HTTP Fingerprinting and Advanced Assessment Techniques

Saumil Shah
Director, Net-Square

Author: “Web Hacking - Attacks and Defense”

BlackHat Asia 2003, Dec 18, Singapore
The Web Hacker’s playground
The Evolution of Web Defense

- Tight web server configuration.
- Web server plug-in filters.
- Secure coding.
- Security by obscurity.
Security by obscurity

- Who is running IIS? ... Not me!
- Web server target acquisition:
  - largely by banner grabbing

$ nc 192.168.7.247 80
HEAD / HTTP/1.0

HTTP/1.1 200 OK
Server: Microsoft-IIS/5.0
Content-Location: http://192.168.7.247/Default.htm
Date: Fri, 01 Jan 1999 20:09:05 GMT
Content-Type: text/html
Accept-Ranges: bytes
Last-Modified: Fri, 01 Jan 1999 20:09:05 GMT
ETag: W/"e0d362a4c335be1:ae0"
Content-Length: 133
Security by obscurity

• Patch web server binaries to change server banner.
  • e.g. “Microsoft-IIS/5.0” rewritten to be “Apache/1.3.26”
• If source is available, recompile with different server banner.
  • e.g. “Apache/1.3.26” rewritten to be “WebSTAR”
• Works well in defeating certain automated attacks / script kiddies.
Security by obscurity

• Web server configuration rules / plug-ins to disguise the server header.
• Re-order HTTP header fields, change cookie names, filter certain responses, etc.

$ nc 192.168.7.247 80
HEAD / HTTP/1.0

HTTP/1.1 200 OK
Date: Fri, 01 Jan 1999 20:06:24 GMT
Server: Apache/1.3.19 (Unix) (Red-Hat/Linux) mod_ssl/2.8.1 OpenSSL/0.9.6 DAV/1.0.2 PHP/4.0.4pl1 mod_perl/1.24_01
Content-Location: http://192.168.7.247/Default.htm
Last-Modified: Fri, 01 Jan 1999 20:06:24 GMT
ETag: W/"e0d362a4c335be1:ae0"
Accept-Ranges: bytes
Content-Length: 133
Content-Type: text/html

with ServerMask 2.0
HTTP Fingerprinting

- Objective: To accurately determine the underlying web server platform.
- Also attempt to uncover any plug-ins, app servers, etc.
- Based on implementation assumptions / peculiarities of the HTTP protocol spec.
HTTP Fingerprinting

- Fingerprinting logic
  - Decision-tree based methods
  - Statistical methods
  - Neural Network based methods
- Fingerprinting engine
  - Set of test cases, carefully chosen
  - Response-tree
  - Weight vectors
HTTP Fingerprinting Techniques

- Deviation from HTTP RFCs
- Behaviour not specified by the HTTP RFCs
- Default behaviour
- Header field order
- Implementation peculiarities
- Error analysis
- Cookie strings
- … similar to OS fingerprinting
HTTP Fingerprinting - Accuracy

- Choice of test cases
- Decision-trees are hard to scale
- Choice of result weights
- Scoring system
- Training input set (for neural networks)
httprint - advanced web fingerprinting

- Identifies web servers, despite obfuscation.
  - Does not rely on banner strings.
  - No string comparison, whatsoever.
  - Overcomes the customisable features of HTTP.
- Paper:
httprint – Features

- Available in GUI and command-line
  - Windows, Linux, FreeBSD and Mac OS X
  - Download from:
    http://net-square.com/httprint/
- Can easily add server signatures.
- Can import from nmap XML output - selects known web server ports automatically.
- New confidence rating technique.
httpprint - Confidence Ratings

- Improved Fuzzy Logic method for picking the best match.
- Allows us to identify false positives.
- Indicates if a server signature is not present within the set of known signatures.
httprint - example 1

Banner Reported: Apache-AdvancedExtranetServer/2.0.44 (Mandrake Linux/11mdk) mod_perl/1.99_08 Perl/v5.8.0 mod_ssl/2.0.44 OpenSSL/0.9.7a PHP/4.3.1

Banner Deduced: Apache/2.0.x

<table>
<thead>
<tr>
<th>Web Server</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apache/2.0.x:</td>
<td>126</td>
<td>81.29</td>
</tr>
<tr>
<td>Apache/1.3.[4-24]:</td>
<td>118</td>
<td>64.83</td>
</tr>
<tr>
<td>Apache/1.3.27:</td>
<td>117</td>
<td>62.94</td>
</tr>
<tr>
<td>Apache/1.3.26:</td>
<td>116</td>
<td>61.09</td>
</tr>
<tr>
<td>Apache/1.3.[1-3]:</td>
<td>113</td>
<td>55.74</td>
</tr>
<tr>
<td>Apache/1.2.6:</td>
<td>113</td>
<td>55.74</td>
</tr>
<tr>
<td>Agranat-EmWeb:</td>
<td>72</td>
<td>10.29</td>
</tr>
<tr>
<td>Stronghold/4.0-Apache/1.3.x:</td>
<td>66</td>
<td>7.13</td>
</tr>
<tr>
<td>Netscape-Enterprise/4.1:</td>
<td>59</td>
<td>4.28</td>
</tr>
<tr>
<td>Com21 Cable Modem:</td>
<td>56</td>
<td>3.31</td>
</tr>
<tr>
<td>Oracle Servlet Engine:</td>
<td>55</td>
<td>3.02</td>
</tr>
<tr>
<td>Microsoft-IIS/6.0:</td>
<td>55</td>
<td>3.02</td>
</tr>
<tr>
<td>Microsoft-IIS/5.1:</td>
<td>55</td>
<td>3.02</td>
</tr>
<tr>
<td>Microsoft-IIS/5.0 ASP.NET:</td>
<td>55</td>
<td>3.02</td>
</tr>
</tbody>
</table>

No obfuscation.
Verification of testing.
Banner Reported: WebSTAR
Banner Deduced: Apache/2.0.x

<table>
<thead>
<tr>
<th>Server Type</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apache/2.0.x</td>
<td>80</td>
<td>51.61</td>
</tr>
<tr>
<td>Com21 Cable Modem</td>
<td>79</td>
<td>49.15</td>
</tr>
<tr>
<td>Apache/1.3.26</td>
<td>78</td>
<td>46.76</td>
</tr>
<tr>
<td>Apache/1.3.27</td>
<td>78</td>
<td>46.76</td>
</tr>
<tr>
<td>Apache/1.3.[4-24]</td>
<td>77</td>
<td>44.44</td>
</tr>
<tr>
<td>Apache/1.3.[1-3]</td>
<td>72</td>
<td>33.94</td>
</tr>
<tr>
<td>SMC Wireless Router 7004VWBR</td>
<td>72</td>
<td>33.94</td>
</tr>
<tr>
<td>dwhttpd (Sun Answerbook)</td>
<td>72</td>
<td>33.94</td>
</tr>
<tr>
<td>EMWHTTPD/1.0</td>
<td>68</td>
<td>26.76</td>
</tr>
<tr>
<td>Microsoft-IIS/5.1</td>
<td>68</td>
<td>26.76</td>
</tr>
<tr>
<td>Microsoft-IIS/5.0 ASP.NET</td>
<td>68</td>
<td>26.76</td>
</tr>
<tr>
<td>Netscape-Enterprise/4.1</td>
<td>64</td>
<td>20.60</td>
</tr>
<tr>
<td>AOLserver/3.5.6</td>
<td>62</td>
<td>17.87</td>
</tr>
<tr>
<td>Apache/1.2.6</td>
<td>62</td>
<td>17.87</td>
</tr>
<tr>
<td>TightVNC</td>
<td>59</td>
<td>14.20</td>
</tr>
<tr>
<td>Microsoft-IIS/5.0</td>
<td>58</td>
<td>13.09</td>
</tr>
</tbody>
</table>

Recompiled Apache banner patching.
Easy to tell.
Banner Reported: Apache/1.3.23 (Unix)
Banner Deduced: Microsoft-IIS/5.0

Microsoft-IIS/5.0: 124 80.00
Microsoft-IIS/5.1: 119 69.34
Microsoft-IIS/5.0 ASP.NET: 119 69.34
Microsoft-IIS/4.0: 92 27.43
Microsoft-IIS/6.0: 82 17.67
Apache/1.3.26: 79 15.26
Apache/1.3.27: 78 14.50
Apache/1.3.[1-3]: 77 13.77
Apache/1.2.6: 77 13.77
Apache/1.3.[4-24]: 77 13.77
Netscape-Enterprise/4.1: 70 9.27
Apache/2.0.x: 69 8.72
Agranat-EmWeb: 67 7.67
Oracle Servlet Engine: 63 5.81
Microsoft-IIS/URLScan: 56 3.24
Com21 Cable Modem: 55 2.94

Servermask:
80% confidence that the underlying server is Microsoft-IIS/5.0
httprint – Win32 GUI interface
Slick HTML reports!

Now in CSV format as well. Import at will!

XML reports in enterprise version.
httprint - Applications

- Proper target selection for assessment.
- Detect 802.11 APs from the wired network.
- Web enabled devices (printers, storage, etc)
httpprint - Real Disguised Servers

- Report below shows five real servers:

<table>
<thead>
<tr>
<th>host</th>
<th>port</th>
<th>ssl</th>
<th>banner reported</th>
<th>banner deduced</th>
<th>icon</th>
<th>confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.walmart.com">www.walmart.com</a></td>
<td>80</td>
<td></td>
<td>Microsoft-IIS/5.0</td>
<td>Apache/2.0.x</td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://www.foundstone.com">www.foundstone.com</a></td>
<td>80</td>
<td></td>
<td>WebSTAR</td>
<td>Apache/2.0.x</td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://www.port80software.com">www.port80software.com</a></td>
<td>80</td>
<td></td>
<td>Yes we are using ServerMask</td>
<td>Microsoft-IIS/5.1, Microsoft-IIS/5.0 ASP.NET, Microsoft-IIS/4.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://www.ubizen.com">www.ubizen.com</a></td>
<td>80</td>
<td></td>
<td>web server</td>
<td>Apache/2.0.x</td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://www.datek.com">www.datek.com</a></td>
<td>80</td>
<td></td>
<td>Ameritrade Web Server</td>
<td>Netscape-Enterprise/4.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- In accordance with Netcraft results.
HTTP Response Codes

- Customised error pages.
- A non-existent page should return an HTTP 404 code.
- Many servers return:
  - 301/302 - redirect to some starting page
  - 200 OK - to fool crawlers
  - ...and other customised codes.
Page Signatures

- **Objective:** To accurately identify proper HTTP response codes.
- Minimize false positives.
- Greatly helps in automated testing.
- Can be extended beyond error detection
  - e.g. group similar pages together
Page Signatures

- Each HTTP response has a page signature.
- Content independent.
- Ability to overlook random content.
- Constant length.
- Computation time: $O(n)$
- Comparision time: $O(k)$
Normal error page

Not Found

The requested URL /junk was not found on this server.

Apache/1.3.26 Server at 192.168.7.70 Port 8222

$ nc 192.168.7.70 8222
GET /junk HTTP/1.0

HTTP/1.1 404 Not Found
Date: Tue, 04 Feb 2003 06:22:00 GMT
Server: Apache/1.3.26 (Unix) mod_perl/1.26 mod_ssl/2.8.9 OpenSSL/0.9.6e
Connection: close
Content-Type: text/html; charset=iso-8859-1

<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
<html><head>
title="404 Not Found"
</head>
<body>
<h1>Not Found</h1>
The requested URL /junk was not found on this server.<p>
</body></html>
Customised error page

Sorry!

Random number: 1198361.67040379

The link you requested http://192.168.7.2/junk was not found

Please contact the site administrator at root@dev.null if you feel this is in error. Alternately, try searching with Google

In 1 minute, you will be refreshed back to the main page

Google™

$ nc 192.168.7.70 8222
GET /junk HTTP/1.0

HTTP/1.1 200 OK
Date: Tue, 04 Feb 2003 01:41:06 GMT
Server: Apache-AdvancedExtranetServer/1.3.19 (Linux-Mandrake/3mdk) mod_ssl/2.8.2
OpenSSL/0.9.6 PHP/4.0.4p11
Connection: close
Content-Type: text/html; charset=ISO-8859-1

<html><body><H1>Sorry!</H1><p>Random number: 318405.070147527<br>The link you requested http://192.168.7.2/junk was not found<br>Please contact the site administrator at root@dev.null if you feel this is in error. Alternately, try searching with Google<br>In 1 minute, you will be refreshed back to the main page<br><p><FORM method=GET action=http://www.google.com/search><IMG SRC=http://www.google.com/logos/Logo_40wht.gif border=0 ALT=Google align=absmiddle><INPUT TYPE=text name=q size=15 maxlength=255><INPUT type=submit name=btnG VALUE=Search></FORM></body></html>
Dealing with random content

- Page signatures are independent of content.


- All of the above are 404 pages.
- Though their content may change, their signature doesn’t.
Application Fingerprinting

- apprint - prototype application server fingerprinting tool.
- Identify the underlying Application Server plugged in to the front-end web server.
- e.g. Weblogic running on IIS, Tomcat running on Apache, etc.
- Based on Page Signatures.
Reverse Proxy Servers

- Web proxy servers may work both ways!
- Typically meant to allow users from within a network to access external web sites.
- May end up proxying HTTP requests from the outside world to the internal network.
- e.g. Compaq Insight Manager
- Usually happens when the front end web server proxies requests to back end app servers.
Reverse Proxying

GET http://10.0.1.2/ HTTP/1.0
Port Scanning through Proxies

- Issue multiple GET requests to the proxy:
  - GET http://10.0.0.3:21/ HTTP/1.0
  - GET http://10.0.0.3:25/ HTTP/1.0
  - GET http://10.0.0.3:135/ HTTP/1.0
  - GET http://10.0.0.3:139/ HTTP/1.0
- Use Page signatures to identify accurately if a port is open on an internal host.
Better CONNECTivity

- HTTP CONNECT can be used to open up a bi-directional TCP connection.
- Originally intended for SSL traffic.
- Often overlooked.
- Ability to tunnel arbitrary TCP data over an HTTP proxy connection.
- Once CONNECTed, the proxy simply passes the TCP data back and forth.
CONNECTivity through proxies

The problem: to connect to an internal Term Server

We know www.acmetravel.com:8001 allows reverse proxying.
CONNECT http://10.0.1.2:3389/ HTTP/1.0
RDP data passes through the established connection.
Automated Web Security Assessment

• The need for overcoming HTTP’s customisable aspects:
  • Server banner strings
  • Response codes
• Improving accuracy
• Using core concepts to extend assessment techniques
Closing Thoughts

- “You cant patch (or hide) carelessness”.
- Web Hacking: Attacks and Defense
  Saumil Shah,
  Shreeraj Shah,
  Stuart McClure
Thank you!

saumil@net-square.com

http://net-square.com/