

一种面向软件定义网络的安全态势感知方法

郑忠斌¹, 黎 聪², 王朝栋¹

(1. 工业互联网创新中心(上海)有限公司, 上海 201303;

2. 同济大学, 上海 200092)

摘要:随着互联网基础设施的飞速发展和新应用的不断涌现,拒绝服务攻击作为网络层攻击的一个典型始终是安全人员重点防范的对象,而新兴的SDN网络在控制层和基础设施层同样也有受到拒绝服务攻击的风险。在现有的安全技术中,网络安全态势感知技术独特的优势使之能同时有效地用于传统网络与SDN网络。研究了应用于网络安全态势感知的相关算法,并实现了一个基于JDL多传感器数据融合的网络安全态势感知模型,并将其应用于SDN网络中对拒绝服务攻击的感知和评估。实验结果显示与单纯使用IDS的方法相比,融合决策的误报率和漏报率均有所下降,且输出的态势值具有一定的准确度,符合系统安全管理人员直观评估的结果。

关键词:网络安全;态势感知;软件定义网络;数据融合;拒绝服务攻击

中图分类号:TP393

文献标识码:A

DOI: 10.19358/j.issn.2096-5133.2020.04.002

引用格式:郑忠斌,黎聪,王朝栋.一种面向软件定义网络的安全态势感知方法[J].信息技术与网络安全,2020,39(5):5-12.

A security situation awareness method for software defined network

Zheng Zhongbin¹, Li Cong², Wang Chaodong¹

(1. Industrial Internet Innovation Center (Shanghai) Co., Ltd., Shanghai 201303, China; 2. Tongji University, Shanghai 200092, China)

Abstract: With the rapid development of Internet infrastructure and the constant emergence of new applications, denial-of-service attack as a typical example of network layer attack is always the object that security personnel focus on preventing. Emerging as it is, SDN network also risks of getting denial-of-service attack in its control layer and infrastructure layer. Among the current security technologies, the unique advantages of Network Security Situational Awareness technology have rendered NSSA competitive in applying both to traditional network model and SDN. Hence this paper studies the application of algorithms in the network security situational awareness, and implements a JDL-multi-sensor-data-fusion-based network security situational awareness model, applied to SDN network perception and evaluation of Denial of Service attacks. The experimental results show that the false alarm rate (negative positive) and non-response rate (positive negative) of fusion decision have both remarkably decreased compared with the IDS method, and the situational value of the output is of certain accuracy, which is in line with the intuitive evaluation results of system administrator.

Key words: network security; situational awareness; software defined network; data fusion; denial of service attacks

0 引言

随着互联网基础设施的飞速发展和新应用的不断涌现,网络在规模和拓扑上都日趋扩大化、复杂化,各种层出不穷、更新换代的网络攻击给安全管理者带来了巨大的挑战。在诸多网络攻击手段中,拒绝服务攻击作为历史最为久远、造成财产损失最为严重的攻击手段之一,始终是政企以及军方网络安全管理人员重点防范的对象。自1999年第一次分布式拒绝服务攻击(Distributed Denial of

Service, DDoS)出现以来,DDoS攻击经历了探索期、组织攻击期、国家网络战期几个阶段,直至今天已成为高度普及化、成熟化、组织化的攻击方式。从2008年至今,智能终端设备的全面普及产生的大量的僵尸网络更是为不法分子开展DDoS攻击提供了便利。2018年3月出现的MemcacheUDP以TB级的带宽攻击了Github,引发了网络安全从业者的警惕,意味着目前的DDoS攻击制造出TB乃至PB级的攻击已不是难事。鉴于目前的服务器的带宽

压力承载量大多不足以应对此种级别的攻击,一次精心预谋的 DDoS 攻击将不仅会造成网络资产损失,也会造成企业乃至国家政府机构的职能瘫痪,产生的后果难以估量。

另一方面,日益臃肿的网络架构催生了新型网络结构“软件定义网络”(Software Defined Network, SDN)的广泛应用。SDN 将发包过程和路由过程分隔在数据层(Data Plane)和控制层(Control Plane),并通过南北向接口为开发者提供了良好的可编程性。其目的是为了简化日益复杂的网络软硬件结构模式,并降低传统网络的控制与全局部署的难度,从而解决传统网络的拓扑复杂、去中心化和难以定位故障的缺陷。

然而这并不意味着 SDN 与传统的 TCP 五层模型相比就具有更好的安全性,由于转发机制的不同,许多应用于传统网络架构的较为成熟的安全管理手段并不能直接套用于 SDN 网络。在目前 SDN 大规模商业部署的背景下,针对 SDN 网络安全管理技术的研究意义重大。SDN 的安全管理者需要运用有效的安全管理手段,对安全威胁和发生的攻击做出及时、精确、有效的响应与决策。从现有的安全管理手段上看,脆弱性检测技术、恶意代码检测技术、入侵检测技术都试图从不同的角度发现、理解并向管理员报告网络中可能存在的安全问题。尽管这些方法各有所长,但是从实时全面地应对威胁的角度上看,均不能充分地满足系统管理员全方位、直观化的需求。

相较于上述的几种安全管理技术,网络安全态势感知技术独特的优势使之能同时有效地用于传统网络或 SDN 网络:相较于单传感器和单数据源的 IDS,一个合格的网络态势感知系统能将不同安全传感器(如 IDS、交换机日志、系统日志、原始流量数据收集器等)中收集到的异质信息进行融合,网络中的安全要素知识进行综合的分析,以合理的指标评估网络安全状况或者对未来的状况做出预测,并以可视化的方式将网络全局的安全状况展现给管理人员,使其能够采取有效的应对措施对安全问题进行及时响应^[1-3]。因此本文选择了网络安全态势感知技术作为 SDN 网络中拒绝服务攻击的解决方案。

本文研究了应用于网络安全态势感知相关算法,实现了一个基于 JDL 数据融合的网络安全态势感知模型,并应用于 SDN 网络中对拒绝服务攻击

(Denial of Service, DoS)的感知和评估。

(1) 研究了几种网络安全态势感知系统的功能模型与算法,比较了其中的数据融合与态势评估算法、态势预测算法的理论适用场景。

(2) 基于(1),针对态势感知模型中的核心模块数据融合和态势评估部分,选择了 SVM 向量机作为主要训练算法,贝叶斯方法作为融合方法,层次化量化评估作为评估方法,基于 SDN 网络中的模拟攻击场景设计并实现了一个基于多传感器融合的网络态势感知模型。实验结果显示与单纯使用 IDS 的方法相比融合决策的性能有所提升,且输出的态势值具有一定的准确度,较吻合安全管理人员直观评估的结果。

本文首先介绍了相关工作已有的研究成果;然后阐述了与本文设计方法直接相关的背景知识;接着详细描述了具体的 SDN 受到 DDoS 攻击这一场景下产生训练数据集,实现数据融合与态势评估的实验设计思路,具体配置与实验流程以及结果分析;最后阐述了实验设计的结论。

1 相关工作

1.1 针对多传感器融合态势感知的研究

盖伟麟^[4]等对态势感知中的数据融合和决策方法作了综述,给出用于态势感知的算法分类和决策方式分类。赵耀南^[5]介绍了针对网络流量数据优化的自离散化算法与常见的连续值离散化算法。刘效武等^[6-7]提出了“融合引擎”的概念并使用支持向量机作为融合引擎,引入了态势评估符合度的指标来表征评估的态势值与实际态势值的符合程度。李志东^[8]针对多源数据融合存储开销较高的问题,提出了基于统计空间映射的多源告警融合决策模型。Wang Huiqiang^[9]等使用了一个多层前馈神经网络,对 Snort 和 NetFlow 收集到的真实流量数据进行融合。在特征约简方面使用了较为简单快速的方法。在态势评估采用了权值因子分布的预警聚合算法来计算态势值。

1.2 针对一般网络态势感知算法和指标的研究

诸葛建伟等^[10]详述了 D-S 证据理论应用于网络异常检查的方法,并主要针对 DDoS 攻击在保证较低误报率的前提下,达到了较高的检测率。陈秀真等^[11]基于 IDS 的报警信息,结合主机的重要性和网络的组织结构,对服务和主机的重要性进行加权和计算,建立了层次化安全威胁态势量化评估模

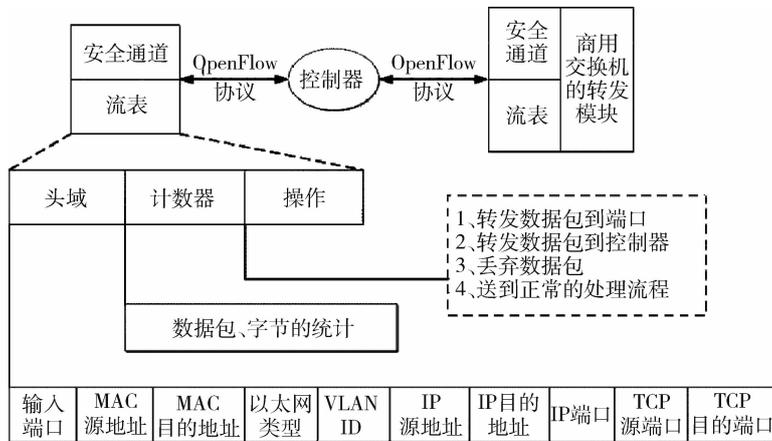


图2 Openflow 协议

提供所需服务或者使服务质量降低。而如果处于不同位置的多个攻击者同时向一个或多个目标发起拒绝服务攻击,或者一个或多个攻击者控制了位于不同位置的多台机器并利用这些机器对受害者同时实施 DoS,这种 DoS 则成为了更为常见、危害更大的 DDoS,又称带宽攻击或风暴型 DoS 攻击。

DDoS 可分为直接风暴型和间接风暴型两种。本文实验中采取的 DoS 攻击方式是直接型的^[19]。

2.4 基于决策层融合和数学模型评估的态势感知模型

如前所述,多传感器数据融合是针对一个系统中使用多个(同质)以及多种传感器这一而提出的处理方法,目的是基于各个独立传感器的观测数据,运用数据融合算法导出更为精准的有效信息,获得最佳的协同效果,并一定程度上消除单传感器的局限性与不可靠性。基于相关工作和模拟环境的综合考量,在实验中选择传感器为处于 HIDS 模式的 Snort、OVS 交换机流表信息以及通用流量监控工具 sFlow-rt 以实现基于多源融合的态势感知。选用的数据融合模型为 JDL 模型,由于态势感知工程较为庞大,在实验中选择性地实现了决策层数据融合与态势评估两个核心模块。

2.4.1 数据预处理

选择了 Snort-IDS、OVS 流表信息和 sFlow-rt 作为多源传感器融合中的传感器。

(1) Snort-IDS

Snort 是一个开源的入侵检测系统,可进行实时的流量分析和数据包捕获与日志记录;一般地,部

署 IDS 有基于主机(Host-based IDS)与基于交换机两种方式,本实验中使用 Snort 的 IDS 模式。

(2) OVS 流表信息

流表在 SDN 架构中类似于现今传统网络路由器中的路由表,由控制层中的控制器下发给转发层中的物理设备如交换机等。OpenFlow 流表包括了包头域(Header)、计数器(Counter)、动作(Action)。通常提取 OVS 流表信息有两种方法:调用 SDN 控制器(OpenDayLight, Ryu 等)提供的北向 API 获取流表信息,或用 `ovsctl` 命令在控制台中显示出即时流表信息并重定向到文件中。

(3) sFlow-rt

sFlow-rt 是基于 sFlow 协议开发的可视化网络流量检测应用。sFlow 监控工具 sFlow-rt 由 sFlow Agent 和 sFlow Collector 两部分组成。Agent 作为客户端,一般内嵌于网络转发设备(如交换机、路由器),通过获取本设备上的接口统计信息和数据信息,将信息封装成 sFlow 报文,当 sFlow 报文缓冲区满或是在 sFlow 报文缓存时间超时后,Agent 将 sFlow 报文发送到指定的 Collector。Collector 作为远端服务器,负责对 sFlow 报文分析、汇总、生成流量报告。

2.4.2 融合过程

如图 3 所示,经过数据采集和预处理后存在原始数据库的数据分为 Snort、OVS 和 sFlow 三个部分。经过对机器算力和决策精度的综合考量,数据融合算法选择了决策层融合,由于 OVS 与 sFlow 得到的特征属于多分类问题,而标准的 SVM 算法适用于二分类问题,此处选择了一对多法多分类 SVM 作为这传感器特征数据的训练模型。

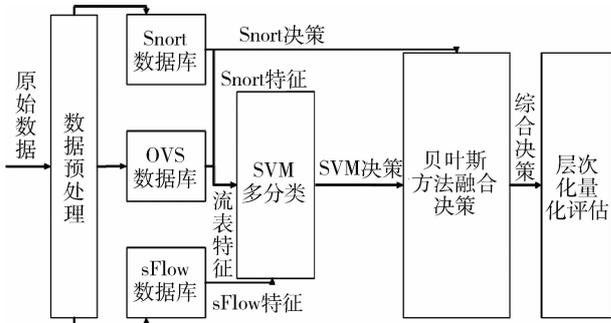


图3 实验模型设计图

在特征约简方面采用了刘效武等^[6]的迭代训练约简法,即将设原始特征向量维度为 n ,则每次去除特征 $c_i (1 \leq i \leq n)$ 使用 $n-1$ 个向量输入 SVM 进行训练。若此轮训练中的精度 A 小于人为设定的阈值 ε ,则将特征 c_i 加入约简集 Selection_set 中。进行 n 轮训练后得到最终的向量集合 Selection_set。

最终,由 Snort 数据的特征部分、OVS、sFlow 的特征经过特征选择训练得出的决策,与 SnortIDS 得出的决策再进行决策层融合,采用贝叶斯方法,得到综合的决策值。

2.4.3 评估过程

综合得出的决策其输入层次化量化评估模型,

分别得到网络级、主机级、服务级的态势评估值 V_1 、 V_2 、 V_3 。规定每一级态势符合度 SF_i 以及平均态势符合度 $SF_{average}$ 满足

$$SF_i = PF_i / AF_i, i = 1, 2, 3 \quad (1)$$

$$SF_{average} = (SF_1 + SF_2 + SF_3) / 3 \quad (2)$$

其中 PF_i 是通过决策得出的分类结果计算出某一级的态势值, AF_i 是由已知样本计算出的真实态势值。

3 实验部分

3.1 实验配置

实验软硬件配置如表 2 所示,网络拓扑配置如图 4 所示。

表 2 实验软硬件配置表

配置项目	名称与版本
操作系统	Ubuntu 16.04 LTS
模拟 SDN 网络	Mininet2.2
SDN 控制器	OpenDayLight
IDS	Snort-2.9.9
IDS 规则库	Snort-2.9.9.rules(部分)
IDS 连接数据库插件	Baynyard2
数据库管理系统	MySQL5.7 + MySQLWorkbench

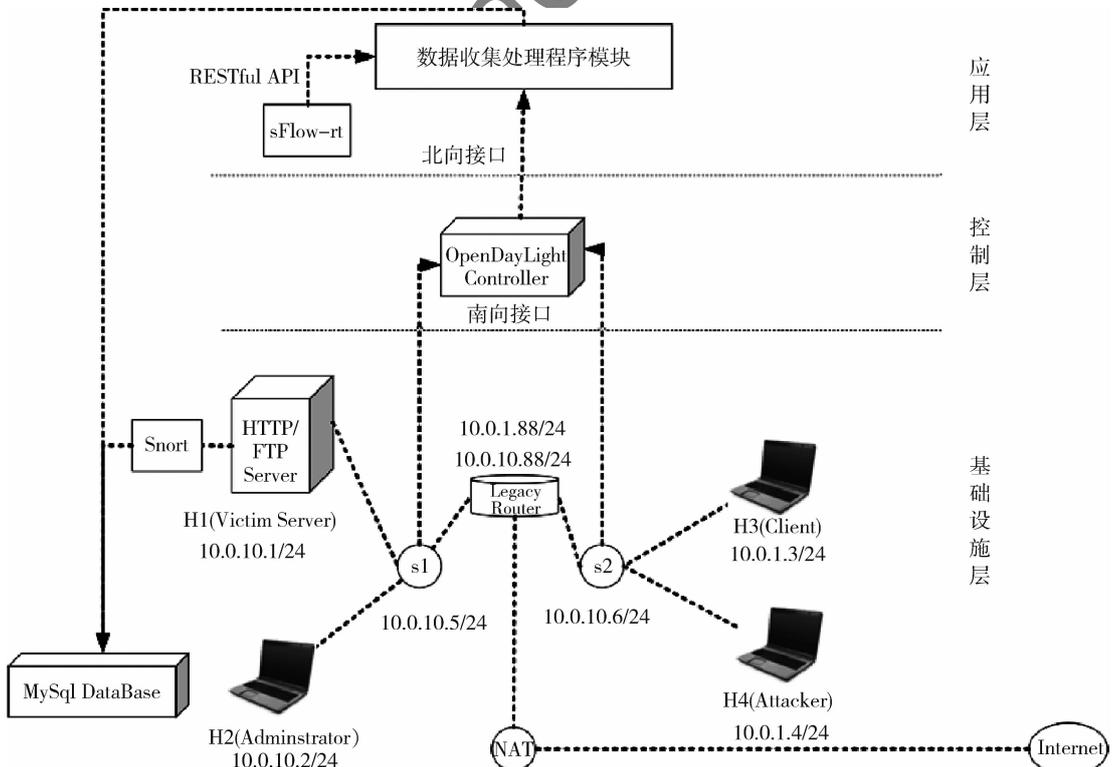


图4 网络拓扑图

3.2 实验流程

3.2.1 数据集获取方式

目前能获取到的公开数据集最大的缺陷在于其都是基于传统网络中交换机、路由器产生的数据与基于 SDN 网络与 SDN 控制器、交换机的数据有所不同。所以,本文采取了通过在 Mininet 网络中用网络工具模拟攻击,自主生成训练数据集的方式。

3.2.2 模拟攻击过程

在 24 h 的时间内,h1 Server 始终开启服务,用网络工具 hping3、LOIC(低轨道离子炮)模拟 DoS/DDoS 攻击的过程。同时,内网 client 通过 wget 等命令在随机的时间点访问 h1 server 产生内网背景流量,整个内网通过 nat 服务连接因特网产生外网背景流量。

由于运行 mininet 的性能限制,发送过多的包将导致机器崩溃,每次 DDoS 的时长与真实攻击相比均较短,同时也对发包数量进行了限制。攻击时段表如表 3 所示。

表 3 攻击时段表

攻击种类	发动次数	平均攻击时长/min	攻击时段/h
TCP SYN Flood	15	5 ~ 15	0 ~ 6
UDP Flood	15	5	6 ~ 12
ICMP Flood	5	5 ~ 15	12 ~ 18
HTTP DoS	10	5	18 ~ 24

在去除了无效的数据之后,得到可供训练的样本数统计如表 4 所示。

表 4 样本统计

样本	总计
正常	38 518
ICMP DoS	2 533
TCP SYN DoS	6 597
UDP DoS	5 498
HTTP DoS	1 873
总计	55 019

3.2.3 Snort 的决策层数据

事实上 Snort 的目前分析可疑数据包的协议范围有一定的局限性,限于 TCP、UDP、ICMP、IP 四种,没有 ARP 等协议的记录。但对于本研究中的 DoS 攻击类别四种来说已经足够。Snort 的 event 表中的

签名字段和时间戳字段标识了 Snort 决策得到的当前报警信息。将在其后的步骤中输入决策层融合模块。

3.2.4 特征训练

Snort 的特征数据、OVS 流表数据、sFlow 提取后将相应字段和数据存入数据库。从 Snort-HIDS 中取得的数据和从 ODL、sFlow 的 RESTful API 中得到的数据存在冗余或无用数据,如果将原始数据直接导入作为输入特征值将影响训练的精度,一些与网络层数据无关的特征在最终的数据集选择中去除了之后,剩余的特征约简后得到输入向量中的关键特征集 Selection_set 共 10 个向量,如表 5 所示。

表 5 用于训练的特征

特征序号	字段名	来源	描述
1	FlowAveragePackets	OVS	平均包数目
2	FlowRate	OVS	流速率
3	SourceAddrEnt	sFlow	源地址熵
4	SourcePortEnt	sFlow	源端口熵
5	DstPortEnt	sFlow	目的端口熵
6	SnortAlertNum	Snort	时间窗内 Snort 报警数目
7	tcpPercent	Snort	TCP 包个数占总数百分比
8	udpPercent	Snort	UDP 包个数占总数百分比
9	IcmpPercent	Snort	ICMP 包个数占总数百分比
10	InFlowRate	Snort	单位时间内流入量的最大变化率

使用贝叶斯方法对 Snort 和多分类支持向量机的决策进行融合,融合决策结果如表 6 ~ 8 所示。

表 6 二分类融合决策结果

样本种类	实际正常 AP	实际攻击 AN	总计
识别正常 RP	34 107	2 420	36 527
识别攻击 RN	4 411	14 081	18 942
总计	38 518	16 501	55 019

表 7 二分类统计

准确率 A (Accuracy)/%	精确率 P (Precision)/%	召回率 R (Recall)/%	F1-Score (2PR/(P+R))
87.6	93.3	88.5	0.90

表 8 多分类融合决策结果

样本种类	样本数目	Snort 识别正确样本数	Snort 识别准确率/%	融合后识别正确样本数	融合后识别准确率/%	相对准确率提升/%
正常	38 518	31 515	81.8	34 107	88.5	+6.7
ICMP DoS	2 533	2 037	80.4	2 305	91.0	+10.6
TCP SYN DoS	6 597	5 503	83.4	5 497	83.3	-0.1
UDP DoS	5 498	4 630	84.2	4 709	85.6	+1.4
HTTP DoS	1 873	1 634	87.2	1 670	89.2	+2.0
总计	55 019	45 369	82.5	48 188	87.6	+5.1

由表 6 可观察到融合决策二分类的 F1-score 达到了 0.90,从实际应用的角度来说属于可以接受的范畴。由表 7 可观察到精确率比准确率高。由表 8 可观察到,除了对 TCP SYN 的分类性能存在误差容许范围内的下降之外,在使用同样的 Snort 规则集的情况下,与 Snort 单源决策相比,加入了 OVS 流表和 sFlow-rt 信息的 SVM 训练决策结果的总决策效果对不同种类的 DoS 攻击判断都有明显的改善。

3.2.5 态势评估与效果检验

根据评估公式可得到 24 h 的态势评估值。选取了从 h4 发动的攻击从 7:07 开始,7:16 结束。从中取 7:00 ~ 7:30 之间的内绘制出如图 5 所示的网段 10.0.10.0/24 的态势曲线图。可见其基本符合直观预期。同时图 6 中显示了 24 h 内以 1 h 为采样时长计算的平均态势符合程度 $SF_{average}$ 亦符合预期。

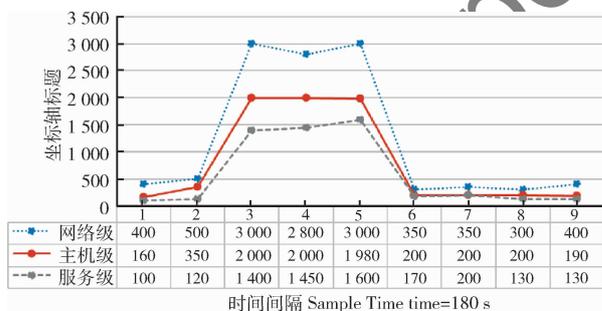


图 5 7:00 ~ 7:30 时段内态势评估折线图

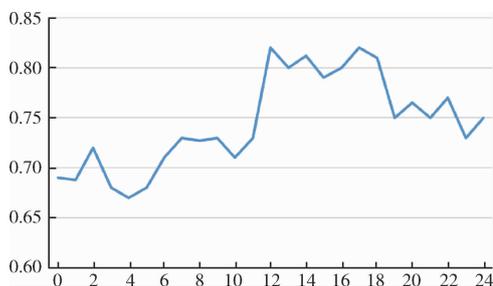


图 6 24 h 内的平均态势符合度 SFaverage 曲线

3.3 与现有方法的对比

由于使用的网络数据集、网络拓扑结构以及选择的评估量化指标不同,本文的分类结果严格说来与文献中的结果不具有横向的可比性,故此不列出。另一方面从实用意义上来说融合模型低于 90% 的分类准确率仍有较大的提升空间。但作为研究与实现的验证性模型来说,本实验目的已经达到。

3.4 实验的不足之处

- (1) 选择的训练算法中 SVM 多分类向量机还应该进行参数优化;
- (2) 对最为常见的 TCP SYN Flood 攻击的分类精度较低,应改进特征选择算法;
- (3) 设置的数据集、攻击方式和网络拓扑过于简单粗糙;
- (4) 限于时间,未实现态势预测算法和实时性可视化模块。

4 结论

本文研究了应用于网络安全态势感知的相关算法,实验验证了在受到 DoS 的情况下,基于 JDL 的多源传感器融合的网络安全态势感知模型做出的决策总体上优于单源传感器决策,且能较为准确地对当前网络、主机、服务的态势做出评估。实现了一个基于 JDL 多传感器数据融合的网络安全态势感知模型,并将其应用于 SDN 网络中对 DoS 的感知和评估。实验结果显示与单纯使用 IDS 的方法相比,融合决策的误报率和漏报率均有所下降,且输出的态势值具有一定的准确度,符合系统安全管理人员直观评估的结果。

参考文献

- [1] JAJODIA S, Liu Peng, SWARUP V, et al. Cyber situational awareness[M]. Springer New York Dordrecht Heidelberg London, 2010.

- [2] 龚俭,臧小东,苏琪,等. 网络安全态势感知综述[J]. 软件学报,2017,28(4):1010-1026.
- [3] 陶源,黄涛,张墨涵,等. 网络安全态势感知关键技术研究及发展趋势分析[J]. 信息安全,2018,18(8):79-85.
- [4] 盖伟麟,辛丹,王璐,等. 态势感知中的数据融合和决策方法综述[J]. 计算机工程,2014,40(5):22-25.
- [5] 赵耀南. 基于多源数据融合的网络安全态势感知技术研究[D]. 北京:北京邮电大学,2016.
- [6] 刘效武,王慧强,禹继国,等. 基于多源融合的网络安全态势感知模型[J]. 解放军理工大学学报,2012,13(4):404-407.
- [7] 刘效武,王慧强,赖积保,等. 基于多源异质融合的网络安全态势生成与评价[J]. 系统仿真学报,2010,22(6):1411-1415.
- [8] 李志东. 基于融合决策的网络安全态势感知技术研究[D]. 哈尔滨:哈尔滨工程大学,2012.
- [9] Wang Huiqiang, Liu Xiaowu, Lai Jibao, et al. Network security situation awareness based on heterogeneous multi-sensor data fusion and neural network[C]. Second International Multisymposium on Computer and Computational Sciences, 2007.
- [10] 诸葛建伟,王大为,陈昱,等. 基于 D-S 证据理论的网络异常检测方法[J]. Journal of Software, 2006, 17(3):463-471.
- [11] 陈秀真,郑庆华,管晓宏,等. 层次化网络安全威胁态势量化评估方法[J]. Journal of Software, 2006, 17(4):885-897.
- [12] 吴建台,乔翌峰,朱赛凡. 基于 HMM 的网络安全态势评估与预测方法[J]. 导航与控制, 2018, 17(2):10-17,31.
- [13] 周长建,司震宇,邢金阁. 基于 DeepLearning 网络态势感知建模方法研究[J]. 东北农业大学学报,2013,44(5):144-149.
- [14] 张勇,谭小彬,崔孝林. 基于 Markov 博弈模型的网络安全态势感知方法[J]. Journal of Software, 2011, 22(3):495-508.
- [15] 张玉杰. 基于 SDN 的网络流量异常检测模型设计和实现[D]. 郑州:河南大学,2016.
- [16] 李传煌,吴艳,钱正哲. SDN 下基于深度学习混合模型的 DDoS 攻击检测与防御[J]. 通信学报,2018,39(7):177-187.
- [17] 李可. 基于 SDN 的网络安全态势感知关键技术研究[D]. 哈尔滨:哈尔滨理工大学,2018.
- [18] BENZEKKI K, FERGOUGUI A E, ELALAOUI A E. Software-defined networking (SDN): a survey [J]. Security Communication Networks, 2016, 9:5803-5833.
- [19] Liu Yalong, Dong Mianxiong, OTA K, et al. Deep reinforcement learning based smart mitigation of DDoS flooding in software-defined networks[J]. 2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), 2018.

(收稿日期:2020-03-31)

作者简介:

郑忠斌(1979-),男,硕士,高级工程师,主要研究方向:移动通信、无线接入和设备网络管理等。

黎聪(1997.11-),男,学士,主要研究方向:网络流量检测与安全态势评估。

王朝栋(1981-),男,硕士,主要研究方向:网络与信息安全战略规划、政策标准制定、技术产业开发。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《信息技术与网络安全》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《信息技术与网络安全》编辑部
中国电子信息产业集团有限公司第六研究所