(Aster)-picking through the pieces of short URL services

An investigation into the maliciousness of short URLs

Robert Diepeveen & Peter Boers
2016
Motivation

• Obfuscation
• Brute force
• Uniform sample
• Contributions:
  – Comparison between services
  – Observation of locality based adware network
Research questions:

• What portion of the short URL services are used for malicious purposes and what does the abuse look like?
  - Which service provides proportionally the most short URLs flagged as malicious?
  - What properties can be observed in encountered malicious sites?
Which services are looked into?

- Previous work found the most popular services
- Alexa.com
- “Well known”
  - TinyURL
  - bitly
  - goo.gl
- t.co, not investigated
How do you classify a site as malicious?

- Google Safe Browse
  - Malware
  - Phishing
  - “Unwanted”

- DNSBL
  - Domain blacklist
  - IP blacklist

- Other methods:
  - PhishTank
What else is interesting to know about the URLs that are online?

- Short URLs
  - Creation date
  - Clicks
  - Referrers
- Long URLs
  - SSL info
  - Malicious classification
  - Server Headers (Last Modified, Server, Status Code)
  - Script links
  - Page Size
Uniform sampling

- Key space approximates and hash lengths:
  - Bitly: 3.5 trillion, max 7
  - TinyURL: 80 billion, max 7
  - Goo.gl: 58 billion, max 6
- Random number generator to base conversion
- [0-9A-Za-z]
- **Keyspace is not fully used**
Setup

- 12 VMs
- 4 days of data gathering
- 96 threads per service
  - Except goo.gl
- 4 short URLs inserted in MongoDB per second
- Average traffic:
  - 8.52 Mbit/s out
  - 2.44 Mbit/s in
The numbers

- Approx 1.4 million short URLs encountered
  - TinyURL: 1,39 million visited.
  - Bitly: +/- 6 K visited.
  - Goo.gl: +/- 4K visited.

- Malware – undetected hits
  - TinyURL: 946
  - Bitly: 2
  - Goo.gl: 0
## The numbers (2)

<table>
<thead>
<tr>
<th>Service</th>
<th>Undetected</th>
<th>Detected</th>
<th>Total</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>TinyURL</td>
<td>946</td>
<td>70,302</td>
<td>71,248</td>
<td>5.17%</td>
</tr>
<tr>
<td>Bitly</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>+/- 0.05%</td>
</tr>
<tr>
<td>Goo.gl</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>+/- 0.01%</td>
</tr>
<tr>
<td>Totals</td>
<td>948</td>
<td>70,307</td>
<td>71,255</td>
<td></td>
</tr>
<tr>
<td>Domain</td>
<td>Count</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://www.asterpix.com">www.asterpix.com</a></td>
<td>495</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>video.asterpix.com</td>
<td>113</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://www.tagvn.com">www.tagvn.com</a></td>
<td>75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://www.filelodge.com">www.filelodge.com</a></td>
<td>57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>keyknowhow.com</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hurl.content.loudeye.com</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>static.zangocash.com</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://www.perfectporridge.com">www.perfectporridge.com</a></td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://www.content.loudeye.com">www.content.loudeye.com</a></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small counts (&lt;= 4)</td>
<td>137</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What is asterpix.com?

- Origins in 2006 as a video sharing site
- Short URLs are created during that period
  - `video.asterpix.com/v/<ID>/<Title>/`
  - `www.asterpix.com/console/?avi=<ID>`
- 2009: links and short URLs “die”
- 2015: malware registered
Taxonomy

- Encountered a Dutch site during first visit.
- How does locality influence redirection?
  - Asia
  - America
  - Europe
- Three phases
  - Entry
  - Redirection
  - Hand off
The phases

• Entry
  – Where is the visitor from?
  – Has he visited in the past?

• Redirection
  – Typical JS redirection to obfuscate paths
  – All over the world and at least 4 hops
  – Depending on location of visitor

• Hand Off
  – Catered to the visitor in language and offering
What was observed?

- One known entry point
- Two known non malicious landing pages
- Eight known malicious landing pages
  - Surveys
  - “Free” money
  - Vouchers
- Overlapping redirect chains
  - park.above.com
  - bidr.trellian.com
  - z[a-z].zeroredirect.com
1. 架设服务器
2. 服装cad下载
3. 阿尔山自助游
4. 外国服务器
5. 健康体检中心
6. 美国ro89
7. 秦皇岛违章查
8. 注册域名
9. 宠物小精灵网
10. 北理工珠海学
11. 软片网站
12. 影视后期制作
13. 看大片的网站
14. 未婚女性排名
15. 姐妹同性排
Jaarlijks bezoekersonderzoek 2016

Browser: Gebruikersonderzoek

Gefeliciteerd!

Je bent persoonlijk geselecteerd om deel te nemen aan ons jaarlijkse bezoekersonderzoek 2016! Vertel ons wat je denkt van Firefox en als “dank” geven we je een kans om een iPhone 6S te winnen!

Vraag 1 van de 4:
Hoe vaak gebruik je Firefox?
- Altijd
- Soms
- Nooit

Volgende...
its natural moisture levels, leaving your face smooth, healthy and radiant.

Regular Price: $89.99
You Pay: $5.95
Quantity Left: (1)

Get a Galaxy S6!
Experience vivid colors and dramatic contrast. This is our brightest, most vibrant screen you've ever seen.

Regular Price: $899.00
You Only Pay: $7
Quantity Left: (3)

Get a new iPad Mini

Regular Price: $600.00
You pay: $5.00
Quantity Left: (2)
Conclusion/Discussion

- Significant amount of malicious sites TinyURL
- Undetected rate more or less the same over the services.
- Proportionally more malicious long URLs at TinyURL in total.
- Sites change over time, short URLs remain active
  - Unable to see if this is actively abused
- Locality based redirection observed
  - Block secondary/tertiary redirectors.
Future work

- The “repurposing” of short URLs and its abuse
- The effectiveness of blocking underlying redirectors
- A further case study into locality based adware networks to find commonalities
- Optimization of the search for bitly and goo.gl
- Look into smaller, lesser known providers.