文章编号：2095-2163(2023)11-0268-07 中图分类号：TP393 文献标志码：A

SDN 网络架构下应用层 DDos 攻击分类及检测方法研究

陈 晔
(常州纺织服装职业技术学院 信息服务中心，江苏 常州 213164)

摘 要：DDoS 攻击手段主要是针传输层的 TCP-SYN、UDP,以及网络层的 ICMP 泛洪等,早期 DDos 的攻击很容易就被更先进的检测技术、如机器学习和深度学习技术检测出来,于是就出现了更复杂和针对性更强的 DDoS 攻击,即应用层攻击。本文将 SDN 架构中的各种组件在遭受 DDoS 攻击后按受到攻击的影响范围和攻击强度进行分类,同时使用轻量化工具 Mininet 构建模拟测试环境,应用 AdaBoost 机器学习模型,通过对数据流的分析,区分正常和恶意的数据流量,进一步提高检测的准确率,对 SDN 网络架构全面实现自动化防御具有现实意义。

关键词：软件定义的网络(SDN)；DDoS；OpenFlow；Mininet；RFE 算法；AdaBoost

Research on the classification and detection methods of application layer DDos attacks under SDN network architecture

CHEN Ye
(Information Service Center, Changzhou Vocational Institute of Textile and Garment, Changzhou Jiangsu 213164, China)

Abstract: DDoS attacks are mainly targeted at TCP-SYN and UDP at the transport layer, and ICMP flooding at the network layer, etc. Early DDos attacks are easily detected by more advanced detection techniques such as machine learning and deep learning techniques, and thus more sophisticated and targeted DDoS attacks, i.e., application layer attacks, emerge. In this paper, various components in the SDN architecture are classified according to the scope of impact and attack intensity after being subjected to DDoS attacks, while a simulation test environment is constructed using the lightweight tool Mininet, and the AdaBoosting machine learning model is applied to further improve the accuracy of detection by distinguishing normal and malicious data traffic through the analysis of data flows, which is useful for the SDN network. It is of practical significance to fully realize automated defense for SDN network architecture.

Key words: software-defined networking(SDN); DDoS; OpenFlow; Mininet; RFE algorithm; AdaBoost

0 引 言

传统网络设备的控制和转发是紧密结合的,一般是由同一台设备实现控制和转发,软件定义网络(SDN)已经成为一种新型的网络架构,其核心内容是控制平面与数据平面分离[1]。SDN 的架构如图 1 所示,通常由 3 层平面组成:数据平面、控制平面和应用平面[2]。其中,数据平面一般是指负责数据转发的交换机。控制平面包含一个或多个 SDN 控制器,集中管理网络中的转发设备。应用平面主要是面向业务应用 API。在交换机和控制器之间使用开放的接口协议:OpenFlow 协议。这种网络集中管理方式不仅简化了网络架构,而且使网络具有灵活性、可编程性。OpenFlow 协议主要解决转发设备(如 OpenFlow 交换机)和控制器之间的通信问题[3]。OpenFlow 交换机包含一个或多个由匹配规则、计数器和动作字段组成的流表,再根据指定流量的匹配规则(控制信令)完成数据转发,如图 2 所示。

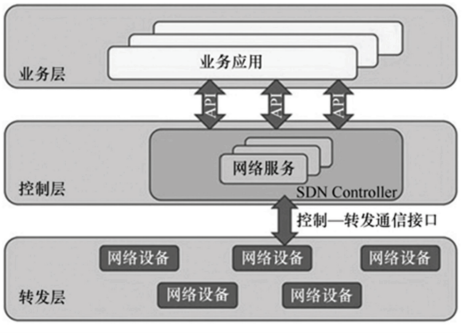

图 1 SDN 的基本架构

Fig. 1 The basic architecture of SDN

基金项目：江苏省现代教育技术研究 2021 年度课题(2021-R-88294)。

作者简介：陈 晔(1981-),男,硕士,高级工程师,主要研究方向:计算机网络技术、网络安全。Email:18676159@qq.com

收稿日期：2022-11-21

SDN 因为架构原因，自身存在如劫持 、中毒、配置错误、拒绝服务和跳板攻击等安全隐患。跳板攻击因为非常容易执行，又很难被发现，所以被黑客所青睐，黑客只要通过 wireshark 之类的抓包工具，监听网络流量，很容易就能将控制器和 OpenFlow 交换机之间的控制信令捕获到。

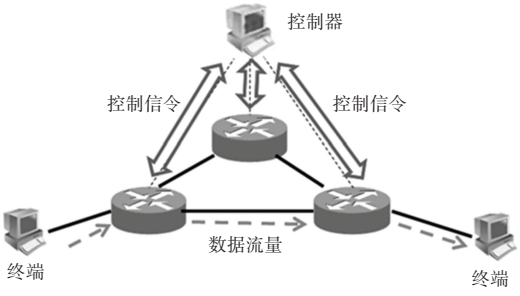

图 2　数据和控制分离

Fig. 2　Separation of data and control

1　Openflow 转发

OpenFlow 协议最初是在 2008 年作为斯坦福大学的一个研究项目部署在校园网，研究 OpenFlow 的目的是对传统网络进行变革，通过一个或多个控制器对多个交换机进行灵活控制，可按规则进行端口转发，简化了网络配置与管理，并通过可编程性实现网络层及应用层的创新。目前 OpenFlow 协议标准已经发布到最新的 1.5 版本，国内很多数通厂商在使用更为稳定的 1.3 版本（如锐捷、华三），OpenFlow 协议所控制的交换机有 2 种运行模式：主动模式和反应模式。其中，主动模式是指，匹配规则在业务流量到达之前提前部署到交换机，而反应模式是指先有流量，在流量无匹配规则的情况下，交换器向控制器请求匹配规则再转发数据，所以整个过程控制器都参与其中，反应模式更智能，这个过程被称为“OpenFlow 转发”。图 3 用一系列的步骤说明这个过程。参照图 3，对比流程步骤可详述如下。

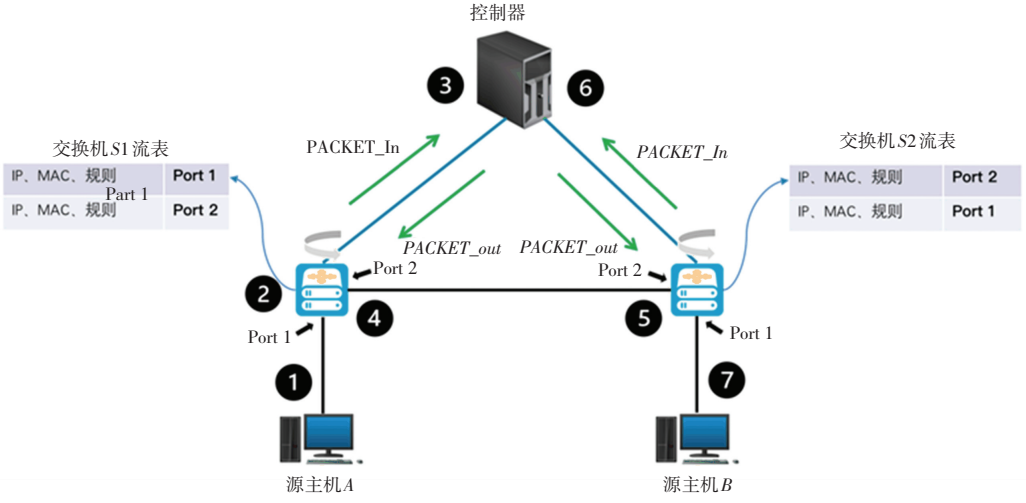

图 3　OpenFlow 协议转发（反应模式）

Fig. 3　OpenFlow protocol forwarding（reactive mode）

（1）源主机“A”将数据包转发到交换机（S1）的通信端口 port1。

（2）在收到数据包后，交换机在其流表中进行查找该数据包的匹配规则。如果数据包在交换机（*S*1）中没有匹配的流条目，那么根据默认，交换机通过 OpenFlow（南向）接口把 *OF_PACKET_IN* 消息转发给控制器。

（3）控制器收到 *OF_PACKET_IN* 消息后，根据自定义程序 ，把匹配规则 *OF_PACKET_OUT* 消息发回给交换机（*S*1）。

（4）交换机（*S*1）更新流表，将数据包从指定端口转发到下一个节点交换机（*S*2）。

（5）同样，交换机（*S*2）收到数据包后，也会在流表中进行查找，如果没有找到匹配规则，也会把 OF_PACKET_IN 消息转发给控制器。

（6）控制器根据自定义程序向交换机（*S*2）回复 *OF_PACKET_OUT* 消息。交换机（*S*2）更新流表后，按规则把数据从 port2 送到 port1。

（7）源主机“B”收到数据包。

（8）源主机“A”到“B”的数据交换过程都是由控制器和交换机完成，数据包前后各字段不做改变，除非是网络发生拓扑变化，对应目的地址、MAC 发

生变化。

反应模式虽然更智能,但是缺点是 OpenFlow 交换机及控制器极易受到 DDoS 攻击。比如应用层 DDoS 攻击中会使用大量欺骗性 *OF_PACKET_IN* 报文占用交换机和控制器之间的 OpenFlow 通道的带宽,降低回复效率,造成合法的规则匹配请求被拒绝。

2 基于 SDN 架构的应用层 DDoS 攻击的分类方法

本节对在 SDN 架构下的应用层 DDoS 攻击进行分类,并说明了 SDN 架构的各种组件遭受攻击的原因。研究可知,分类则包括针对交换机漏洞的分类,以及按不同的攻击类型进行分类。

基于 SDN 架构的 DDos 攻击分类如图 4 所示。由图 4 可知,基于 SDN 架构的 DDos 攻击主要包括:攻击数据平面的交换机;攻击交换机横向数据通道;攻击交换机的控制单元;攻击交换机的流表;攻击交换机的数据包缓冲器;攻击南向 OpenFlow 接口;攻击 SDN 控制器 ;攻击控制器横向接口;攻击北向 API 接口。

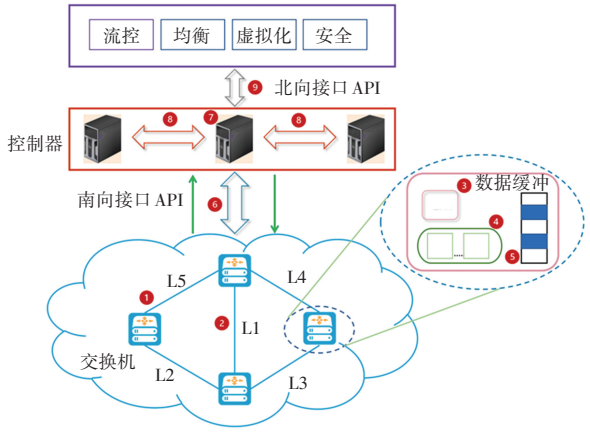

图 4 基于 SDN 架构的 DDos 攻击分类

Fig. 4 Classification of DDos attacks based on SDN architecture

2.1 按交换机漏洞分类

(1)控制单元过载。交换机的转发由控制器控制,这依赖于交换机控制单元。因为交换机控制单元处理和转发数据包的数量是有限的,当 DDoS 攻击发送 *Packet_In* 消息的速度超过控制单元可以转发到控制器的速度,那么就会造成控制单元过载。所以一旦泛洪,交换机的整体性能就会下降,甚至当机。

(2)数据包缓冲区溢出。当交换机收到新数据包时,就会加载到数据包缓冲区,而后使用Packet_In message 将数据包的头转发给控制器。在 DDoS 攻击下,缓冲区很快就会溢出,根据 OpenFlow 协议,ofp_action_output 中的 max_len 字段要转换成 OFPCML_NO_BUFFER,因此交换机必须将数据包完整地转发给控制器,由此南向接口会产生大量的数据包,造成控制信道的带宽和控制器的资源枯竭,使控制器的匹配规则无法及时送达,造成交换机处理数据转发时间增加,最后引起合法用户丢包。

(3)流条目持续时间。OpenFlow 交换机对每个流表都会设置流条目的超时机制、即流条目的在交换机内的持续时间。当 *idle_timeout* 非零时,如果没有收到流量,流条目会在指定的 *idle_timeout* 值之后过期;当 *hard_timeout* 为非零时,流条目在指定 *hard_time out* 值之后过期,与入口的数据包是否到达无关。有一种应用层 DDos 攻击以最小的持续时间连续发送攻击流量,造成交换机流条目超时溢出,合法的流表被覆盖,使交换机数据包转发失败。

2.2 按攻击类型分类

应用层 DDoS 攻击主要分为 2 种类型:带宽饱和攻击和资源饱和攻击。其中,带宽饱和攻击的目的是通过发送大量的欺骗性数据包,消耗其通道的带宽能力,攻击 SDN 架构的控制通道(南向 API)、交换机横向数据通道、控制器横向通道和北向通道。另外,资源饱和攻击消耗的是 SDN 网络设备的性能资源,如 CPU、内存。

(1)交换机的资源饱和度。交换机是转发设备,使用 OpenFlow 通道与控制器进行通信。OpenFlow 交换机最多支持几百到几千个流条目。与控制器每秒可处理的流量请求数量相比,交换机的处理能力也是一个瓶颈。DDoS 攻击者可以利用流条目的反应性规则安装机制瞬间使交换机流条目达到峰值。

(2)控制器的资源饱和度。因为控制器的包处理能力远高于交换机,所以黑客往往优先攻击交换机,当交换机的数据包缓冲区溢出,因为 OFPCML_NO_BUFFER 的原因,同步攻击控制器,就会大量消耗控制器的处理能力(CPU)和物理内存(RAM)。当主控制器的资源耗尽出现问题,就会影响整个网络,造成高延迟和长响应时间,使合法的网络服务完全退化和不可用。现在成熟的 SDN 网络都会部署多个控制器做负载均衡,从而提高控制器的处理能力。

(3)控制通道饱和度。南向 API,也被称为“控制通道”,用于连接交换机和控制器。通过

OpenFlow 协议在交换机和控制器之间提供了一个接口。控制器实时与交换机保持连接，为交换机提供路由和控制网络流量（Qos）的决策。黑客可以利用 OpenFlow 在数据平面和控制平面之间的这种可扩展性缺陷，发起基于新流表的分布式拒绝服务攻击，也就是通过发送大量的欺骗性 IP 地址攻击数据包，使交换机和控制器之间的 OpenFlow 协议接口饱和，导致控制器无法控制交换机。

（4）交换机横向数据通道饱和度。交换机横向数据通道是指 2 个以上 OpenFlow 交换机之间的通信链接。是在交换机之间转发网络流量，受到 DDos 攻击后，交换机之间的一些数据通道或链接将被恶意数据包占据，那么交换机之间将中断连接。一旦数据通道存在瓶颈，交换机根本无法转发任何数据包，这种情况也会造成整个网络的瘫痪。

（5）控制器横向通道饱和度。在 SDN 架构中，横向通道指的多个控制器之间的接口通道。当单一的集中式控制器由于网络交换机数量的增加而无法处理网络流量时，多个控制器通过横向绑定的 API 接口提供负载均衡以保证整个网络可靠性，DDos 可以从任何一个方向攻击控制器，导致横向通道的接口被大量进入的攻击数据包所占用，控制器如果无法及时解决负载问题，也会造成整个网络瘫痪。

（6）北向通道饱和度。控制器依靠可编程应用的北向 API 来保证整个网络系统灵活度。软件开发人员使用这个接口对网络控制实现可编程。与南向 API 不同，每个平台都有自己的北向 API，缺少统一的标准。这种标准化的缺失也是一种安全威胁，所以黑客也可以通过北向 API 针对控制器进行攻击，导致北向接口的拥堵。

3　基于 SDN 架构的 DDOS 攻击检测技术研究

SDN 网络一直受到 DDoS 攻击，研究者掌握了许多检测 DDoS 攻击的方法：

（1）一种基于时间特征的 DDoS 攻击检测方法，提取攻击的时间并记录，使用时间特征快速有效地检测和防御 DDoS 攻击。有的黑客利用控制器处理新网络数据包的反应时间差，在这个时间窗口内向控制器发送大量的请求，从而对 SDN 控制器发起攻击，利用这个特点，可以迅速发现此类攻击并定位到源 IP 进行有效阻断，这个检测方法的缺点是仅针对长快频的攻击有效，对低速低频攻击无效。

（2）一种过滤请求的新方法可以通过 OpenFlow 主动模式将所有的新数据包直接发送到安全网关，而不是使用控制器来降低熵值（参见图 5），通过熵值法来检测 DDoS 攻击，及时生成新规则更新交换机流表，交换机对符合规则的源目地址做丢包处理，这种方法必须抓取 3 个特征值：协议、源 IP 地址和目的 IP 地址。该检测方法需要消耗时间来处理新的数据流，同时也会消耗系统性能。

（3）在 OpenFlow 中采用基于熵的轻量级 DDoS 泛滥攻击检测方法，减少对控制器的流量收集负载，减轻控制器因为频繁的流量收集带来性能消耗，使交换机更智能地主动检测交换机上的 DDoS 攻击，这种方式缺点就是会降低控制器和交换机之间的通信频率。

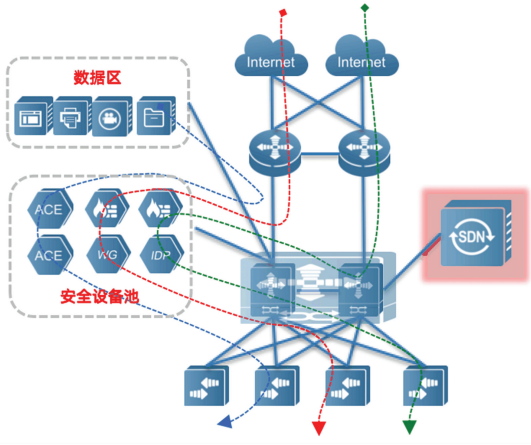

图 5　主动模式下的 DDos 防御

Fig. 5　DDos defence in active mode

4　Mininet 创建模拟检测环境

4.1　Mininet

Mininet 是一个基于 Linux 的轻量级虚拟化工具，本文使用 Mininet 快速创建 SDN 网络模型（见图 6），该软件可以虚拟添加交换机、主机和控制器，也可以改变或修改网络结构和连接。创建步骤如下：

（1）将 Mininet 安装在 Ubuntu Linux 上。

（2）命令行：

mn -topo single,3 -mac -switch ovsk -controller remote

创建 3 个虚拟主机，在内核中创建 1 个具有 3 个端口的 OpenFlow 交换机，为每个主机设置 MAC 和 IP 地址，配置交换机连接到控制器。控制器在本地运行，与模拟器 Mininet 所运行的硬件相同。

（3）命令行：

dpctl dump-flows tcp:127.0.0.1:6634

连接到交换机并显示安装的流量表。

（4）命令行：

dpctl add-flow tcp:127.0.0.1:6634 in_port=0, actions=output:1

创建一个规则,所有到达交换机端口 0 的数据包将被转发到端口 1。

(5)安装开源的 Floodlight OpenFlow 控制器。

(6)安装网络分析软件 sFlowRT,该软件内嵌在 OpenFlow 控制器中,可以实时监控 OpenFlow 交换机和控制器之间的通信流量。

(7)安装 sFlow-RT 分析器的命令:

sudo mn-controller=remote,ip=172.0.0.1,port=6653-topo = single,3。

(8)连接 OpenFlow 到 sFlow-RT 分析器的命令:

sudo ovs-vsctl--id=@ sflow create sflow agent=eth0 target= 172.0.0.1:6643,ampling=10 polling=20--set bridge s1 sflow=@ sflow,设置完成后,手动启动 sFlow 分析器,通过./sFlow-rt/start 命令收集的数据样本,同时设置访问地址:http://localhost:8080/ui/pages/index.html。

(9)安装 WEKA 机器学习软件。

(10)测试期间控制器不可人为断开,否则测试将失败。

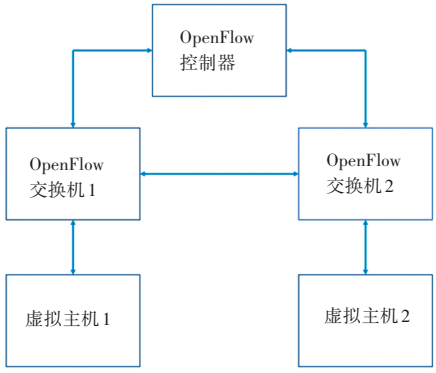

图 6 简单 SDN 网络模型

Fig. 6 Simple SDN network model

4.2 采集数据

本次测试使用了 2 台主机、2 台 OpenFlow 交换机和 1 个控制器。WEKA 是一款开源的机器学习以及数据挖掘软件,该软件可以在同一数据集上建立多个机器学习模型,本文使用 WEKA 机器学习软件建立模型并测试模型的准确性,同时对正常和 Dos 攻击的数据建立数据集,包括 TCP、UDP、ICMP、ARP、IPv4 和 SSH 等 6 种协议。在 DDos 攻击期间,1 台主机作为受害者,1 台主机作为攻击者,主动采集每个交换机及控制器上的流量情况。正常情况下控制器上数据流量如图 7 所示,DDos 环境下控制器数据流量如图 8 所示。

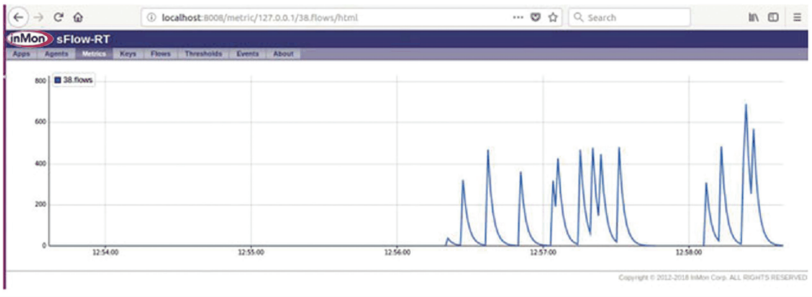

图 7 正常情况下控制器上数据流量

Fig. 7 Data traffic on the controller under normal conditions

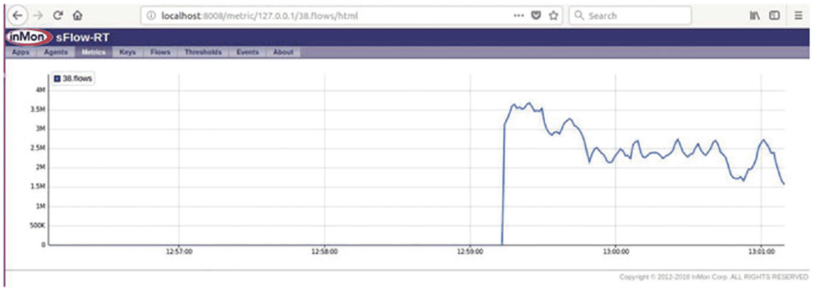

图 8 DDos 环境下控制器数据流量

Fig. 8 Controller data traffic in the DDos environment

5　AdaBoost 机器学习算法

由于数据集实例不是很大，为了减少偏差，在数据集中加入了 25% 的噪声数据。本文使用 AdaBoosting 机器学习算法，将决策树作为一个弱分类器来建立网络的分类器模型。

5.1　数据分类

为了降低复杂性，减少了特征的数量，使用递归特征消除（RFE）进行递归删除特征，并在剩余的特征上构建模型。RFE 根据数据集中所有特征在实例分类中对重要特征进行排序。RFE 算法需要 2 个参数，一是要保留的特征数量，二是在评估特征重要性的过程中要使用的模型，经测试选择的特征数量最终确定为 7 个，分别是：目的地端口；Bwd 包长度平均值；Bwd 包长度标准；Bwd 包/s；包长度平均值；最大包长度；平均 Bwd 段大小。

5.2　建立分类模型

（1）AdaBoost 是自适应模型，这个模型从弱分类器中自我学习，性能随着后续的分类器而提高。分类方程为：

$$f(x) = sign\left(\sum_{m=1}^{m} (\theta mfm(x))\right) \tag{1}$$

（2）决策树评估了特征的重要性和准确性。推得的公式为：

$$RFfi_i = \frac{\sum_j norm\, f_{ij}}{\sum_{j \in allfeatures, k \in alltrees} norm\, fi_{jk}} \tag{2}$$

（3）多层感知器模型。模型公式可写为：

$$y = \varphi\left(\sum_{i=1}^{n} w_i x_i + b\right) = \varphi(w^t x + b) \tag{3}$$

5.3　建立后流程

模型建立后，该模型与 SDN 控制器相连接，控制器只转发那些被分类器模型过滤的流量，涉及 2 次网络流量过滤，第一次由 SDN 控制器本身执行，利用自定义网络配置过滤掉普通的攻击流量；第二次，由 SDN 控制器通过机器学习分类器再次过滤掉攻击流量，该分类器只通过机器学习算法中被标记为良性的数据流量。WEKA 数据集概述见图 9。

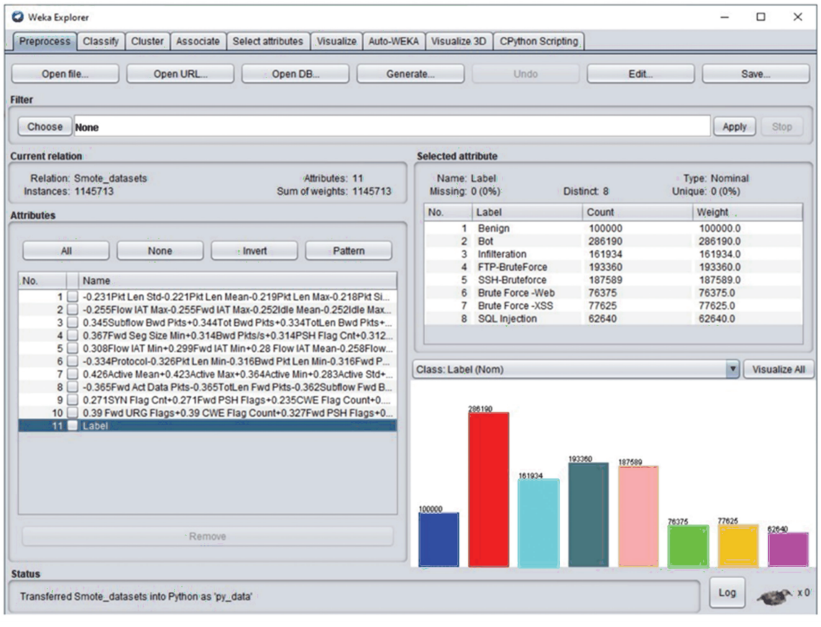

图 9　WEKA 数据集概述

Fig. 9　WEKA dataset overview

5.4　指标分析

（1）真阳性（*TP*）：攻击时产生的异常数据；假阳性（*FP*）：将正常数据分类为攻击；假阴性（*FN*）：异常数据被分类为正常数据；真阴性（*TN*）：被分类为正常的数据。

（2）指标定义及计算公式

① 精度。是指正确分类为攻击的数据与分类为攻击的总数据之比。具体数学公式为：

$$\text{精度} = \frac{TP}{TP + FP} \tag{4}$$

② 召回。是指正确分类数据的比率。具体数学公式为：

$$\text{召回} = \frac{TP}{TP + FN} \tag{5}$$

③ F - 衡量。是指精确度和召回率的加权平均值。具体数学公式为:

$$F\text{ - 衡量} = 2 \times \frac{\text{召回} \times \text{精度}}{\text{召回} + \text{精度}} \tag{6}$$

(3)对多个机器学习模型计算得出结果进行比较,AdaBoost 精度相对更准确些,多层感知器模型,属于人工神经网络,但是执行效率非常低,见表 1。

表 1 不同机器学习技术的性能指标分析

Tab. 1 Analysis of performance metrics of different machine learning techniques

模型	精度	召回	F - 衡量
多层感知器	0.833	0.833	0.833
AdaBoost (以决策树作为弱分类器)	0.943	0.943	0.943
随机森林	0.839	0.835	0.835
J48 决策树	0.912	0.911	0.900

6 检测 SDN 中 DDoS 攻击遇到的困难

综上所述,SDN 将控制平面与数据平面分离,使管理更加便捷,同时实现可编程[4-6]。尽管国内外学者在 SDN 环境下针对 DDoS 攻击方面的研究有了很多成果,但 DDoS 攻击面仍在扩大,而且技术更新迭代很快。以下是 SDN 环境下检测 DDoS 攻击所面临一些实际困难。

6.1 数据统计

大多数 DDoS 攻击检测技术需要从 OpenFlow 交换机中收集数据构建规则,例如提取数据报头的特征来检测异常行为的方法。但是在低速率 DDoS 攻击时,从流量中收集统计数据就很困难,另外采用负载均衡技术用多个分布式交换机来收集数据,所收集的数据精度达不到要求,同时收集难度也会加大。

6.2 算法选择

DDoS 攻击行为的多样化使 SDN 环境中的异常流量检测变得复杂。因此,许多算法已经转向人工神经网络、贝叶斯分类法、模糊逻辑等来检测 DDoS 攻击行为。但是没有一种算法能够真正应对所有 DDoS 攻击。

6.3 响应速度

及时响应是 SDN 控制器的关键,被 DDoS 攻击后,控制器要处理大量的流量,这会耗尽其性能,从而削弱响应合法用户请求的能力。现有的 DDoS 攻击检测方法存在许多问题,包括控制器在短时间内处理大量入口数据包的响应速度;无法检测低速率的 DDoS 攻击;高网络带宽的消耗;无效的数据包带来的处理负担等都会导致攻击检测的延迟或无效。

7 结束语

本文使用 Adaboost 和决策树作为弱分类器,发现数据集的 DDoS 攻击检测准确率可以达到 94%。用机器学习模型的优点是 SDN 控制器通过自定义规则,利用分类器的输出,可以有效阻断那些特定的攻击类型,在一定程度上增加了安全性。但是随着 DDoS 攻击的复杂程度不断提高,尤其针对数据中心和网络基础设施的攻击级别不断提高,基于应用层的 DDoS 攻击给 SDN 架构带来了许多安全挑战和威胁,研究人员希望使用一种检测防御方法来解决所有的 DDos 的攻击问题显然是不可能的。个人认为 SDN 中集中控制的特性应该是 DDoS 攻击的核心,而分布式控制器的设计可以提供更好的负载分配、处理能力和可靠性,最大限度地降低因为单个控制器的通信故障而造成整网瘫痪的风险,同时机器学习在未来的网络应用中具有很大的研究潜力,如何提高检测精度,使检测过程完全自动化,减少人为干预也是未来的研究方向。

参考文献

[1] 张凯. 基于机器学习的数据中心网络建模与优化算法研究[D]. 南京:东南大学,2021.
[2] 左志斌. 基于密码标识的软件定义网络数据面安全关键技术研究[D]. 郑州:战略支援部队信息工程大学,2020.
[3] 曾志豪. 基于多控制器架构下的 SDN 网络关键技术研究[D]. 成都:电子科技大学,2022.
[4] 周颖超. 面向确定性时延应用的资源管理研究与实现 [D]. 北京:北京邮电大学,2021.
[5] 周佳敏,高泽华,马蓁,等. 基于 OpenFlow 的 SDN 扩展方案研究[J]. 数据通信, 2017(01):1-4,18.
[6] 李保星. SDN 环境下的 L7 路由选择方案研究与实现 [D]. 北京:北京邮电大学,2021.