深聊waf那些事儿（一）——安全小课堂第二十五期

京东安全应急响应中心 2016-09-02

安全小课堂第二十五期

waf是web应用防火墙（Web Application Firewall）的简称，对来自Web应用程序客户端的各类请求进行内容检测和验证，确保其安全性与合法性，对非法的请求予以实时阻断，为web应用提供防护，也称作应用防火墙，是网络安全纵深防御体系里重要的一环。本期我们来聊一聊短信验证码安全。

本期邀请到
携程安全专家张亮
唯品会安全专家yy
阿里安全专家破见
大家欢迎~

1

豌豆妹

请问什么是waf？

柴可夫斯基

waf是web应用防火墙（Web Application Firewall）的简称，对来自Web应用程序客户端的各类请求进行内容检测和验证，确保其安全性与合法性，对非法的请求予以实时阻断，为web应用提供防护，也称作应用防火墙，是网络安全纵深防御体系里重要的一环。waf属于检测型及纠正型防御控制措施。waf分为硬件waf、软件waf（ModSecurity）、代码级waf。

2

豌豆妹

waf的原理是什么？

葫芦娃 

waf对请求的内容进行规则匹配、行为分析等识别出恶意行为，并执行相关动作，这些动作包括阻断、记录、告警等。

哆啦A梦 

waf工作在web服务器之前，对基于HTTP协议的通信进行检测和识别。通俗的说，waf类似于地铁站的安检，对于HTTP请求进行快速安全检查，通过解析HTTP数据，在不同的字段分别在特征、规则等维度进行判断，判断的结果作为是否拦截的依据从而决定是否放行。

小丸子 

^_^，都说得很全了。补充一点点，从事前、事中、事后来看，waf是在攻击进行时，用于阻断攻击的安全产品。

3

豌豆妹

绕过waf防御都有哪些技巧？

小新

（1）请求真实ip绕过waf：部分waf部署架构的特性，部分waf并不是直接串在目标站点线路上，而是通过DNS解析的形式部署，此时可以先探测到目标站点的真实ip，直接请求ip以此绕过waf的检测；
（2）检测内容范围绕过：waf性能限制，检测特定内容前几k或几十K的内容，然后在此特定内容段内填充物用数据，payload放于物用数据后，以此绕过检测；
（3）协议盲区绕过：waf根据自己的防御策略所支持的协议特性，针对该协议内的请求进行检查，但是存在一些协议检测或协议运行机制上的缺陷导致被绕过，例如协议未覆盖、协议解析不正确、协议解析遗漏等；
（4）检测规则绕过：waf工程师规则编写经验、规则覆盖面等问题，来绕过检测，例如利用MySQL对一些特殊字符处理的特性、语法特性绕过；
（5）文件包含绕过：相对路径、绝对路径。

葫芦娃

在规则绕过里，常见的就是各种编码（urlencode、unicode、多重编码、大小写、换行符等），填充符号（填充注释符、空格换加号、\x00等特殊字符），比较特色的就是sqlmap的tamper脚本。缺陷绕过的话，一般有填充无用数据、特殊HTTP请求方法、耦合白名单等绕过方式。其实，绕waf就是一句话，想别人之所想。人写的规则难免会有遗漏，去考虑安全工程师会有那些遗漏，往往很有效。

柴可夫斯基

前段时间我写的一篇文章，介绍了很多思路和案例：
从架构层Bypass WAF ；
从资源限角度bypass WAF；
从协议层面bypass WAF；
从规则缺陷bypass WAF；
http://weibo.com/ttarticle/p/show?id=2309404007261092631700
大家有什么问题可以边看边思考。

4

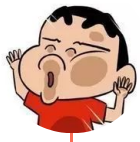
豌豆妹

都有哪些经典的waf漏洞案例呢？

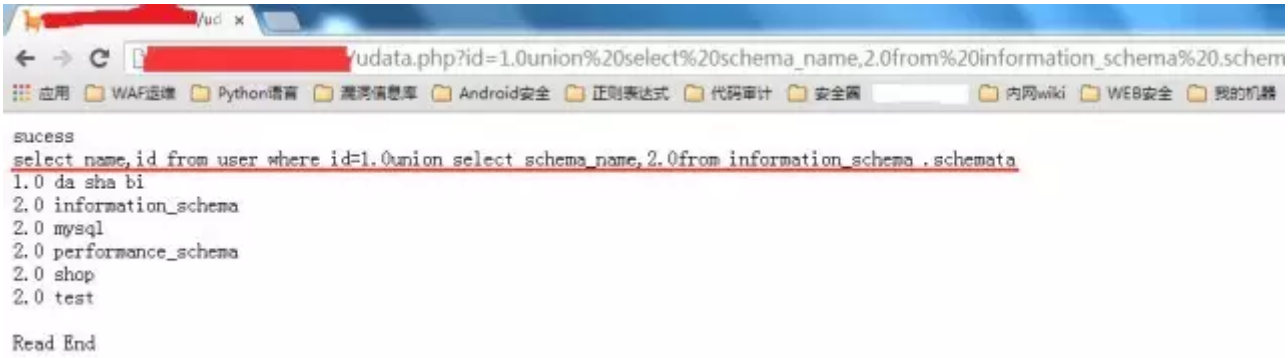
小丸子

CVE-2007-1359漏洞里，mod_security就被特殊字符\x00绕过了，算是比较经典。PHP DOS漏洞的新利用：CVE-2015-4024 Reviewed。利用漏洞进行bypass的案例，当时秒杀所有WAF的。

小新



waf的案例倒是有一个，XXX的SQL注入：



5

豌豆妹

那处理waf漏报、误报的问题都有哪些成熟的方案呢？

哆啦A梦

这是两个非常大的课题，也是waf体系建设必不可少的。
在互联网企业，waf每天拦截的请求量都是千万级别的，发现waf误报的运营工作对waf建设很重要，量很大，不可能一条条看，必定需要集合自动化。
发现waf的漏报，更是在亿级别的请求海量数据发现waf未拦截的请求。而且有个悖论：waf未拦截，说明攻击特征有问题，所以不能从攻击特征的角度去发现waf的漏报。如果从攻击特征描述，怎么保证这次的特征不被绕过。
（1）waf漏报的发现思路
* 对于漏报来说，一个是要靠完整的协议和边界测试，观察waf在各种HTTP请求的绕过，变形下是否能正常工作；
* 另一个方法可以采用只检测不拦截模式上一些比较粗的规则，整理其中比较好的规则策略再上到拦截模式。
（2）waf误报的发现思路：
* 首先，进行人机识别，把人的请求和机器请求区分，机器的请求都不是误报；
* 其次，其次对人的请求进行攻击指纹识别；
* 最后，对剩下的拦截日志进行HMM模型(隐马尔可夫模型)分析。

豌豆妹

做waf的时候除了规则，有没有使用算法和机器学习的实践？

哆啦A梦

有使用过算法和机器学习。一开始算法是在离线环境跑，看看效果。

豌豆妹

使用的什么算法呢？

哆啦A梦

HMM模型(隐马尔可夫模型)分析，用于误报分析和漏报分析。

豌豆妹

哈~干货多多，意犹未尽呢！别走开~下周五我们继续深入的聊一聊waf那些事儿！(。·∀·)ﾉ

安全小课堂往期回顾：

1、论安全响应中心的初衷；
2、安全应急响应中心之威胁情报探索；
3、论安全漏洞响应机制扩展；
4、企业级未授权访问漏洞防御实践；
5、浅谈企业SQL注入漏洞的危害与防御；
6、信息泄露之配置不当；
7、XSS之攻击与防御；
8、电商和O2O行业诈骗那些事儿（上）；
9、电商和O2O行业诈骗那些事儿（下）；
10、CSRF的攻击与防御；
11、账户体系安全管理探讨；
12、远程代码执行漏洞的探讨；
13、服务器安全管控的探讨；
14、畅谈端口安全配置；
15、谈一谈github泄露；
16、撞库攻击是场持久战；
17、url重定向攻击的探讨；
18、聊聊弱口令的危害（一）；
19、聊聊弱口令的危害（二）；
20、聊聊XML注入攻击；
21、聊聊暴力破解；
22、谈谈上传漏洞；
23、浅谈内网渗透；
24、聊聊短信验证码安全。

京东安全应急响应中心