Limitations of spoofing user-agent strings to download malware

April 4, 2010
rationallyparanoid.com

Wget is a non-interactive command-line web browser that is often used as a safer method for retrieving malware samples from malicious web sites. However malware authors are aware of the use of this tool and certain malicious web sites are configured to either send you a benign file or not respond properly unless they see a specific user-agent string from the web browser making the request (they also sometimes use the referrer string or examine the IP address making the request to decide whether or not to send the malicious file).

Most security researchers are aware of these techniques and make sure to spoof the user-agent string when using Wget in order to fool the web site into thinking that the request is coming from the targeted victim web browser. However as we will demonstrate below a HTTP request using Wget still contains tell-tale characteristics that a web site owner could use to identify the web browser making the request. Therefore we will show you a better method to impersonate another web browser.

The tests were done on Ubuntu Linux 9.10 with the following versions of the software:

Wget 1.11.4
Curl 7.19.5
Firefox 3.5.8

The web server in this example is Apache 2.2.12 and for demonstration purposes the malware will be located at http://example.com/update.exe

If you were to use Firefox 3.5.8 on Ubuntu 9.10 to retrieve the file http://example.com/update.exe, the web server in its logs would identify you as the following:

```
192.0.2.1 - - [28/Mar/2010:12:26:38] "GET /update.exe HTTP/1.1" 200 6186 "-"
"Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10
(karmic) Firefox/3.5.8"
```

If you were to use Wget instead of Firefox to make the same request, the logs would be as follows:

```
192.0.2.1 - - [28/Mar/2010:12:29:54] "GET /update.exe HTTP/1.0" 200 6186 "-" "Wget/
1.11.4"
```

The ability to spot Wget as the browser making the request is trivial, and this should come to no surprise for anybody who has viewed web server logs in the past. Fortunately with Wget you can use the --user-agent parameter to send a forged user-agent string to the web server. Below we will instruct Wget to spoof the user-agent to that of Firefox 3.5.8 on Ubuntu 9.10 while fetching http://example.com/update.exe:

```
user@linux:~$ wget --user-agent="Mozilla/5.0 (X11; U; Linux i686; en-US;
rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10 (karmic) Firefox/3.5.8"
http://example.com/update.exe
```

In the web server's logs the request from Wget is now identical to Firefox':

```
192.0.2.1 - - [28/Mar/2010:12:33:42] "GET /update.exe HTTP/1.0" 200 6186 "-"
"Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10
(karmic) Firefox/3.5.8"
```

But here is what some people may not know: Although the logs have Wget and Firefox looking the same, the HTTP requests sent by the two web browsers remain quite different. Below is the HTTP packet for Wget when spoofing the user-agent string:

In ASCII

```
GET /update.exe HTTP/1.0
User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214
Ubuntu/9.10 (karmic) Firefox/3.5.8
Accept: */*
Host: example.com
Connection: Keep-Alive
```

In Hex

```
0000   47 45 54 20 2f 75 70 64 61 74 65 2e 65 78 65 20  GET /update.exe
0010   48 54 54 50 2f 31 2e 30 0d 0a 55 73 65 72 2d 41  HTTP/1.0..User-A
0020   67 65 6e 74 3a 20 4d 6f 7a 69 6c 6c 61 2f 35 2e  gent: Mozilla/5.
0030   30 20 28 58 31 31 3b 20 55 3b 20 4c 69 6e 75 78  0 (X11; U; Linux
0040   20 69 36 38 36 3b 20 65 6e 2d 55 53 3b 20 72 76   i686; en-US; rv
0050   3a 31 2e 39 2e 31 2e 38 29 20 47 65 63 6b 6f 2f  :1.9.1.8) Gecko/
0060   32 30 31 30 30 32 31 34 20 55 62 75 6e 74 75 2f  20100214 Ubuntu/
0070   39 2e 31 30 20 28 6b 61 72 6d 69 63 29 20 46 69  9.10 (karmic) Fi
0080   72 65 66 6f 78 2f 33 2e 35 2e 38 0d 0a 41 63 63  refox/3.5.8..Acc
0090   65 70 74 3a 20 2a 2f 2a 0d 0a 48 6f 73 74 3a 20  ept: */*..Host:
00a0   65 78 61 6d 70 6c 65 2e 63 6f 6d 0d 0a 43 6f 6e  example.com..Con
00b0   6e 65 63 74 69 6f 6e 3a 20 4b 65 65 70 2d 41 6c  nection: Keep-Al
00c0   69 76 65 0d 0a 0d 0a                              ive....
```

Now contrast this with Firefox:

In ASCII

```
GET /update.exe HTTP/1.1
Host: example.com
User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214
Ubuntu/9.10 (karmic) Firefox/3.5.8
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300
Connection: keep-alive
```

In Hex

```
0000   47 45 54 20 2f 75 70 64 61 74 65 2e 65 78 65 20  GET /update.exe
0010   48 54 54 50 2f 31 2e 31 0d 0a 48 6f 73 74 3a 20  HTTP/1.1..Host:
0020   65 78 61 6d 70 6c 65 2e 63 6f 6d 0d 0a 55 73 65  example.com..Use
0030   72 2d 41 67 65 6e 74 3a 20 4d 6f 7a 69 6c 6c 61  r-Agent: Mozilla
0040   2f 35 2e 30 20 28 58 31 31 3b 20 55 3b 20 4c 69  /5.0 (X11; U; Li
0050   6e 75 78 20 69 36 38 36 3b 20 65 6e 2d 55 53 3b  nux i686; en-US;
0060   20 72 76 3a 31 2e 39 2e 31 2e 38 29 20 47 65 63   rv:1.9.1.8) Gec
0070   6b 6f 2f 32 30 31 30 30 32 31 34 20 55 62 75 6e  ko/20100214 Ubun
```

```
0080   74 75 2f 39 2e 31 30 20 28 6b 61 72 6d 69 63 29  tu/9.10 (karmic)
0090   20 46 69 72 65 66 6f 78 2f 33 2e 35 2e 38 0d 0a   Firefox/3.5.8..
00a0   41 63 63 65 70 74 3a 20 74 65 78 74 2f 68 74 6d  Accept: text/htm
00b0   6c 2c 61 70 70 6c 69 63 61 74 69 6f 6e 2f 78 68  l,application/xh
00c0   74 6d 6c 2b 78 6d 6c 2c 61 70 70 6c 69 63 61 74  tml+xml,applicat
00d0   69 6f 6e 2f 78 6d 6c 3b 71 3d 30 2e 39 2c 2a 2f  ion/xml;q=0.9,*/
00e0   2a 3b 71 3d 30 2e 38 0d 0a 41 63 63 65 70 74 2d  *;q=0.8..Accept-
00f0   4c 61 6e 67 75 61 67 65 3a 20 65 6e 2d 75 73 2c  Language: en-us,
0100   65 6e 3b 71 3d 30 2e 35 0d 0a 41 63 63 65 70 74  en;q=0.5..Accept
0110   2d 45 6e 63 6f 64 69 6e 67 3a 20 67 7a 69 70 2c  -Encoding: gzip,
0120   64 65 66 6c 61 74 65 0d 0a 41 63 63 65 70 74 2d  deflate..Accept-
0130   43 68 61 72 73 65 74 3a 20 49 53 4f 2d 38 38 35  Charset: ISO-885
0140   39 2d 31 2c 75 74 66 2d 38 3b 71 3d 30 2e 37 2c  9-1,utf-8;q=0.7,
0150   2a 3b 71 3d 30 2e 37 0d 0a 4b 65 65 70 2d 41 6c  *;q=0.7..Keep-Al
0160   69 76 65 3a 20 33 30 30 0d 0a 43 6f 6e 6e 65 63  ive: 300..Connec
0170   74 69 6f 6e 3a 20 6b 65 65 70 2d 61 6c 69 76 65  tion: keep-alive
0180   0d 0a 0d 0a                                      ....
```

It is easy to see that they are quite different, and some web sites may use this information to identify you as somebody trying to impersonate another browser. The table below compares the two:

Firefox	**Wget**
GET /update.exe HTTP/1.1 Host: example.com User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10 (karmic) Firefox/3.5.8 Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8 Accept-Language: en-us,en;q=0.5 Accept-Encoding: gzip,deflate Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7 Keep-Alive: 300 Connection: keep-alive	GET /update.exe HTTP/1.0 User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10 (karmic) Firefox/3.5.8 Accept: */* Host: example.com Connection: Keep-Alive

Wget supports the --header parameter to set additional headers, so we can use these to instruct Wget to spoof all of the same headers as Firefox. You can specify the --header parameter multiple times:

```
user@linux:~$ wget --header="User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US;
rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10 (karmic) Firefox/3.5.8" --header="Accept:
text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8" --header="Accept-
Language: en-us,en;q=0.5" --header="Accept-Encoding: gzip,deflate"
--header="Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7" --header="Keep-Alive:
300" http://example.com/update.exe
```

Below is the HTTP packet for this command:

```
GET /update.exe HTTP/1.0
User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214
Ubuntu/9.10 (karmic) Firefox/3.5.8
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Host: example.com
Connection: Keep-Alive
```

```
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300
```

It may look like we are closer to our goal, but the order of the headers are not the same, and the HTTP/1.0 operation from Wget also gives us away. Below in red are the areas where the two browsers differ:

Firefox	**Wget**
GET /update.exe HTTP/1.1 Host: example.com User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10 (karmic) Firefox/3.5.8 Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8 Accept-Language: en-us,en;q=0.5 Accept-Encoding: gzip,deflate Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7 Keep-Alive: 300 Connection: keep-alive	GET /update.exe HTTP/1.0 User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10 (karmic) Firefox/3.5.8 Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8 Host: example.com Connection: Keep-Alive Accept-Language: en-us,en;q=0.5 Accept-Encoding: gzip,deflate Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7 Keep-Alive: 300

Even if we were to hardcode the host header as the first parameter, the order would still not be the same and the HTTP/1.0 string would still give us away:

```
user@linux:~$ wget --header="Host: example.com" --header="User-Agent: Mozilla/5.0
(X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10 (karmic)
Firefox/3.5.8" --header="Accept:
text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8" --header="Accept-
Language: en-us,en;q=0.5" --header="Accept-Encoding: gzip,deflate"
--header="Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7" --header="Keep-Alive:
300" http://example.com/update.exe
```

This produces the same results as the previous attempt:

```
GET /update.exe HTTP/1.0
User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214
Ubuntu/9.10 (karmic) Firefox/3.5.8
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Host: example.com
Connection: Keep-Alive
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300
```

The solution for this is to use Curl instead of Wget. As explained here (http://daniel.haxx.se/docs/curl-vs-wget.html) Curl uses HTTP/1.1 for its operations whereas Wget use HTTP/1.0, and Curl supports an extensive list of options to craft your packets. We begin by creating a .curlrc file in our home directory to include the list of headers that we wish to spoof:

user@linux:~$ **nano .curlrc**

header = "User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10 (karmic) Firefox/3.5.8"
header = "Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8"
header = "Accept-Language: en-us,en;q=0.5"
header = "Accept-Encoding: gzip,deflate"
header = "Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7"
header = "Keep-Alive: 300"
header = "Connection: keep-alive"

Now we issue the following command to fetch the malware binary and save it locally to a file called update.exe:

```
user@linux:~$ curl http://example.com/update.exe -o update.exe
```

Below is the HTTP packet from curl:

In ASCII

```
GET /update.exe HTTP/1.1
Host: example.com
User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214
Ubuntu/9.10 (karmic) Firefox/3.5.8
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300
Connection: keep-alive
```

In Hex

```
0000   47 45 54 20 2f 75 70 64 61 74 65 2e 65 78 65 20  GET /update.exe
0010   48 54 54 50 2f 31 2e 31 0d 0a 48 6f 73 74 3a 20  HTTP/1.1..Host:
0020   65 78 61 6d 70 6c 65 2e 63 6f 6d 0d 0a 55 73 65  example.com..Use
0030   72 2d 41 67 65 6e 74 3a 20 4d 6f 7a 69 6c 6c 61  r-Agent: Mozilla
0040   2f 35 2e 30 20 28 58 31 31 3b 20 55 3b 20 4c 69  /5.0 (X11; U; Li
0050   6e 75 78 20 69 36 38 36 3b 20 65 6e 2d 55 53 3b  nux i686; en-US;
0060   20 72 76 3a 31 2e 39 2e 31 2e 38 29 20 47 65 63   rv:1.9.1.8) Gec
0070   6b 6f 2f 32 30 31 30 30 32 31 34 20 55 62 75 6e  ko/20100214 Ubun
0080   74 75 2f 39 2e 31 30 20 28 6b 61 72 6d 69 63 29  tu/9.10 (karmic)
0090   20 46 69 72 65 66 6f 78 2f 33 2e 35 2e 38 0d 0a   Firefox/3.5.8..
00a0   41 63 63 65 70 74 3a 20 74 65 78 74 2f 68 74 6d  Accept: text/htm
00b0   6c 2c 61 70 70 6c 69 63 61 74 69 6f 6e 2f 78 68  l,application/xh
00c0   74 6d 6c 2b 78 6d 6c 2c 61 70 70 6c 69 63 61 74  tml+xml,applicat
00d0   69 6f 6e 2f 78 6d 6c 3b 71 3d 30 2e 39 2c 2a 2f  ion/xml;q=0.9,*/
00e0   2a 3b 71 3d 30 2e 38 0d 0a 41 63 63 65 70 74 2d  *;q=0.8..Accept-
00f0   4c 61 6e 67 75 61 67 65 3a 20 65 6e 2d 75 73 2c  Language: en-us,
0100   65 6e 3b 71 3d 30 2e 35 0d 0a 41 63 63 65 70 74  en;q=0.5..Accept
0110   2d 45 6e 63 6f 64 69 6e 67 3a 20 67 7a 69 70 2c  -Encoding: gzip,
0120   64 65 66 6c 61 74 65 0d 0a 41 63 63 65 70 74 2d  deflate..Accept-
0130   43 68 61 72 73 65 74 3a 20 49 53 4f 2d 38 38 35  Charset: ISO-885
0140   39 2d 31 2c 75 74 66 2d 38 3b 71 3d 30 2e 37 2c  9-1,utf-8;q=0.7,
0150   2a 3b 71 3d 30 2e 37 0d 0a 4b 65 65 70 2d 41 6c  *;q=0.7..Keep-Al
0160   69 76 65 3a 20 33 30 30 0d 0a 43 6f 6e 6e 65 63  ive: 300..Connec
0170   74 69 6f 6e 3a 20 6b 65 65 70 2d 61 6c 69 76 65  tion: keep-alive
0180   0d 0a 0d 0a                                      ....
```

Is it the same as Firefox? Let's compare:

Firefox	Curl
GET /update.exe HTTP/1.1 Host: example.com User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10 (karmic) Firefox/3.5.8 Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8 Accept-Language: en-us,en;q=0.5 Accept-Encoding: gzip,deflate Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7 Keep-Alive: 300 Connection: keep-alive	GET /update.exe HTTP/1.1 Host: example.com User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.1.8) Gecko/20100214 Ubuntu/9.10 (karmic) Firefox/3.5.8 Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8 Accept-Language: en-us,en;q=0.5 Accept-Encoding: gzip,deflate Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7 Keep-Alive: 300 Connection: keep-alive

The two certainly look quite similar now!

For this demonstration Firefox on Linux was used, however in practice you would likely try to impersonate Internet Explorer, Firefox, or even Java on a Windows computer. Below as reference are sample HTTP packets from IE6 and Firefox 3.0.18 on a Windows XP machine while fetching http://example.com/update.exe:

Internet Explorer 6:
GET /update.exe HTTP/1.1
Accept: image/gif, image/x-xbitmap, image/jpeg, image/pjpeg, application/x-shockwave-flash, application/vnd.ms-excel, application/vnd.ms-powerpoint, application/msword, application/x-ms-application, application/x-ms-xbap, application/vnd.ms-xpsdocument, application/xaml+xml, */*
Accept-Language: en-us
Accept-Encoding: gzip, deflate
User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 1.0.3705; .NET CLR 1.1.4322; .NET CLR 2.0.50727; .NET CLR 3.0.4506.2152; .NET CLR 3.5.30729)
Host: example.com
Connection: Keep-Alive

Firefox 3.0.18:
GET /update.exe HTTP/1.1
Host: example.com
User-Agent: Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.0.18) Gecko/2010020220 Firefox/3.0.18 (.NET CLR 3.5.30729)
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300
Connection: keep-alive