Security Reputation Metrics for Hosting Providers

@CSET’15
10 Aug. 2015

Arman Noroozian, Maciej Korzcyński,
Samaneh Tajalizadehkoob, Michel van Eeten
Reputation Metrics are Hard!

... to make and interpret properly
Why Metrics?

- The “Lemons Market” Problem
  - Information Asymmetry

- Consumer / Policy maker / Law enforcement officer
  - Which provider is better/worse in security?

- The provider (intermediary) itself doesn’t know either!

- Erodes incentives to invest in security
Hosting Providers

- Legitimate hosting provider types

<table>
<thead>
<tr>
<th>Hosting Type</th>
<th>Hardware</th>
<th>Operating System</th>
<th>Software</th>
<th>Abuse Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedicated</td>
<td>Provider</td>
<td>Customer</td>
<td>Customer</td>
<td>Customer</td>
</tr>
<tr>
<td>Managed</td>
<td>Provider</td>
<td>Provider</td>
<td>Provider</td>
<td>Provider or Customer</td>
</tr>
<tr>
<td>Reseller</td>
<td>Provider or Customer</td>
<td>Customer or its Client</td>
<td>Customer or its Client</td>
<td>Customer or its Client</td>
</tr>
<tr>
<td>Shared</td>
<td>Provider</td>
<td>Provider</td>
<td>Provider</td>
<td>Provider and Customer</td>
</tr>
<tr>
<td>Unmanaged</td>
<td>Provider and Customer</td>
<td>Customer</td>
<td>Customer</td>
<td>Customer</td>
</tr>
<tr>
<td>Virtual Private Server</td>
<td>Provider</td>
<td>Provider</td>
<td>Customer</td>
<td>Customer</td>
</tr>
</tbody>
</table>


- Bulletproof Hosting!
Concentration of Abuse

- Attractive Pressure points
  - Remediation
  - Policy making

Source: McAfee Threats Report Q2 2012
### Concentrations of Abuse (Cont.)

<table>
<thead>
<tr>
<th>STOP BADWARE (SITES)</th>
<th>F.I.R.E. (COMPOSITE)</th>
<th>PHISHTANK</th>
<th>ZEUSTRACKER (TOP ZEUS C&amp;Cs)</th>
<th>MALWARE DOMAIN LIST (JAN.1-TODAY)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHINANET BACKBONE (AS14035)</td>
<td>PAH Inc GoDaddy.com (AS26496)</td>
<td>MetroRED Telecom Services (AS13391)</td>
<td>HANARO Telecom (AS9318)</td>
<td>OVH - OVH (AS16276)</td>
</tr>
<tr>
<td>PAH Inc:GoDaddy.com (AS26496)</td>
<td>OVH - OVH (AS16276)</td>
<td>RAPIDSWITCH-AS (AS29131)</td>
<td>TTNET (AS59121)</td>
<td>DXTNET BEIJING (AS17964)</td>
</tr>
<tr>
<td>N/A</td>
<td>BLUEHOST-AS (AS11798)</td>
<td>CENTROHOST-AS (AS41112)</td>
<td>GR-Vertical Ltd (AS45135)</td>
<td>VEST-EN (AS47560)</td>
</tr>
<tr>
<td>eNom Inc. (AS6151)</td>
<td>IPNAP - GigeNET (AS23522)</td>
<td>ThePlanet.com (AS21844)</td>
<td>CHINANET BACKBONE (AS14035)</td>
<td>TTNET (AS59121)</td>
</tr>
<tr>
<td>Google Inc. (AS15169)</td>
<td>EcomD-Coloquest/GigeNet (AS32181)</td>
<td>iWeb Technologies Inc. (AS32613)</td>
<td>ThePlanet.com (AS21844)</td>
<td>DIRECT-NET2 - Data Tex (AS4229)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARBOR TOP ASN THREATS</td>
<td>EMERGING THREATS COMPROMISED IPS</td>
<td>EMERGING THREATS BBN</td>
<td>SHADOWSERVER (BOT C&amp;CS)</td>
<td>GOOGLE SAFE BROWSING</td>
</tr>
<tr>
<td>NJ INTL INTERNET EXCHANGE (AS18812)</td>
<td>Softlayer Technologies (AS36851)</td>
<td>net-Dx2a-ukraine (AS48587)</td>
<td>ThePlanet.com (AS21844)</td>
<td></td>
</tr>
<tr>
<td>CHINANET BACKBONE (AS14035)</td>
<td>Deutsche Telecom (AS3329)</td>
<td>CHINA TELECOM (AS4114)</td>
<td>Webalta Networks (AS1947)</td>
<td>PAH Inc GoDaddy.com (AS26496)</td>
</tr>
<tr>
<td>ThePlanet.com (AS21844)</td>
<td>SRC Internet Services (AS7132)</td>
<td>NETDIRECT-D. (AS28753)</td>
<td>Network Operations Center (AS21788)</td>
<td>MESH GmbH (AS25074)</td>
</tr>
<tr>
<td>OVH - OVH (AS16276)</td>
<td>PROXAD Free SAS (AS12322)</td>
<td>Hivelocity Ventures Corp (AS29802)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COLUMBUS-NA-P (AS10297)</td>
<td>Telecom Sao Paolo (AS27999)</td>
<td>Leasewebs (AS16265)</td>
<td>CHINA TELECOM (AS4114)</td>
<td>HETZNER ONLINE (AS24940)</td>
</tr>
<tr>
<td>Softlayer Technologies (AS36851)</td>
<td>China Network Comm. (AS48837)</td>
<td>HETZNER ONLINE (AS24940)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpal (AS16138)</td>
<td>HANARO Telecom (AS9318)</td>
<td>NIX (AS19318)</td>
<td>Interactive3D (AS49544)</td>
<td>VPLS INC (AS39088)</td>
</tr>
<tr>
<td>HINET (AS462)</td>
<td>National Internet Backbone (AS5829)</td>
<td>Layered Tech (AS25276)</td>
<td>SBC Internet Services (AS7132)</td>
<td></td>
</tr>
<tr>
<td>AMAZON (AS14618)</td>
<td>CHINANET-BJ-AS-169 (AS4908)</td>
<td>OVH - OVH (AS16276)</td>
<td>Ecatel Network (AS59073)</td>
<td>18.1 Internet Ag (AS58396)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>root eSolutions (AS5577)</td>
</tr>
</tbody>
</table>

Hoster Size Matters

Top 50 Hosts

A list of the 50 ASes with the highest HE indexes i.e. the highest observed concentrations of malicious activity.

<table>
<thead>
<tr>
<th>HE Rank</th>
<th>HE Index</th>
<th>ASN</th>
<th>Name</th>
<th>Country</th>
<th>IPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>291.22</td>
<td>11042</td>
<td>Landis Holdings, Inc.</td>
<td>US</td>
<td>28,416</td>
</tr>
<tr>
<td>2</td>
<td>289.06</td>
<td>28347</td>
<td>New Dream Network, LLC</td>
<td>US</td>
<td>156,928</td>
</tr>
<tr>
<td>3</td>
<td>287.71</td>
<td>25182</td>
<td>HostDime.com, Inc.</td>
<td>US</td>
<td>78,548</td>
</tr>
<tr>
<td>4</td>
<td>245.64</td>
<td>31034</td>
<td>Aruba S.p.A.</td>
<td>IT</td>
<td>145,564</td>
</tr>
<tr>
<td>5</td>
<td>242.00</td>
<td>29182</td>
<td>iSPSystem</td>
<td>RU</td>
<td>44,544</td>
</tr>
<tr>
<td>6</td>
<td>239.48</td>
<td>47583</td>
<td>Hostinger International</td>
<td>US</td>
<td>13,568</td>
</tr>
<tr>
<td>7</td>
<td>219.72</td>
<td>13335</td>
<td>CloudFlare, Inc.</td>
<td>US</td>
<td>258,360</td>
</tr>
<tr>
<td>8</td>
<td>211.48</td>
<td>12624</td>
<td>home.pl</td>
<td>PL</td>
<td>204,800</td>
</tr>
<tr>
<td>9</td>
<td>191.78</td>
<td>25532</td>
<td>Masterhost</td>
<td>RU</td>
<td>77,824</td>
</tr>
<tr>
<td>10</td>
<td>191.71</td>
<td>24995</td>
<td>GoDaddy.com, LLC</td>
<td>US</td>
<td>1,758,192</td>
</tr>
<tr>
<td>11</td>
<td>187.04</td>
<td>8563</td>
<td>1&amp;1 Internet AG</td>
<td>DE</td>
<td>372,224</td>
</tr>
<tr>
<td>12</td>
<td>182.24</td>
<td>16276</td>
<td>OVH Systems</td>
<td>FR</td>
<td>1,079,552</td>
</tr>
<tr>
<td>13</td>
<td>180.30</td>
<td>34619</td>
<td>Cugi Telekomunikasyen</td>
<td>TR</td>
<td>30,269</td>
</tr>
<tr>
<td>14</td>
<td>179.01</td>
<td>25384</td>
<td>Vaurion Rechenzentrum AG</td>
<td>DE</td>
<td>22,784</td>
</tr>
<tr>
<td>15</td>
<td>169.96</td>
<td>46608</td>
<td>Unified Layer</td>
<td>US</td>
<td>648,360</td>
</tr>
<tr>
<td>16</td>
<td>168.71</td>
<td>27823</td>
<td>Dattatec.com</td>
<td>AR</td>
<td>1,758</td>
</tr>
<tr>
<td>17</td>
<td>166.51</td>
<td>38553</td>
<td>World4You Internet Services</td>
<td>AT</td>
<td>5,460</td>
</tr>
<tr>
<td>18</td>
<td>163.94</td>
<td>16265</td>
<td>LeaseWeb B.V.</td>
<td>NL</td>
<td>397,824</td>
</tr>
<tr>
<td>19</td>
<td>162.89</td>
<td>29763</td>
<td>Eutelsat Network</td>
<td>NL</td>
<td>1,2800</td>
</tr>
<tr>
<td>20</td>
<td>161.04</td>
<td>48034</td>
<td>Confluence Networks Inc</td>
<td>VG</td>
<td>16,128</td>
</tr>
<tr>
<td>21</td>
<td>161.00</td>
<td>48159</td>
<td>Telecommunication Infrastructure</td>
<td>IR</td>
<td>385,728</td>
</tr>
<tr>
<td>22</td>
<td>160.00</td>
<td>24540</td>
<td>Hetzner Online AG</td>
<td>DE</td>
<td>702,280</td>
</tr>
<tr>
<td>23</td>
<td>159.48</td>
<td>43146</td>
<td>Agava Ltd.</td>
<td>RU</td>
<td>20,736</td>
</tr>
</tbody>
</table>

Measures of Size

Advertised IP Space

Hosted 2nd level domains
## Indicators of Abuse

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Why</th>
<th>Challenge</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Occurrence of Abuse</strong></td>
<td>Signs network hygiene and vulnerability</td>
<td>Hard to isolate provider efforts from other factors</td>
</tr>
<tr>
<td><em>(How often abused?)</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Uptime of abuse</strong></td>
<td>Signs effectiveness of abuse handing</td>
<td>Hard to measure at scale</td>
</tr>
<tr>
<td><em>(How long abused?)</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sensitivity of Metrics

- Choice of abuse data
- Biases and errors in abuse data
- Errors in mapping abuse data
- Biases and errors in size estimation data
A Dutch Case Study

Dutch Police:

“Who are the worst hosting providers in our jurisdiction?”
Data Sources

- Abuse
  - StopBadware
  - Shadowserver
    - Compromised servers
    - Outbound malware connections
  - Zeustracker C&Cs (Abuse.ch)
  - Mutual Legal Assistance Treaty (MLAT) requests
  - Dutch child pornography hotline
  - Phishtank
  - Anti-Phishing Working Group

- IP Routing Data
  - Python pyasn library

- Passive DNS (pDNS)
  - DNSDB from Farsight Security
  - 750 million unique 2LDs
  - 93 million unique IPv4 Addresses
Our Methodology
Step 1+2: Mapping

Abuse Feeds
- Shadow Server Compromise
- Shadow Server Sandbox URL
- Zeustracker C&Cs
- MLAT requests
- PhishTank
- APWG
- Child Pornography Hotline

p-DNS / IP Routing
- Farsight Security p-DNS Data
- Internet IP Routing Data

Abuse Mapping
- # Unique Abuse / AS

Abuse Maps
- PhishTank
  - AS#1 $\rightarrow$ 100
  - AS#2 $\rightarrow$ 200
- MLAT
  - AS#1 $\rightarrow$ 50
  - AS#2 $\rightarrow$ 73

Size Mapping
- # Advertised IPs
- # IPs in p-DNS
- # Domains Hosted

Size Maps
- Advertised IPs
  - AS#1 $\rightarrow$ 256
  - AS#2 $\rightarrow$ 1024
- Domains Hosted
  - AS#1 $\rightarrow$ 23
  - AS#2 $\rightarrow$ 1232
Step 3: Normalization

**Abuse Maps**
- PhishTank
  - AS#1 $\rightarrow$ 100
  - AS#2 $\rightarrow$ 200
- MLAT
  - AS#1 $\leftarrow$ 50
  - AS#2 $\leftarrow$ 73

**Size Maps**
- Advertised IPs
  - AS#1 $\rightarrow$ 256
  - AS#2 $\rightarrow$ 1024
- Domains Hosted
  - AS#1 $\leftarrow$ 23
  - AS#2 $\leftarrow$ 1232

**Normalized Abuse**
- PhishTank / Advert. IPs
  - AS#1 $\rightarrow$ 0.39
  - AS#2 $\rightarrow$ 0.19

- PhishTank / Domains Hosted
  - AS#1 $\rightarrow$ 4.34
  - AS#2 $\rightarrow$ 0.16

- MLAT / Advert. IPs
  - AS#1 $\rightarrow$ 0.19
  - AS#2 $\rightarrow$ 0.07

- MLAT / Domains Hosted
  - AS#1 $\rightarrow$ 2.17
  - AS#2 $\rightarrow$ 0.05

**Normalization**
- # Abuse / Size
Step 4: Ranking

Normalized Abuse

PhishTank / Advrt. IPs
AS#1 ➔ 0.39
AS#2 ➔ 0.19

PhishTank / Domains Hosted
AS#1 ➔ 4.34
AS#2 ➔ 0.16

MLAT / Advrt. IPs
AS#1 ➔ 0.19
AS#2 ➔ 0.07

MLAT / Domains Hosted
AS#1 ➔ 2.17
AS#2 ➔ 0.05

Abuse Ranking

PhishTank Ranking 1
AS#1 ➔ 834
AS#2 ➔ 833

PhishTank Ranking 2
AS#1 ➔ 834
AS#2 ➔ 833

MLAT Ranking 1
AS#1 ➔ 235
AS#2 ➔ 234

MLAT Ranking 2
AS#1 ➔ 235
AS#2 ➔ 234
Step 5: Aggregation

Abuse Ranking

*PhishTank Ranking 1*
AS#1 ↔ 834
AS#2 ↔ 833

*PhishTank Ranking 2*
AS#1 ↔ 834
AS#2 ↔ 833

*MLAT Ranking 1*
AS#1 ↔ 235
AS#2 ↔ 234

*MLAT Ranking 2*
AS#1 ↔ 235
AS#2 ↔ 234

Combine Ranks

Borda Count

Abuse Ranking

*Overall Ranking*
AS#1 ↔ 1
AS#2 ↔ 0.92
AS#3 ↔ 0.87
AS#4 ↔ 0.86
Security Reputation Metrics

20 worst Dutch hosting providers

Abuse Rate vs Cleanup Rate

Occurrence Metric Top 20 Worst ASes

Comparison of Occurrence and Uptime
Abuse Metrics are Hard

- How to measure abuse and remediation
  - What abuse can be observed
  - What does it tell us about remediation efforts

- How to associate it with hosting providers
  - What is a hosting provider
  - How to identify them at scale

- How to control for differences among providers and interpret metric(s)
  - How to take size into account
  - How to take different business models into account

- How to aggregate indicators into a comprehensive metric (set of metrics)?
Towards better metrics

- How to measure abuse and remediation
  - Increase coverage, add different global abuse feeds
  - Add uptime data (e.g. phishing)

- How to associate it with hosting providers
  - Identify hosting providers from IP ownership data (WHOIS) instead of AS-level routing data (BGP)

- How to control for differences among providers and interpret metric(s)
  - Extract ‘profiles’ from pDNS data (size, shared hosting, dedicated, non-webdomain)

- How to aggregate indicators into a comprehensive metric (set of metrics) ?
  - More sensitivity analysis of aggregation methods
Thank you for attention

Questions?