

基于专家经验的网络异常流量 自动检测平台设计与实现

黄自力^{1,2†}, 潘孝闻^{1,2}, 杨阳^{1,2}

(1. 电子商务与电子支付国家工程研究中心, 上海 201201; 2. 中国银联股份有限公司, 上海 201201)

摘 要: 随着网络流量规模的日益增大, 异常流量越来越多。传统基于规则的异常流量检测效果不尽理想, 而专家的人工分析又会消耗大量的时间与极高的人力成本, 为此, 本文基于专家经验设计并实现了一个网络异常流量自动检测平台, 以此降低人工分析的工作量。检测平台通过机器学习、词法分析的方法, 自动收集流量特征, 同时对正常与恶意 HTTP 流量进行分析, 从而有效地检测出异常流量。实验结果表明, 相比于传统 WAF, 本文设计的自动检测平台检出率更高, 误判率更低。

关键词: 流量检测; 专家经验; 机器学习; 词法分析

中图分类号: TP39

文献标识码: A

Design and Implementation of Automatic Detection Platform for Network Abnormal Traffic Based on Expert Experience

HUANG Zili^{1,2†}, PAN Xiaowen^{1,2}, YANG Yang^{1,2}

(1. National Engineering Research Center of Electronic Commerce and Electronic Payment,
Shanghai 201201, China; 2. China UnionPay Co., Ltd., Shanghai 201201, China)

Abstract: With the increasing scale of network traffic, more and more abnormal traffic is generated. Traditional rule-based abnormal traffic detection is not ideal, and manual analysis by experts will consume a lot of time and high labor costs. Therefore, this paper designs and implements an automatic detection platform for network abnormal traffic based on expert experience to reduce workload of manual analysis. By means of machine learning and lexical analysis the detection platform collects traffic features automatically and analyzes normal and malicious HTTP traffic respectively so as to detect abnormal traffic effectively. The experiment results comparing with traditional web application firewall shows that the automatic detection platform proposed by this paper can reduce workload of manual analysis and has higher positive rate and lower misjudgment rate.

Key words: traffic detection; expert experience; machine learning; lexical analysis

随着攻防技术的升级, 面对层出不穷的新型、持续性、高威胁攻击时, 传统以防御为主的安全策略越来越难以有效的阻断。基于规则的安全策略严重依赖于人工分析, 专业分析人员难以处理与日俱增的海量数据, 网络空间^[1]的威胁风险不可预估。

近年来, 由于人工智能^[2]技术的兴起, 基于机器学习^[3]分析恶意流量的方法为提高网络安全检测与防御能力开辟了新的思路, 更适合在复杂的网络环境中满足实时准确的需求^[4]。基于有监督学习的检测方法中, 蒋鸿玲等^[5]通过大量的僵尸网络和正常流量样本训练 BP 神经网络分类器, 使其学

收稿日期: 2023-11-10

基金项目: 上海移动电子商务工程技术研究中心资助项目(19DZ2284800)

作者简介: 黄自力(1991—), 男, 上海人, 中级工程师, 硕士, 研究方向: 信息安全, 人工智能。

† 通信联系人, E-mail: huangzili@unionpay.com

会辨认僵尸网络的流量,自动记忆僵尸流量特征,从而有效检测出被感染的主机。Chiba 等^[6]通过将恶意 HTTP 流量中的 URL path,URL query 和 User-agent 生成模型,再对比 HTTP 请求与模型来检测恶意软件的 HTTP 请求。基于无监督学习的检测方法中,Perdisci 等^[7]通过提取恶意流量的 HTTP 请求总数,GET 请求数量,POST 请求数量,URL 长度,平均参数数量等行为特征,再通过聚类来检测恶意软件活动。Riesi 等^[8]的模型是基于 Haar 小波^[9]的网络流量模型,小波模型功率谱在理论上可以任意接近幂律,而功率谱满足幂律的随机过程是自相似的。

除了对各类恶意 Web 攻击流量的分析,还有较少一部分对正常流量开展了初步研究。邹柏贤^[10]通过方差分析的方法,对实际网络中非单播包的观测值时间序列平衡化,建立自回归滑动平均模型预测网络流量及其变化趋势,并检测网络异常情况的发生。

为提升网络流量的检测效率,形成网络空间安全态势感知^[11-12],减轻专业分析人员的分析负担,利用专家经验来完善网络安全保障体系已成为业界共识。在此背景下,构建网络异常流量自动检测平台,研究基于恶意流量与正常流量特征的自动检测技术,具有重要的现实意义。

1 基于专家经验的检测技术

基于专家经验的检测技术,是指根据安全专家对可疑行为的分析经验而形成的可信推理规则。在此基础上设计相应的攻击检测系统,在检测过程中,系统将已知的攻击手段抽象成规则,构成一颗判定树,然后所有的网络流量都经过特征提取,最后输入到专家系统得到判定结果^[13-15]。

安全专家在分析可疑行为时,通常的做法包括:

(1)常识判定:根据业务逻辑,预估普通请求中字段的参数值范围,当超过这个范围,则极可能不是正常的请求响应流量,从而发现异常请求。

(2)统计分析:根据专家经验,通过分析一段时间的统计值,计算固定指标、参数均值、多变量协方差、状态转移概率等,从而发现可疑事件。

(3)行为分析:对于系统中的用户行为模式进行分类与识别,专家创建用户的行为特征轮廓,分析其行为的可变性,从而发现异常用户。

(4)知识分析:专家提前记录曾发生过的入侵行为,将每种网络攻击的显著特征保存在知识库

中,将所分析的行为与之一一比较,从而发现可疑行为。

(5)挖掘分析:收集大量的、不完全的、噪声日志,专家采用数据挖掘的方法对其进行分析,以此分析异常的活动模式。

通过专家经验的人工分析,还容易识别变形绕过攻击,因而得到的分析结果通常可认定是正确的结论,但此方法会消耗大量的时间与极高的人力成本,不适用于海量的流量检测。因此,亟待一种自动化的方法处理海量数据,将专家经验汇总并集成在检测平台中,从而高效、可靠地检测恶意与正常流量。

2 网络异常流量检测平台

网络异常流量检测平台整合了专家分析与判定网络流量的技术经验,通过自动化的方式进行检测。平台由基于恶意流量特征的检测模块与基于正常流量特征的检测模块两部分组成,分别对恶意和正常两种流量展开分析,对流量的多重检测可达到更好的检测效果。如图 1 所示,基于恶意流量特征的检测模块包含了各类 Web 攻击的检测引擎,其中以 SQL 注入,XSS 攻击与 WebShell 为典型场景。基于正常流量特征的检测模块中无需对攻击形式进行划分,可应对多类型攻击,模块中分为了日志初分引擎,正常流量日志学习引擎以及日志检测引擎三个部分。当流量进入时,两个模块分别进行检测,最后对异常日志进行汇总。

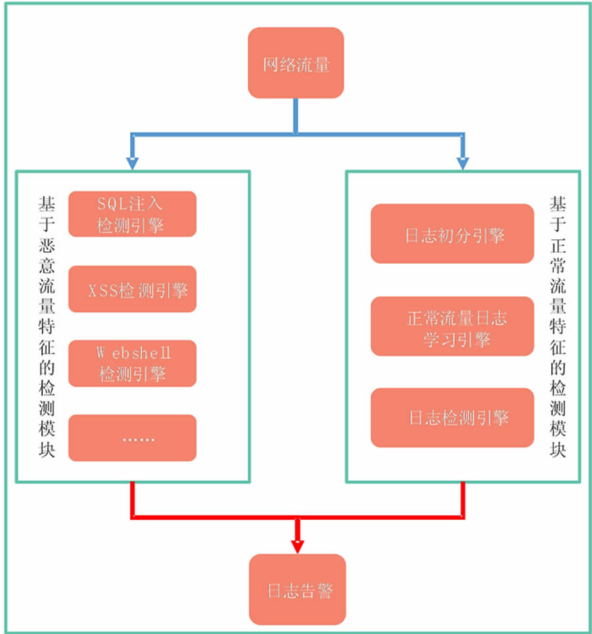


图 1 网络异常流量检测平台

2.1 自动检测方案设计

如图 2 所示,网络流量自动检测方案包含了三层检测。首先在快速过滤层,基于黑白样本强特征将绝大部分的初始网络流量与日志过滤。然后,将过滤后的网络流量传到智能检测层,通过机器学习、关联分析等人工智能方法,再次进行检测与筛选。最后,对于仍难以判定黑白的流量,通过事后审计层,由技术专家来判断,将分析结果反馈给前两层,迭代优化前两层的过滤规则与模型参数。

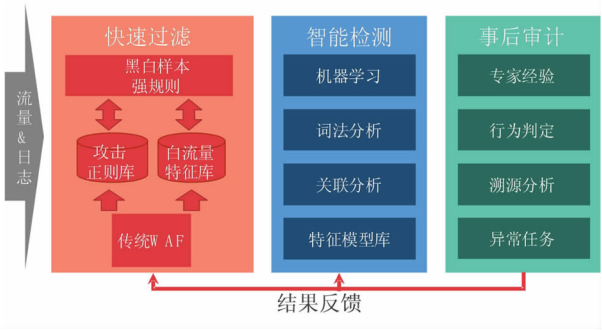


图 2 自动检测方案

2.2 基于恶意流量特征的检测模块

基于恶意流量特征的检测模块包含了各种类型的 Web 攻击检测,由于 SQL 注入、XSS 攻击与 WebShell 这三类攻击最为常见,将其作为典型场景详细说明。其总体流程图如图 3 所示,数据流经过预处理后进行缓存,分别发送到三个检测引擎,三个引擎再分别检测该流量,若存在攻击则直接报出,否则将结果汇总到综合判断,最后输出结果。

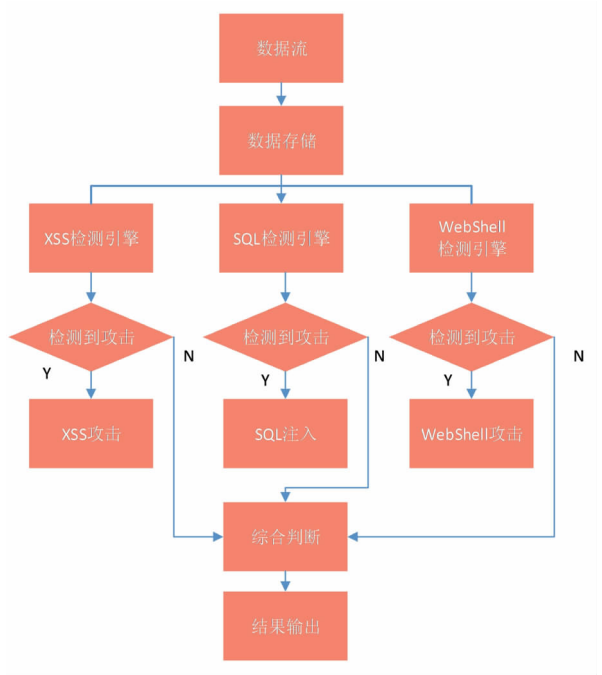


图 3 基于恶意流量特征的检测整体流程

2.2.1 SQL 注入检测引擎

SQL 注入检测引擎按照功能划分,可分为词法分析模块、语法重写模块、指纹比对模块。

(1)词法分析模块

词法分析模块主要是基于词法分析将流量数据转换成 token 序列。将每个词根据词法分析的映射关系分为不同小类的词性。

(2)语法重写模块

语法重写模块主要功能是对原始 token 序列进行基于语法的重写,然后对结果序列抽取所有的 3~6 元组作为其指纹,主要包括同类合并、大类化简、数学表达式归并等。

(3)指纹比对模块

指纹比对模块是将提取到的指纹和指纹库进行比对,如果指纹在指纹库中,则认为是 SQL 注入。

2.2.2 XSS 攻击检测引擎

XSS 攻击检测引擎按照功能划分,可分为引擎初始化模块、训练模块、测试模块。

(1)初始化模块

引擎的配置信息以文本格式存储,其中规定了 XSS 检测引擎的工作模式,包括训练模式、预测模式、待检测文件路径等信息。

(2)训练模块

首先从初始化模块获取配置信息,读取训练数据路径和模型保存路径。然后调用特征提取模块,对训练数据进行特征提取,形成特征向量。最后,对数据进行训练建模,得到模型结果。

(3)测试模块

首先从初始化模块获取配置信息,读取测试数据。然后调用特征提取模块,对测试数据进行特征提取,形成特征向量。最后加载模型文件,对测试数据进行预测,输出模型测试性能。

2.2.3 WebShell 攻击检测引擎

WebShell 攻击检测引擎按功能划分,可分为前置过滤模块、语法模块。

(1)前置过滤模块

语法引擎的计算复杂度相对较高,如果分析全部流量,则语法分析性能无法接受。现在根据字段与关键词做前置过滤,滤出可能的攻击语句,再送到语法引擎进行分析。

(2)语法模块

由于网站服务器可能由 asp、jsp、php 等编程语言编写,不同类型的 WebShell 也由不同的语法构成。要将可能的攻击语句同时送入对应的语法模块中,进行攻击检测,并根据匹配度进行判断其是 WebShell 攻击命令的可能性。

2.3 基于正常流量特征的检测模块

基于正常流量特征的检测模块的总体框架图如图 4 所示,分为日志初分引擎,正常日志学习引擎以及日志检测引擎三个部分。

首先,将流量同时分发到日志过滤及判断。日志过滤中,将流量与日志过滤规则库对比,若存在且相似,则将正常流量日志传入正常日志学习引擎,否则需由人工分析。人工分析后若认为是正常的,则将其提取规则并添加到日志过滤规则库中,以便下次对比过滤,否则直接过滤。在日志判断中,从模型数据库中搜索模型,检验其模型的存在性。

然后,对正常 HTTP 流量日志按字段提取特征,将特征形成对应模型存入数据库。当新的正常流量日志进入时,将其学习到的特征与现有的模型数据库中的特征进行比较,若不一致则对模型及时更新。

最后,提取 HTTP 日志的所有字段键值对,从规则知识库与学习引擎中的模型数据库中学习的特征进行对比检测。对于不符合模型特征的日志给出警告。

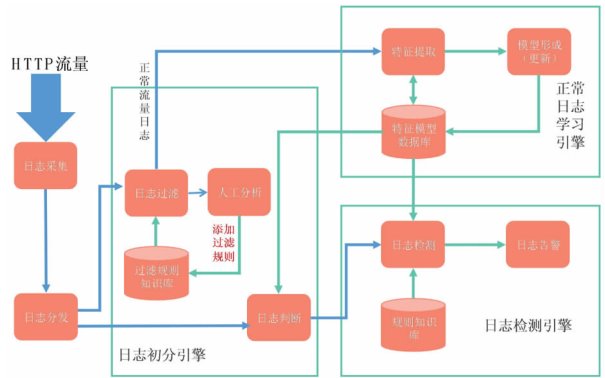


图 4 基于正常流量特征的检测模块的总体框架图

2.3.1 日志初分引擎

由于 HTTP 流量日志中存在恶意流量日志,因此,需要对恶意流量日志过滤,筛选正常的流量日志并对其特征提取。筛选方法可使用基于规则的传统 WAF,筛选后的剩余日志中还包含了很多无效的非网络攻击流量日志。为了筛选掉无效的非网络攻击流量日志,可使用响应码,URL 路径及参数名出现频率以及按照流持续时间,负载,IP 及包长等特征对有效及无效流量日志进行分析,形成有效流量日志特征模式串,通过特征模式串匹配的方法对日志过滤。对于部分不能判断的日志,可以将其记录,人工分析其中不属于无效流量日志的正常流量日志,将类似的日志都判定为正常并将这些正常日志提取规则放入日志过滤规则知识库中。

再使用过滤规则库进一步判断过滤日志,符合规则的加入正常流量日志。

2.3.2 正常日志学习引擎

对过滤后得到的正常 HTTP 流量日志,需要对日志按照字段进行特征提取。本文中分析的字段有如 URL, Cookie, POST 请求体等。文中根据字段值的类型对字段值分类。主要类型包括数值类、字符串类、特殊格式类、文本类与 User-Agent 类。对于数值类型可以使用机器学习中的聚类方法聚类出其数值范围;对于字符串类,提取字符串长度,特殊符号、小写字母及数据占比等特征,并通过聚类获得特征值范围;对于特殊格式类,提取其正则表达式并将其存在规则知识库中;对于请求体中的如 json, xml 的文本格式类,使用机器学习对文本内容进行分析;对于 User-Agent 类,使用 Ratcliff-Obershelp 相似度算法计算原始 User-Agent 值之间的相似度,筛选后得到 User-Agent 数据库。最后,将每一个 URL 后提取到的每个字段的特征组成对应 URL 的模型存在数据库中。

2.3.3 日志检测引擎

对于新的 HTTP 流量日志,将每一条日志都按照字段拆分,将字段键值对分别与对应模型的特征进行匹配。若日志的每一类字段都符合相应模型的特征范围则判断该日志为正常。若日志的某一类字段或者多类字段都不符合相应模型的特征,则给出日志警告。如数值类型值中,若超出数值聚类范围,则给出警告;字符串类值中,若超出各个特征值的聚类范围则给出警告;特殊格式类中,若匹配不成功则给出警告;文本类中,若文本特征不符合则给出警告;User-Agent 类中。若与数据库中的所有值进行相似度匹配后最大值小于一定阈值则给出警告。

3 实验结果与分析

3.1 实验结果

本文选取了 1446 条公开 SQL 注入、3382 条公开 XSS 攻击、2421 条公开 WebShell、384 条其他类型及其变形攻击、5000 余条正常的流量数据进行测试,测试时,为了更加直观地展示自动检测平台的效果,使用了传统 WAF 与纯人工智能(AI)算法引擎作为对照。

同时,本文从误报率、检出率对检测实验结果进行了衡量,检出率是指真实标签为黑的样本中,系统预测为黑的比例;误报率是指真实标签为白的样本中,系统判定为黑的比例;一般而言,检出率越高,说明系统对攻击的拦截效果越好,误报率越低,

则说明系统对正常业务的影响越小。

本文针对自动检测平台、传统 WAF、纯 AI 算

法引擎进行了对比测试,测试结果汇总在表 1 中。

表 1 自动检测平台与其他方式的检测结果对比情况

数据类型	SQL 注入		XSS 攻击		WebShell		其他类型	
计算方式	误报率	检出率	误报率	检出率	误报率	检出率	误报率	检出率
传统 WAF	18.70%	91.10%	5.30%	80.15%	4.80%	77.32%	9.20%	82.81%
纯 AI 算法引擎	6.90%	98.72%	3.66%	97.96%	2.91%	93.89%	24.46%	96.36%
自动检测平台	0.10%	98.93%	0.10%	98.19%	0.10%	94.11%	0.60%	97.85%

3.2 检出率分析

从实验结果来看,对于各类流量攻击,自动检测平台的检出率均高于传统 WAF 与纯 AI 算法引擎,这说明三层检测方案的设计实现是有效的,可以提升流量攻击的检测效果。其中的原因主要有两点,首先是针对不同类型的攻击类型,设计特定的检测方案与模块实现,可发现纯规则匹配无法发现的可疑流量,从而检测出常见攻击的变种;其次,多种智能算法模块在流量检测过程中,及时将信息与结果共享,相互补足检测盲点,形成多维度、更立体的流量检测体系。

3.3 误报率分析

从实验结果来看,自动检测平台的误报率远远低于传统 WAF 与纯 AI 算法引擎,这使得运维人员的分析工作量大大降低,提升了日常运维的工作效率。其中的原因是:传统 WAF 是基于常见攻击特征添加了若干条正则表达式,以此来检测网络流量黑白。纯 AI 算法则是通过各类攻击流量特征,分别建立并训练模型,用以最终检测。而自动检测平台中具备正常流量检测模块,通过对于大量日常的业务流量学习,避免了一些具有攻击特征但实际是业务特性的流量误判,从而在整体上降低了对流量的误判率。

4 结 论

介绍了基于机器学习的恶意流量与正常流量的检测方法,分析了常见的专家判定恶意流量与可疑行为的手段与思路,并基于此,设计了一种基于专家经验的网络异常流量自动检测平台。该平台根据专家经验,采用词法分析、机器学习等技术,分别提取正常流量特征与恶意攻击特征,由两个相辅相成的检测模型所组成,可有效检测网络行为中的异常流量。基于专家经验转化的自动检测系统不仅大大提升检测准确性,也节省人工成本。同时,由于特征的泛化性与模型的通用性,本文所提出的方法可复用于其他类似的安全检测场景,为加强网

络安全防范能力提供了新契机。

不过现有平台中分析的恶意流量攻击类型还较少,正常流量分析的字段还不够全面,检测的性能较低,日后需逐步完善。总体而言,本文为未来的网络安全智能化体系防御提供了新思路,具有战略指导意义以及落地的实用性。

参考文献

[1] 王世伟. 论信息安全、网络安全、网络空间安全[J]. 中国图书馆学报, 2015, 41(2): 72—84.

[2] 张妮, 徐文尚, 王文文. 人工智能技术发展及应用研究综述[J]. 煤矿机械, 2009, 30(2): 4—7.

[3] 唐政治、曾学文、陈君, 等. 基于机器学习的网络流量分析综述[J]. 网络新媒体技术, 2020, 9(5): 8.

[4] 陈胜, 朱国胜, 祁小云, 等. 基于机器学习的网络异常流量检测研究[J]. 信息通信, 2017(12): 39—42.

[5] 蒋鸿玲, 邵秀丽. 基于神经网络的僵尸网络检测[J]. 智能系统学报, 2013, 8(2): 113—118.

[6] CHIBA D, YAGI T, AKIYAMA M, et al. Botprofiler: profiling variability of substrings in http requests to detect malware-infected hosts[C]//2015 IEEE Trustcom/BigDataSE/ISPA. IEEE, 2015, 1: 758—765.

[7] PERDISCI R, LEE W, FEAMSTER N. Behavioral clustering of http-based malware and signature generation using malicious network traces[C]//NSDI. 2010, 10: 14.

[8] RIEDI R H, CROUSE M S, RIBEIRO V J, et al. A multi-fractal wavelet model with application to network traffic[J]. IEEE transactions on Information Theory, 1999, 45(3): 992—1018.

[9] STANKOVIC R S, FALKOWSKI B J. The Haar wavelet transform: its status and achievements[J]. Computers & Electrical Engineering, 2003, 29(1): 25—44.

[10] 邹柏贤. 网络流量正常行为模型的建立[J]. 计算机应用, 2002, 22(7): 3.

[11] 龚正虎, 卓莹. 网络态势感知研究[J]. 软件学报, 2010, 21(7): 1605—1619.

[12] 李硕, 戴欣, 周渝霞. 网络安全态势感知研究进展[J]. 计算机应用研究, 2010(9): 3227—3232.

[13] 郭瑜, 孙毅, 万芳. 一个基于专家系统的入侵检测系统的实现[J]. 系统工程与电子技术, 2001, 23(12): 84—85.

[14] 鲁宏伟, 罗钢. 基于专家系统的入侵检测方法[J]. 武钢技术, 2003, 41(1): 43—48.

[15] 杨慧. 基于专家系统的入侵检测技术[J]. 中国科技信息, 2008(15): 81—82.