

基于 SVM 的 Web 攻击检测技术

吴少华 程书宝 胡 勇
(四川大学电子信息学院 成都 610065)

摘 要 针对各种变形 Web 攻击难以检测的问题,分析 SQL 注入和跨站攻击特征的选择和提取的一般方法,利用人工挑选和数学统计概括出 6 个特征,将原始攻击载荷转换成固定维数的特征向量,标记特征选择和提取后的样本数据,进行支持向量机算法的训练和分类。借助机器学习工具 Weka 验证了该检测方案的可行性。
关键词 SQL 注入,跨站脚本,Web 攻击检测,特征选择与提取,支持向量机
中图法分类号 TP393.08 文献标识码 A

Web Attack Detection Method Based on Support Vector Machines

WU Shao-hua CHENG Shu-bao HU Yong
(School of Electronics & Information Engineering, Sichuan University, Chengdu 610065, China)

Abstract Web attack detection is a kind of dynamic Web security protection technology, but the intruder can use different coding schemes, mixed case, alternative statements and other skills, bypassing defense mechanism. For the particularity of web security and the shortage of the existing detection technology, we took SQL injection and cross site scripting attacks as an example. Firstly, the thesis studies the feature selection and extraction of SQL injection and cross site scripting attacks, and uses the artificial selection and mathematical statistical methods to covert the original payload into fixed dimension feature vector. Secondly, it marks the sample data after feature selection and extraction, and performs support vector machine training and classification. Finally, using the Weka, it verifies the feasibility and effectiveness of the approach. The experimental results show that features after selection and extraction can reflect the nature of the original data and this method has higher detection rate.

Keywords SQL injection, Cross site scripting, Web attack detection, Feature selection and extraction, Support vector machine

1 引言

Web 安全已成为信息安全领域的研究热点之一。实时检测 Web 攻击是目前 Web 安全防护的必要手段。国内外一些安全公司推出的 Web 安全防护产品多是通过一些固定的通信特征来识别 Web 攻击。但如果攻击者对提交的攻击载荷使用不同的编码方案、大小写变换以及替代性语句等技巧,就可绕过检测,实施各种变形攻击。如果对这些变形攻击都添加特征规则,会造成特征库臃肿,影响检测性能。

针对基于通信特征的 Web 攻击检测技术不能检测未知或变形的攻击、误检率较高以及实时性差等问题,文献[1,2]提出分析 Web 日志来发现攻击,但攻击行为在分析日志前已经发生,属事后检测,并且该方法无法检测变形的 Web 攻击。文献[3]提出把攻击检测与 Web 应用集成,但该方法只适用于 Apache Web 服务器,通用性不足,也无法检测未知变形攻击。文献[4]提出利用 ID3 算法在训练阶段构建一棵决策树用以分类检测 Web 攻击,但 ID3 算法有偏向于取值较多的属性的缺点,因此检测率较低。文献[5]提出基于免疫原理的 Web 攻击检测方法,对 Web 攻击数据进行编码,形成抗原集

合,然后将抗原集合提交给免疫学习系统进行学习,产生免疫细胞来检测 Web 攻击,该方法具有高检测率和实时性好等特点,但克隆和变异过程复杂,收敛速率慢。文献[6]提出一种自适应的 Web 攻击异常检测方法,使用多隐马尔可夫模型对 HTTP 请求样本进行分类处理,并根据分类样本集的离散性分析,自适应地区别出正常行为,但需要进行较长时间的学习。文献[7]分析针对二进制和 N 元字母表的基于编码的 XSS 攻击,在已有客户端跨站脚本攻击检测技术的基础上,给出一种动态访问控制的防范方法。文献[8]提出一种基于确定有限自动状态机语法的网络攻击检测方法。正常的网络行为符合一定的语法规则,异常的行为会偏离正常的语法规则。通过对正常行为样本的学习得到基于 DFA 的语法,用学习得到的 DFA 模型检测针对网络服务器的应用层攻击。文献[9]提出基于序列比对的 SQL 注入攻击检测方法,但该方法在执行过程中需要对应用程序的每个 SQL 语句调用验证函数,对应用程序改动较大。

本文通过分析变形 Web 攻击的各种形式,提出基于支持向量机的 Web 攻击检测技术。利用支持向量机从训练样本中自动分析获取规律,并通过这些规律对未知数据进行判断

本文受国家计算机网络与信息管理中心 242 课题资助。

吴少华(1977—),男,博士,副教授,主要研究方向为网络安全,E-mail:2353241@qq.com;程书宝(1990—),男,硕士,主要研究方向为信息系统安全;胡 勇(1973—),男,博士,副教授,主要研究方向为网络安全,E-mail:huyong@scu.edu.cn(通信作者)。

来检测可疑的 Web 攻击。检测主要包括：特征选择与提取、SVM 训练阶段以及分类阶段，流程如图 1 所示。

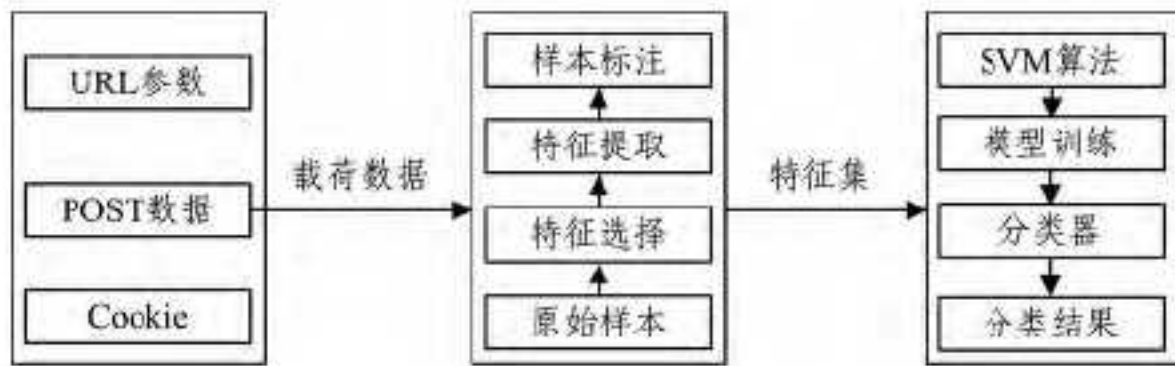


图 1 Web 攻击检测模型

2 特征选择与提取

从 HTTP 请求数据包中提取的载荷数据是非结构化的，需进行结构化处理，从原始特征中挑选出数量为 n 的一组最优特征，在不降低分类准确率的前提下降低原始特征空间维数，将原始载荷数据转换为固定维数的特征向量，作为 Web 攻击检测算法的输入数据。

通过分析漏洞验证阶段、漏洞利用阶段及各种绕过情景下的 SQL 注入和 XSS 攻击语句的各种形式，采用人工挑选和数学统计相结合的方式对原始载荷数据进行特征选择，概括出 6 个特征：特征关键字、各种闭合、截断等特殊字符频率（特殊字符个数/字符数）、特殊前缀字符频率、大写字母字符频率、数字字符个数频率和空格字符频率。特征的汇总和具体含义如表 1 所列。

表 1 特征名称与含义

特征名称	特征含义
特征关键字	相应类型攻击的特征关键字
特殊字符频率	各种闭合、截断等特殊字符频率
特殊前缀字符频率	&#、&#x、\、\x、\u、% 字符频率
大写字母字符频率	大写字母(A-Z)频率
数字字符频率	数字(0-9)频率
空格字符频率	空格字符频率

其中，SQL 注入攻击语句中常见的特征关键字有：and、or、xor、sysobjects、msysobjects、version、substr、substring、len、length、exists、mid、asc、inner join、xp-cmdshell、exec、having、group by、back up、union select、order by、information schema、load-file、load data infile、into outfile、into dumpfile。XSS 攻击语句中常见的特征关键字有：script、alert、prompt、location、hash、src、href、@import、eval、XMLHttpRequest、ActiveXObject。常见的特殊字符有：“！#%&'；<>=?@[]\|{}|\$, * + -。”。

3 支持向量机

SVM 是从线性可分情况下的最佳分类面发展而来，其基本思想可用图 2 所示的两维情况解释。设有给定的训练集 $D = \{(X_1, y_1), (X_2, y_2), \dots, (X_n, y_n)\}$ ，其中 X_i 是原始样本的特征向量， y_i 是相关联的类标号，在二分类中，每个 y_i 取二值之一（即 $y_i \in \{+1, -1\}$ ），表示是否属于这个类。如果有线性函数能将两类样本完全分开，就称为线性可分，否则为非线性可分。

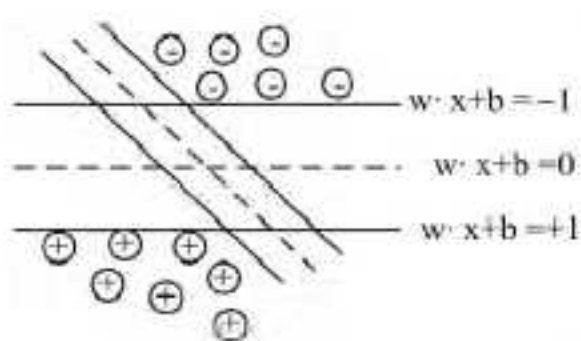


图 2 支持向量机最佳超平面

图 2 中的两类样本是线性可分的，可以找到一条直线将 $+1$ 类和 -1 类分开，这样的直线有无数条，所谓最优分类线要求分类线能将这两类样本正确分开，而且分类间隔最大。SVM 通过搜索使分类间隔最大的那一条来完成对样本的分类，最优分类线可以用方程 $w \cdot x + b = 0 (w \in R^n, b \in R)$ 表示^[10]。推广到 n 维空间，SVM 通过搜索最佳超平面来完成分类过程，SVM 如何找出最佳超平面等价于求解被约束的优化问题，公式表示为^[11]：

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i$$

s. t. $y_i(w \cdot x_i + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, \dots, N$

其中， $C > 0$ 为惩罚参数，表示对离群点的重视程度。松弛变量 ξ_i 是对离群点离群程度的度量。

Web 攻击检测等价于 SVM 多类分类过程。常用的 SVM 多类分类思路是分解重构，将多类问题分解为多个二类问题，构造出多个对应的 SVM 二类分类器，然后按一定的策略重构分类器实现多个类别的判决。论文采用“一对一”SVM 多类分类算法。在训练阶段，从这 n 种类别中任意挑选两个类别构造二类分类器，共构成 $n(n-1)/2$ 个 SVM 二类分类器。在分类阶段，分别利用这 $n(n-1)/2$ 个分类器对输入样本进行分类判断，每次将样本判决为某一类，最终得到的判决次数最多的那一类为样本数据的最终分类结果^[12]。

4 系统仿真实验

为验证该 Web 攻击检测方法的有效性，首先证明选择的 6 个特征能反映 HTTP 请求中载荷数据的本质特征。然后，借助机器学习工具 Weka 验证该检测方案对未知 Web 攻击的识别能力。

实验数据集由 3 部分组成：正常 HTTP 请求样本集、SQL 注入样本集和 XSS 样本集。通过分析 Web 服务器日志提取出正常访问资源的请求样本，从漏洞提交网站 XSSED^[13]、HA、CKKERS^[14] 与 exploit-db^[15] 收集 XSS 与 SQL 注入攻击样本。通过前期工作共收集 1032 条样本数据，包含 400 条正常 HTTP 请求样本、309 条 SQL 注入攻击样本、323 条 XSS 攻击样本。表 2—表 4 列出了部分原始样本数据，可以看出原始数据无法作为 SVM 输入，需按第 1 节给出的特征选择与提取方法将其转换为固定维数的特征向量^[16]。

表 2 正常 HTTP 请求样本

780778624&otype=json&callback=videoCount&ran=0.6194237943566041
TwSMYE&type=1&of=1&uid=13973551297432480448940863754139
[32773,30622,32749,43151,38357,29525]&requestToken=-19374601&-
rtk=5a48095
http%3A%2F%2Fwww.test.com%2FicsonBanner.html%3Fsize%3D320x180%26time
17030-34976-58,30905-39101-58,36471-15380-58,40365-16244-58&cb=bds.base64.cbr

表 3 SQL 注入攻击样本

%20and%20substring(version(),1,1)=5
sid=140; backup database gns to
disk='C:\FTP\gns\wormadmin\1.bak';--
-1+union+select+concat(email,0x3a,code)+from+clf-ads--
1 and 0x0=0x1 union select 1,
CONCAT(username,0x3a3a3a,password),3,4from tbladmin#
' AND 3086=BENCHMARK(5000000,MD5(0x454a5a64))
AND 'qjLM'='qjLM

表4 XSS攻击样本

```

/ ><script>alert('Xss By Atm0n3r')</script><script>
%22%3E%3C/script%3E%3Cscript%3Ealert('XSS')%3C/script%3E0xAli
+-+XSSSED%3C!--
/ ><script>alert(1)</script>&imageField.x=31&imageField.y=22
%22/%3E%3Cbody%20onload=alert%28document.cookie%29%3E
%22%3E%3Cscript%3Ealert%28%22XSS%22%29%3C%2Fscript%
3E&x=0&y=0

```

良好的特征应具有可区分性、稳定性及独立性。可区分性指不同类型系统的特征有差别,稳定性指同一类型系统特征应接近,独立性指各个特征之间应不相关^[17]。表5—表7给出了提取后的部分样本特征。特征₁表示载荷数据中是否出现相应类型攻击的特征关键字,特征₂表示各种闭合、截断等特殊字符出现的频率,特征₃表示&#、&#x等字符串出现的频率,特征₄表示大写字母字符个数百分比,特征₅表示数字字符个数百分比,特征₆表示空格字符个数百分比。

表5 正常样本特征

正常样本	特征 ₁	特征 ₂	特征 ₃	特征 ₄	特征 ₅	特征 ₆
1	0	0.1105	0	0.0404	0.2941	0.0043
2	0	0.1030	0	0.0446	0.3577	0
3	0	0.1228	0	0.0657	0.3348	0
4	0	0.1288	0	0.0394	0.3135	0
5	0	0.1095	0	0.0446	0.3348	0.0018

表6 SQL注入样本特征

SQL注入	特征 ₁	特征 ₂	特征 ₃	特征 ₄	特征 ₅	特征 ₆
1	-1	0.1621	0	0.0822	0.1584	0.0652
2	-1	0.1610	0	0.0845	0.1667	0.0526
3	-1	0.1586	0	0.0667	0.1333	0.0769
4	-1	0.1796	0	0.0714	0.1944	0.0594
5	-1	0.1228	0	0.0632	0.1736	0.0645

表7 XSS样本特征

XSS样本	特征 ₁	特征 ₂	特征 ₃	特征 ₄	特征 ₅	特征 ₆
1	1	0.2295	0	0.2340	0.0476	0.0244
2	0	0.2028	0	0.2432	0.0435	0.0204
3	1	0.1921	0	0.2647	0	0.0247
4	1	0.2202	0	0.2469	0.0476	0.0370
5	1	0.1990	0	0.2280	0.0654	0.0204

上述3个表所列3类样本中,同一类样本的特征值比较接近,而不同类样本的特征值差别较大,这表明提取的样本特征具有可区分性、稳定性,而不同特征之间互不相关,表明样本特征具有独立性。每一个样本经特征选择与提取后对应一个特征向量,该特征向量能反映HTTP请求中载荷数据的本质特征,并作为SVM分类器的输入。

为了评估该方法的正确率,采用真正类率(true positive rate, TPR)和负正类率(false positive rate, FPR)描述该分类器。真正类率刻画的是分类器所识别出的正实例占所有正实例的比例,负正类率描述的是分类器错认为正类的负实例占所有负实例的比例。将经过预处理之后的特征向量标记类别后保存成csv格式文件,加载到数据挖掘工具Weka进行分类识别。在实验过程中,选用序列最小优化(SMO)算法,核函数选择“PolyKernel”,设置惩罚参数 $C=1$,并使用交叉验证法来分析分类器的检测能力。实验采取三重交叉验证法,三重交叉验证将数据集平均分为3等份,其中两份用来训练,另外一份用来测试,作为测试集的样本数据,对于分类器而言即为未知类型的请求。表8是3种类型请求的分类结果情况。从表8可以看出,该方法对XSS攻击与SQL注入攻击均有较高的真正类率和较低的负正类率。ROC^[18,19]曲线是真正类率和负正类率的综合指标。ROC曲线下的面积值(ROC are-

a)越接近1,表明对该类样本识别准确率越高,表8中ROC area均在0.9以上,说明该分类器对未知类型Web攻击有较高的识别率。

表8 分类结果

类别	TP Rate	FP Rate	ROC Area
正常请求	0.920	0.051	0.936
XSS攻击	0.876	0.058	0.909
SQL注入攻击	0.900	0.041	0.933

结束语 本文给出了基于支持向量机的Web攻击检测方案。实验结果表明,经选择和提取后的特征能反映原始载荷数据的本质特征,提出的方法对未知Web攻击有较高的检测率。本文主要是建立在实验的基础上,在实际应用中需进行改进和完善。

参考文献

- [1] Adeva J J G, Atxa J M P. Intrusion detection in web application using text mining[J]. Engineering Applications of Artificial Intelligence, 2007, 20(4): 555-566
- [2] Almgren M, Debar H, Dacier M. A lightweight tool for detecting Web server attacks[C]// Proceedings of Network and Distributed Systems Security. 2000: 157-170
- [3] Almgren M, Lindqvist U. Application-integrated data collection for security monitoring[C]// RAID 2001, LNCS 2212. Berlin: Springer-Verlag, 2001: 22-36
- [4] Garcia V H, Monroy R, Quintana M. Web attack detection using ID3[OL]. <http://homepage.cem.itesm.mx/raulm/pub/id3-ids>, 2013-12
- [5] 温凯, 郭帆, 余敏. 自适应的Web攻击异常检测方法[J]. 计算机应用, 2012, 32(7): 2003-2006, 2014
- [6] 张伟, 吴源, 邹鄂路. 针对基于编码的跨站脚本攻击分析及防范方法[J]. 小型微型计算机系统, 2013, 34(7): 1615-1619
- [7] 杨晓峰, 孙明明, 胡雪蕾. 一种基于DFA的网络攻击检测算法[J]. 计算机工程, 2010, 36(13): 149-150, 153
- [8] 孙义, 胡雨霏, 黄皓. 基于序列比对的SQL注入攻击检测方法[J]. 计算机应用研究, 2010, 27(9): 3525-3528
- [9] 曾金全, 赵辉, 刘才铭, 等. 受免疫原理启发的Web攻击检测方法[J]. 电子科技大学学报, 2007, 36(6): 1215-1218
- [10] 张博锋. 面向内容安全的文本分类研究[D]. 长沙: 国防科学技术大学, 2007
- [11] Joachims T. Text categorization with support vector machines: learning with many relevant features[C]// 10th European Conference on Machine Learning. 1998: 137-142
- [12] 张晓惠, 林柏钢. 基于特征选择和多分类支持向量机的异常检测[J]. 通信学报, 2009, 30(10A): 68-73
- [13] XSSSED[OL]. <http://xssed.com>, 2014. 1
- [14] XSS(Cross Site Scripting) Cheat Sheet[EB/OL]. <http://ha.ckers.org/xssAttacks.xml>, 2014. 1
- [15] exploit-db[EB/OL]. <http://www.exploit-db.com/webapps>, 2014. 1
- [16] 程书宝. 基于支持向量机的Web攻击检测技术[D]. 成都: 四川大学, 2014
- [17] 甘俊英, 张有为. 一种基于奇异值特征的神经网络人脸识别新途径[J]. 电子学报, 2004, 32(1): 170-173
- [18] (美) Han Jia-wei, Kamber M. 数据挖掘概念与技术(第3版)[M]. 范明, 孟小峰, 译. 北京: 机械工业出版社, 2012
- [19] 杨挚诚. 基于机器学习的文本分类算法研究[D]. 桂林: 广西大学, 2007