Detecting Malicious Activity with DNS Backscatter

Kensuke Fukuda • John Heidemann

Proc. of ACM IMC '15, pp. 197-210, 2015.

Presented by Xintong Wang and Han Zhang

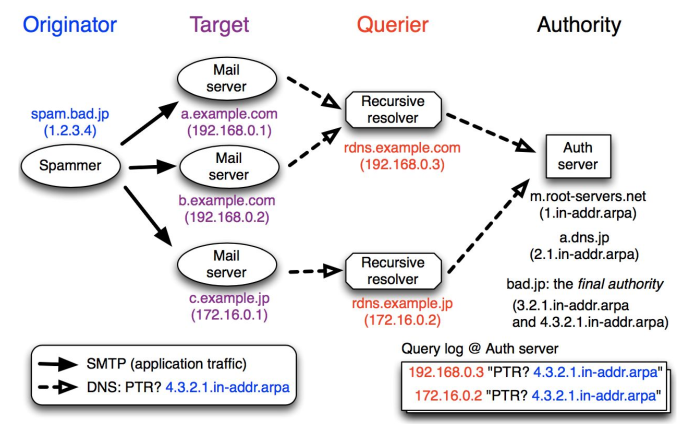
Challenges in Network Monitoring

- Need a better monitoring service for network-wide activities
  - Malicious activity: Spammer, scanner
  - Non-malicious activity: Ad tracker, CDN
- Hard to achieve: Decentralized nature
- Reverse DNS (DNS Backscatter) provides a centralized strategic point

Reverse DNS

```
zhanghan@Koffing:~/Desktop
$ host tcprst.us
tcprst.us has address 52.54.234.153
tcprst.us mail is handled by 15 eforward4.registrar-servers.com.
tcprst.us mail is handled by 10 eforward3.registrar-servers.com.
tcprst.us mail is handled by 10 eforward2.registrar-servers.com.
tcprst.us mail is handled by 10 eforward1.registrar-servers.com.
tcprst.us mail is handled by 20 eforward5.registrar-servers.com.
zhanghan@Koffing:~/Desktop
$ host 52.54.234.153
153.234.54.52.in-addr.arpa domain name pointer tcprst.us.
```

DNS Backscatter Sensor



DNS Backscatter

- DNS backscatter is the set of reverse DNS queries observed by a DNS authority
- Cache happens at all layers
- Final authority vs. root authority
  - Final authority sees all queries for a specific originator
  - Root authority should see all originators, if not cached

Privacy Concerns over DNS traffic

- Get approval from IRB (though sometimes an IRB review is not enough)
- Reasons to address privacy concerns in this case:
  - Caching and shared cache mask individual traffic, focusing on prevalent network activity instead
  - Authorities have little interaction with targets due to recursive resolvers
  - Mostly automated traffic, not human traffic, in reverse DNS

Methodology - Datasets

- Collected at authorities
  - One national authority managing .jp country TLD, two root servers (B, M) out of 13
  - And a final authority? (Not clear in the paper)
- Format: (originator, querier, authority) tuple

type	dataset	operator	start (UTC)	duration	sampling	queries ($\times 10^9$) (all)	queries ($\times 10^9$) (reverse)	qps ($\times 10^3$) (all)	qps ($\times 10^3$) (reverse)
ccTLD	JP-ditl	JP-DNS	2014-04-15 11:00	50 hours	no	4.0	0.3	22	1.8
root	B-post-ditl	B-Root	2014-04-28 19:56	36 hours	no	2.9	0.04	22	0.2
root	B-long	B-Root	2015-01-01	5 months	no	290*	5.14	22*	0.39
root	M-ditl	M-Root	2014-04-15 11:00	50 hours	no	8.3	0.06	46	0.3
root	M-ditl-2015	M-Root	2015-04-13 11:00	50 hours	no	9.9	0.07	55	0.4
root	M-sampled	M-Root	2014-02-16	9 months	1:10	36.2	1.5	1.6	0.07

Table 1: DNS datasets used in this paper.

Methodology - Features

- Derive static features from querier's domain name (mail.google.com)
  - mail, ns, firewall, cdn, nxdomain, etc
- Dynamic features from query patterns
  - queries per querier, unique ASes, unique countries, etc
- Classes for originator
  - ad-tracker, cdn, cloud, mail, spam, etc
- Manually label originators for training

Constraints in Backscatter

- Limited information about targets
  - Based only on querier domain name
- Backscatter is spread over multiple authorities due to anycast
- Could be tricked by careful spammer. Only increase the cost at certain degree

```
$ host google.com
google.com has address 172.217.4.238
$ host 172.217.4.238
238.4.217.172.in-addr.arpa domain name pointer ord30s31-in-f238.1e100.net.
```

Outline

- DNS Backscatters
- Methodology
- Validation
- Evaluation

Validation

- Select appropriate features
- Label ground truth
- Choose learning algorithm
- Validate through cross-validation

Select Appropriate Features

- Static features to distinguish different classes of originators

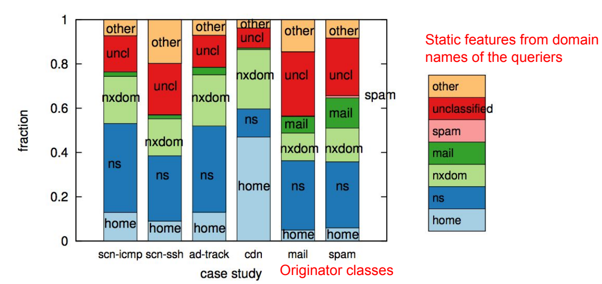


Figure 2: Static features for case studies, derived from querier domain names. (Dataset: JP-ditl.)

Select Appropriate Features

- Dynamic features to distinguish different classes of originators

case	queries/ querier	global entropy	local entropy	queriers/ country
scan-icmp	3.3	0.83	0.92	0.006
scan-ssh	4.7	0.84	0.96	0.006
ad-track	2.3	0.85	0.94	$\boxed{0.017}$
cdn	4.4	$\boxed{0.48}$	0.97	$\boxed{0.018}$
mail	$\boxed{1.7}$	0.71	0.94	$\boxed{0.009}$
spam	$\boxed{3.4}$	0.85	0.95	0.005

Table 2: Dynamic features for case studies.

Label Ground Truth

- Generate moderate to large lists of potential IP addresses in each application class from external sources;
- Intersect with the top-10000 originators in dataset by the number of queries;
- Manually verify intersection

dataset	ad-track	cdn	cloud	crawler	dns	mail	ntp	p2p	push	scan	spam	update	total
JP-ditl	15	8	-	-	26	44	10	37	-	25	64	6	**235**
B-post-ditl	13	29	16	17	16	46	5	-	12	29	35	-	**214**
M-ditl	13	36	16	16	17	50	8	-	12	33	43	-	**240**
M-sampled	54	81	82	35	52	111	-	-	73	124	136	-	**746**

Table 3: Number of examples of each application class in labeled ground-truth, per dataset.

Choose Learning Algorithm

Learning algorithms:

- Classification And Regression Tree (CART)
- Random Forest (RF)
- Kernel Support-Vector Machines (SVM)

Metrics:

- Accuracy: (tp + tn) / all
- Precision: tp / (tp + fp)
- Recall: tp / (tp + fn)
- F1-score: 2tp / (2tp + fp + fn)

Classification Accuracy

dataset	algorithm	accuracy	precision	recall	F1-score
JP ditl	CART	0.66 (0.05)	0.63 (0.08)	0.60 (0.06)	0.61 (0.06)
	RF	**0.78** (0.03)	**0.82** (0.05)	**0.76** (0.06)	**0.79** (0.05)
	SVM	0.73 (0.04)	0.74 (0.05)	0.71 (0.06)	0.73 (0.05)
B post-ditl	CART	0.48 (0.05)	0.48 (0.07)	0.45 (0.05)	0.46 (0.05)
	RF	**0.62** (0.05)	**0.66** (0.07)	**0.60** (0.07)	**0.63** (0.07)
	SVM	0.38 (0.11)	0.50 (0.14)	0.32 (0.13)	0.39 (0.13)
M ditl	CART	0.53 (0.06)	0.52 (0.07)	0.49 (0.06)	0.51 (0.06)
	RF	**0.68** (0.04)	**0.74** (0.06)	**0.63** (0.05)	**0.68** (0.05)
	SVM	0.60 (0.08)	0.68 (0.10)	0.52 (0.08)	0.59 (0.09)
M sampled	CART	0.61 (0.03)	0.65 (0.04)	0.58 (0.04)	0.61 (0.04)
	RF	**0.79** (0.02)	**0.82** (0.02)	**0.77** (0.03)	**0.79** (0.02)
	SVM	0.72 (0.02)	0.76 (0.03)	0.70 (0.03)	0.73 (0.02)

- Benchmark: 0.08 accuracy for randomly guessing
- Roots are attenuated (B post-ditl & M ditl)

Discriminative Features

- Gini Impurity $I_G(f)=\sum_{i=1}^{J} f_i(1-f_i)=\sum_{i=1}^{J}(f_i-{f_i}^2)=\sum_{i=1}^{J} f_i-\sum_{i=1}^{J} {f_i}^2=1-\sum_{i=1}^{J} {f_i}^2=\sum_{i\neq k} f_i f_k$

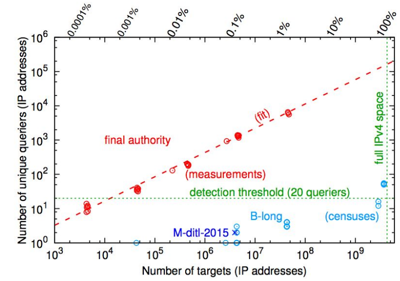
  Larger Gini values indicate features with greater discriminative power.

	JP-ditl		M-ditl	
rank	**feature**	**Gini**	**feature**	**Gini**
1	mail(S)	8.4	mail(S)	12.5
2	home(S)	7.9	ns(S)	8.3
3	spam(S)	6.3	unreach(S)	7.0
4	nxdomain(S)	6.2	query rate(D)	6.2
5	unreach(S)	5.2	home(S)	6.0
6	global entropy(D)	5.0	nxdomain(S)	5.8

Table 5: Top discriminative features. Classifier: RF.

Evaluate DNS Caching

- Backscatter is highly attenuated due to disinterested targets and DNS caching.

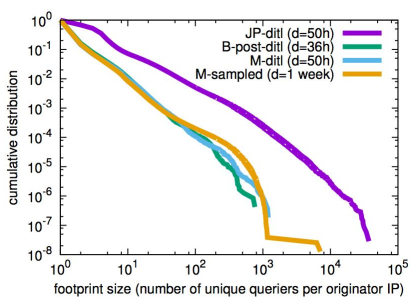


Roughly 1 querier per 1000 targets

Figure 3: Size of footprint of random network scans at the final authority. (Datasets: B-long and M-ditl.)

Results - Size of Originator Footprints

- There are hundreds of originators that touch large parts of the Internet



(Log-log scale)

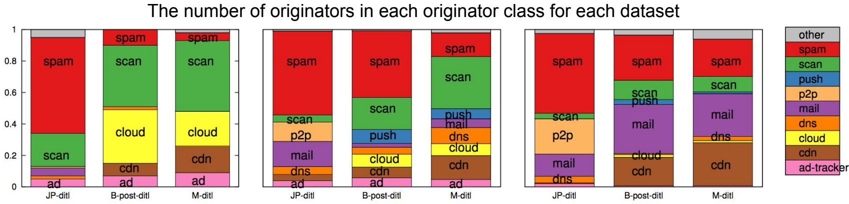
Classification of Top Originators

- Focus on the originators with the largest footprints;
- Understand the type of activity and the aggressiveness of activity.

Results - Trends of Network-wide Activities

data	ad-track	cdn	cloud	crawl	dns	mail	ntp	p2p	push	scan	spam	update
JP-ditl	210	49	-	-	414	1412	237	2235	-	355	**5083**	6
B-post-ditl	72	1782	168	361	76	**3137**	8	-	318	1228	2849	-
M-ditl	76	**2692**	135	557	258	**2750**	67	-	119	983	2353	-
M-sampled	1329	17,708	2035	885	1202	14,752	-	-	3652	**47,201**	34,110	-

The number of originators in each originator class for each dataset



(a) Top 100 (b) Top 1000 (c) Top 10,000

Big footprints are often unsavory activities!

Figure 6: Fraction of originator classes of top-N originators. (Dataset: JP-ditl, B-post-ditl, M-ditl; classifier: RF.)

Results - Trends of Network-wide Activities

- Fluctuations of originators may be explained by reactions to network security events.

Public announcement of the Heartbleed vulnerability
(a security bug in the OpenSSL cryptography library)

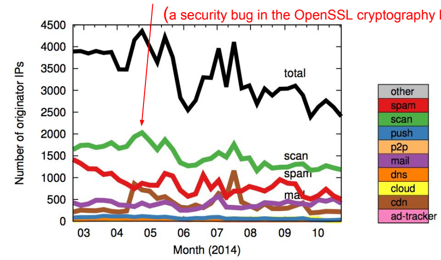


Figure 7: Number of originators over time. (Dataset: M-sampled; classifier: RF.)

Results - Trends of Network-wide Activities

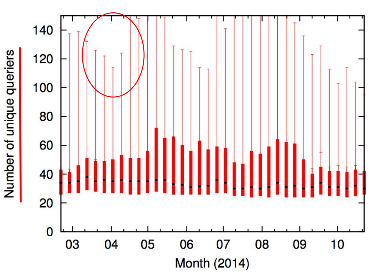


Figure 8: Box plot of originator footprint (queriers per scanner) over time; whiskers: 10%ile/90%ile. (Dataset: M-sampled.)

Very large scanners come and go.

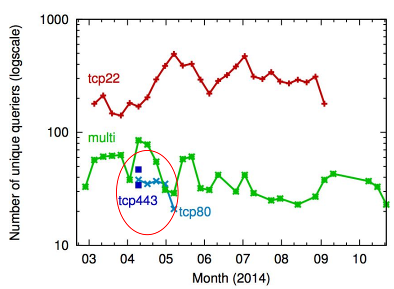


Figure 9: Three example originators with application class *scan*. (Dataset: M-sampled with darknet.)

Results - Trends of Network-wide Activities

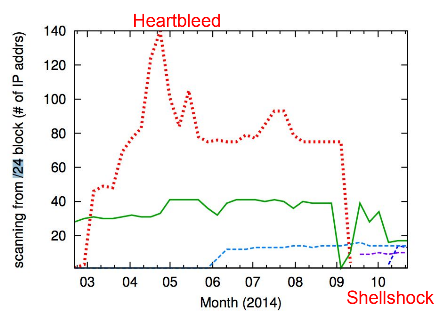


Figure 10: Five example blocks originating scanning activity. (Dataset: M-sampled.)

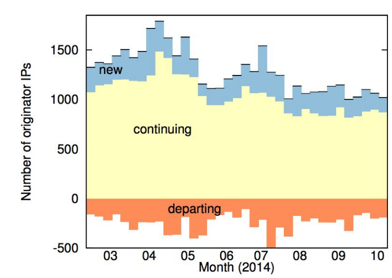


Figure 11: Week-by-week churn for originators of class *scan*. (Dataset: M-sampled.)

Contributions

- Identify DNS backscatter as a new source of information about benign and malicious network-wide activity;
- Keep in mind of privacy and address any potential related issues in paper;
- Collect trainable dataset with ground truth label;
- Understand the type and trend of network-wide activity based on classifications;

Discussions

- Adoption of botnets to circumvent the system
  - Intentionally camouflage network traffic at each originator
- The possibility of other prominent features
- The possibility of other classifiers
- Limited training data
  - The number of data points in some application classes is too small