A Taxonomy of JavaScript Redirection Spam

Kumar Chellapilla and Alexey Maykov
Microsoft Live Labs

Redirection

- "Automatically" move browser from one web page to another

- Legitimate uses
  - Migrating users from old site to new site
  - Slideshows, Stock tickers, weather map, etc

- Questionable uses (spam?)
  - Domain forwarding, Doorway pages, others

Redirection Spam

- Present one web page to the crawler but "automatically" redirect the browser to a different web page
- It is a hiding technique (a.k.a cloaking)
- Types of Redirection Spam
  - HTTP Status Codes
  - META Refresh
  - JavaScript
  - Flash/ActiveX/BHO
- Arms race between Spammers and Search Engines

HTTP and META Refresh

- HTTP Status code returned by a web request
  - 300(Multiple Choices)   303(Redirect Method)
    301(Moved Permanently)   307(Temporary Redirect)

    302(Redirect)
  - Page redirects efficiently, no page content is exchanged
  - Web server setting
- META Refresh
  - <meta http-equiv="refresh" content="0;url=http://www.destination.com/">
  - Head part of page exchanged
- Detection is easy

JavaScript Redirection

- Web page is downloaded, script is executed in stages
  - onload, onunload, onchange, onsubmit, onreset, onselect, onblur, onfocus
  - onkeydown, onkeypress, onkeyup
  - onclick, ondblclick, onmousemove, onmousedown, onmouseover, onmouseout, onmouseup
- Dynamic redirection
  - Browser type, referral URL, delay, time of day, random, …
- Detection is difficult

Data set

- Popular pages
  - Top 5000 queries, top 200 results
  - 782,937 URLs (deduped)
- Blog pages
  - Top 100 monetizable keywords
  - grep from a large set of blogspot URLs
    <name>.blogspot.com
  - 934,876 blogspot URLs

- Automated browser with JS enabled/disabled to analyze URLs
  - Inter-domain JavaScript redirects = Spam (naïve)
  - 0.35% of popular pages were classified as Spam
  - 0.77% of blogspot pages were classified as Spam

Categorization

- Popular
  - Sampled 175 / 2,712 spam pages
- Blogspot
  - Sampled 175 / 7,196 spam pages

- Each of the 350 pages was manually examined for the following techniques
  - Plain,
  - Unescape, Eval, String Manipulation, Decode
  - Add-Element, Add-Script, Add-Form, Click/Submit

Redirection Features of JS

- Location property

```
<script type="text/JavaScript">
   window.location = http://www2007.org/
</script>
```

- Time delay

```
<head><script>
function delayed_redirect() {
          window.location = "http://redirect.org"
}
</script></head>
<body onLoad="setTimeout('delayed_redirect()', 5000)">
…
</body>
```

Hiding Features of JS

- String manipulation
  - concatenation
- Obfuscation
  - UrlDecode
  - Custom decoding
- Eval
- Dynamic code and event injection
  - Script, Form, Link

JavaScript Redirection Types

- Plain
- String manipulation
- Unescape and Decode
- Referrer
- Add script, form, element & click/submit

- Categories: Popular pages, Blog pages

EXAMPLES

1. Plain

```
var1=24;  var2=var1;
if(var1==var2)
        document.location="http://www.topsearch10.com/
search.php?aid=59731&q=bad+credit+auto+loan";
```

http://bad-credit-auto--loan.blogspot.com/

2. String Manipulation + Eval

```
var
a1="win", a2="dow.", a3="loca", a4="tion.", a5="replace",
a6="('http://www.partypoker.com/index.htm?wm=250106
8')";
var i,str="";
for(i=1;i<=6;i++)
{
   str += eval("a"+i);
}
eval(str);
```

http://party-poker-bonus.cjb.net/

3. Unescape

```
var s =
'%5CBEOD%5C%05GDHJ_BDE%16%0CC__%5B%11%04%04
%5C%5C%5C%05SMYNNFD%5DBNX%05HDF%04%0C';
var e = '', i;
eval(unescape('s%3Dunescape%28s%29%3Bfor%28i%3D0%
3Bi%3Cs.length%3Bi%2B%2B%29%7Be%2B%3DString.from
CharCode%28s.charCodeAt%28i%29%5E43%29%3B%7D%3
Beval%28e%29%3B'));
```

**http://freegayporntodays.blogspot.com/
2006_10_01_freegayporntodays_archive.html**

3. Unescape – Stage 1

```
var s =
'%5CBEOD%5C%05GDHJ_BDE%16%0CC__%5B%11%04%04
%5C%5C%5C%05SMYNNFD%5DBNX%05HDF%04%0C';
var e = '', i;
eval(unescape('s%3Dunescape%28s%29%3Bfor%28i%3D0%
3Bi%3Cs.length%3Bi%2B%2B%29%7Be%2B%3DString.from
CharCode%28s.charCodeAt%28i%29%5E43%29%3B%7D%3
Beval%28e%29%3B'));
```

**http://freegayporntodays.blogspot.com/
2006_10_01_freegayporntodays_archive.html**

3. Unescape – Stage 2

```
s=unescape(s);
for(i=0;i<s.length;i++)
{
        e += String.fromCharCode(s.charCodeAt(i)^43);
};
eval(e)
```

3. Decode – Stage 3

```
s=unescape(s);
for(i=0;i<s.length;i++)
{
        e += String.fromCharCode(s.charCodeAt(i)^43);
};
eval(e)

window.location='http://www.xfreemovies.com/'
```

http://freegayporntodays.blogspot.com/ 2006_10_01_freegayporntodays_archive.html

4. Decode + Script Inj. (1)

```
var s,q,e,d,i;   s=String.fromCharCode;    q='script'+s(62);
var e =
unescape('%BD%AD%BF%EE%BA%ED%F3...%EE%A1%F0');
var d = '';
for(i=0;i<e.length;++i)
        d+=s(e.charCodeAt(i)^(((i%10)+203)&255));
document.write(s(60)+q+d+s(60)+'/'+q);
```

http://pori-chudai.igotclicks.com/taktaz

4. Decode + Script Inj. (2)

```
<script>var
u="http://www.veracitek.com/adTracker/?source=1976
&w=http%3A%2F%2Fpori-chudai.star-
gossip.com%2Ftaktaz”,
e=escape,d=document;
d.location=u;
</script>
```

http://pori-chudai.igotclicks.com/taktaz

5. Link Injection + Click

```
var lnk='<a id="rdr"
href="http://fairsearch.net/rd/find.php?q=';lnk+=kwd;
lnk+='"> </a>';
document.write(lnk);
var obj = document.getElementById("rdr");
try{ obj.click();}
catch (MyError){
  obj.click = function() {document.location.href = this.href; }
  obj.click();
}
```

http://live-sex-52817.blogspot.com/2006/09/free-live-sex-chat-nude-strip-cam.html

6. Form Injection + Click

```
document.write('<form id="f" method=post
action="h’+'t’+'t’+'p://‘+'qblz.c’+'om/pc/l’+'n.cgi?2&param
eter=free%20ringtones" style=display:none>
<input type=submit name="xlt2"></form>');
document.forms.f.xlt2.click();
```

http://downloadfree-ringtones-.blogspot.com/

7. Referrer

```
var r=document.referrer,    t="",    q;
if(r.indexOf("google.")!=-1) t="q";
if(r.indexOf("msn.")!=-1) t="q";
if(r.indexOf("yahoo.")!=-1) t="p";
if(r.indexOf("altavista.")!=-1) t="q";
if(r.indexOf("aol.")!=-1)t = "query";
if(r.indexOf("ask.")!=-1)t = "q";
if(t.length&&((q=r.indexOf("?"+t+"="))!=-1 ||
                    (q=r.indexOf("&"+t+"="))!=-1)) {
window.location="http://pharma-online.org/search.php?q=tramadol"}
else {
window.location="http://pharma-online.org/search.php?q2=tramadol"
}
```

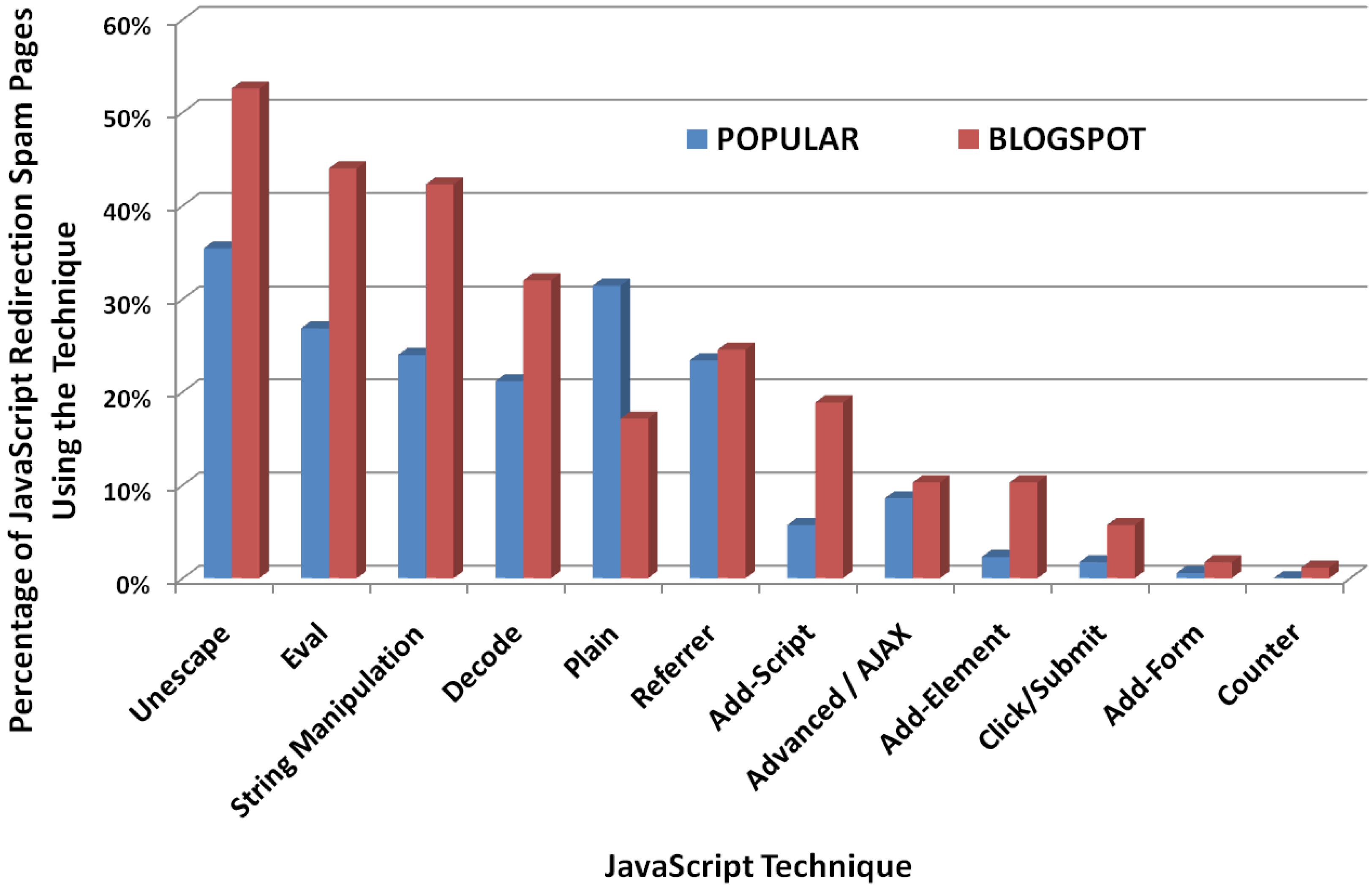
http://us-tramadol-hcl.blogspot.com/

7. Referrer

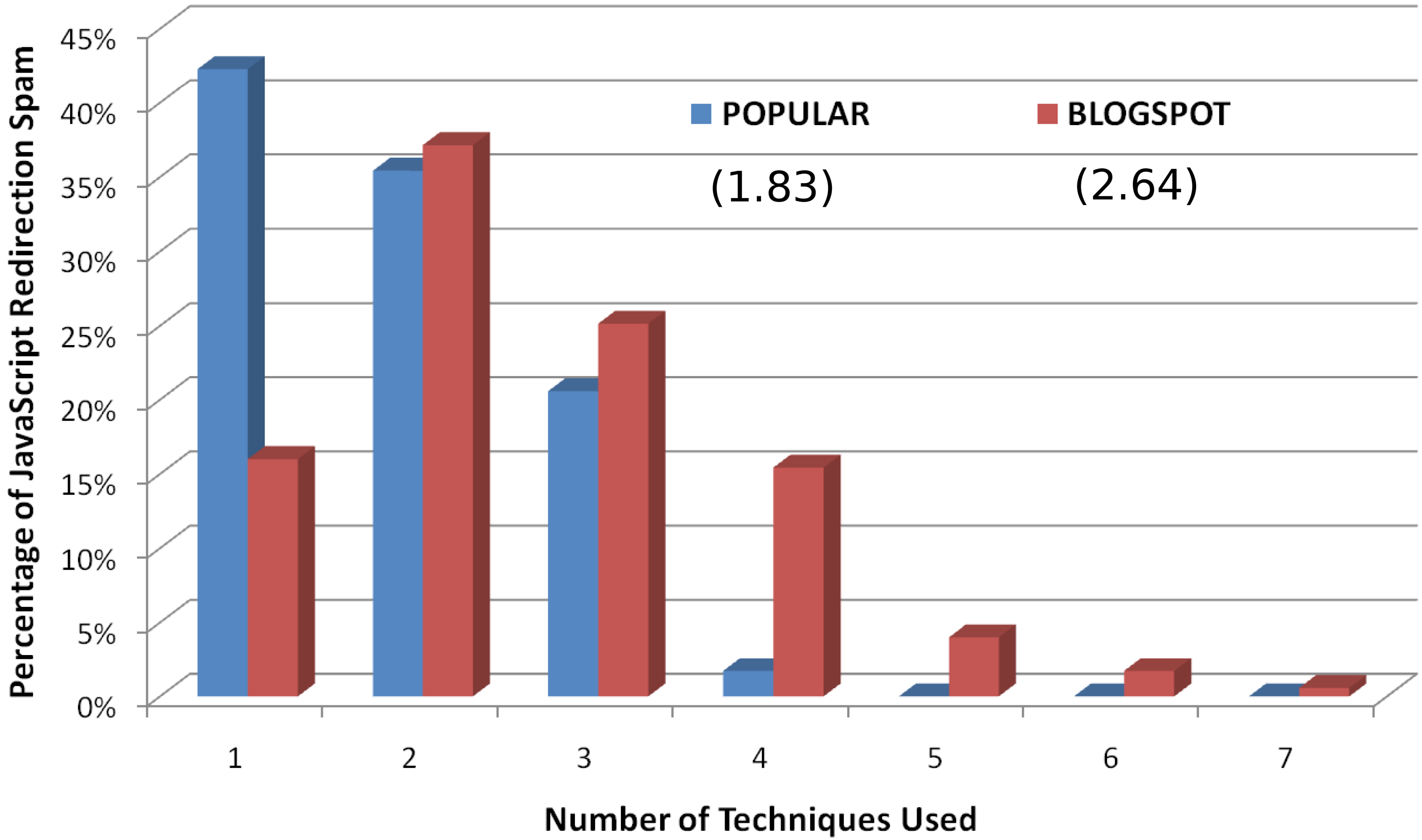
```
var r=document.referrer,   t="",   q;
if(r.indexOf("google.")!=-1) t="q";
if(r.indexOf("msn.")!=-1) t="q";
if(r.indexOf("yahoo.")!=-1) t="p";
if(r.indexOf("altavista.")!=-1) t="q";
if(r.indexOf("aol.")!=-1)t = "query";
if(r.indexOf("ask.")!=-1)t = "q";
if(t.length&&((q=r.indexOf("?"+t+"="))!=-1 ||
                (q=r.indexOf("&"+t+"="))!=-1)) {
window.location="http://pharma-online.org/search.php?q=tramadol"}
else {
window.location="http://pharma-online.org/search.php?q2=tramadol"
}
```

http://us-tramadol-hcl.blogspot.com/

Distribution of Techniques



Number of Techniques Used



Observations

- Popular pages 2x more likely to use plain redirection
- Obfuscation is very prevalent (44%-62%)
- 25% of all JavaScript Spam redirects examine referrer property
- 10-25% use dynamic code injection
- Less than 10% use AJAX/Advanced JavaScript