Dynamics of Online Scam Hosting Infrastructure

Maria Konte, Nick Feamster
Georgia Tech

Jaeyeon Jung
Intel Research

Online Scams

October 8, 2008 4:28 PM PDT

How botnets use 'bullet-proof' domains

by Robert Vamosi

Font size Print

Botnets are proving to more resilient and harder to shut down.

That's largely due to an increased use of methods people use to obscure the domain by constantly mapping to different bots within the network, according to a recently released study **(PDF)**.

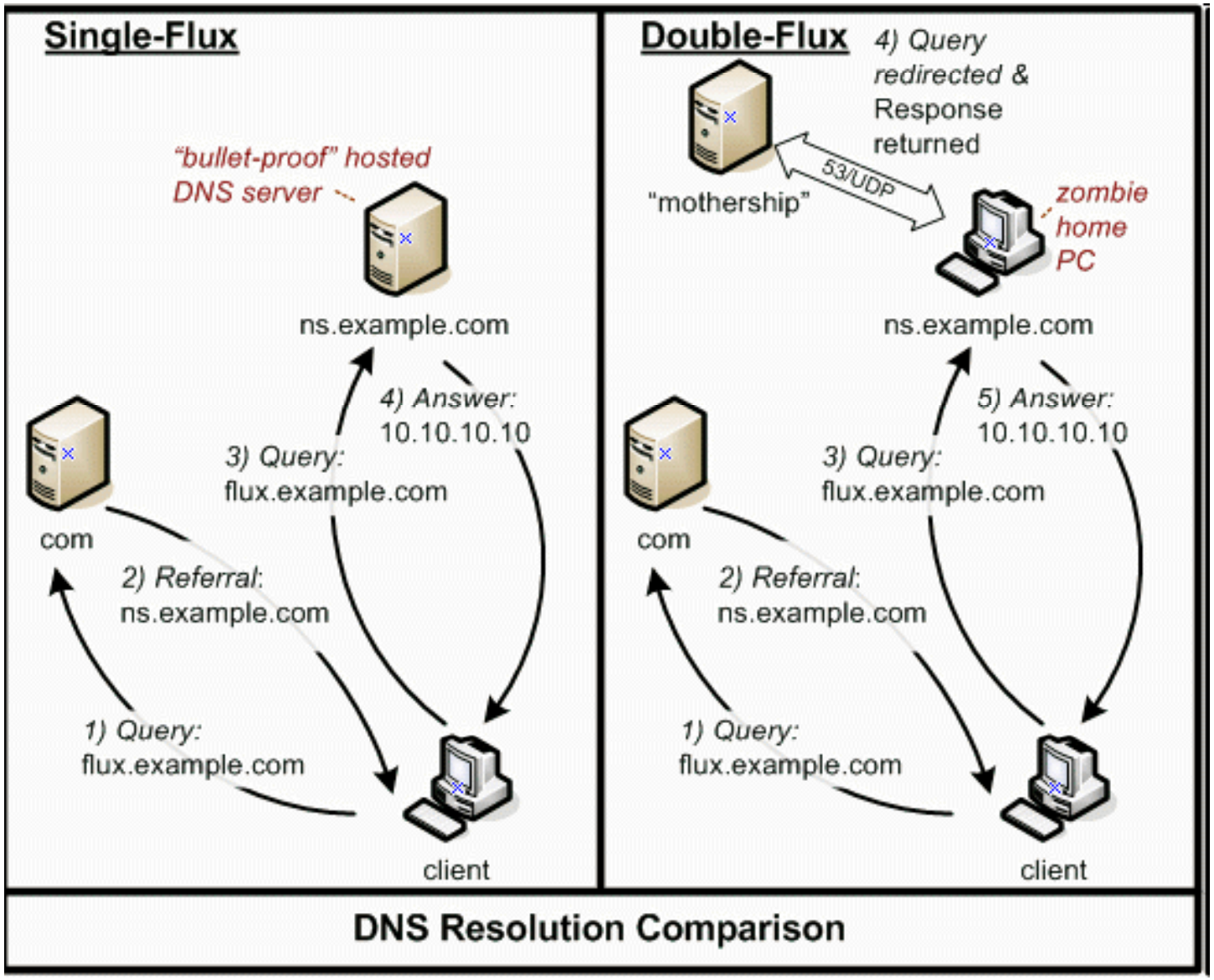
- Often advertised in spam messages
- URLs point to various point-of-sale sites
- These scams continue to be a menace
  - As of August 2007, **one in every 87 emails** constituted a phishing attack
- Scams often hosted on bullet-proof domains

- **Problem:** Study the dynamics of online scams, as seen at a large spam sinkhole

Online Scam Hosting is Dynamic

- The sites pointed to by a URL that is received in an email message may point to different sites
- Maintains agility as sites are shut down, blacklisted, etc.
- One mechanism for hosting sites: **fast flux**

Overview of Dynamics



Source: HoneyNet Project

Why Study Dynamics?

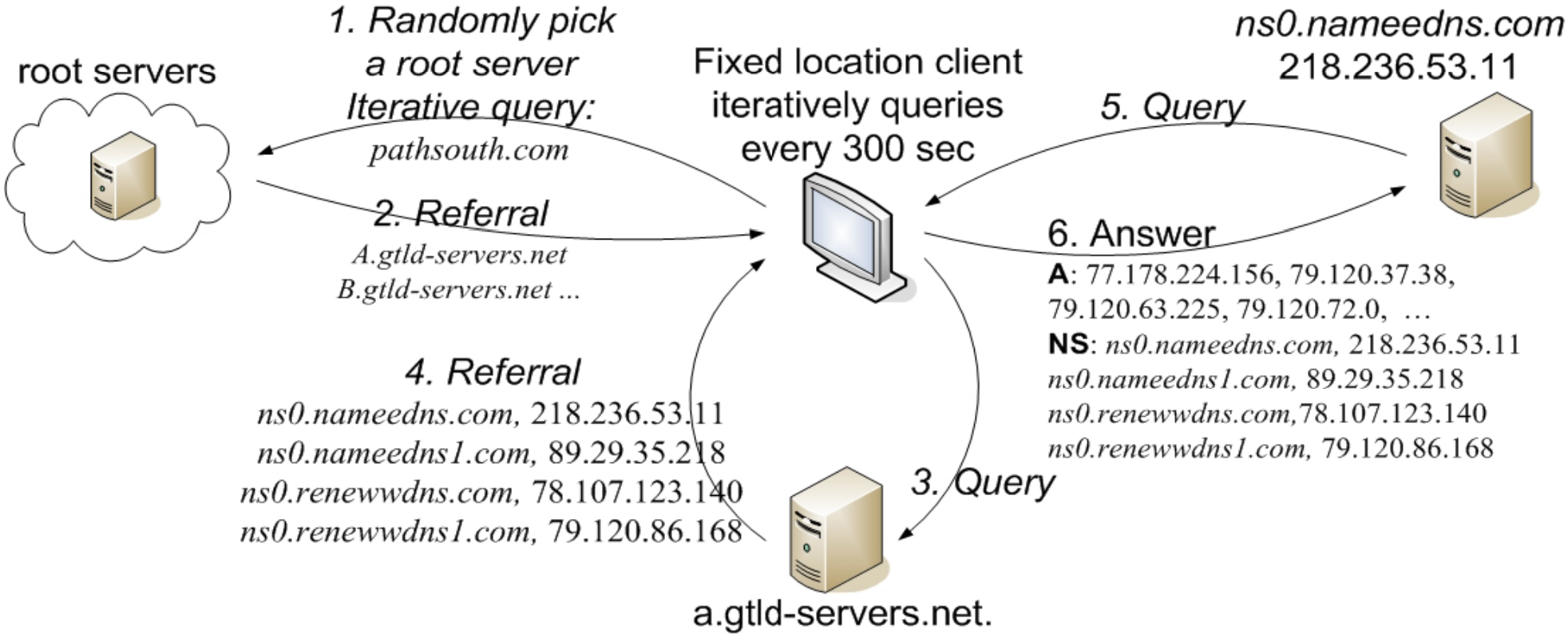
- Understanding
  - What are the possible invariants?
  - How many different scam-hosting sites are there?

- Detection
  - **Today:** Blacklisting based on URLs
  - **Instead:** Identify the network-level behavior of a scam-hosting site

Summary of Findings

- What are the *rates and extents* of change?
  - Different from legitimate load balance
  - Different cross different scam campaigns
- How are dynamics *implemented*?
  - Many scam campaigns change DNS mappings at all three locations in the DNS hierarchy
    - A, NS, IP address of NS record
- **Conclusion:** Might be able to detect based on monitoring the dynamic behavior of URLs

Data Collection



- **One month of email spamtrap data**
  - 115,000 emails
  - 384 unique domains
  - 24 unique spam campaigns

Top 3 Spam Campaigns

Campaign	Spam emails	Spam advertising IPs	Campaign domains	Fluxing Domains	IPs of A rec	IPs of NS rec	IPs of both A+NS rec
Pharmacy-A	18459	11670	149	149	9448	2340	9705
Watch-A	40681	30411	34	30	1516	225	1572
Watch-B	454	427	43	19	1204	219	1267
All campaigns	115198	77030	465	384	9521	2421	9821
Alexa data set				500	1048	852	1877

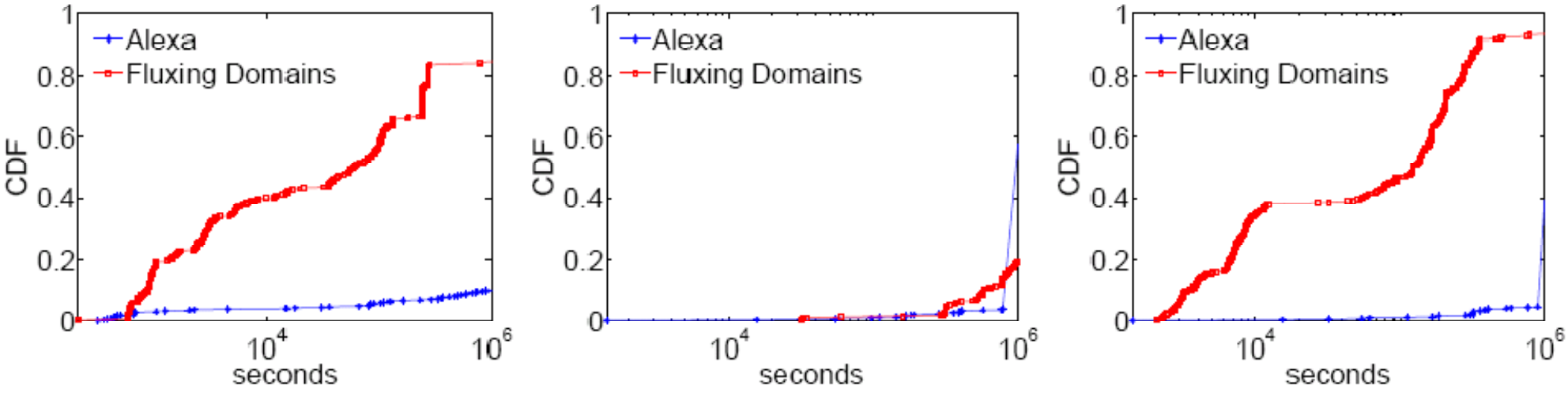
- Some campaigns hosted by thousands of IPs
- Most scam domains exhibit some type of flux
- Sharing of IP addresses across different roles (authoritative NS and scam hosting)

Time Between Changes

- **How quickly do DNS-record mappings change?**

- Scam domains change on shorter intervals than their TTL values
- Domains within the same campaign exhibit similar rates of change

Rates of Change



(a) A records (b) NS records (c) IP of NS records

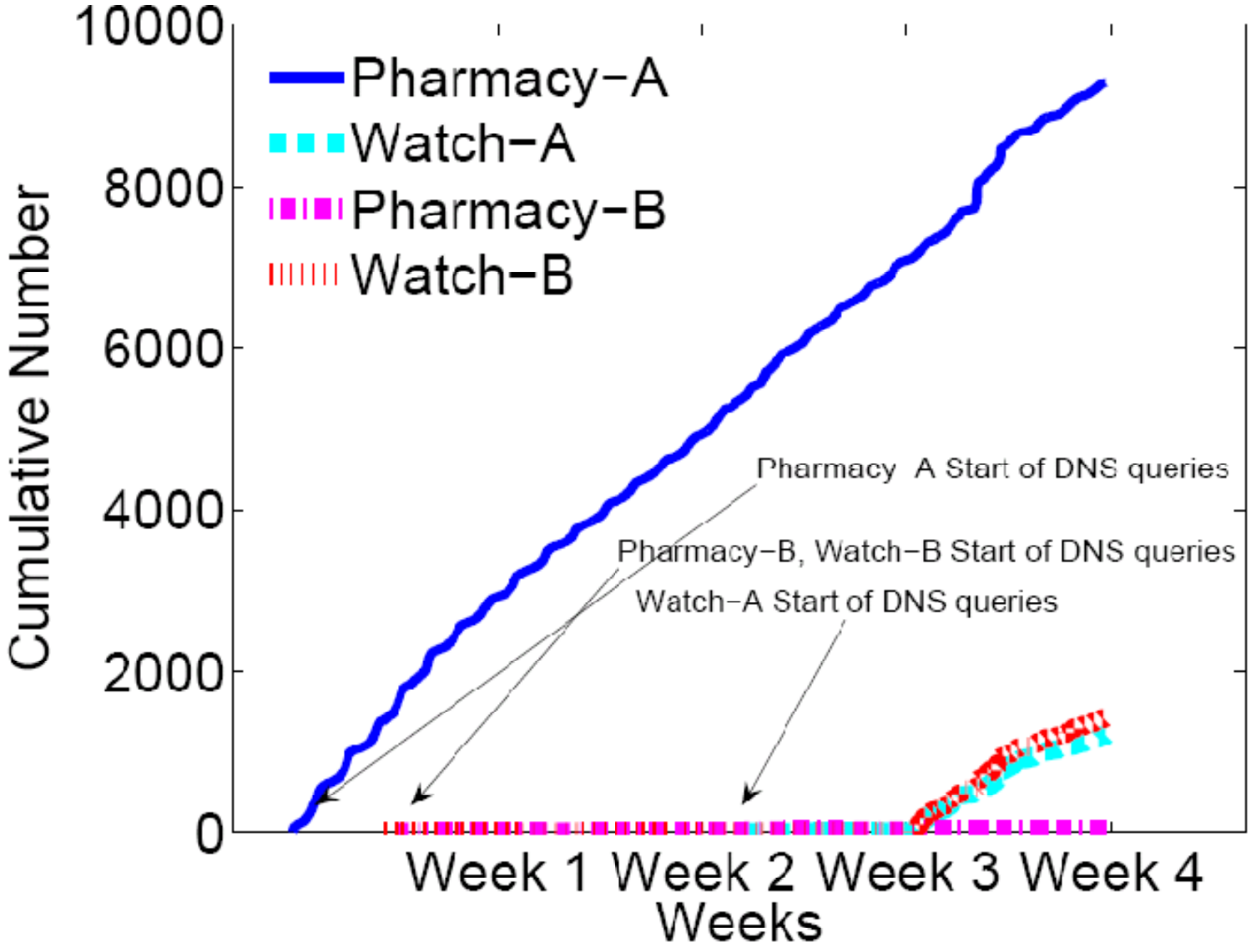
- Domains that exhibit fast flux change more rapidly than legitimate domains
- Rates of change are inconsistent with actual TTL values

Rates of Accumulation

- **How quickly do scams accumulate new IP addresses?**

- Rates of accumulation differ across campaigns
- Some scams only begin accumulating IP addresses after some time

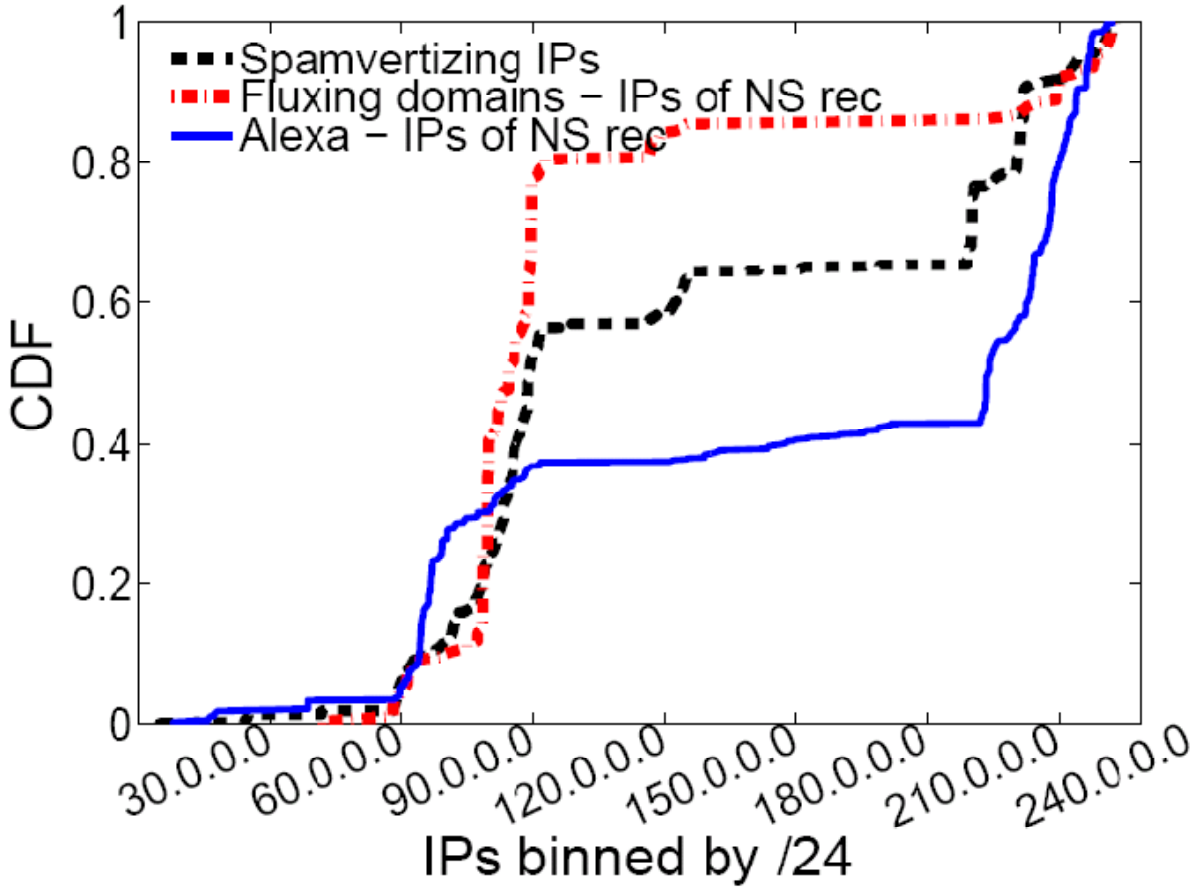
Rates of Accumulation



Location of Change in Hierarchy

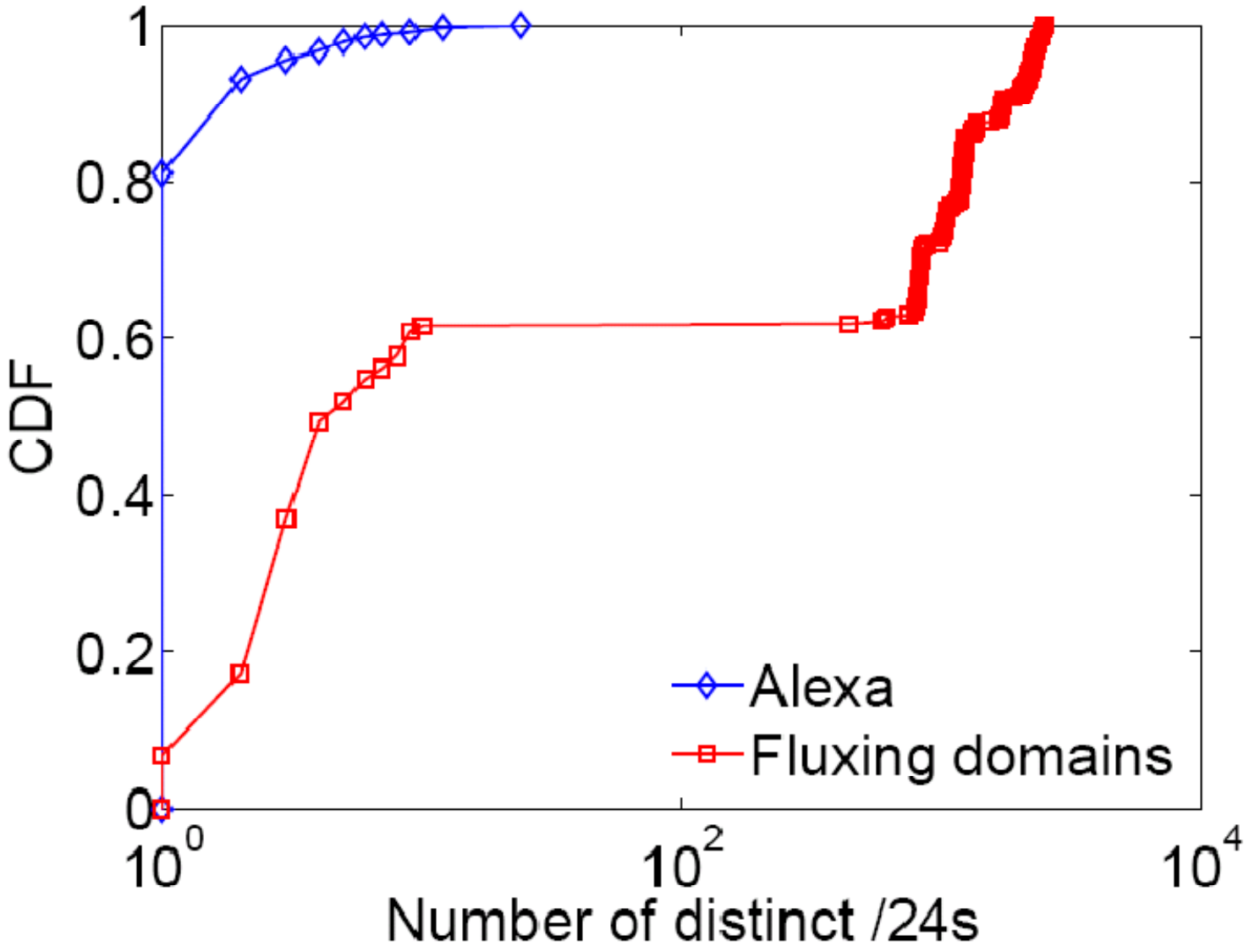
- Scam networks use a different portion of the IP address space than legitimate sites
  - 30/8 – 60/8 --- lots of legitimate sites, no scam sites
- DNS lookups for scam domains are often more widely distributed than those for legitimate sites

Location in IP Address Space



- Scam campaign infrastructure is considerably more concentrated in the 80/8-90/8 range

Distribution of DNS Records



Registrars Involved in Changes

Registrar	Country	Domains	Registrar	Country	Domains
dns.com.cn	China	180 (46.9%)	leadnetworks.com	India	3 (0.8%)
paycenter.com.cn	China	65 (16.9%)	coolhandle.com	US	2 (0.5%)
todaynic.com	China	12 (3.1%)	webair.com	US	1 (0.3%)
signdomains.com	India	7 (1.8%)	stargateinc.com	US	1 (0.3%)
total active domains: 271 (70.6%)					

- About 70% of domains still active are registered at eight domains
- Three registrars responsible for 257 domains (95% of those still marked as active)

Conclusion

- Scam campaigns rely on a dynamic hosting infrastructure
- Studying the dynamics of that infrastructure may help us develop better detection methods
- Dynamics
  - Rates of change differ from legitimate sites, and differ across campaigns
  - Dynamics implemented at all levels of DNS hierarchy
- Location
  - Scam sites distributed more across IP address space

http://www.cc.gatech.edu/research/reports/GT-CS-08-07.pdf