Freie Universität Berlin

UNIVERSITY OF TWENTE.

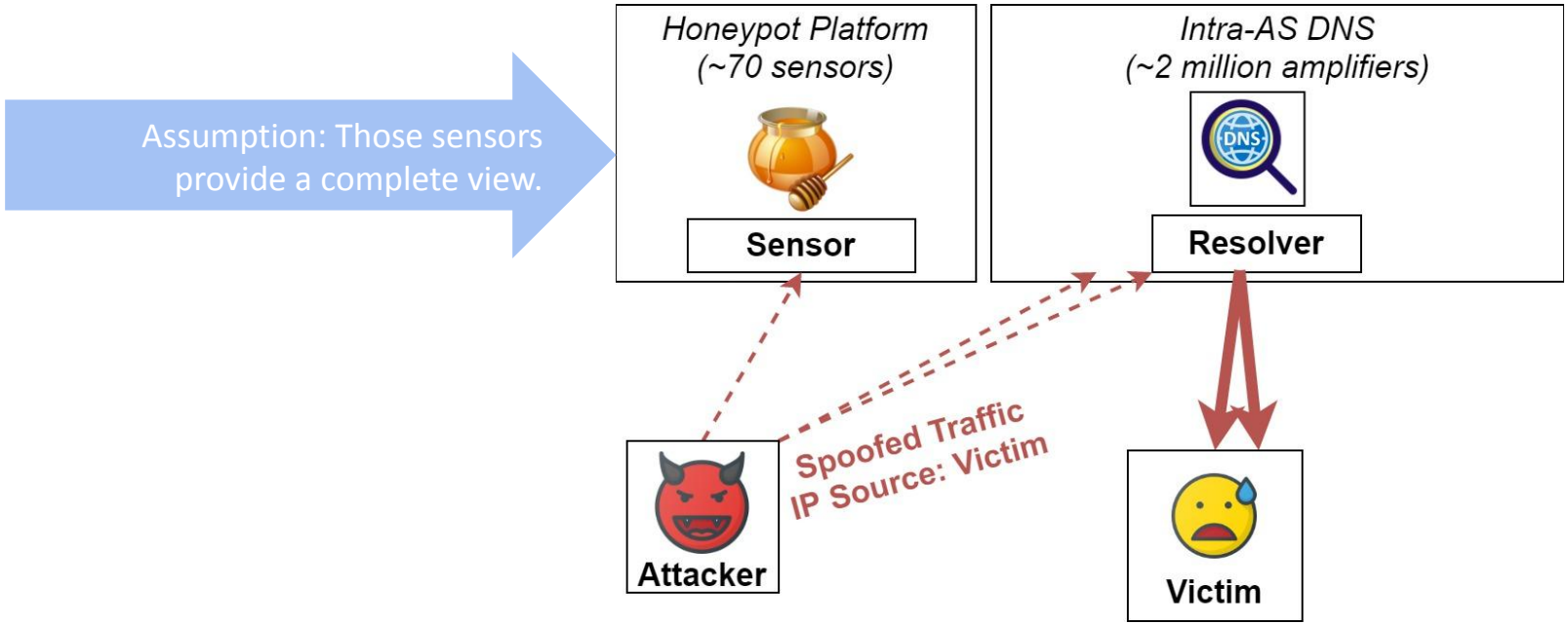
Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

The Far Side of DNS Amplification: Tracing the DDoS Attack Ecosystem from the Internet Core

Marcin Nawrocki, Mattijs Jonker, Thomas C. Schmidt, Matthias Wählisch

{marcin.nawrocki, m.waehlisch}@fu-berlin.de
m.jonker@utwente.nl, t.schmidt@haw-hamburg.de

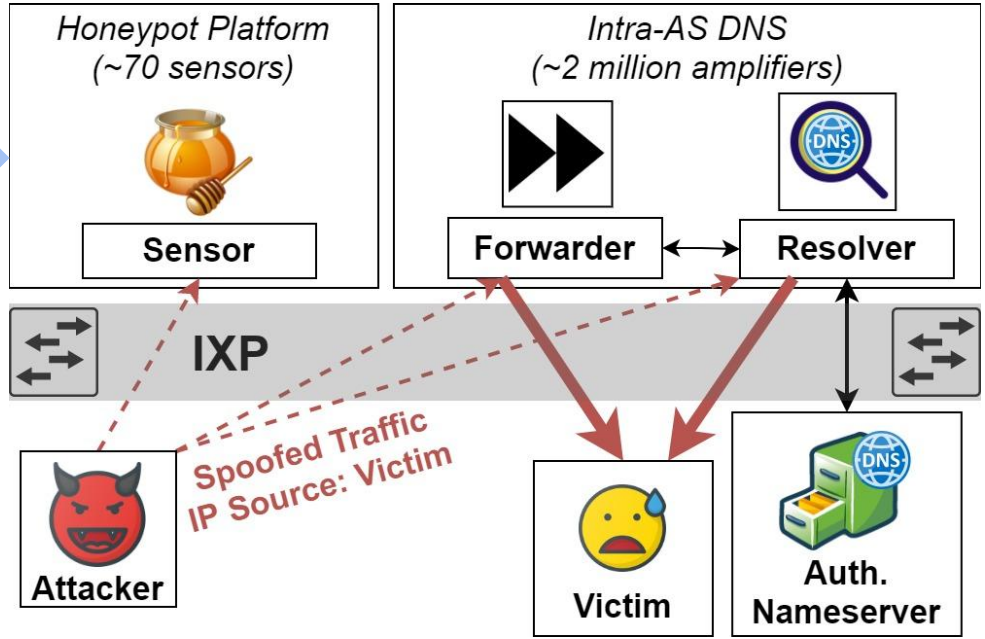
DNS amplification attacks and a common assumption



DNS amplification attacks and a common assumption

Assumption: Those sensors provide a complete view.

?



What is this talk about?

Does an IXP observe additional DNS amplification attacks?

Does an IXP contribute new insights into the efficiency of attacks?

Is DNSSEC fully exploited by an attacker?

Our contributions

A **method** to infer DNS amplification attacks in sampled IXP flow data.

Comparative measurement study using complementary data from Internet core and edge.

Unveiling of **new DNS attack practises**.

What is this talk about?

Does an IXP observe additional DNS amplification attacks?

Does an IXP contribute new insights into the efficiency of attacks?

Is DNSSEC fully exploited by an attacker?

Our vantage points for comprehensive observations

IXP

[sampled flow data]

[captured traffic]

Open INTEL

[DNS records]

[IPv4 scans]

How to detect attacks at an IXP? Identify misused names.

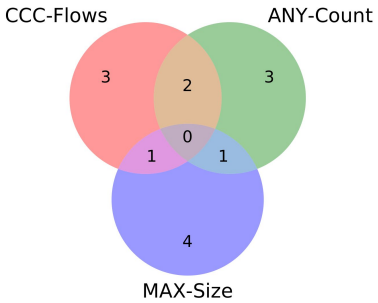
Key assumption:

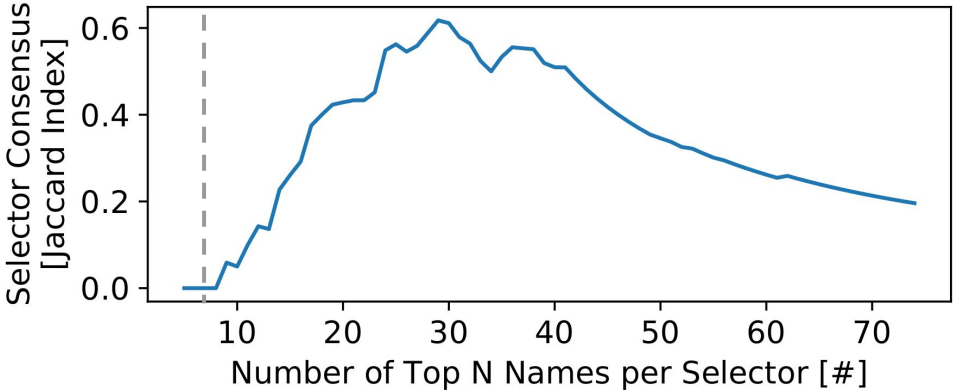
Attackers are likely to reuse names that lead to large responses.

Selectors:

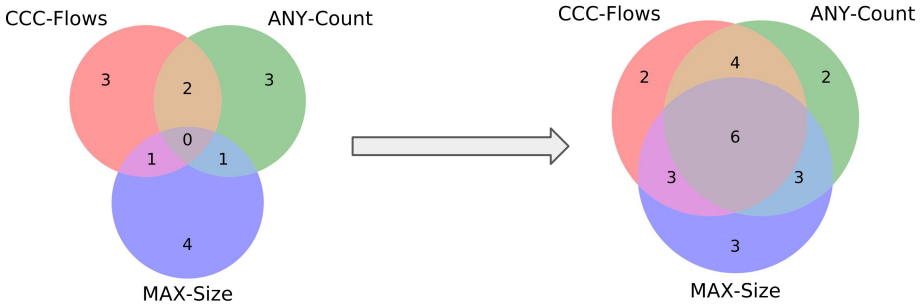
Maximum response size	# `ANY` packets	CCC ground truth traffic

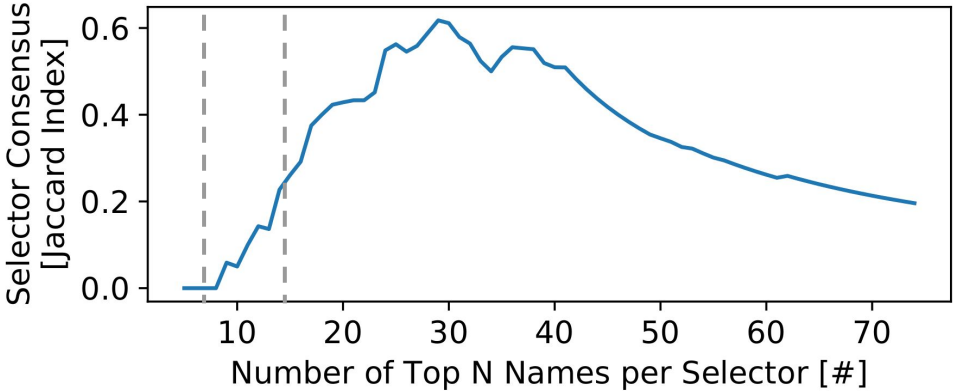
How many names per selector? Selector consensus.



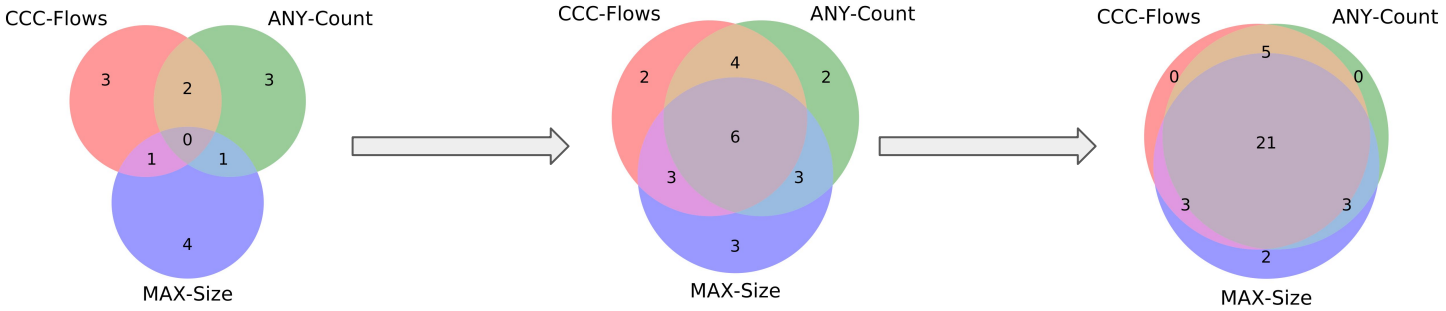


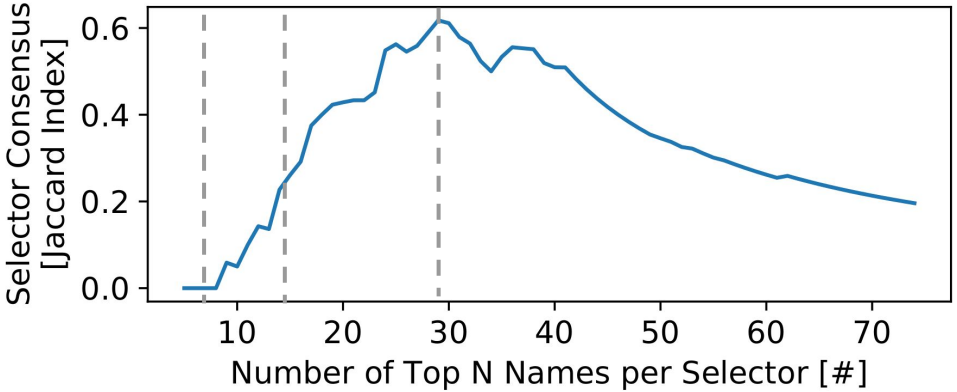
How many names per selector? Selector consensus.





How many names per selector? Selector consensus.





How to detect attacks at an IXP? Identify irregular DNS behavior.

Assumption:

A host is under attack if it exchanges *many* DNS queries or responses with misused names.

Approach:

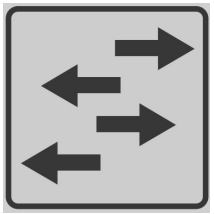
Apply thresholds to verify misused name candidates.

IXP detects attacks unseen by a large honeypot platform

Honeypots: 31k attacks

4%

of attacks observed by both vantage points



IXP: 26k attacks

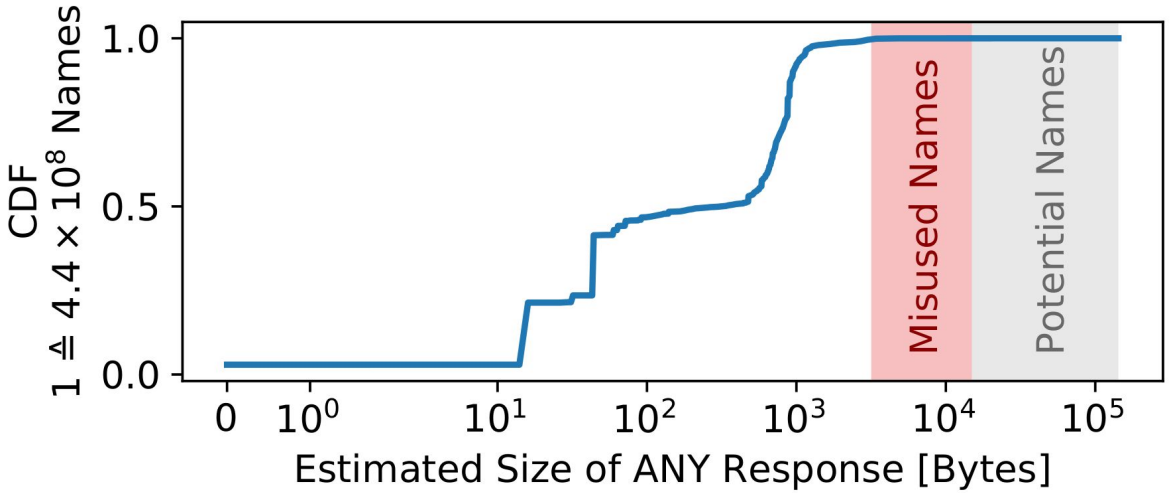
What is this talk about?

Does an IXP observe additional DNS amplification attacks?

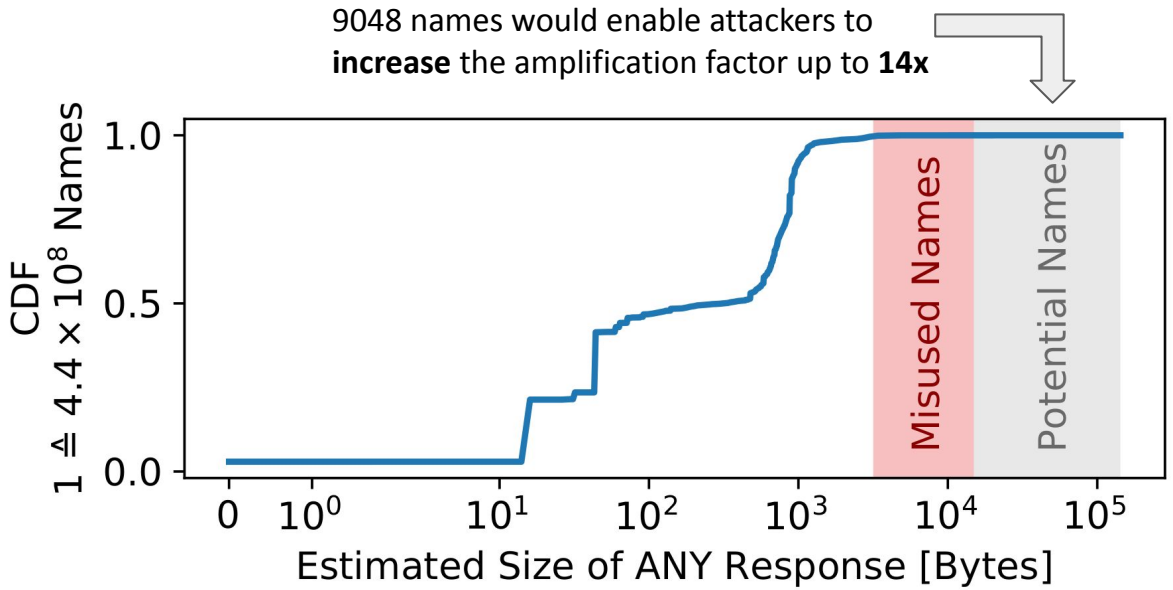
Does an IXP contribute new insights into the efficiency of attacks?

Is DNSSEC fully exploited by an attacker?

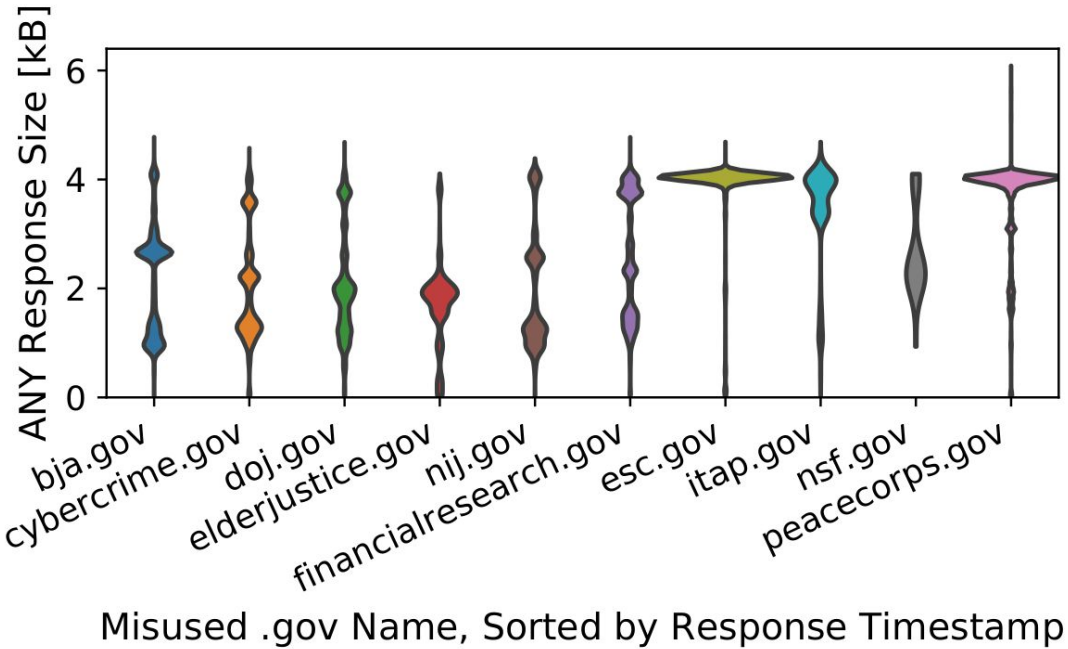
Estimating ANY response sizes based on OpenINTEL data



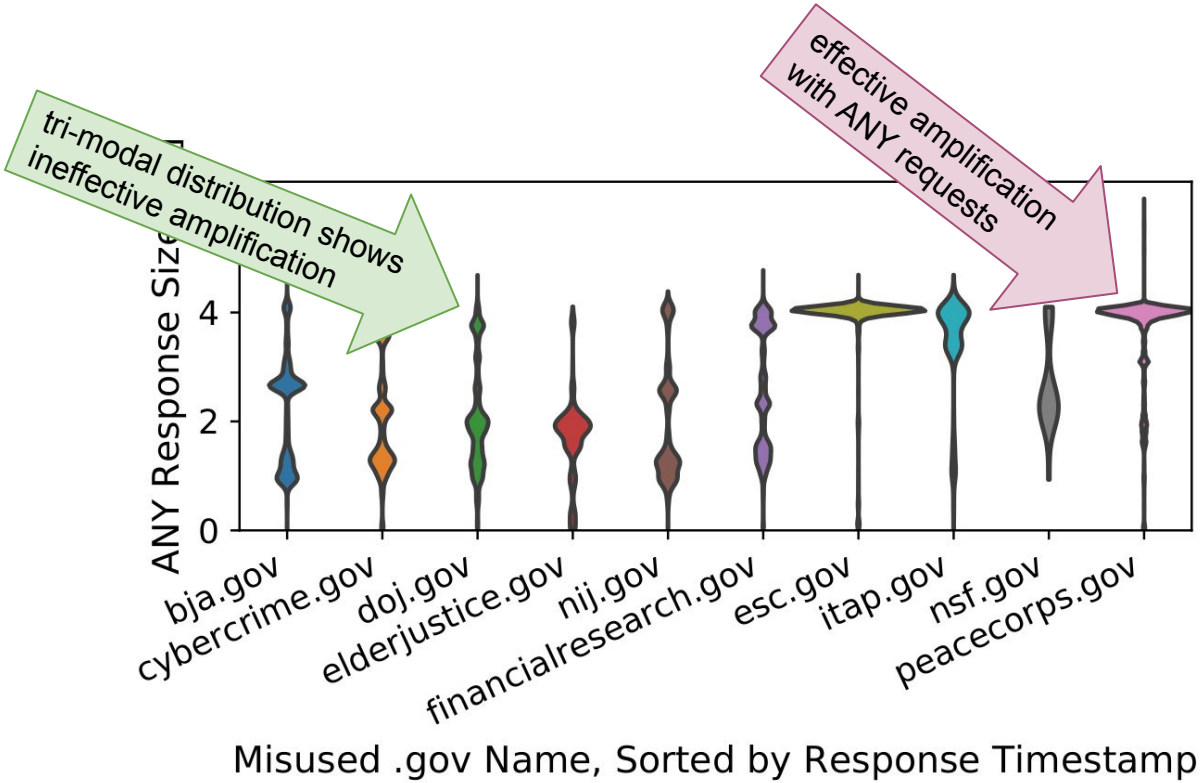
Estimating ANY response sizes based on OpenINTEL data



Amplifiers react to `ANY`, observed in real IXP traffic



Amplifiers react to ANY, observed in real IXP traffic



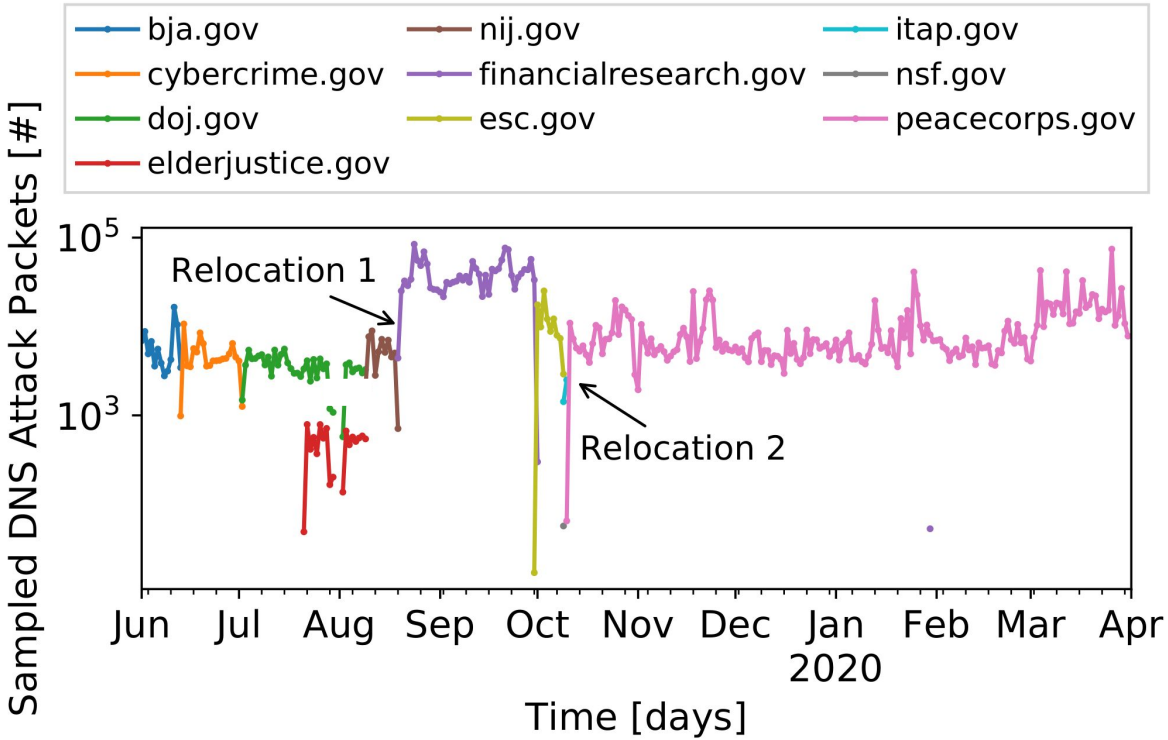
What is this talk about?

Does an IXP observe additional DNS amplification attacks?

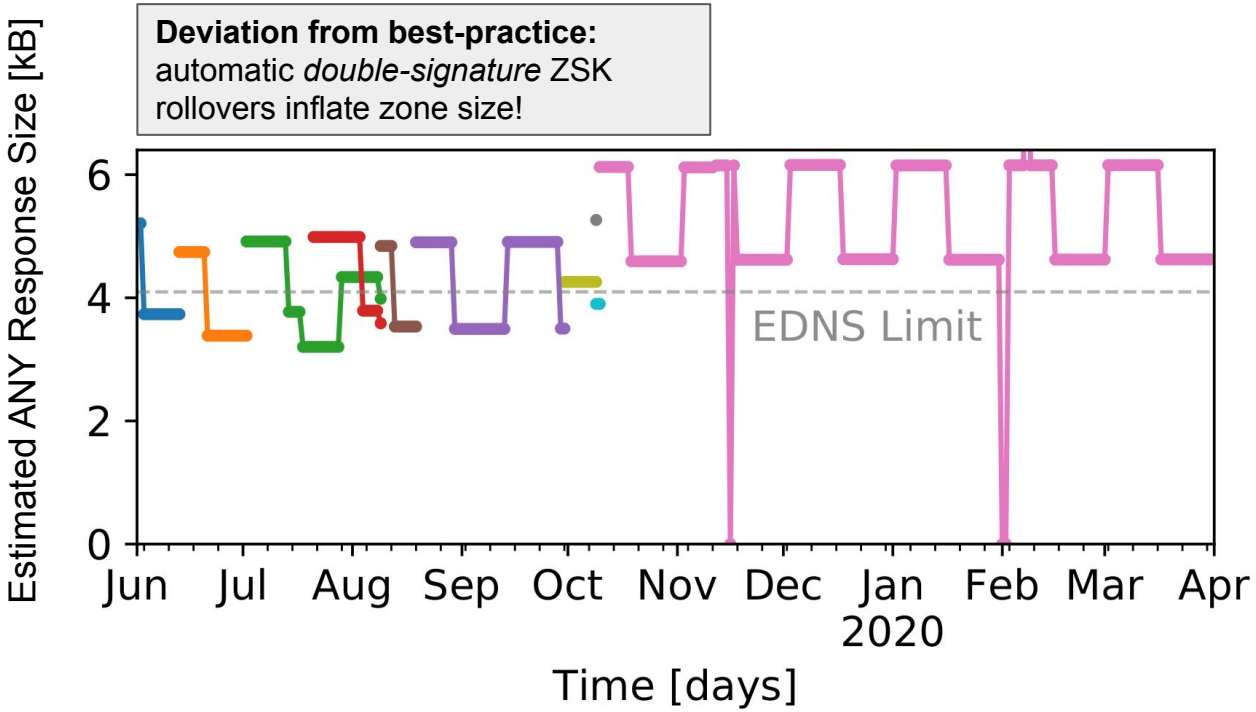
Does an IXP contribute new insights into the efficiency of attacks?

Is DNSSEC fully exploited by an attacker?

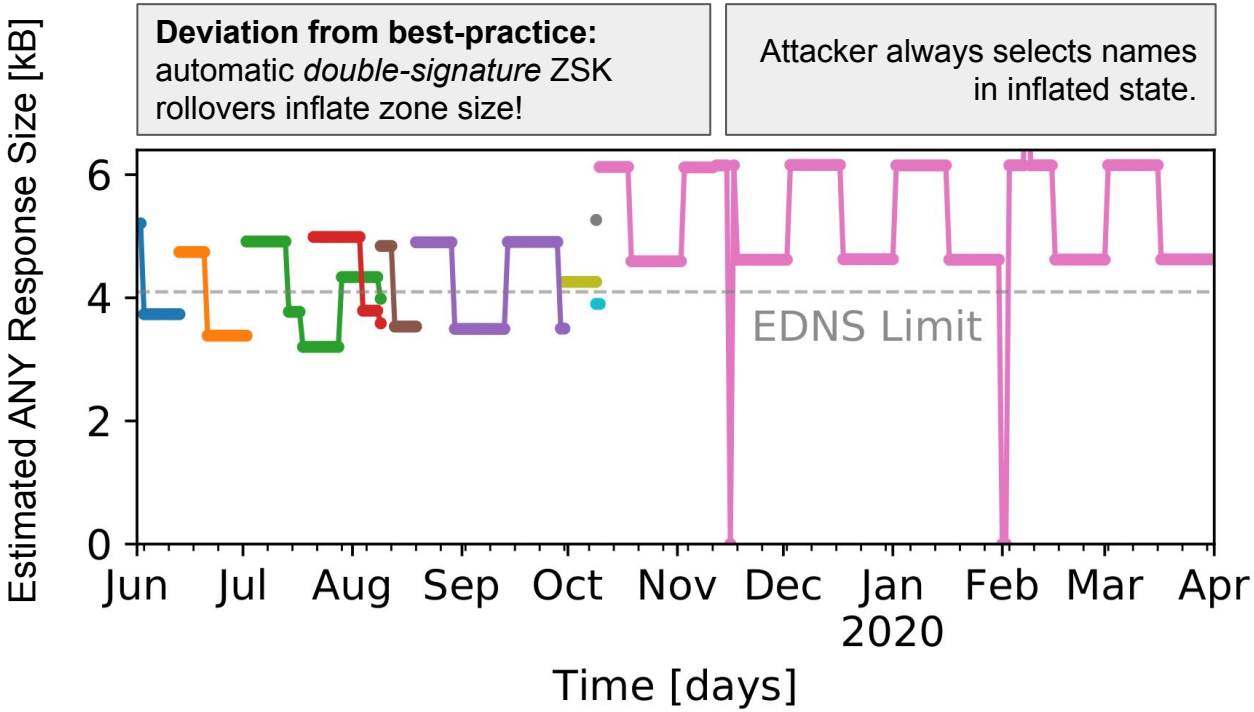
Clear transitions of misused names expose a new attacker



Attackers select inflated DNS zones

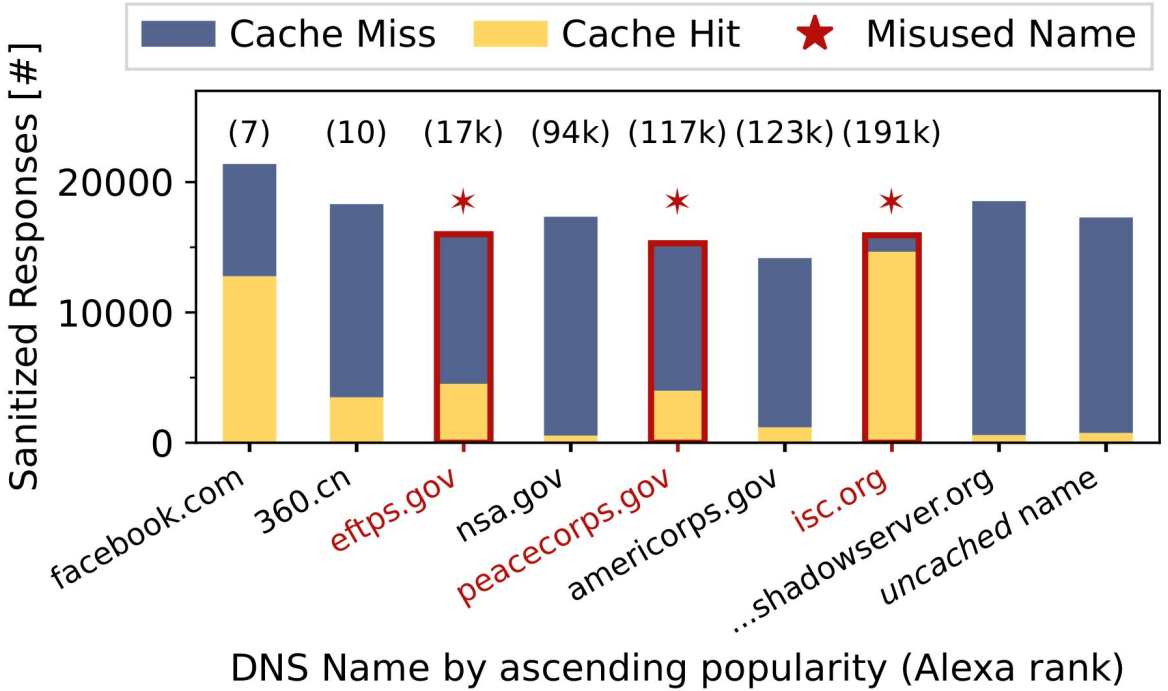


Attackers select inflated DNS zones



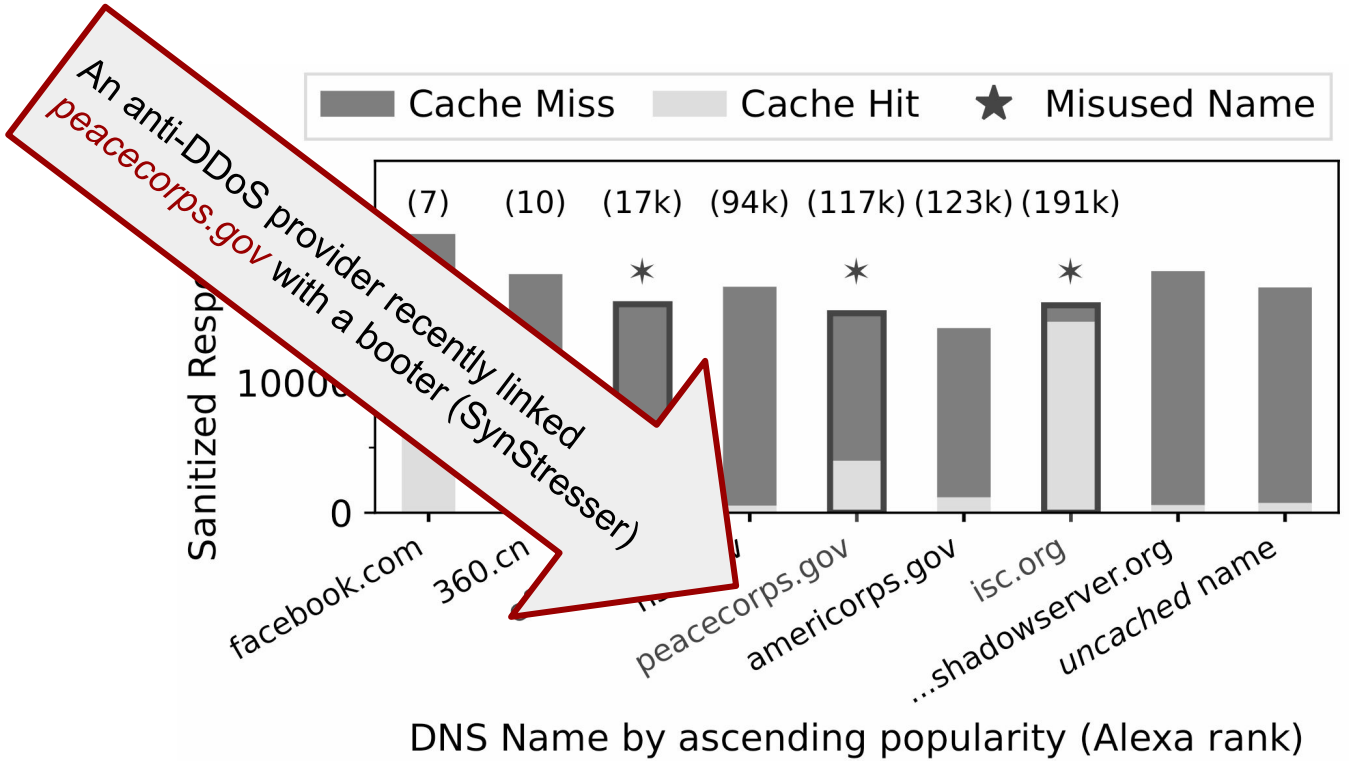
Brief Validation

Misused names have low web popularity but high cache hits, indicating frequent usage due to other reasons.



Misused names have low web popularity but high cache hits,

indicating frequent usage due to other reasons.



What this talk was about.

Does an IXP observe additional DNS amplification attacks? ✓

Only 4% overlap compared to honeypots.

Does an IXP contribute new insights into the efficiency of attacks? ✓

`ANY` queries still effective. Attackers could launch larger attacks.

Is DNSSEC fully exploited by an attacker? ✓

Bad DNSSEC key rollover practices misused.

++ Backup Slides ++

DNSSEC Key Rollovers

Pre-Publish

- introduces only the new key in stand-by mode, i.e., not yet used to sign RRsets, until everyone learns about it
- prone to race-conditions, still the recommended best practice

Double-Signature

- two active ZSKs and two (redundant) RRSIG records signatures, “old” key is then retired after a timeout
- valid rollover (RFC 6781), but doubles the number of signatures in a zone

Why don’t honeypots observe everything?

1. Honeypots integrate into an ecosystem of amplifiers. How many other amplifiers exist and how are they misused? This might be different for each protocol!

2. Honeypots can be identified as such because they apply rate limiting (ethics & liability) and often only emulate a vulnerable service, which leads to specific fingerprints.

3. Amplification honeypots have to distinguish between scans and attacks, which they do by thresholds. What are *good* thresholds?

4. Honeypots are deployed in very specific networks, usually cloud or universities. Bias?

Why only DNS?

Next to NTP, DNS is still one of the most-commonly misused protocols, *since years.*

Also, compared to other protocols, the DNS amplification ecosystem is very special:

- There is not a single attack query but many various ways to trigger attacks (--> memcache).
- Attacks tend to misuse legitimate names, this might have an adverse effect on third parties, which are actually not involved and completely unaware.
- We have a lot of additional, research-based data sources (e.g. OpenIntel) which really help to understand the more complex observations.
- (DNS amplifiers have the highest churn rates. We wanted to see whether attackers adapt.)

Why do we need to “select” names? Why 3 selectors?

Using the selectors allows us to focus the analysis on very suspicious parts of the DNS traffic. This leads to a huge performance boost. Again, it’s an IXP, so we have a lot of data.

This allows us a deployment at the IXP which detects attacks with a minimal delay.

The nice thing here is that the selectors are easily extendable, if you want you could throw in any new selector and see how it performs.

Do the auth. nameservers with larger zones allow `ANY`?

We did not specifically check this. However, attacks are also possible without `ANY`:

1. Sometimes individual resource records like the `RRSIG` or `TXT` are already large enough to be attractive for attackers.
2. It is more important whether the amplifiers allow `ANY`, and they do.
   95% of amplifiers are DNS forwarders, which forward queries to a resolver. For example, we observed a resolver responsible for 40k forwarders. This means that the attacker could fill up the cache of this specific resolver with a couple of requests for individual resource records and than use all these forwarders as amplifiers.
   This is possible because `ANY` is not `ALL`, so there will be answers with a large subset.

Are authoritative nameservers used for attacks?

1. Authoritative servers should not recursively resolve DNS queries, which makes them less attractive amplifiers. That’s why we only found a couple of name servers.

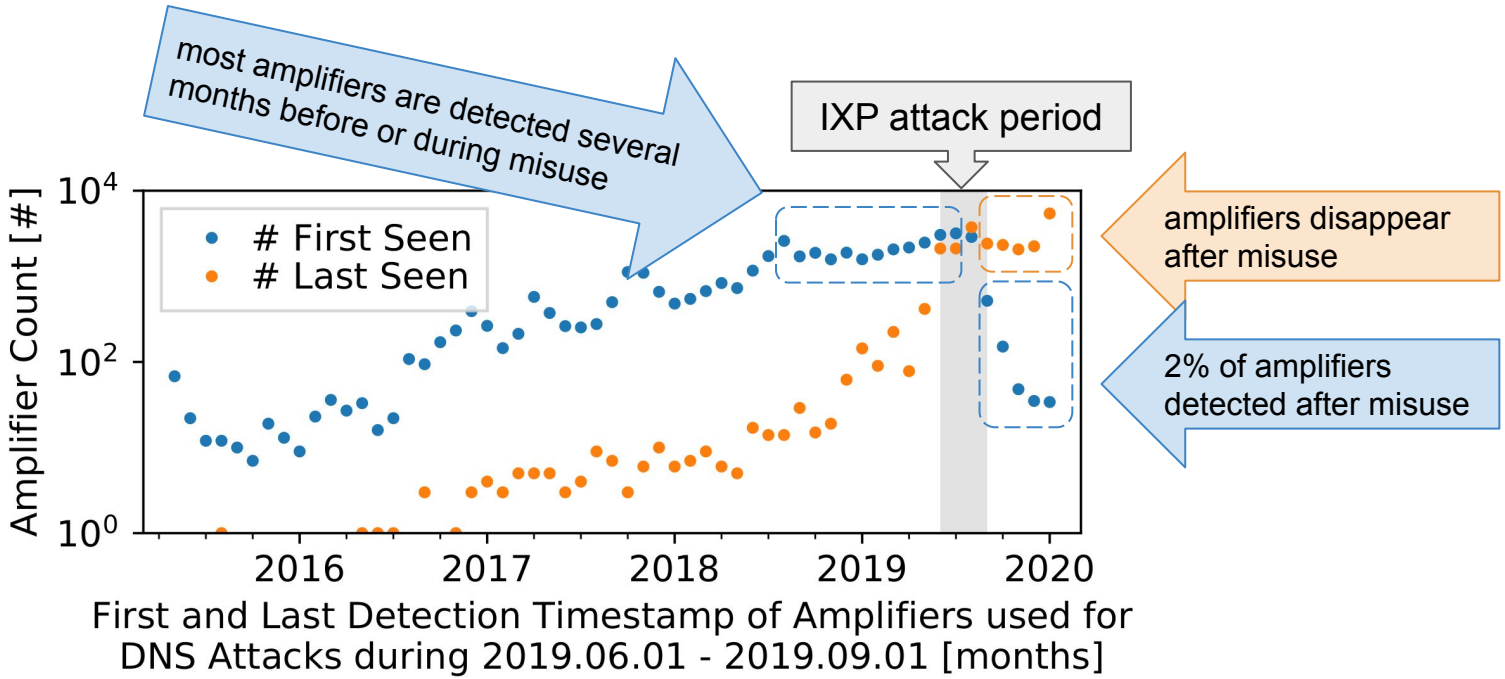
2. But there is one special case: Attacks which use the root-name for amplification are 4x more likely to use authoritative nameservers. This is because misconfigured nameservers actually answer with the root hint-files, which are quite large, if they receive queries for the root-name.

How do honeypots detect attacks?

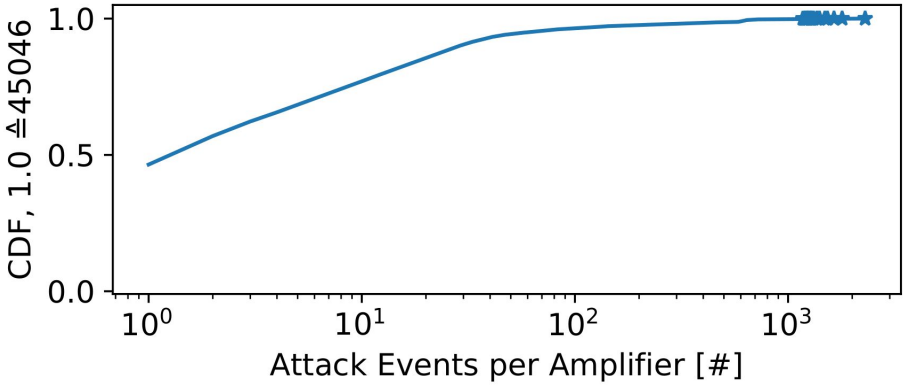
CCC applies a sensor-based attack threshold. Attack if:

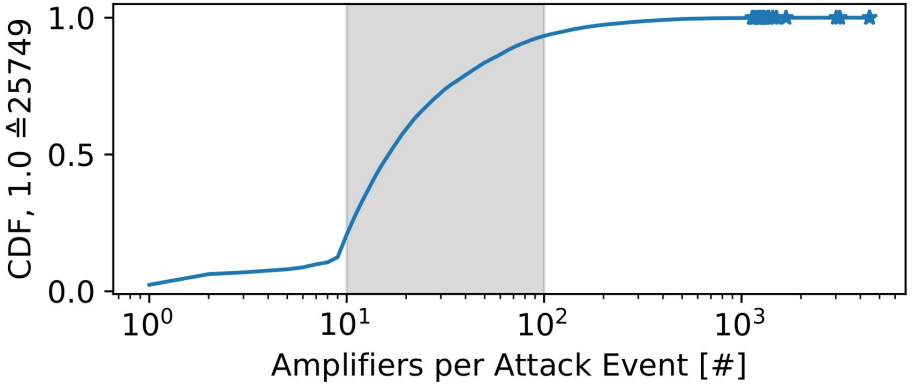
```
>=5 DNS requests per sensor, idle timeout of 900 seconds
```

Shodan: Unveiling the lifecycle of amplifiers



How are the amplifiers used?





* These values are not extrapolated by the sampling rate (`1:16k`).

IXP attack thresholds with misused names

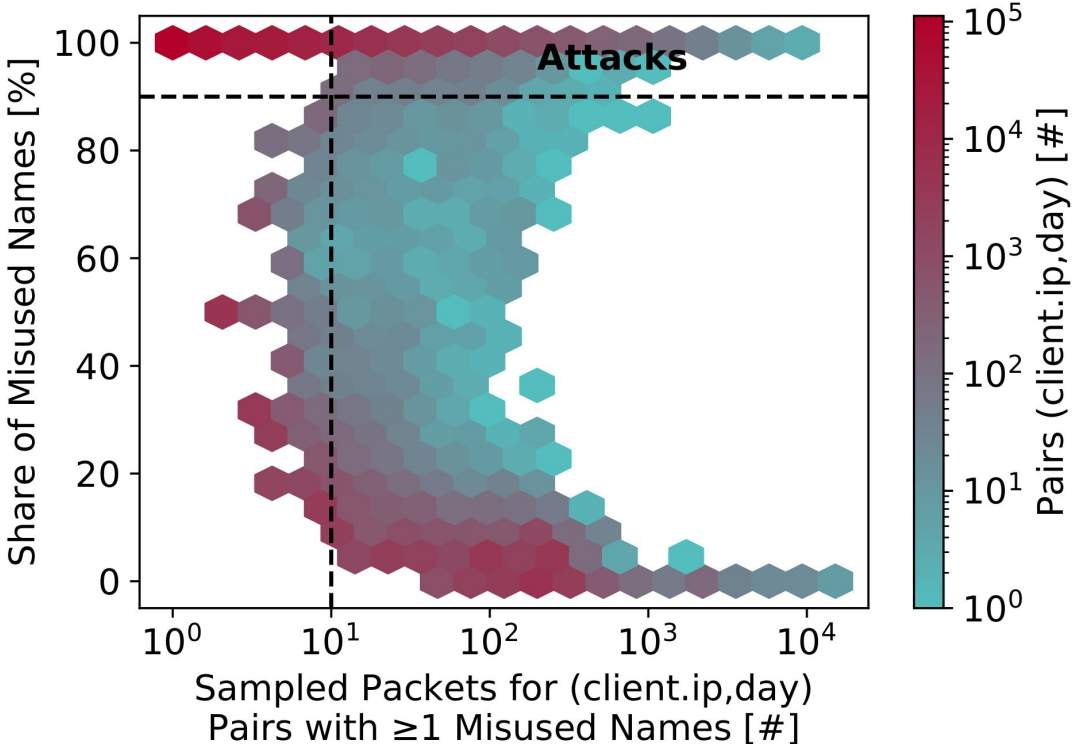
A DNS client is under attack if:

1. ≥10 sampled DNS queries or responses
2. share of 90% of misused names

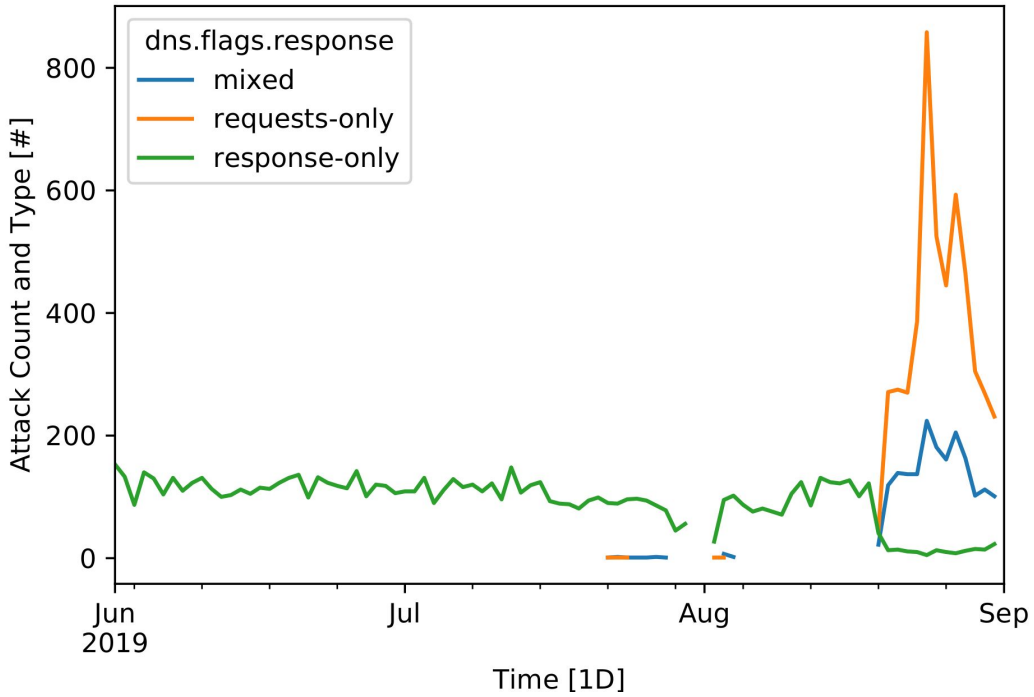
Extrapolating by sampling rate, this corresponds to ≥144k packets with misused names. No legitimate client needs so many requests, especially with caches.

In total, 34 candidate names. For 32 of these names (94%), we detect attacks.
Our candidates are clearly misused for attacks.

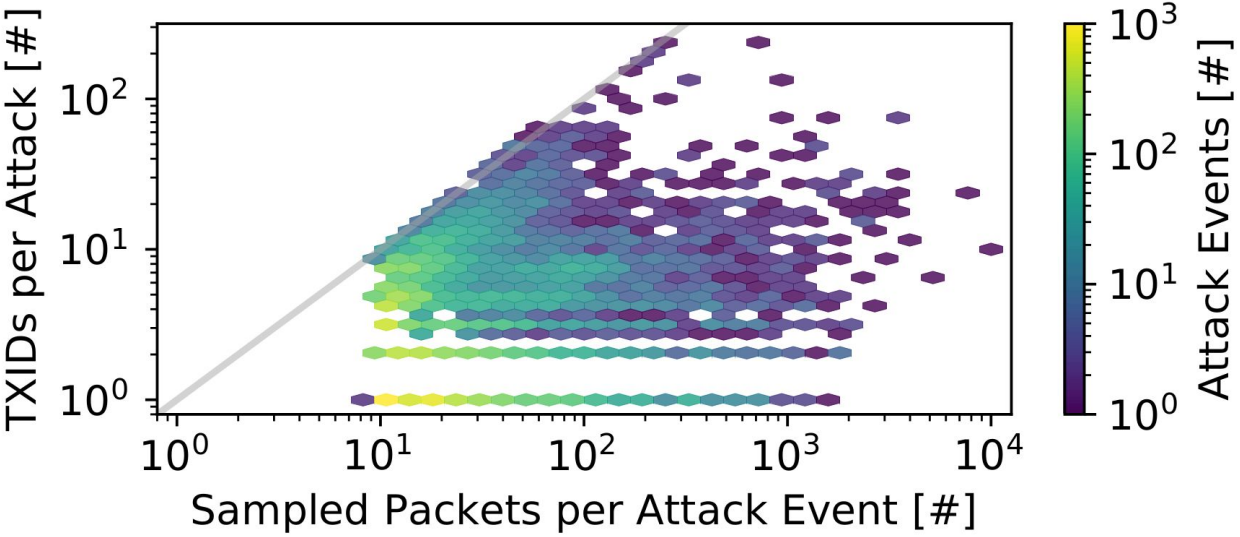
What is the influence of your thresholds?



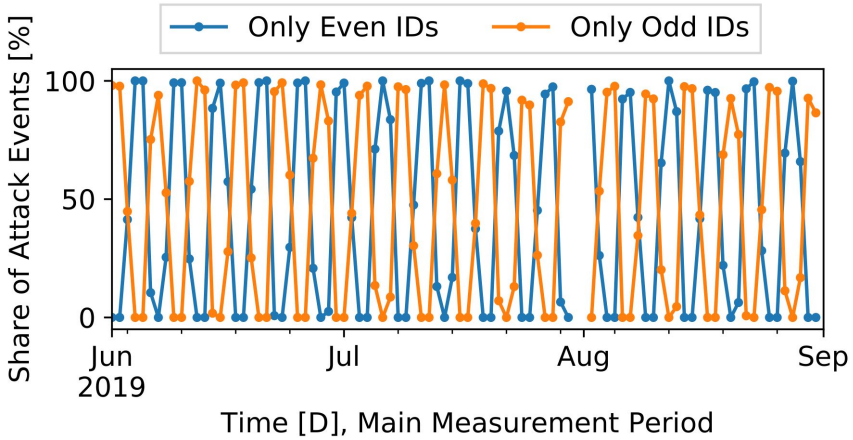
The attacker relocated into the customer cone of an IXP member, which increased attack visibility (queries now visible).

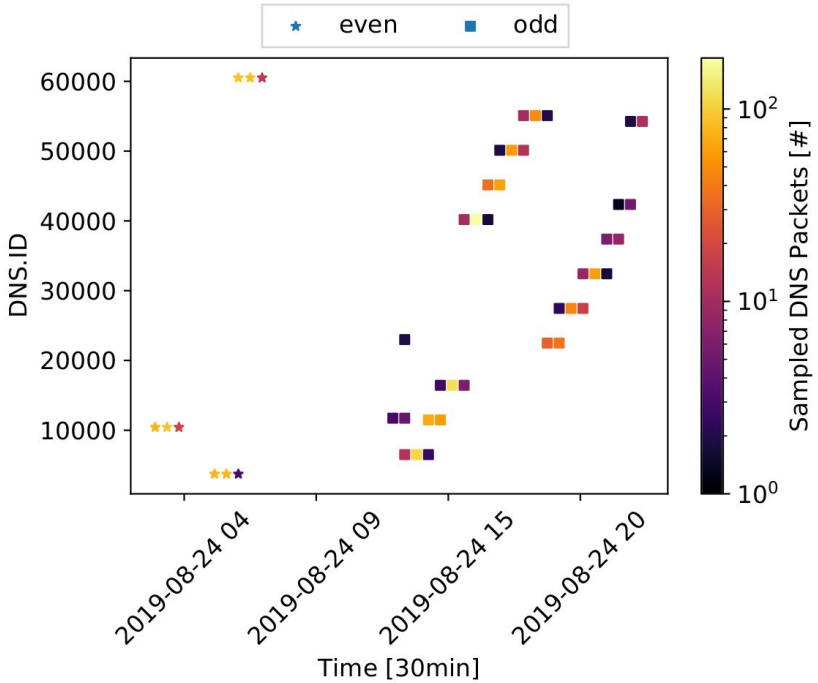


More fingerprints for major attack entity?
DNS TXIDs are not random!

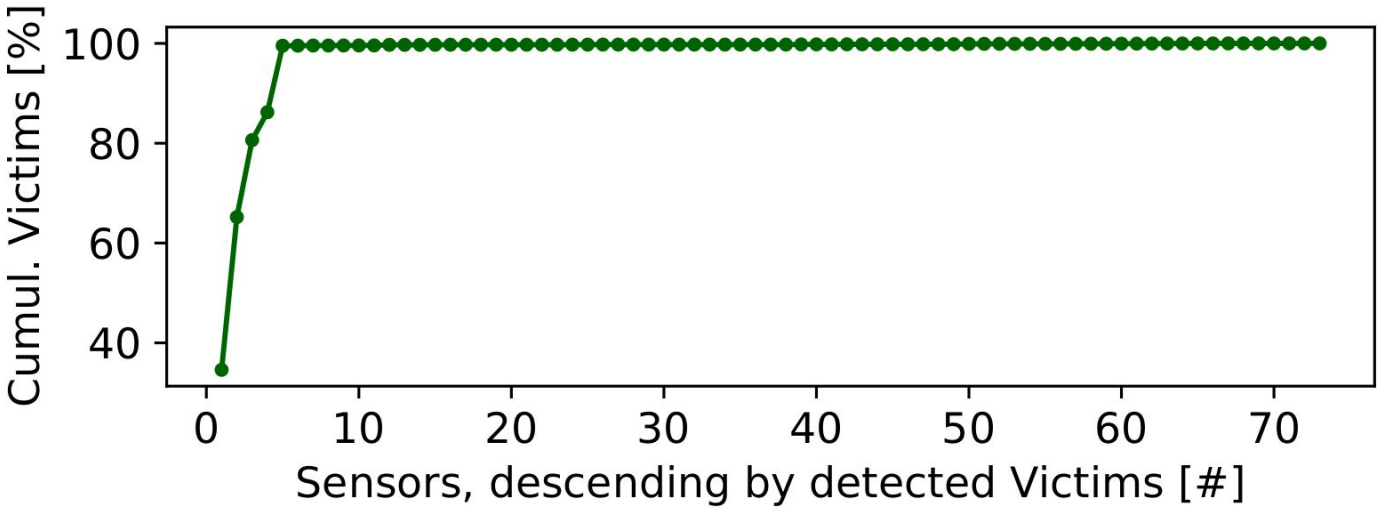


… but are alternating between odd and even.

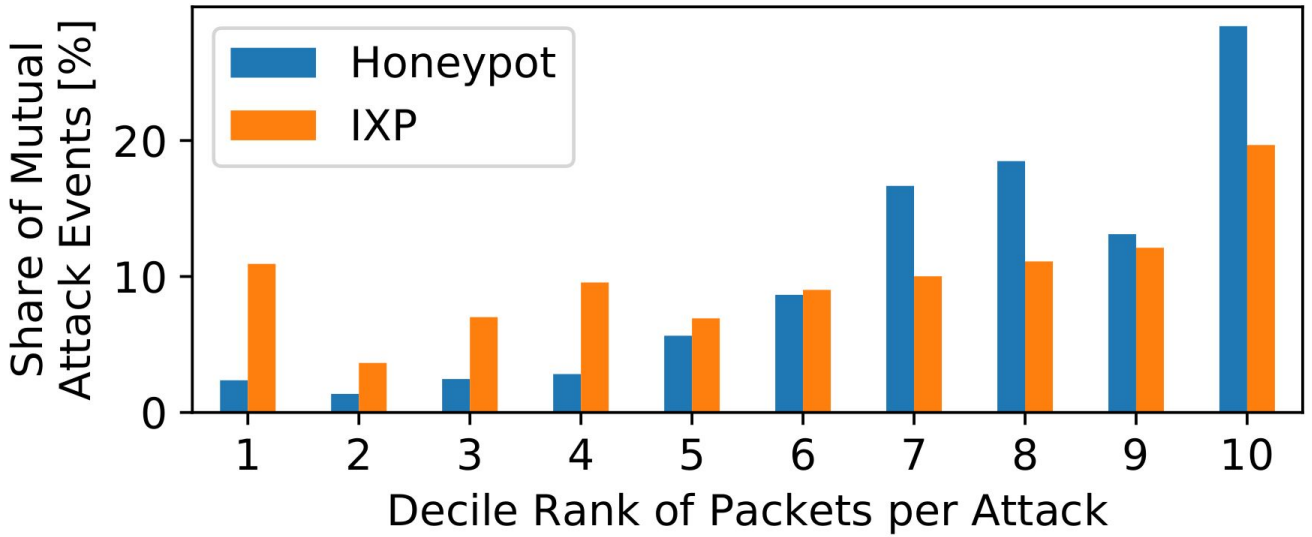




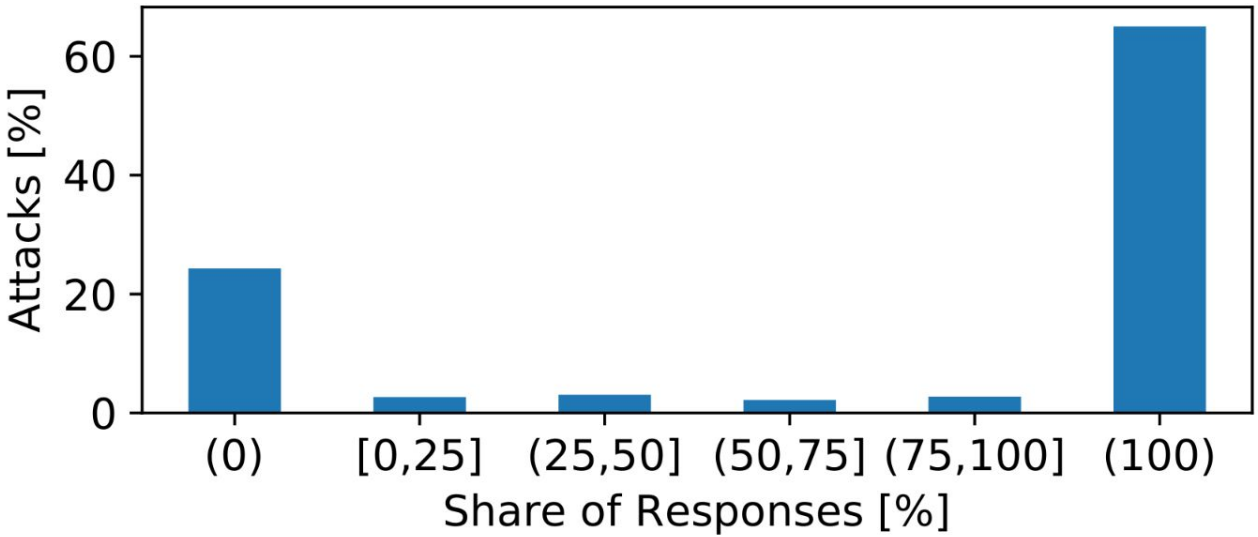
Honeypots convergence: The more the better is not true.



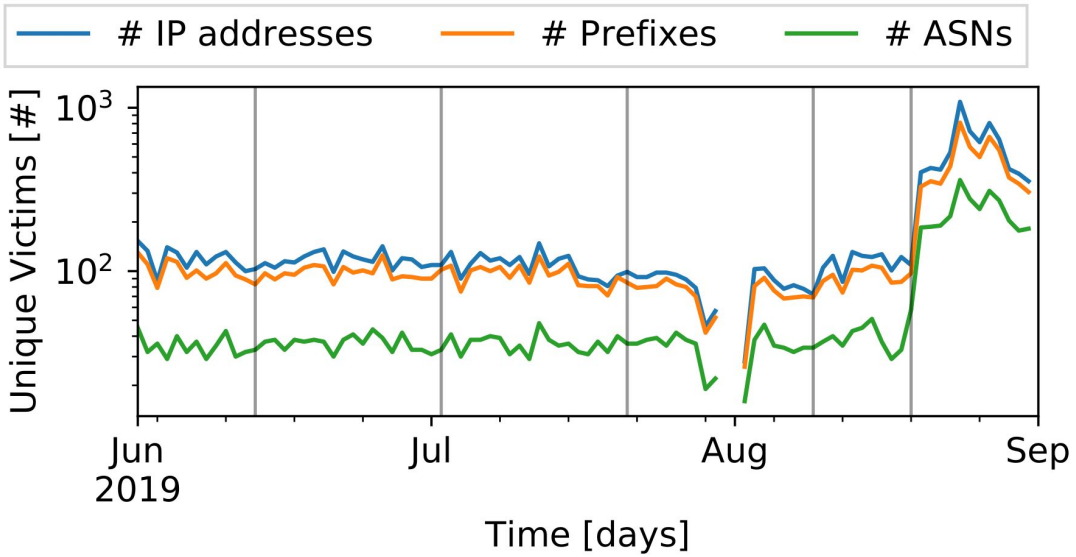
What about mutual attack events?



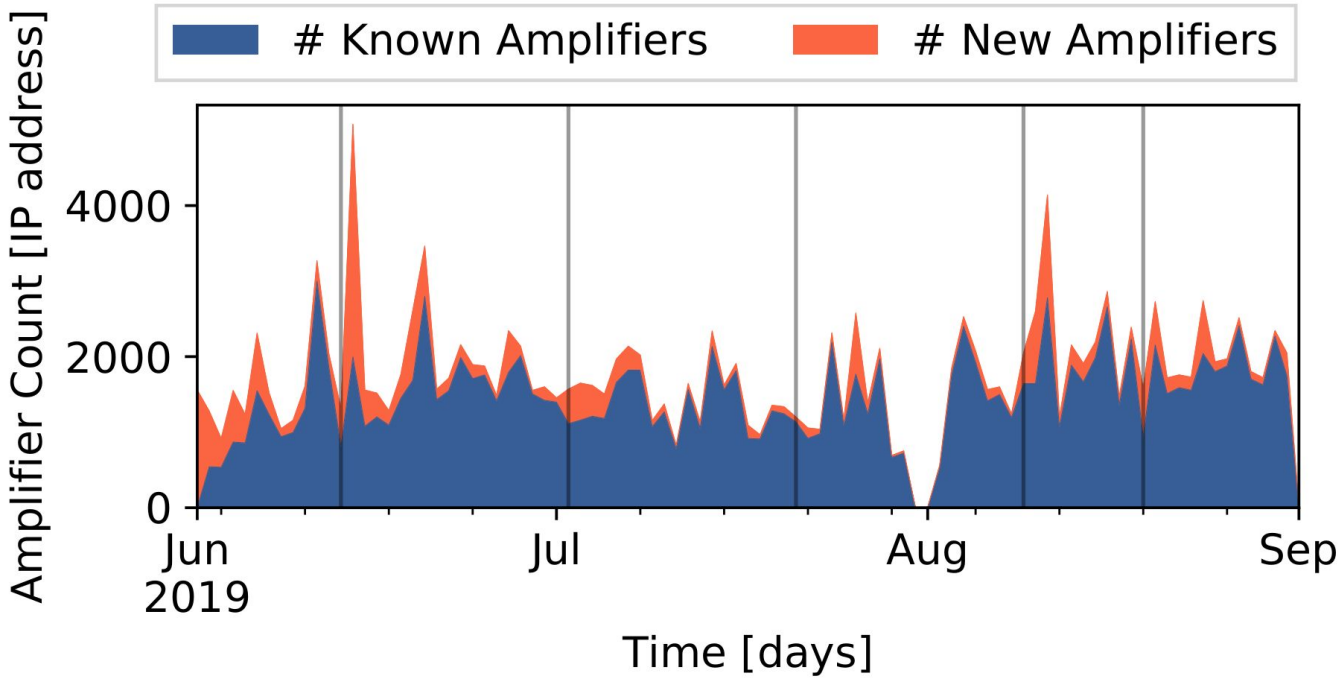
How asymmetric is the spoofed attack traffic?



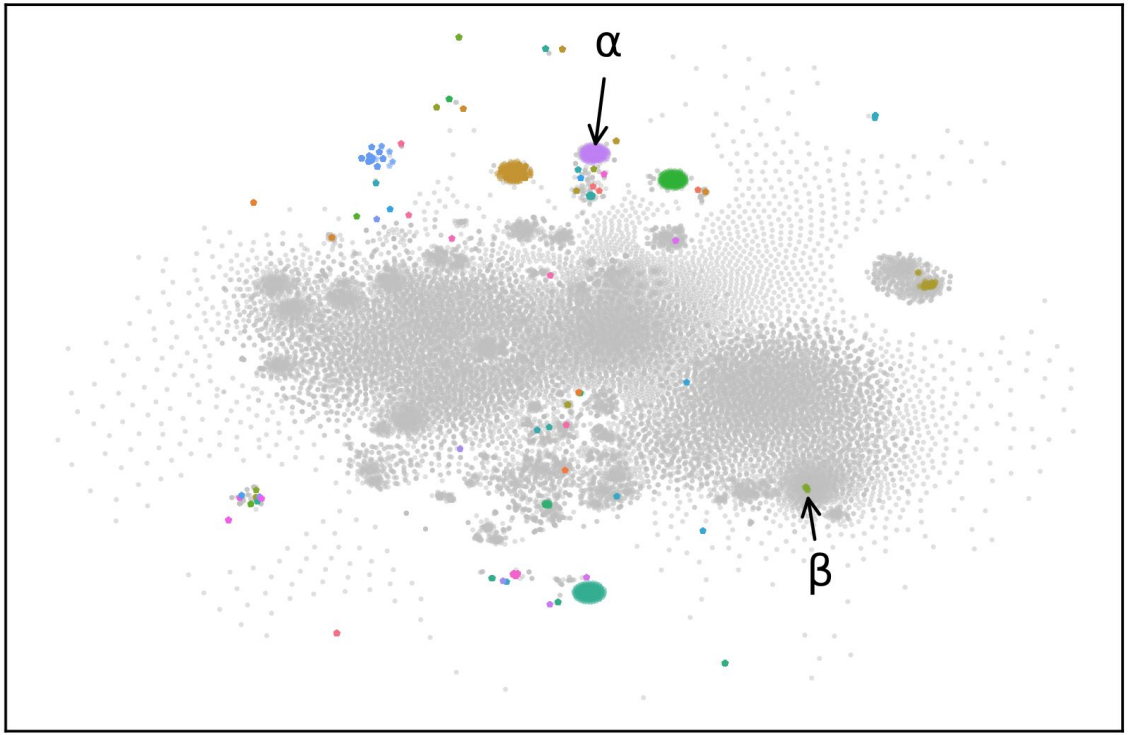
Hot distributed are the attacks by the attack entity?



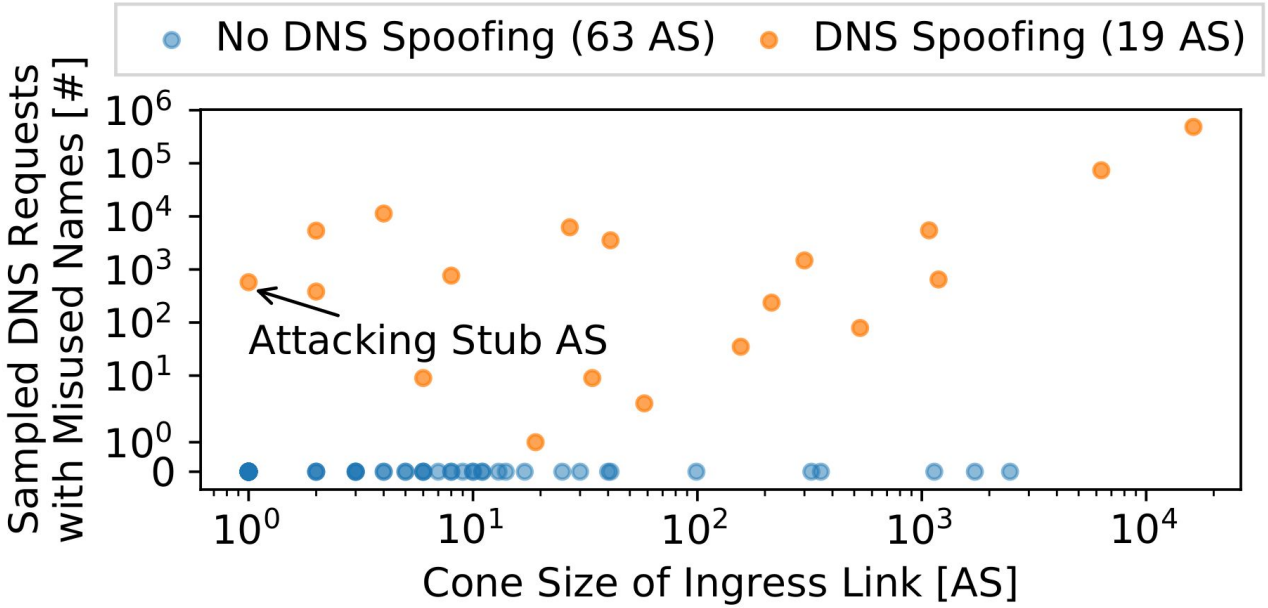
New amplifier lists by major attack entity?



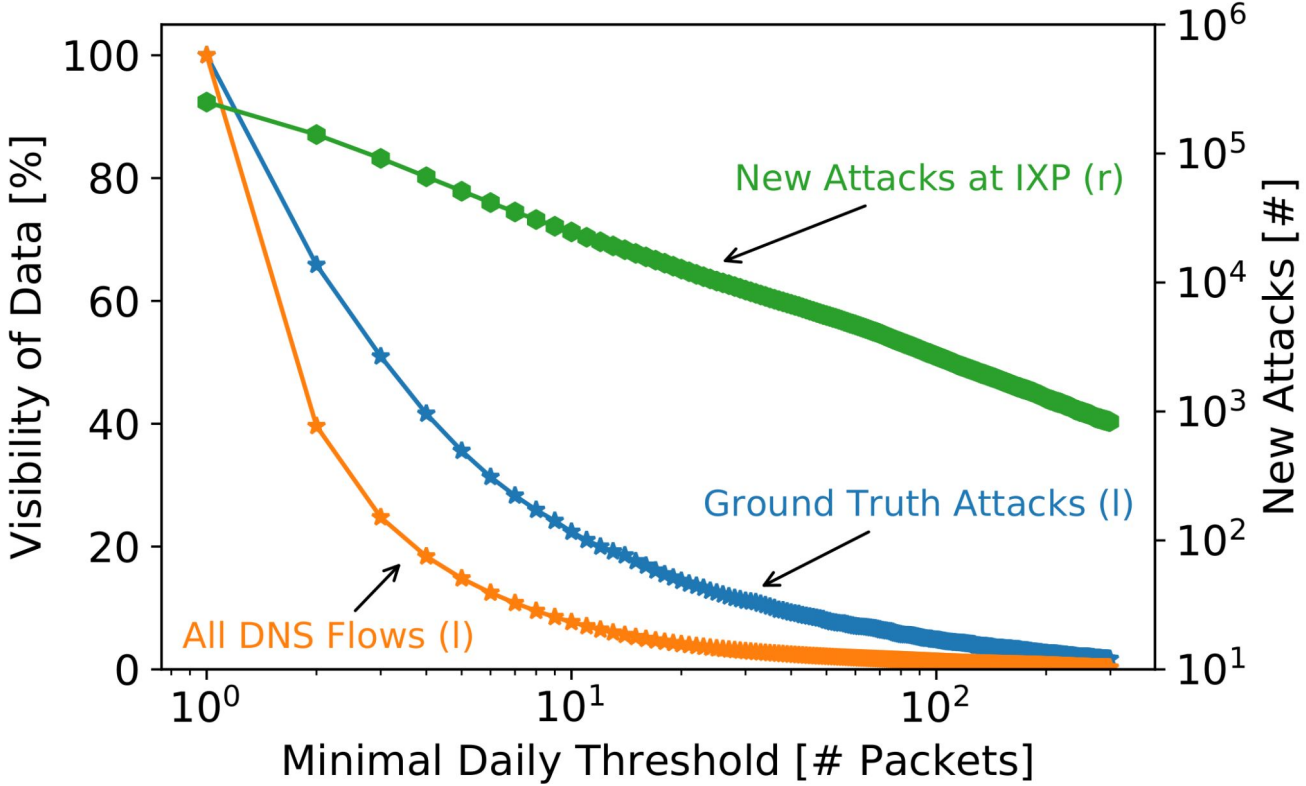
Unsupervised clustering of amplifier lists?
T-SNE & DBSCAN find almost no clusters.



Can you attribute attacks at the IXP?



How many flows do you see per client?



Is the major attack entity still active?

No. This summer changed a lot. It is unlikely that the attack entity is currently active.