WordPress Security Plugins vs. WAF Services

A Comparative Test of WAF Accuracy Rates in Security Solutions
## Contents

Abstract ........................................................................................................................................... 3  
Introduction .................................................................................................................................... 3  
WAF Evaluation Criteria .................................................................................................................. 4  
Results ............................................................................................................................................... 5  
Pattern Analysis ............................................................................................................................... 6  
  True Positives ................................................................................................................................. 6  
  False Positives ............................................................................................................................... 8  
Conclusion ......................................................................................................................................... 9  
Appendix .......................................................................................................................................... 10
Abstract

This report contains the results of a comparative penetration test, the practice of testing a system in the same way a hacker would to identify security weaknesses, conducted by the research team at Cloudbric against two WordPress security plugins, Wordfence Security ("Wordfence") and All In One WP security & Firewall ("All In One"), and Cloudbric’s website protection service. Unlike other web application security tests that generate only attack traffic for testing, this test generated not only attack traffic but also legitimate traffic in order to test the ability of the WAF to detect web attacks and distinguish malicious traffic from legitimate traffic. Through testing, the research team concluded that Cloudbric has the highest true positive rate and the lowest false positive rate when tested against these security plugins. The results show that Cloudbric has superior detection capabilities in terms of both depth and accuracy. This is because Cloudbric uses a “logic-based detection engine” to understand the structure and meaning of the attack by analyzing traffic. This report also provides a pattern analysis for purpose to figure out why these patterns causes WAFs to generate false negatives or false positives.

Introduction

There is no doubt that WordPress is the CMS (Content Management System) of choice for most people. Because of its open source nature and ease of use, it has become a popular target for hackers. Therefore, many WordPress site operators have begun installing security plugins that offers WAF (Web Application Firewall) features as an attempt to secure their website. However, the question that remains is - can one protect a website from dangerous web attacks with a simple WAF provided by the plugin?

In order to investigate the quality of WAF security provided by security plugins, the research team at Cloudbric conducted a comparative penetration test, the practice of testing a system in the same way a hacker would to identify security weaknesses. This report will show the results of a test comparing the WAF detection capabilities of popular security plugins Wordfence Security ("Wordfence") and All In One WP Security & Firewall ("All In One"), against Cloudbric’s website protection service.

Wordfence & All In One

Wordfence is one of the most downloaded WordPress security plugins with one million downloads and a rating of 4.9/5. It provides security features including WAF, malware scanner, and login security. Though a free plugin, Wordfence also offers a premium version covering country blocking, remote scans and more. For this particular test, Wordfence plugin free version 6.2.5 was used.

All In One is another popular WordPress security plugin with over 500,000 downloads. It differs from Wordfence in offering greater customizability of security deployed. By enabling more security features than the default, users can increase site security levels. While All In One’s default security setting for firewall was ‘OFF’, all options in ‘Basic Firewall Rules’, ‘Additional Firewall Rules’, ‘6G Blacklist Firewall Rules’, ‘Internet Bots’ and ‘Prevent Hotlinks’ tab of ‘Firewall Settings’ were set to ‘ON’ for the test (refer to figure below). All In One plugin version 4.1.9 was used.
Cloudbric, is a cloud-based security service offering WAF, SSL and DDoS protection. A comprehensive list of premium security features is provided to all users regardless of pricing as usage is priced on monthly website traffic. In any case, the test website used for this research qualifies for free Cloudbric usage.

**WAF Evaluation Criteria**

Before testing the plugins, there are two criteria that should be considered when choosing a WAF service: True Positives and False Positives. True positives are cases in which the WAF correctly identifies attack traffic and blocks it. On the other hand, if the WAF failed to identify and block attack traffic, this instance is referred to as a false negative. Therefore, the true positive rate represents the WAF’s detection ability. The higher the true positive rate, the more kinds of attacks the WAF is capable of blocking.

False positives are cases in which the WAF incorrectly identifies legitimate traffic as an attack. The opposite case, where legitimate traffic passes through as it should, is called True Negative (or often termed “Normal”). If the false positive rate is too high, it means the WAF is not sensitive enough to distinguish legitimate traffic from malicious traffic.
There is a trade-off between false negatives and false positives. The more restrictive the rules configured in a WAF, the more likely malicious traffic will be blocked (low false negatives). Consequently, however, more legitimate traffic may be mistakenly blocked as well (high false positives). On the contrary, the more tolerant the rules configured in a WAF, the more freely legitimate traffic can access a site (low false positives). Similarly, malicious traffic from hackers will less likely be blocked (high false negatives).

In conclusion, an accurate WAF is one that prevents more web attacks while blocking fewer legitimate users. In other words, a WAF with a higher true positive rate and lower false positive rate is ideal.

Therefore, it is important to consider both True Positives and False Positives to evaluate the performance of a WAF. However, most existing web application security tests focus only on attack detection abilities. Unlike those tests, this test generated legitimate traffic along with attack traffic to test the WAF’s ability to correctly detect web attacks and distinguish malicious traffic from legitimate traffic.

Results

This test was conducted using 1524 web attack patterns published by Exploit-DB (“EDB”) and 661 legitimate patterns developed by Cloudbric (see Appendix).

EDB (www.exploit-db.com) is a comprehensive database of exploit codes maintained by Offensive Security, an information security training company that provides various information security certifications as well as high end penetration testing services.

The true positive rate achieved against attack patterns published by EDB was found to be significantly higher through Cloudbric than through the two plugins. Cloudbric detected 87% of the attacks, on average, a 41% higher true positive rate against EDB attack patterns than the two plugins. Cloudbric also had 1% false positives compared to 3% for Wordfence and 17% for All In One. While both Cloudbric and Wordfence demonstrated few false positives, All In One blocked legitimate traffic 113 times, causing customer inconvenience.

![True Positive vs. False Positive Rates](image)
Pattern Analysis

In order to figure out why some patterns cause false negatives or false positives results, patterns which leads false negatives or false positives are analyzed. The following patterns used method that pass data in the query portion of the requested URL.

True Positives

While there were several attack patterns that only Cloudbric was able to correctly detect, below is a thorough explanation of two particular attack patterns that the other WordPress security plugins’ WAF failed to detect.

Basic LFI attack pattern

Pattern 1: Basic pattern including NULL character
http://test.com/pages/../../../../../../../../../../etc/passwd%00

Pattern 1 is a prime example of avoiding detection by adding meaningless characters to a basic LFI (Local File Inclusion) attack pattern.

LFI is an attack that targets files that are already locally present on the server. In the attack pattern example above, hackers can read the “/etc/passwd” file after accessing the top level directory using ‘./’(move to parent directory). This file contains basic information necessary for hackers to gain entry into the system and access user account information. Though it is regarded as a basic pattern among LFI attack patterns, Wordfence and All In One did not detect this attack.

SQL injection attack pattern

SQL injection is another common web attack, and can be fatal, causing all data to be leaked or deleted. In the worst- case scenario, it can completely compromise the application system. Despite its potentially catastrophic consequences, the principle behind SQL injection is quite simple.

For example, in a standard login process, the user enters an ID and password on the site, and then a server verifies whether this information matches the information that was stored in the database.
However, in login processes affected by SQL injection, a hacker enters SQL statements that will behave abnormally in the database, rather than normal data such as an ID or password. For example, if "or 1=1 --" is inserted in the ID or password field, the server will recognize the condition as true even if the password is not correct (because 1=1 is always true), and will output the information requested.

Within the test, there were several SQL injection patterns that Cloudbric was able to detect correctly, but the other plugins could not detect.

**Pattern 2: Basic SQL Injection attack pattern**

http://test.com/admin/admin.php?username=admin%27+and+substring%28user%28%29,1,4%29=%27root%27+--&session=c8d7ebc95b9b1a72d3b54eb59bea56c7*

Decoding this pattern:

http://test.com/admin/admin.php?username=admin’ and substring(user(),1,4) =’root’ -- &session=c8d7ebc95b9b1a72d3b54eb59bea56c7*

In the above pattern, ‘and substring(user(),1,4)=’root’ -- &session=c8d7ebc95b9b1a72d3b54eb59bea56c7*’ is inserted in the username field. When this pattern works, the code after the ' --' is commented out, thus preventing it from being interpreted, and the following query is executed into MySQL:

```
Select * from users where id=’admin’ and substring(user(),1,4)=’root'
```

This is a SQL SELECT statement with the common structure “Select [Field(column)] from [Table_name] where [Condition]”. It can be understood as a command to retrieve information about a user whose ID is 'admin' and 'substring (user (), 1,4)' is 'root' from the 'users' table. Therefore, the database server that receives this query will look for a user according to those requirements. With this attack pattern, ‘admin’ can be replaced with any other username registered on the site. This pattern is an attack because of the 'substring (user (), 1,4) =' root '' syntax.

To break down this pattern for easier understanding, substring (str, start, length) is a MySQL default method that returns a portion of a string, extracting 'length' number of consecutive characters, beginning with the character in the 'start' position. For example, substring ('abcde', 2, 3) returns 'bcd' because it will return three characters from the character in the second position within the string 'abcde'. Next, user(), which is used as an argument of substring() in the above pattern, is a function of MySQL that returns the username of the database. Combined, substring (user (), 1, 4) means the string from the first to the fourth character of the database username. If the database username is root, substring (user (), 1, 4) will also be 'root.'

The purpose of this attack is to confirm if the database username is 'root'. This is because the most commonly used database user name in SQL is 'root'. If the database user name is not 'root', the hacker would receive a pop up informing him that there is no matching information. If the database username is indeed 'root', the hacker will be prompted to enter a password. In this way, hackers can gain information about the site by using the fact that the response of the site varies according to the input value. This attack is a prelude to a full-blown SQL injection attack.
Pattern 2 is a relatively well-known SQL injection basic pattern. Pattern 3 is a modified SQL injection attack pattern with complex parentheses as shown below:

Pattern 3: Modified SQL Injection attack pattern with multiple parentheses
http://test.com/40/PetRatePro/index.php?cmd=4tand(select1from(selectcount(*),concat((selectconcat((CHAR(52),CHAR(67),CHAR(117),CHAR(121),CHAR(82),CHAR(65),CHAR(101),CHAR(74),CHAR(100),CHAR(109),CHAR(55)) from information_schema.tables limit 0,1),floor(rand(0)*2))x frominformation_schema.tables group by x)a)and'

Pattern 3 attempts SQL injection using double, triple complex parentheses. The key code hidden in multiple parentheses is simply:

```
select concat(CHAR(52), ... ,CHAR(109),CHAR(55)) from information_schema.tables limit 0,1
```

This syntax consists of a `select A from B` SQL SELECT statement, which extracts the 'A' column from the 'B' table. Within ‘A’, ‘concat’ is used to add strings together and CHAR(52) ... CHAR(55), ASCII code, are used as those strings. It is intended to represent column names in a complex way by using ASCII code. On the other hand, ‘information_schema.tables’, which is equivalent to ‘B’ in the syntax, is the default database for MySQL and has the entire structure and content of the database. When this table is exposed, manipulating or releasing data on the website is facilitated. Therefore, attacks of this pattern attempt to subtract some information from the information_schema.tables. Because of complex parentheses and the obfuscation of column names via ‘A’, it is difficult to detect this type of attack pattern.

False Positives

Below is a thorough explanation of two legitimate patterns that WordFence or All In One blocked.

WordFence and All In One False Positive Patterns

The above pattern is a legitimate pattern that only Cloudbric has properly identified. The decoded pattern is as below:

```
http://test.com/865/sendmsg.jsp?subj=&text=There are many possible encodings for data: UTF-7, UTF-8, base64 and so on&Username=Anonymous &x=32&y=7
```

WordFence and All In One identified this pattern as an attack due to the value in text, suggesting that "There are many possible encodings for data: UTF-7, UTF-8, base64 and so on" in the pattern. This is a simple sentence, but the WAF considers this sentence as [There ... Base 64] and [so on], in other words, an attack with ‘A’ and ‘B’ structures. The ‘A and B’ structures are especially common for SQL injection attacks. In this case, the pattern corresponding to ‘A’ contains several words frequently used in attack patterns such as 'encodings', ';', 'UTF-7' and 'UTF-8', triggering false positives. In reality, these patterns are harmless when embedded as text; hence, it should not have been blocked. While traffic may contain patterns commonly associated with attacks, Cloudbric does not block it unless such patterns are determined to accompany a malicious intent to attack the website. In pattern 4, the text input triggering the plugins’ false positives were not intended as an attack, and so Cloudbric allowed it to access the website. WAFs that block the above pattern are likely to generate many false positives for websites with forum postings or discussions about coding.
All In One False Positive Patterns

Pattern 5: All In One False Positive Patterns
http://test.com/board.php?text=lsof%20is%20a%20command%20meaning%20"list%20of%20open%20files%22%20which%20is%20used%20in%20many%20Unix-like%20systems%20to%20report%20a%20list%20of%20all%20open%20files%20and%20the%20processes%20that%20opened%20them.

This pattern is a legitimate pattern that only the All In One has incorrectly detected. The decoded pattern is:

http://test.com/board.php?text=lsof is a command meaning "list open files", which is used in many Unix-like systems to report a list of all open files and the processes that opened them.

The element triggering the false positive is "Isol". Isol is a system command that prints all the information about the files utilized by the system. This command is useful when managing a system, but it can lead to an attack that leaks information about all files when manipulated. Some WAFs register Isol as a signature and block the request containing this word. For All In One, it has identified this particular pattern as an attack. However, when blocking a pattern containing Isol, WAF needs to make sure that this command is intended to be an attack, which is only when an additional filename or directory name follows Isol. Looking back at the above pattern, Isol does not have any effect on the system because no filename or directory name follows. Therefore, the above pattern is a legitimate pattern that should not have been identified as an attack.

Conclusion

The penetration test has shown that the WAF of WordPress security plugins performed poorly in terms of both true positives and false positives as compared to Cloudbric. All In One, especially, showed the lowest true positive rate and the highest false positive rate. The results demonstrated that Cloudbric’s WAF provides more effective and accurate security, delivering more powerful attack detection while retaining a low false positives rate.

Cloudbric uses not only a “signature-based detection method” which identifies attacks by finding a match against a database of existing attack patterns, but also a “logic-based detection engine” which analyzes traffic through a set of 26 preset security policies. This is the main reason why there were several patterns which only Cloudbric was able to identify correctly, while the other plugins failed to detect or even mistakenly blocked. Utilizing a logic-based detection engine to understand the structure and meaning of the attack enables more accurate detection of hidden or modified attacks.
Appendix

The following patterns (legitimate) were used to test false positives. If you want a complete list of test patterns, then please contact support@cloudbric.com.

<table>
<thead>
<tr>
<th>No.</th>
<th>Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>/board.php?id=1%20UNION+ON+SELECT+1,2,3,4,5,6,7--</td>
</tr>
<tr>
<td>2</td>
<td>/sendmsg.jsp?subj=&amp;text=select+one+item+from+these+funny+things+%3D%29&amp;Username=Anonymous&amp;x=32&amp;y=7</td>
</tr>
<tr>
<td>3</td>
<td>/sendmsg.jsp?subj=&amp;text=message+is+not+good%3B+drop+the+item+you+are+holding&amp;Username=Anonymous&amp;x=32&amp;y=7</td>
</tr>
<tr>
<td>4</td>
<td>/sendmsg.jsp?subj=&amp;text=I+have+selected+an+instrument+%28a+new+one%29+for+today&amp;Username=Anonymous&amp;x=32&amp;y=7</td>
</tr>
<tr>
<td>5</td>
<td>/sendmsg.jsp?subj=&amp;text=send+me+a+check+or+100000%24+cash&amp;Username=Anonymous&amp;x=32&amp;y=7</td>
</tr>
<tr>
<td>6</td>
<td>/sendmsg.jsp?subj=&amp;text=There+are+many+possible+encodings+for+data%3A+UTF-7%2C+UTF-8%2C+base64+and+so+on&amp;Username=Anonymous&amp;x=32&amp;y=7</td>
</tr>
<tr>
<td>7</td>
<td>/sendmsg.jsp?subj=&amp;text=An+additional+important+file+is+boot.ini,+which+contains+boot+configuration+(if+missing,+NTLDR+will+default+to+Windows+on+the+first+partition+of+the+first+hard+drive).</td>
</tr>
<tr>
<td>8</td>
<td>/sendmsg.jsp?subj=&amp;text=Apache+is+a+marvelous+web+server+that+offers+.htpasswd+and+.htgroup+for+controlling+restricted+access+to+your+website.</td>
</tr>
<tr>
<td>9</td>
<td>/Extension_Filtering/test.exe</td>
</tr>
<tr>
<td>10</td>
<td>/Extension_Filtering/test.png</td>
</tr>
<tr>
<td>11</td>
<td>/Extension_Filtering/test.ico</td>
</tr>
<tr>
<td>12</td>
<td>/Extension_Filtering/test.asf</td>
</tr>
<tr>
<td>13</td>
<td>/Extension_Filtering/test.tif</td>
</tr>
<tr>
<td>14</td>
<td>/Extension_Filtering/test.zip</td>
</tr>
<tr>
<td>15</td>
<td>/Extension_Filtering/test.aspx</td>
</tr>
<tr>
<td>16</td>
<td>/Extension_Filtering/test.rar</td>
</tr>
<tr>
<td>17</td>
<td>/board.php?text=An%20additional%20important%20file%20is%20boot.ini%3B%20which%20contains%20boot%20configuration%20(if%20missing%2C%20NTLDR%20will%20default%20to%20%3Findows%20on%20the%20first%20partition%20of%20the%20first%20hard%20drive).</td>
</tr>
<tr>
<td>18</td>
<td>/board.php?text=Apache%20is%20a%20marvelous%20web%20server%20that%20offers%20.htpasswd%20and%20.htgroup%20for%20controlling%20restricted%20access%20to%20your%20website.</td>
</tr>
</tbody>
</table>
Cloudbric Whitepaper

Find out what telnet.exe is doing on computer, whether it's safe, info on related errors and how to remove it.

Many people just have this wrong notion that windows command line FTP.exe can not be run in passive mode.

So Typescript%20compiled%20javascript%20actually%20works%20in%20WSH.exe? Thats so awesome.

Typescript%20compiled%20javascript%20actually%20works%20in%20WSH.exe.

Information about the Unix kill command, including examples and syntax.

Used without parameters, net localgroup displays the name of the server and the names of local groups on the computer.

National Association of Schools of Music. Welcome to the NASM website.

I’m using tcl8.4 and I’m trying to do a little tclsh scripting. I’m trying to execute a shell command using a tcl variable that has embedded spaces.

In computer networking, the Name/Finger protocol and the Finger use information protocol are simple network protocols for the exchange of human-oriented status and user information.

lsof is a command meaning “list open files” which is used in many Unix-like systems to report a list of all open files and the processes that opened them.

Information and examples on the Unix and Linux rm command.

Python is a widely used general-purpose, high-level programming language.

How to perform a traceroute to a domain or website - this will help identify connection problems.

Information about the Unix and Linux chmod command including the full syntax and several different examples.

onSubmit is a scripting event that occurs when the user attempts to submit the form to the CGI.

Tutorials and over 400 free scripts. Also JavaScript objects, properties, and methods reference.

Visual Meta GmbH operates multiple shopping platforms across Europe under the brands LadenZeile and ShopAlike.

The ActiveXObject object is used to create instances of OLE Automation objects in Internet Explorer on Windows operating systems.

VBScript (Visual Basic Scripting Edition) is an Active Scripting language developed by Microsoft that is modeled on Visual Basic.

The onkeypress event occurs when the user presses a key (on the keyboard).

This event triggers when a JavaScript ‘unescape()’ or ‘eval()’ contains the ‘String.fromCharCode()’.

For example, a script could generate a popup alert box message, or provide a dropdown menu.

An application is a program or group of programs designed for end users.

The background CSS property is a shorthand for setting the individual background values in a single place in the style sheet.

I have been using GetSpecialFolder from madshell unit, I looked at the source examples provided with madcollection documentation.

For the most part, customization is different in the new Disqus because we decided to completely re-implement our commenting embed inside of an iframe.

print is not actually a real function (it is a language construct) so you are not required to use parentheses with its argument list.
Review the categories of user that might want to use with Citrix, and how to configure for concurrent multi-users.

Locate finds files and directories based on file and folder names stored in a database.

The EXP function returns the exponential of the value specified by NumericExpression. The parameter can be any built-in numeric data type.

CURRENT_TIMESTAMP returns a TIMESTAMP value representing the current date and local time.

Returns a new string with some or all matches of a pattern replaced by a replacement.

The @XmlType annotation can be used with the following program elements.

The CHARACTER_LENGTH function returns the length of the first argument in the specified string unit.

CAST addresses issues of animal sciences, food sciences and agricultural technology, plant and soil sciences, and plant protection sciences with input from throughout the scientific and legal community.

A varchar or Variable Character Field is a set of character data of indeterminate length. 

Look up position in Wiktionary, the free dictionary. Position refers to the spatial location (rather than orientation) of an entity.

The DB2 Administration Tool panel ADB21TAB "Alter Table" does not qualify the SELECT from SYSIBM.SYSCOLUMNS with the TBCREATOR.

Remove WEIGHT_STRING() from parser (where it does not belong). If someone wants to they can reimplement this as a straight function.

A microsecond is an SI unit of time equal to one millionth (10^-6 or 1/1,000,000) of a second. Its symbol is . One microsecond is to one second as one second is to 11.574 days.

The SYSTEM_USER special register specifies the authorization ID of the user that connected to the database.

In a SQL Server 2005 trace, I've got lots of "exec sp_execute" statements. I know they are connected to a corresponding "exec sp_prepare" statement which specifies the actual SQL.

I found the command "SELECT CONNECTION_ID()" in mysql which returns the connection ID with the server. Is there any equivalent command in PostgresSQL?

The encryption key to use is chosen based on the second argument to DES_ENCRYPT(), if one was given.

It's burdened by a predictable, overly melodramatic story, but The Greatest benefits from strong performances by its talented cast.

Reading and Writing Server-side Files. UTL_FILE is a package that has been welcomed warmly by PL/SQL developers.

If the first and only parameter is an array, min() returns the lowest value in that array.

Cloudbric Whitepaper | 13
<table>
<thead>
<tr>
<th>Line</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>96</td>
<td>This article mentions How to Enable &amp; Disable XP_CMDSHELL using SP_CONFIGURE.</td>
</tr>
<tr>
<td>97</td>
<td>This will allow you to call sp_executesql with @eStatus as a parameter instead of embedding it into the SQL.</td>
</tr>
<tr>
<td>98</td>
<td>THE EUROPEAN EXTREMELY LARGE TELESCOPE (&quot;ELT&quot;) PROJECT.</td>
</tr>
<tr>
<td>99</td>
<td>The new Oracle built-in package DBMS_JAVA gives you access to and the ability to modify various characteristics of the Aurora Java Virtual Machine.</td>
</tr>
<tr>
<td>100</td>
<td>Serene is a truly beautiful blogging theme with post format support.</td>
</tr>
</tbody>
</table>