Pope: Infrastructure for DNS Scanning

Manos Antonakakis[1] David Dagon[1] Luo “Daniel” Xiapu[1]

[1]{manos@cc,dagon@cc,xluo7@mail}.gatech.edu
Georgia Institute of Technology
Information Security Center
Atlanta, Georgia

OARC 2009 - Altanta, Georgia

Outline

Georgia Tech Campus
(Cross Sectional View)

based on joint work with:

- David Dagon
- Luo “Daniel” Xiapu
- Artem Dinaburg
- Justin Bellmor,Chis Lee and the GT-OIT stuff for the abuse handling
- Many thanks to Paul, Duane from ISC and Norm, Matthew from CIRA for the DNS scan points

Outline

- Challenges in DNS scanning
- Pope components
- Data analysis
- Poisoning Detection

Challenges in DNS scanning

- Abuse - you don’t want to stop scanning
- How can you scan efficiently and safe (for the network you’re in)
- You need to loop over “interesting” open-recursive servers
- Can we easily shift the scan on a certain AS/CIDR/geographical region to increase our confidence on a detection heuristic like RTT of IP reputation possible associated with an attack?

Challenges in DNS scanning

- Abuse - you don’t want to stop scanning
- How can you scan efficiently and safe (for the network you’re in)
- You need to loop over “interesting” open-recursive servers
- Can we easily shift the scan on a certain AS/CIDR/geographical region to increase our confidence on a detection heuristic like RTT of IP reputation possible associated with an attack?

Challenges in DNS scanning

- Abuse - you don’t want to stop scanning
- How can you scan efficiently and safe (for the network you’re in)
- You need to loop over “interesting” open-recursive servers
- Can we easily shift the scan on a certain AS/CIDR/geographical region to increase our confidence on a detection heuristic like RTT of IP reputation possible associated with an attack?

Challenges in DNS scanning

- Abuse - you don’t want to stop scanning
- How can you scan efficiently and safe (for the network you’re in)
- You need to loop over “interesting” open-recursive servers
- Can we easily shift the scan on a certain AS/CIDR/geographical region to increase our confidence on a detection heuristic like RTT of IP reputation possible associated with an attack?

Pope components

- Pope Scan Control Center
  - Open-recursive lists
  - Domain name lists
  - Do-not scan lists
- Pope Scan Engine
  - Categories of scanning
  - ICMP errors logging (3,{1,2,3,9,10,11,12})
  - dnscap based network trace
- Pope Data Analysis
  - DB “scan”
  - Network trace file aggregation point and analysis
  - Reports

Pope components

- Pope Scan Control Center
  - Open-recursive lists
  - Domain name lists
  - Do-not scan lists
- Pope Scan Engine
  - Categories of scanning
  - ICMP errors logging (3,{1,2,3,9,10,11,12})
  - dnscap based network trace
- Pope Data Analysis
  - DB “scan”
  - Network trace file aggregation point and analysis
  - Reports

Pope components

- Pope Scan Control Center
  - Open-recursive lists
  - Domain name lists
  - Do-not scan lists
- Pope Scan Engine
  - Categories of scanning
  - ICMP errors logging (3,{1,2,3,9,10,11,12})
  - dnscap based network trace
- Pope Data Analysis
  - DB “scan”
  - Network trace file aggregation point and analysis
  - Reports

Pope Scan Control Center

- python xml-rpc based
- Responsible for blacklist checking for all new tasks
- Can provide a random list of active open recursive servers
- Responsible for r-syncing the trace files from the scan points to the archive server

Pope Scan Engine

- Build on-top of dnspython
- Requires as input an ORDNS list, a domain name list and the type of scan operation
- What type of scan operation we support:
  - Simple single A query (lightweight poison scan)
  - Pope type of scanning (1 A com + 3 A DN + 3 A <RND>.DN + 1 {NS,MX,AAAA})
  - SPR - gtexperimentdns.{org,biz,info}
  - Discovery of new ORDNS per CIDR (requires list of CIDRs and not IPs)
- Additional possible actions:
  - Stop/pause/restart scanning
  - Remove IP from all active scans - update do-not-scan lists
  - Retrieve ICMP-errors logs - adjust do not scan list
  - Adjust scan rate for a specific scan point

Pope Scan Engine

- DB scan
  - Holds the last 30 days DNS transactions
  - Capable to provide information for:
    - RTT deltas / zone / AS / CIDR / ORDNs
    - DN-NS-TTL mappings / zone / AS / CIDR / ORDNs
    - DN-IP-TTL mappings / zone / AS / CIDR / ORDNs
    - Historic SPR for each active ORDNs
    - IP reputation
    - Based on a white list and IP frequencies we can automatically identify highly likely new poisoning cases
  - Automatically inserts any new captured dnscap files when dropped into the archive server

Heuristics: Using RTTs

- cons
  - We need to first build a profile for delta RTTs
  - Keep track of NXDOMAINs: is hard to be sure that NXDOMAIN was originated from the malicious ANS
  - We need to ask more than one zones at the same time to exclude network lag
  - The malicious ANS needs to have significant network distance than the legit ANS
- pros
  - No state tracking necessary
  - The RTT profile is relatively easy to maintain for the most phisable domain names
  - Once you've build the normal profile some cases of poisoning will be easily detectable
  - Theoretical easy to enforce this in transit layer

Heuristics: Using IP reputation and frequencies

- IP frequency for each zone
  - Works extremely well for domain names that hold their ANSs in few CIDR
  - Works very well for domain names that tend to keep deferent ANSs per geographical region, after a small training period
  - If the zone is AKAMAIzed then it is very hard to use this heuristic by itself as detection criterion
  - Not a trivial task to white-list all CIDRs that AKAMAI uses highly sparse and in the majority of the cases they utilize slash 24 networks

DNS Poison cases verified

- More than 6000 cases for 400K open recursive servers that we probe the last 3 months
- The vast majority of these open-recursives uses only port 53
- Most cases in Asia and Europe.
- RFC3330 IPs periodically appear in A records
- Other interesting observations
  - Attackers try to evade "quick look" type of analysis
  - Kaminsky DNS-tunneling (e.g 85.237.87.174) and Vermicelli ToTD (e.g 88.50.108.34,88.50.111.82) needs to be excluded from the analysis
  - Recursive that constantly give us mixed false and legit IPs (e.g 61.67.67.154)
  - Incredible hard to identify all CIDRs owned by AKAMAI

Example Analysis: ebay.com resolved in ASN11305

ip	count	#	ip	count
66.135.195.180	22	#	192.5.6.30	32
66.135.195.181	20	#	192.12.94.30	36
66.135.205.13	520	#	192.26.92.30	44
66.135.205.14	538	#	192.31.80.30	44
66.135.207.137	657	#	192.33.14.30	34
66.135.207.138	657	#	192.35.51.30	42
66.135.215.5	657	#	192.41.162.30	14
66.135.221.10	539	#	192.42.93.30	30
66.135.221.11	540	#	192.43.172.30	28
66.135.223.137	657	#	192.48.79.30	30
66.211.160.87	533	#	192.52.178.30	22
66.211.160.88	543	#	192.54.112.30	32
192.55.83.30	8	#	216.113.167.215	22

Example Analysis: ebay.com resolved in ASN3462

```
       ip         | count | class |      info
------------------+-------+-------+----------------
...
 66.135.205.11    |    32 |     1 | ebay
 66.135.221.10    | 13472 |     1 | ebay
 66.135.221.11    | 13728 |     1 | ebay
 66.135.223.137   | 20496 |     1 | ebay
 66.208.160.87    |     9 |     4 | pois-manos
 66.208.160.88    |     3 |     4 | pois-manos
 66.209.160.87    |     6 |     4 | pois-manos
 66.209.160.88    |     2 |     4 | pois-manos
 66.211.160.87    |  8640 |     1 | ebay
 66.211.160.88    |  8500 |     1 | ebay
 192.12.94.30     |    12 |     1 | verisign
 192.42.93.30     |    12 |     1 | verisign
...
```

Example Analysis: ebay.com resolved in ASN3462

ip	count	class	owner
...			
66.135.205.11	32	1	EBAY - eBay, Inc
66.135.221.10	13472	1	EBAY - eBay, Inc
66.135.221.11	13728	1	EBAY - eBay, Inc
66.135.223.137	20496	1	EBAY - eBay, Inc
66.208.160.87	9	4	KEC-NET - Kentucky Educational Computing Network
66.208.160.88	3	4	KEC-NET - Kentucky Educational Computing Network
66.209.160.87	6	4	ASN-CXA-ALL-CCI-22773-RDC - Cox Communications Inc.
66.209.160.88	2	4	ASN-CXA-ALL-CCI-22773-RDC - Cox Communications Inc.
66.211.160.87	8640	1	EBAY - eBay, Inc
66.211.160.88	8500	1	EBAY - eBay, Inc
192.12.94.30	12	1	FGTLD - VeriSign Global Registry Services
192.42.93.30	12	1	GGTLD - VeriSign Global Registry Services
...			

Small portion of SPR analysis for 50K probed ORDNS

More than 60% of ORDNs scanned does not use more than 200 different ports

```
scan=# SELECT count(*) from sprpcnt where count<=200;
 count
-------
 37823
(1 row)

scan=# SELECT count(*) from sprpcnt where count<=1024;
 count
-------
 38847
(1 row)

scan=# SELECT count(*) from sprpcnt where count>=1024;
 count
-------
   159
(1 row)

scan=# SELECT count(*) from sprpcnt where count<=2;
 count
-------
   817
(1 row)
```

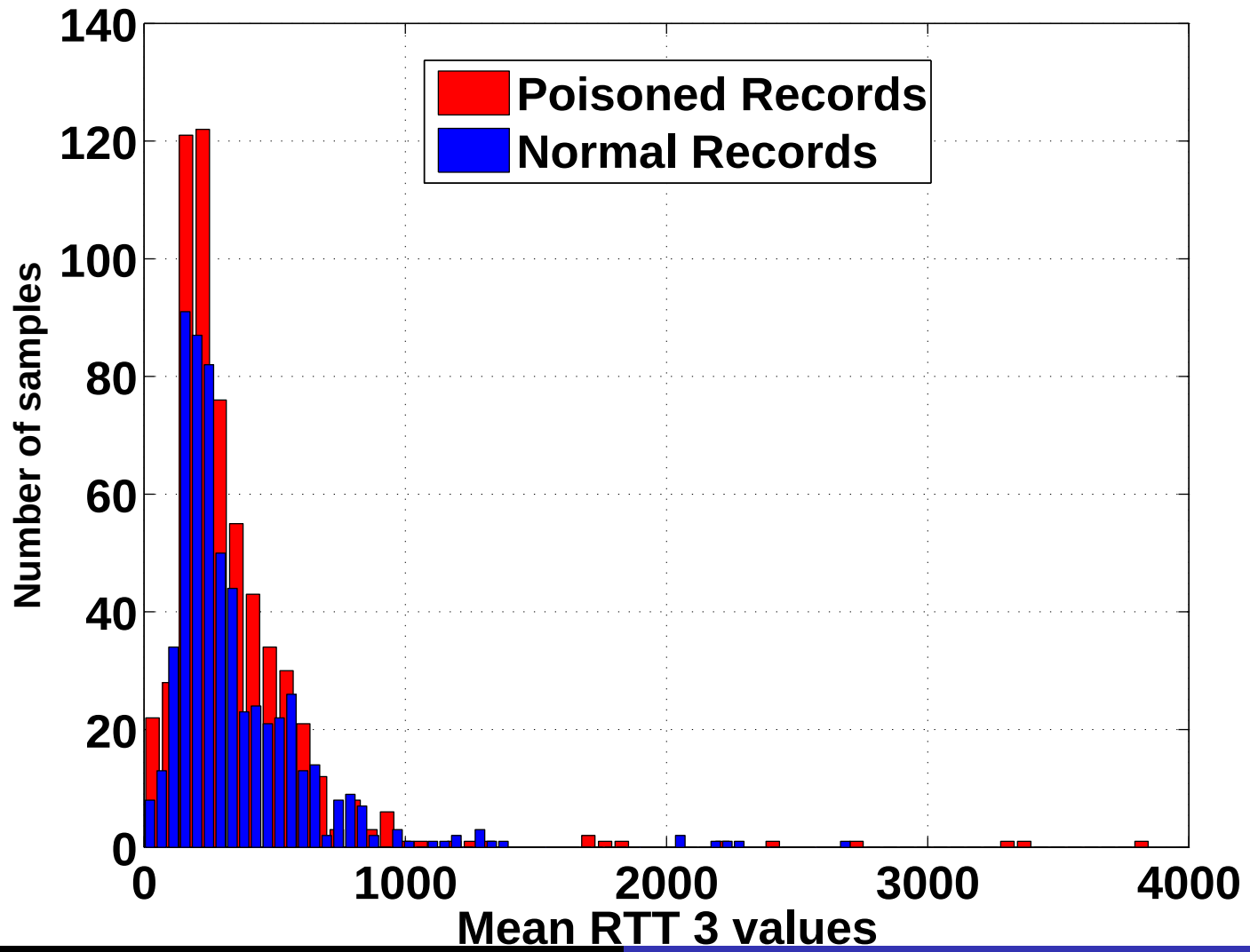
Example Analysis: Chameleon Poisoning

```
#ip             ordns             dn          date
#-------------------------------------------------------------------------
66.208.160.87   60.250.77.70      ebay.com    2009-01-09 03:25:59.955633
66.208.160.87   163.25.24.250     ebay.com    2009-01-09 02:12:56.058606
66.208.160.88   220.128.214.2     ebay.com    2009-01-09 16:16:10.245063
66.208.160.88   220.229.200.248   ebay.com    2009-01-09 04:52:20.48586
66.208.168.209  211.20.98.131     paypal.com  2009-01-13 12:36:18.373829
66.208.168.209  220.128.214.2     paypal.com  2009-01-09 16:12:33.355915
66.208.168.209  163.25.24.250     paypal.com  2009-01-09 02:09:54.372559
66.208.168.209  220.128.251.130   paypal.com  2009-01-13 12:40:04.507792
66.208.168.209  202.85.110.67     paypal.com  2009-01-09 02:04:12.952607
66.208.168.209  60.54.211.150     paypal.com  2009-01-09 02:12:37.644716
212.77.101.134  62.146.123.89     paypal.com  2009-01-09 03:13:43.814504
212.77.101.134  62.146.123.89     paypal.com  2009-01-09 03:13:43.280711
212.77.101.134  62.146.123.89     paypal.com  2009-01-09 03:13:42.970945
```

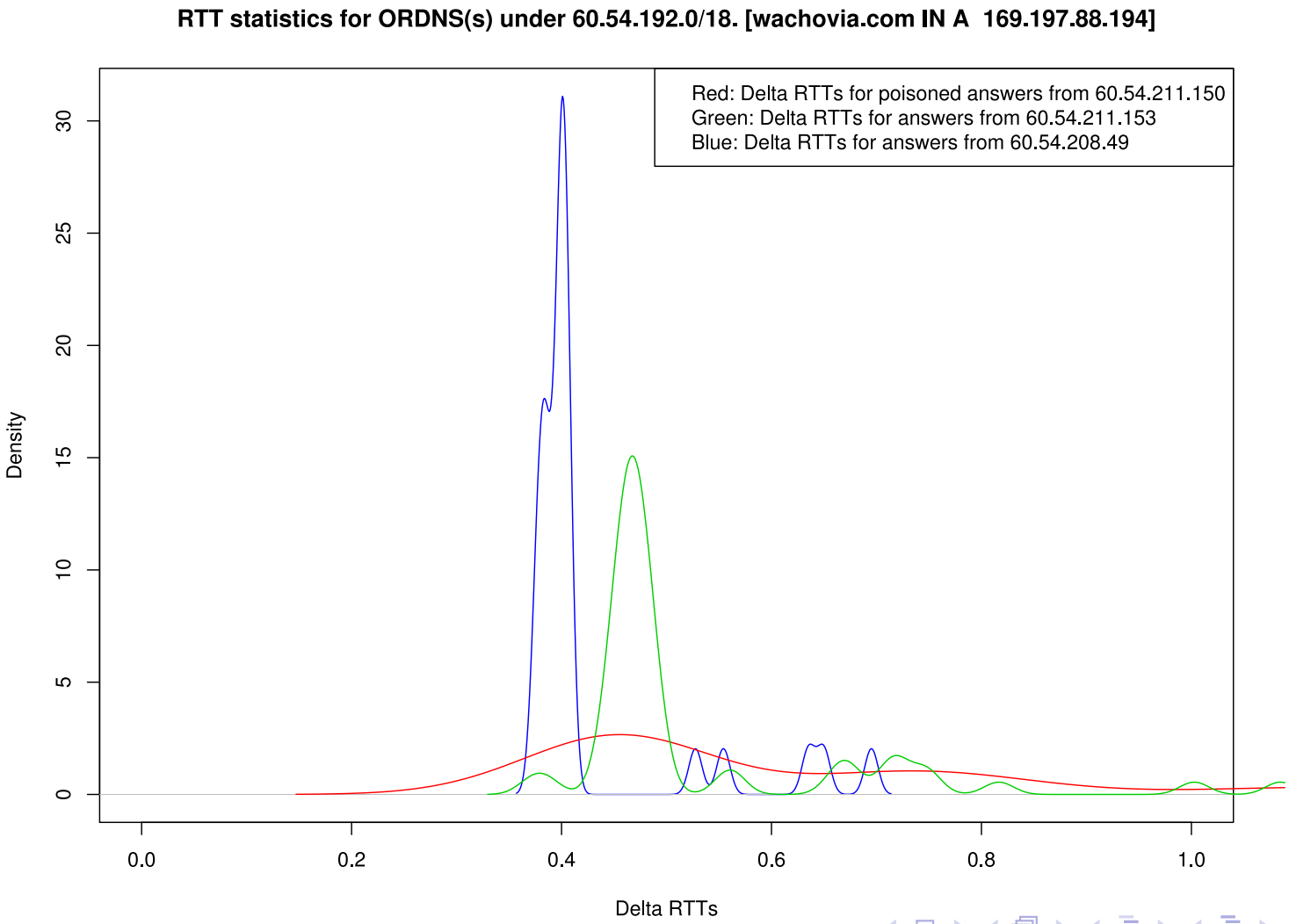
Example Analysis: Chameleon Poisoning

Fake	Real	Domain
66.**208**.168.209	66.**211**.168.209	paypal.com
66.**133**.221.10	66.**135**.221.10	ebay.com
157.**165**.224.26	157.**163**.224.26	cnn.com

Example Analysis: RTT Comparison



Example Analysis: RTT Comparison



Summary and Conclusion

1. Pope: slightly better than King
   - Provides “Poison ex machina” probes of openrec caches
   - Distributed scan infrastructure
   - XML-RPC work units sent to sensors, who probe cross product of ordns work unit ips and domains
     - knobs for throttling rate, recovery of lost work units
     - Sensors in Redwood City and Canada (thanks ISC and CIRA!)
     - Sensors coming to Tokyo, Chicago and Amsterdam
   - Real-time do-not-scan updates
   - Responsible handling of abuse@
     - Soon an update to RFC 1262
2. Data (shortly) available to OARC members