Domain Abuse Reporting Tool (DART)

Dave Piscitello
on behalf of the ICANN Office of the CTO SSR Team
The DART Project

What is the Domain Abuse Reporting Tool?
• A platform for reporting domain name registration and abuse behavior across TLD registries and registrars

How does DART differ from other reporting?
• Studies all TLD registries and registrars for which we can collect zone and registration data
• Employs a very large set of reputation feeds
• Warehouses data for historical studies
• Studies multiple threats: phishing, botnet, malware, spam
• Scientific approach: unbiased, transparent, reproducible
Goals of the Domain Abuse Reporting Tool

*Provide ICANN community with data to support the policy development process*

- Data can be used to
  - Identify threats reported at TLD or registrar level for all TLDs for which we can obtain data
  - Historically track security threats, domain registration activity (adds, deletes) at a TLD or registrar level
  - Help operators understand or consider how to manage their reputations, their anti-abuse programs or their terms of service
  - Study malicious registration behaviors
  - Assist the operational security community by sharing open data or data analyzed by the reporting tools
DART Uses TLD Zone Data

• Collects zones for TLDs for registry analytics
  – Any {new, legacy, cc} from which we can get a zone
  – Currently gTLDs. Some ccTLD expressed interest in being added during ICANN 58, Copenhagen

• Currently, system collects zones from 1236 TLDs
  – Approximately 193 million domains
  – Application rejection or renewal issues with ICANN Centralized Zone Data Service
DART Uses Whois

• Collects registration data to associate delegated domain names in zone files with sponsoring registrars
  – Issues with Whois rate limiting
• DART uses domain names that appear in zones
  – Security threats cannot be executed if a domain name cannot resolve to an IP addresses
DART Uses Reputation Data (Blocklists)

• Uses multiple domain or URL abuse data sets (reputation feeds) to
  – Count spam, phishing, malware host, botnet (C2) domain names, total abuse domains, cumulative abuse domains
  – Create histograms, days in the life views...
  – Search abuse database by argument

• If a domain appears on any list, it is included in the counts (de-duplication is part of process)
DART Uses Many Reputation Data Sets

- DART collects the same abuse data that is reported to industry and Internet users
  - The abuse data that DART collects are used by commercial security systems that protect billions of users daily
  - Academic and industry use and endorse these data sets
  - Studies and industry use show that they have history of accuracy, global coverage, and low false positive rates

- DART reflects how parties external to ICANN community see the domain ecosystem

- Extensible framework
  - Experimenting with doing analyses using subsets of data
Why Multiple Data Sets?

- Expands our abuse data set with low duplication
  
  http://dl.acm.org/citation.cfm?id=2808129

- Research finds that there is little overlap between block lists

- We use data feeds with
  
  - Industry reputation for accuracy, clarity of process
  - Threat classification that matches our purposes
  - Consensus adoption across operational security community, i.e., inclusion in commercial security systems
  - Frequency of citation in academic literature
More on “Why multiple lists?”

• No reputation provider can see all the abuse
  – Each is catching only some (what they see)
• Providers look for different types of abuse, use different methods or infrastructures
• Some lists are big and some are small.
  – The smaller the list, the less % overlap it might have with a larger list
• Experience with our data sets is similar to Metcalfe & Spring’s and Sinha findings
• Purpose of scoring is to assess deviation (distance) from mean scoring
  – Measure the extent to which an operator is a target of malicious actors

• Experimenting with strawman proposal for scoring abuse impacting TLDs and Registrars

• Looking for input
  – Goal is to gain industry-wide acceptance on scoring algorithms
Abuse Score, TLD

• The number of unique, currently listed domains per 100 domains in the zone

\[ \text{SCORE} = \frac{\text{abuse-listed domains in a TLD on a given day}}{\text{domains in the TLD zone on this day}} \times 100 \]

• This shows us the percentage of domains in the zone file that are currently listed on abuse blocklists that we monitor
Abuse Score, Registrar

• The number of unique, currently listed abuse domains per 100 domains that the registrar sponsors.

\[
\text{SCORE} = \frac{\text{abuse-listed gTLD domains sponsored by registrar on a given day}}{\text{gTLD domains sponsored by the registrar on this day}} \times 100
\]

• This shows us the percentage of the domains that the registrar sponsors are currently listed on abuse blocklists that we monitor.
Access to Reporting System

• Currently in Beta, internal use
• Soliciting community input on kinds and frequency of reporting.
  – What should we report?
  – To whom should we report?
  – Order of reporting?
  – Access to our data?
  (Note: may be affected by use licenses)
DART Dashboard

@ TLD Reporting Tools

gTLD and Registrar Statistics for 2017-05-10

- Total gTLDs: 1,238
- Total number of gTLD domains in zones: 194,046,422
- With at least one incident: 454
- Total nTLDs: 1,222

- Total: 705
- With at least one incident: 705

gTLD registrations

- .COM: 66.8%
- .NET: 13.1%
- .ORG: 6.0%
- nTLDs: 5.0%
- Other gTLDs: 5.6%

Top gTLDs

Show 10 entries

- gTLD

Domains In Zone

Listed Domains

Abuse Score

5/11/17
### gTLDs

**Statistics for 2017-05-10**

<table>
<thead>
<tr>
<th>TLD</th>
<th>Domains In Zone</th>
<th>Abuse Domains</th>
<th>Abuse Score</th>
<th>Spam</th>
<th>Phishing</th>
<th>Botnet C&amp;C</th>
<th>Malware</th>
</tr>
</thead>
<tbody>
<tr>
<td>41</td>
<td>92,523</td>
<td>2</td>
<td>0.00</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>42</td>
<td>90,401</td>
<td>13,658</td>
<td>15.11</td>
<td>13,501</td>
<td>20</td>
<td>0</td>
<td>139</td>
</tr>
<tr>
<td>43</td>
<td>89,171</td>
<td>179</td>
<td>0.20</td>
<td>138</td>
<td>101</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>44</td>
<td>88,945</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>80,932</td>
<td>19,482</td>
<td>24.07</td>
<td>19,272</td>
<td>14</td>
<td>184</td>
<td>122</td>
</tr>
<tr>
<td>46</td>
<td>80,925</td>
<td>25,876</td>
<td>31.98</td>
<td>25,837</td>
<td>39</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>47</td>
<td>76,776</td>
<td>21,499</td>
<td>28.00</td>
<td>21,487</td>
<td>9</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>48</td>
<td>74,105</td>
<td>816</td>
<td>1.10</td>
<td>812</td>
<td>9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>49</td>
<td>72,642</td>
<td>1</td>
<td>0.00</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>72,568</td>
<td>172</td>
<td>0.24</td>
<td>158</td>
<td>20</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

Showing 41 to 50 of 1,229 entries

prevous 1 ... 4 5 6 ... 123 next