Correlating Spam Activity with IP Address Characteristics

Chris Wilcox, Christos Papadopoulos
CSU
John Heidemann
USC/ISI

AIMS 2010 - Feb 10, 2010
Introduction

• Common belief: spamming hosts exhibit specific address characteristics:
  – dynamically allocated addresses
  – specific geographical areas
  – more tolerant spam policies
  – less stability, more volatility, shorter uptimes.

Our goal: quantify differences in address characteristics between spammers and legitimate hosts
Approach

• Correlate the results of an IP address visibility study with a commercial IP address blacklist for the same period
  – Quantify differences between address characteristics of spammers and non-spammers
  – Quantify differences in domain names
  – Investigate collateral damage if a /24 is blocked due to presence of spammers
Data Sources

- **Address visibility:**
  - survey of reachable Internet addresses every 3 months.
  - Use active probing (ICMP) over ~24,000 /24 IPV4 blocks (1% of the Internet)

- **Reputation-based block list from eSoft.com**
  - <IP addr, score>, based on sender address verification, sender policy framework, heuristic analysis, reputation filtering, historical averaging, etc…
Visibility Study

- Census: ping every internet address every three months
- Survey: select 1% of /24 subnets and ping each address every 11 mins
- We use surveys
Visibility Metrics

- **Availability (A)** is the fraction of time that an IP address returns positive replies.
- **Volatility (V)** captures the number of transitions from up to down over survey.
- **Uptime (U)** is the median duration of positive replies from an IP address.
- Each statistic computed for IP addresses, then averaged over a /24 subnet.
Spammer List

• Spammer data from eSoft.com
  – Two lists: Block list and Raw list
  – Both delivered to CSU every 30mins
  – (yes we archive and we can share)
• List of IP addresses with spam score per address
  – Score range: -60 to +70
  – Score >30: spam with high confidence (conservative)
• We use eSoft’s Raw List:
  – ~1.25M addresses spanning 400k /24 subnets daily
  – We assume score >= 20 is spammer
eSoft World Coverage

eSoft has pretty good coverage of the world
eSoft List Score Distribution

Distribution of eSoft Raw Scores

Number of Addresses

Raw Spam Scores

$\times 10^4$
Research Methodology

• Correlate ping survey data with eSoft list between Sept. 14-28, 2009
• Intersect data from the survey and eSoft to identify spamming subnets
• **The rest are** Non-spamming subnets, i.e., have no spammers (yes, this might be a weak assumption)
• Study the differences between spamming and non-spamming subnets.
Spammer Distribution

- Most subnets have fewer than 5 spamming hosts
Non-Spamming hosts much more evenly distributed

..but note a large number of subnets that are almost fully populated.
Question 1: Address Characteristics

- Question: Do spammer and non-spammer subnets have different IP characteristics (availability, volatility, uptime)?

- Approach:
  - intersect blacklist and survey subnets and study their characteristics
  - before intersection: 818k blacklist and 20k survey subnets
  - after intersection: 4k spamming and 15k non-spamming subnets.
Address Availability

Spammer vs. Non-Spammer Availability

• 72% of non-spammers but only 50% of spammers have >0.5 availability
• 50% of non-spammers but only 24% of spammers have >0.8 availability
Address Volatility

Spammer vs. Non-Spammer Volatility

- 90% of non-spammers but only 75% of spammers have <0.02 volatility
- 50% of non-spammers but only 28% of spammers have <0.01 volatility
Address Uptime

Spammer vs. Non-Spammer Uptime

- 70% of non-spammers, 42% of spammers have > 14 hour uptime
- 44% of non-spammers, 22% of spammers have > 28 hour uptime
Availability with Spam Score

- 83% of low spammers have > 0.9 availability.
- 14% of high spammers have > 0.9 availability.
Question 2: Domain Names

• Question: How do spammer domain names differ from non-spammer names?
• Approach:
  – resolve all names in intersected subnets using Linux host command
  – categorize based on key strings in the name
Domain Name Comparison

- 2X the spammers in dynamic category, 30.5% vs. 15.3%.
- 3X the non-spammers in static category, 14.1% versus 4.2%.
Question 3: Collateral Damage

• Question: Is blocking the entire /24 subnet a good idea when one or more addresses have been used for sending spam?

**Collateral Damage** consists of legitimate mail servers that are incorrectly blacklisted.

• Approach:

  1) Compute population of spamming hosts versus non-spamming hosts per subnet.
  2) Quantify the number of legitimate mail servers in subnets with spammers.
Collateral Damage: Population

Spammers versus NonSpammers in Intersection

- Many subnets do have spammers (and may get black listed)
- Blue cluster shows high spammer activity
- Diagonal blue cluster shows some highly compromised subnets - negligent or collaborating provider?
Collateral Damage: Results

**TABLE II**
**COLLATERAL DAMAGE STUDY**

<table>
<thead>
<tr>
<th>Description</th>
<th>Domains</th>
<th>Hosts</th>
<th>Subnets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intersected Subnets</td>
<td>646,040</td>
<td>4,126</td>
<td></td>
</tr>
<tr>
<td>Domain Query Timeout</td>
<td>12,899</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domain Query Invalid</td>
<td>175,535</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domain Query Valid</td>
<td>457,606</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unique Domain Names</td>
<td>4,044</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number Mail Servers</td>
<td>6,718</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unique Mail Servers</td>
<td>3,872</td>
<td>2,154</td>
<td></td>
</tr>
<tr>
<td>Collateral Damage</td>
<td>1,377</td>
<td>365</td>
<td></td>
</tr>
</tbody>
</table>

- Collateral damage in 365 subnets out of 4,126 studied (8.8%)
- This seems significant to us
Robustness

• Ping-based address probes undercount the number of responsive addresses
• Spam list may not be complete (depends on eSoft’s customer reach)
• Email volume from servers isn’t considered, some servers may be receive-only
• Spam blacklists vary greatly between vendors, no industry standard for scores
Conclusions

• Significant differences in IP availability, volatility uptime and domain names between spamming and non-spamming hosts
• Network behavior can be used to help identify and mitigate spamming behavior
• Coarse-grained blacklisting of /24 blocks incurs significant collateral damage
Acknowledgements

• Yuri Pradkin, and Xue Cai (USC/ISI) for access to survey data sets.
• Dan Massey, and Steve DiBenedetto (CSU) for help in many areas.
Automatic IP Hit list Generation

Xun Fan and John Heidemann
USC/ISI
Research:
IP Hitlist Generation

- an IP hitlist is a list of representatives for each edge network
- essential input to
  - traceroute mapping (CAIDA’s Skitter, Ark, etc.)
  - routing reachability studies (Bush et al.)
- ideal hitlist: current, complete, stable, reachable

traceroute to www.mit.edu (18.9.22.169), 30 hops max, 60 byte packets
1 router.postel.org (128.9.112.7) 0.624 ms 1.040 ms 1.475 ms
2 198.32.16.30 (198.32.16.30) 0.262 ms 0.307 ms 0.376 ms
3 lax-bpr.cosnetos-bpr.cenic.net (137.164.27.241) 0.781 ms 0.837 ms 0.885 ms
4 bpr-nlr-pn--lax-bpr.cenic.net (137.164.26.150) 1.417 ms 1.416 ms 1.411 ms
5 houz-loosa-87.layer3.nlr.net (216.24.186.31) 32.885 ms 32.901 ms 32.888 ms
6 atla-houa-70.layer3.nlr.net (216.24.186.9) 57.642 ms 57.593 ms 57.561 ms
7 wash-atla-64.layer3.nlr.net (216.24.186.21) 71.317 ms 70.982 ms 71.146 ms
8 newy-wash-98.layer3.nlr.net (216.24.186.22) 77.498 ms 77.511 ms 77.493 ms
9 216.24.184.102 (216.24.184.102) 76.360 ms 76.437 ms 76.480 ms
10 OC11-RTR-1-BACKBONE-2.MIT.EDU (18.168.1.41) 82.744 ms 82.788 ms 82.857 ms
11 * * *

**THE INTERNET: 01.01.00**
Automatic Hitlist Generation

study series of censuses
(data on all reachablity)

look at each /24’s history

addresses

time

to find best representative for each /24 over whole Internet
Hitlist Design Questions

• how much history is needed?
  – A: more is better, 8 censuses (24 months) enough

• what function of history best predicts future?

<table>
<thead>
<tr>
<th>Function</th>
<th>Equation</th>
<th>Input History</th>
<th>Calculation</th>
<th>Score</th>
<th>Predictivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>[ y = \sum_{i=1}^{17} Bi ]</td>
<td>0000000000000001011</td>
<td>0+0+...+1+0+1+1</td>
<td>3</td>
<td>54%</td>
</tr>
<tr>
<td>Linear</td>
<td>[ y = \sum_{i=1}^{17} a_{i} * Bi ]</td>
<td>0000000000000001011</td>
<td>14+16+17</td>
<td>47</td>
<td>55%</td>
</tr>
<tr>
<td>Power</td>
<td>[ y = \sum_{i=1}^{17} 1/(18-i) * Bi ]</td>
<td>0000000000000001011</td>
<td>1/4+1/3+0+1/2+1</td>
<td>1.75</td>
<td>56%</td>
</tr>
</tbody>
</table>
Fundamental Limits of Hitlist Accuracy

- accuracy: will representative be there?
- what accuracy should be expected?
  - best possible hitlist accuracy is ~60%
  - (even with >3 year history!)
- preliminary explanation [work-in-progress!]
  - 40% of the network is *unstable*
  - dynamically addressed or firewalled
  - (confirms: manual hitlists are unmaintainable)