasing web service availability by detecting application-layer DDoS attacks in encrypted traffic[C]// 2016 23rd International Conference on Telecommunications (ICT). IEEE, 2016.
22. Cordero C G , Hauke S , Muhlhauser M , et al. Analyzing flow-based anomaly intrusion detection using Replicator Neural Networks[C]// 2016 14th Annual Conference on Privacy, Security and Trust (PST). IEEE, 2016.
23. Alrawashdeh K , Purdy C . Toward an Online Anomaly Intrusion Detection System Based on Deep Learning[C]// IEEE International Conference on Machine Learning & Applications. IEEE, 2017.
24. Tang T A , Mhamdi L , Mclernon D , et al. Deep Learning Approach for Network Intrusion Detection in Software Defined Networking[C]// The International Conference on Wireless Networks and Mobile Communications (WINCOM'16). IEEE, 2016.
25. Manuel L M , Belen C , Antonio S E , et al. Conditional Variational Autoencoder for Prediction and Feature Recovery Applied to Intrusion Detection in IoT[J]. Sensors, 2017, 17(9):1967-.
26. Al-Qatf M , Lasheng Y , Al-Habib M , et al. Deep Learning Approach Combining Sparse Autoencoder With SVM for Network Intrusion Detection[J]. IEEE Access, 2018, 6:52843-52856.
27. Mirsky Y , Doitshman T , Elovici Y , et al. Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection[J]. 2018.
28. Aygun R C , Yavuz A G . Network Anomaly Detection with Stochastically Improved Autoencoder Based Models[C]// 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud). IEEE Computer Society, 2017.
29. Lin Z , Shi Y , Xue Z . IDSGAN: Generative Adversarial Networks for Attack Generation against Intrusion Detection[J]. 2018.
30. Chuanlong Yin, Yuefei Zhu, Shengli Liu, Jinlong Fei, and Hetong Zhang. An enhancing framework for botnet detection using generative adversarial networks. In 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), pages 228–234. IEEE, 2018.
31. Markus Ring, Daniel Schlor, Dieter Landes, and Andreas Hotho. Flow-based network traffic generation using generative adversarial networks. arXiv preprint arXiv:1810.07795, 2018.
32. Majd Latah. When deep learning meets security. arXiv preprint arXiv:1807.04739, 2018.
33. Yotam Intrator, Gilad Katz, and Asaf Shabtai. Mdgan: Boosting anomaly detection using\\multi-discriminator generative adversarial networks. arXiv preprint arXiv:1810.05221, 2018.
34. Takashi Matsubara, Ryosuke Tachibana, and Kuniaki Uehara. Anomaly machine component detection by deep generative model with unregularized score. In 2018 International Joint Conference on Neural Networks (IJCNN), pages 1–8. IEEE, 2018.
35. Miguel Nicolau, James McDermott, et al. A hybrid autoencoder and density estimation model for anomaly detection. In International Conference on Parallel Problem Solving from Nature, pages 717–726. Springer,2016.
36. Maria Rigaki. Adversarial deep learning against intrusion detection classifiers, 2017.
37. Ritesh K Malaiya, Donghwoon Kwon, Jinoh Kim, Sang C Suh, Hyunjoo Kim, and Ikkyun Kim. An empirical evaluation of deep learning for network anomaly detection. In 2018 International Conference on Computing,Networking and Communications (ICNC), pages 893–898. IEEE, 2018.
38. Donghwoon Kwon, Kathiravan Natarajan, Sang C Suh, Hyunjoo Kim, and Jinoh Kim. An empirical study on network anomaly detection using convolutional neural networks. In 2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS), pages 1595–1598. IEEE, 2018.
39. Ni Gao, Ling Gao, Quanli Gao, and Hai Wang. An intrusion detection model based on deep belief networks. In Advanced Cloud and Big Data (CBD), 2014 Second International Conference on, pages 247–252. IEEE, 2014.
40. Ralf C Staudemeyer. Applying long short-term memory recurrent neural networks to intrusion detection. South African Computer Journal, 56(1):136–154, 2015.
41. Sheraz Naseer, Yasir Saleem, Shehzad Khalid, Muhammad Khawar Bashir, Jihun Han, Muhammad Munwar Iqbal, and Kijun Han. Enhanced network anomaly detection based on deep neural networks. IEEE Access,6:48231–48246, 2018.
42. Blowers M , Williams J . Machine Learning Applied to Cyber Operations[J].
43. Bilge L , Sen S , Balzarotti D , et al. EXPOSURE: A passive DNS analysis service to detect and report malicious domains[J]. ACM Transactions on Information and System Security, 2014, 16(4).
44. Shon T , Moon J . A hybrid machine learning approach to network anomaly detection[J]. Information Sciences, 2007, 177(18):3799-3821.