What’s In A Name?

Wes Hardaker, Haoyu Jiang
October 2019

<hardaker@isi.edu>
What is all this stuff??
Overview

- Research started by USC masters student Haoyu Jiang
- Data analyzed: B-Root’s contributions to DITL 2018
- Results of breaking down DITL root traffic in new-ish ways:
  - By “chrome”
  - By language
  - By length
- Future Work
Chromes Effect On Root Server Traffic

• When chrome starts up, it generates 3 garbage queries
  − To detect “pay-walls” or other DNS rewriting

  
  semantics {
    sender: "Intranet Redirect Detector"
    description:
      "This component sends requests to three randomly generated, and "
      "thus likely nonexistent, hostnames. If at least two redirect to "
      "the same hostname, this suggests the ISP is hijacking NXDOMAIN,"
      "and the omnibox should treat similar redirected navigations as "
      "'failed' when deciding whether to prompt the user with a 'did you "
      "mean to navigate' infobar for certain search inputs."
  
    trigger: "On startup and when IP address of the computer changes."
Chromes Effect On Root Server Traffic

- When chrome starts up, it generates 3 garbage queries
  - To detect “pay-walls” or other DNS rewriting

```cpp
// Start three fetchers on random hostnames.
for (size_t i = 0; i < 3; ++i) {
    std::string url_string("http://");
    // We generate a random hostname with between 7 and 15 characters.
    const int num_chars = base::RandInt(7, 15);
    for (int j = 0; j < num_chars; ++j)
        url_string += ('a' + base::RandInt(0, 'z' - 'a'));
    GURL random_url(url_string + '/');
}```
All these requests end up at the root

• But what to what effect?
• How much of the root traffic is garbage names?
Histogram of the length of records in the root zone
Requests received per single-label length
Total Traffic vs Multiple-Labels vs Single-Label vs Root Name
What are the **single-labels**?
Studying Single Labels by Common Languages
Machine Learning: Here’s a label, what language is it?
Analyzing Traffic by Label Count
Reminder: this is what the total traffic looks like
Traffic per minute for **QNames** with a **single label**
Note: typically 1 > 3 > 4 > 2
Top 25 right-hand 2-labels in and outside the event

Quantity of requests received during 2 minutes within the plateau and just after the plateau
Let's look at these Quantity of requests received during 2 minutes within the plateau and just after the plateau
Results: DGA like queries to multiple TLDs

- novartiscqmsumerhealth.at
- service-novartisadvisory.tm
- novovarrtishealthh.tw
- initiat-ive-familien-ba-nde.tm
- iniitiative-familienbannnde.at
- novartis-c-ibavis-ion.tm
- sandozkerdiyolo1i.at
- wwwadidasgolfcomvn.at
- wwwsandozpolskainth.tm
- n-ova-rtisege-nvard.tm
- authorize.mynet
- sandozkarbjiy0loji.at
- authorize.mynet
- sandozkarbjiy0loji.at
- initiative--familien-ban-de.tw
- car-novartiseyecare.tw

- novartlseomsumerhealth.tm
- dealsconstellation-alcon.tm
- constellation-alconuser.tw
- novartlseomsumerhealth.tm
- dealsconstellation-alcon.tm
- constellation-alconuser.tw
- noovarrtishealthh.tw
- novartis1ntarnationel.tw
- qswmhudwx.mynet
- wwwnovartiseyecarenetnz.tm
- qswmhudwx.mynet
- dvugmdn.mynet
- wwwnovartiseyecarenetnz.tm
- dvugmdn.mynet
- nov-a-r-tisinternational.tw
- novartjsconsumcrhe1th.tw
6 and 7 label traffic looks “sort of normal”

6 was afraid of 7...
8 and 9 labels are much more periodic
Length 8 and 9

- In 100k names:
  - openstacklocal: 34434
  - local: 15551
  - localdomain: 6710
  - net: 4804
  - virtual: 4629
  - com: 4073
  - internal: 2920
  - LOCAL: 2101
openstacklocal

AAAA  IP.bogons.cymru.com.openstacklocal.
AAAA  IP.badconf.rhsbl.sorbs.net.openstacklocal.
A     IP.cblplus.anti-spam.org.cn.openstacklocal.
AAAA  IP.rbl.interserver.net.openstacklocal.
AAAA  IP.cbl.anti-spam.org.cn.openstacklocal.  (99.7% of these were from one ASN)
A     IP.dyna.spamrats.com.openstacklocal.
A     IP.misc.dnsbl.sorbs.net.openstacklocal.
AAAA  IP.cblplus.anti-spam.org.cn.openstacklocal.  Note: real IPv4 → “IP”
A     IP.dnsbl.rangers.eu.org.openstacklocal.
Length 20

• Some queries make sense of course
  - SOA 0.0.0.0.0.9.4.c.6.f.6.f.6.2.0.0.2.ip6.arpa.
  - SOA 0.0.0.0.0.0.a.5.e.6.f.6.f.6.2.0.0.2.ip6.arpa.
  - SOA 0.0.0.0.0.b.5.c.1.3.0.a.a.2.0.0.2.ip6.arpa.

• Some are data leaks or mistakes
  - www.2-17.2-12.2-6.1.noarch.i386.i386.i386.2.0-1.1.2-12.i386.2-6.i386.1-1.i386.kde.rpm.
Length 23 – 80% sophosxl.net

- 11% .arpa
- 80% sophosxl.net
  - 0.0.163.0.0.132.0.0.16.6.3.0.0.0.0.00.04.b2610633ea1476298c681f8016dd3d39d60c43f11c9cd75cfdb2ed7e011f14c.f.00.s.sophosxl.net.
  - 0.0.39d4.0.0.0.0.11b4.0.8d5.0.0.0.00.01.02660438680dbfa8e175d4b48ea1953b0cf91c694a2ea920280cf3d18f6b22a.f.00.s.sophosxl.net.
  - 0.0.7c.0.0.11.0.0.7.0.2.0.0.0.00.01.b2610633ea1476298c681f8016dd3d39d90f6b3ae45b1ef23bb11a5eae2d8da.f.00.s.sophosxl.net.
  - 0.0.76.0.0.61.0.0.2.5.0.0.0.0.5.06.04.61a4b12b2a9c745aaafaa74594443bcfb52a458e69b66f69c4a9948530c4754.f.06.s.sophosxl.net.
The longer lengths start to get random and low-level
A few long lengths have sudden bursts of energy.
Let's end with a look at some of these oddballs
Length 66+ example one: 67%

\255\001\001\001\001\001.A.\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001\001.
Length 66+ example two: 28%

m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.
m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.
m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.
omdok3mjiymjczn泽2otg4ng.
m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.
m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.
m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.
watermax.supopupimage.js
m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.
m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.
m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.
kinobucks.ruodk3mjiymjczn泽2otg4ng.
m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.m.
cryptovpn.netodk3mjiymjczn泽2otg4ng
Length 66+ example three: 4%

- 40.47.67.111.109.111.110.47.115.111.97.112.114.100.45.114.111.45.103.101.116.109.101.109.98.101.114.118.51.45.49.50.49.49.52.95.112.111.110.108.21.47.67.111.109.109.111.110.47. [SNIP].12113.
- 44.47.67.111.109.111.110.47.111.108.110.112.114.100.50.45.103.101.116.109.101.109.98.101.114.99.111.117.110.116.101.114.115.45.49.57.48.49.48.95.112.111.111.108.21.47.67.111.109.109.111.110.47. [SNIP].19010.
- 40.47.67.111.109.111.110.47.119.119.119.46.104.97.114.118.97.114.100.112.105.108.103.114.105.109.45.108.98.114.45.57.52.48.49.95.112.111.108.20.47.67.111.109.109.111.110.47. [SNIP].9411.
Length 66+ example three: 4%

- 40.47.67.111.109.111.110.47.115.111.97.112.114.100.45.114.111.45.103.101.116.109.101.109.98.101.114.118.51.45.49.50.49.49.52.95.112.111.111.108.21.47.67.111.109.109.111.110.47.[SNIP].12113.


- 40.47.67.111.109.111.110.47.119.119.119.46.104.97.114.118.97.114.100.112.105.108.103.114.105.109.45.108.98.114.45.57.52.48.49.95.112.111.111.108.20.47.67.111.109.109.111.110.47.[SNIP].9411.

DOTTED ASCII???
Length 66+: example three decoded

• /Common/soaprd4-getmembercounters-16030_pool/Common/IP
• (/Common/soaprd-ro-getmemberv3-12114_pool/Common/IP
• +/Common/soaprd1-getmemberporttype-6693_pool/Common/IP
Conclusions and Future Directions

- Better DNS name classification mechanisms
- Better analysis of temporal patterning
- Deeper dive into language studies
- Questions?