

Towards detecting DNSSEC validation failure with passive measurements

Kensuke Fukuda
NII/Sokendai
kensuke@nii.ac.jp

Yoshiro Yoneya
Japan Registry Services
yoshiro.yoneya@jprs.co.jp

Takeshi Mitamura
Japan Registry Services
mitamura.takeshi@jprs.jp

Abstract—DNSSEC is a promising technique to enhance the security of the DNS, by authenticating a chain of trusts in the DNS hierarchy. However, the deployment of DNSSEC is still on the way. One of the reasons of this under deployment is due to not enough reliability of DNSSEC validations in real world operations, especially DNSSEC validation failures due to operation misses. Thus, DNS operators require quick and reliable ways to detect such DNSSEC validation failures. The current best practice of this issue is to validate all the DNSSEC available zones periodically. However, this approach has a scalability issue. In this paper, we aim at detecting the validation failures by analyzing passively collected DNS queries at TLD-level DNS authoritative servers in a lightweight way, instead of by actively validating all the zones. In particular, we focus on changes of DNS query patterns at the authoritative server before and after DNSSEC validation failures. We conduct large-scale active and passive controlled measurements with RIPE Atlas probes and a dedicated authoritative server, to show the validity of this approach. We demonstrate that increases in DNSKEY queries are a promising candidate of metrics to detect the failures from passively collected query data. Also, we show that an increase in the number of queries is limited even for short TTL values of DNSSEC queries, thus shorter TTL is beneficial for mitigating caching effect for DNSSEC validation failures.

I. INTRODUCTION

DNSSEC (DNS Security Extensions) [2]–[4] is a technique to increase the reliability of the DNS by authorizing the chain of trusts that spans in the DNS hierarchy. DNSSEC enables users to guarantee that the query results are not modified between authoritative servers and clients, so that it is useful in preventing man-in-the-middle type attacks. However, the deployment of DNSSEC is still on the way. One of the reasons of this low deployment is due to low reliability of DNSSEC operations. According to a website [14], TLD-level DNSSEC validation failures have been reported many times since DNSSEC has been introduced. These failures are mainly because of complexity of key management in the DNS hierarchy. To realize the chain of trusts, each authoritative servers appropriately maintain pairs of DNSSEC related records (DNSKEY and DS). If either of the records in the DNS hierarchy is wrongly set or expired, the DNSSEC validation fails at validators (caching resolvers). In this sense, the impact of failures at a higher location in the hierarchy is more serious in the DNS operation.

Furthermore, DNS (and DNSSEC) ecosystem is very complicated; many types of validations are available at client side.

A typical scenario is validation at ISP's DNS resolvers. In this case home users using these resolvers do not care about the validation process itself. Similarly, users can choose DNSSEC available public DNS service such as google, cloudflare, yandex. Client's web browser extensions [8], furthermore, support DNSSEC validation without external validation resolvers. Thus, this complicated ecosystem yields a difficulty on inferring the impact of the validation failure on Internet users.

As explaining so far, detecting such validation failures is an important issue for network and DNS operators. A simple and current best practice used by DNS operators is to actively send DNSSEC queries to authoritative servers and to validate the query replies. However, the scalability does matter for a larger number of DNSSEC available zones.

For this issue, we focus on macroscopic changes of DNS query behaviors at TLD-level authoritative servers before and after DNSSEC validation failure. One of the advantages of this approach is that DNS operators have already been collecting coarse grained query data (with some tools, e.g., DSC [10]) as their daily task. Also, the passive approach is helpful in understanding the impact of the failures because the query data shows the types of validators (e.g., ISP's resolvers or public resolvers).

As a preliminary work, we conduct active and passive controlled measurements with a dedicated local DNS authoritative server and more than 500 RIPE Atlas probes [20], to investigate a possibility of validity of this approach. We set up some zone settings including several TTL values and DNSSEC validation availability at the local authoritative server to emulate several types of domains (e.g., mainly short TTLs or long TTLs). Then, Atlas probes send DNSSEC queries to the DNS, and capture those replies at JP-DNS (ccTLD servers of “.jp”) as a vantage point.

Our measurement results highlight that the change of DNSKEY queries at the vantage point is a promising candidate to detect the validation failures at TLD-level authoritative servers though DS queries do not change significantly. We also show that shorter TTL values are beneficial for mitigating the impact of DNSSEC validation failures on users, because the increase in the number of queries is limited even for shorter TTL values of DNSSEC related records.

II. RELATED WORK

Since DNSSEC has been introduced, several measurement papers focused on its security, availability, and operational issues.

The first category of the papers quantified security impact of DNSSEC. DNSSEC introduces new resource records that contain cryptographic keys and signatures. Thus, larger size the DNSSEC response may be used for reflection-based amplification DDoS attacks [12], [22]. Bau et al. [5] analyzed several security properties of the DNSSEC protocol with a finite-state enumeration tool.

The second category of the papers highlighted misconfiguration and availability of DNSSEC [6], [9], [15], [19]. They pointed out the difficulty of setting up complicated DNSSEC key settings. The overhead/performance of the validation has also been studied [1], [13], [16]. However, this issue seems not serious due to a recent increase in computation powers. As an operational issue, some works focus on finding DNSSEC validators from passive query data [11], [18], because caching resolvers send DNSSEC queries but most of them do not actually validate those query replies. Furthermore, reasons of the low adaption of DNSSEC are analyzed in terms of DNS registrars [7]. It points out that user interfaces of DNSSEC key management is not well prepared in many registrars.

Regarding the TTL settings of DNS, Moura et al. [17] reported a wide spread of TTL settings for A/AAAA in popular websites (Alexa/Majestic/Umbrella). Also, Moura et al. and Chung et al. [6] showed that the TTL of DNSSEC KEY are mostly 1 hour or 1 day [6].

Compared with the past literature, the originality of our paper is focusing on detecting the DNSSEC validation failures at the authoritative side towards assessing the impact of the failures.

III. METHODOLOGY

A. Controlled measurements

It is not an easy task to measure DNSSEC validation failures by observing DNS queries in general, because such validation errors do not occur so frequently at target zones. Instead of waiting for failures, we rely on scenario-based measurements that we control TTLs of records, query frequency, and DNSSEC settings. Through this paper, we assume that failures occur at the authoritative side (e.g., inconsistency of DS and DNSKEY records), and do not consider the case at the validator (e.g., software bugs at the validator). Also we concentrate on the impact of the validation failures on home users that use ISP's validators (i.e., without public DNS and web browsers).

Figure 1 illustrates an overview of our controlled measurements. The measurement equipment consists of three parts: a dedicated local authoritative server, DNS clients, and a vantage point. We set up a local (3LD) authoritative server that includes both DNSSEC validation successful (e.g., dns-ok1.dnslab.jp) and failed zones (e.g., dns-ng1.dnslab.jp). We prepare several parameter sets of these two types of zones in order to emulate

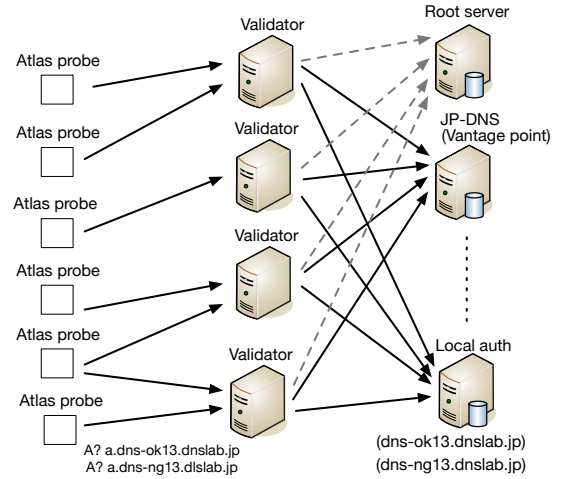


Fig. 1. DNSSEC validation measurements

typical TTL settings of websites. These parameters do not dynamically change during the measurement.

As DNS clients, we use RIPE atlas probes that are a Linux-based small computer for network measurements. We intend to select a set of probes whose issued DNSSEC queries are successfully validated, because simply issuing DNS queries with DO bit (DNSSEC) does not always mean that the queries are validated. In the end, we conduct a following pre-test: (1) We pick up 1200 stable IPv4 probes located at home by tag information provided by the system/users. (2) Then, we set up a user-defined measurement of Atlas that queries “A (random-id).dns-ok1.dnslab.jp” with DO flag (use DNSSEC) for the candidate Atlas probes. (3) After the measurement, we know about a set of probes that use DNSSEC validators by checking the status flag of DNS query reply (i.e., AD flag (Authentic Data)). Probes send queries to caching resolvers followed by their DNS setting (e.g., appeared in resolv.conf in a unix system). We consider mainly three possible cases for selecting caching resolvers: forwarded to ISP's resolvers, forwarded to public DNS, forwarded to local resolvers (at home). Our focus is on estimating influence on DNS clients using customer-side (home, ISPs) DNSSEC validators. In order to exclude public DNSSEC validators (e.g., google, cloudflare), we also check DNSSEC queries at the local authoritative server in this pre-test. If source IP addresses of queries belong to such ASes (e.g., google) at the authoritative server, we do not use these caching resolvers, because the original query is forwarded to the public DNS. Thanks to random prefix of the query (random-id), we can detect probe-validator relationship by matching the query name at the probe and the authoritative server (as shown in the past literature [21]). In our pre-test, we obtained over 600 probes and 900 DNSSEC validators (cache resolvers).

As a vantage point, we measured all DNS queries at JP-DNS (TLD for JP), consisting of eight servers (and more than 20 replications) spread over the world. We conducted

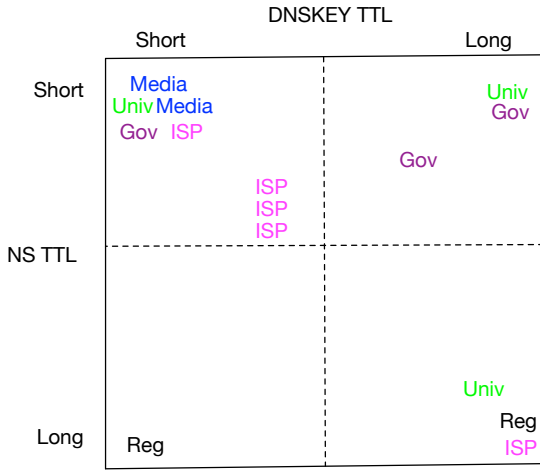


Fig. 2. TTL settings of major Japanese DNSSEC available zones; Media: TV and newspaper sites, Gov: government sites, Univ: universities, Reg: registries, ISP: broadband ISPs.

the active measurements with the probes for 50 hours in April 2018. In this active measurements, the probes repeatedly send A query (e.g., A www.dns-ok1.dnslab.jp) with DO flag, and all the related queries are recorded at the probes, the local authoritative server, and the vantage point. These three query logs for each target queries will be analyzed in the following section.

B. Zone parameters

In order to estimate the effect of DNSSEC validation failure, we need to consider some typical TTL settings for queries. Here, we focus on A, NS, DNSKEY, and DS query types. There are no clear rules for those TTLs so DNS operators set arbitrary TTL values based on their choice [17]. Before the main measurement, we survey TTL settings of some Japanese DNSSEC available websites by sending several query types.

Among those query types, the TTL of DS record is the same value in .jp (7200s). A/AAAA record is typically short for contents providers, but stable sites (less dynamic updates) prefer longer TTLs for cache efficiency (0s - 86400s). In order to visualize the TTLs of the rest types (DNSKEY and NS), we plot their TTLs in Figure 2. Long TTL represents more than 86400s (1day) and short TTL represents less than 300s. Symbols in the figure represent the types of websites: Media: TV and newspaper sites, Gov: government sites, Univ: universities, Reg: DNS registries, ISP: broadband ISPs. At a glance, it is difficult to identify typical settings of the TTLs per type of website, except for Media; they prefer shorter TTLs for dynamic contents update. From these results, we conclude to use four settings of NS and DNSKEY TTLs in the main measurement: (1) short NS and short DNSKEY, (2) short NS and long DNSKEY, (3) long NS and short DNSKEY, and (4) long NS and long DNSKEY.

According to the above survey and criteria, we determine following parameter sets in our measurement (Table I). Here, label "ok" indicates the setting that DNSSEC validation is

Label	Valid	TTL				Interval	Duration
		NS	DKEY	DS	A		
ok1	y	600	600	7200	600	540	8100
ok2	y	600	7200	7200	600	540	8100
ok3	y	7200	600	7200	600	540	8100
ok4	y	7200	7200	7200	600	540	8100
ng1	n	600	600	7200	600	540	8100
ng2	n	600	7200	7200	600	540	8100
ng3	n	7200	600	7200	600	540	8100
ng4	n	7200	7200	7200	600	540	8100

TABLE I

TTL SETTINGS. VALID INDICATES SUCCESSFUL VALIDATOR (Y) AND FAILED VALIDATOR (N).

name	All	A	DKEY	DS	NS
ok1	16003	6777	5512	2705	745
ok2	12938	6973	2117	2834	767
ok3	16018	6841	5675	2858	370
ok4	11424	6469	2003	2622	188

TABLE II

NUMBER OF QUERIES AT LOCAL (VALIDATION SUCCESS)

successful, and "ng" is the case that the validation is always failed due to the mismatch of DNSKEY record to the corresponding DS record. We also prepare four different types of TTL settings based on the above result.

C. Probe parameters

As explained, we set up user-defined measurements on RIPE Atlas probes for sending DNSSEC queries. During a measurement period, all the probes repeatedly query A records (e.g., www.dns-ok1.dnslab.jp). We empirically set the period of one measurement to 8100s and the interval of the queries to 540s, in order to observe an enough number of queries for data processing in the largest zone TTL setting (7200s). The interval of the queries is randomly distributed in 15s to avoid excess of the query load. All the measurement results are available at the RIPE Atlas web page (we will add a link to the raw data after the paper is accepted).

IV. RESULTS

A. Observing at local authority

At first, we analyze query replies at the local authoritative server in order to understand basic query behavior in our experiment. This is because the appearance of the queries at the local authoritative server is less affected by the DNS complicated caching mechanism. In other words, we would confirm more queries at the local authoritative server for the further analysis as the baseline.

Table II lists the number of queries appeared at the local authoritative server with the successful DNSSEC validation

name	All	A	DKEY	DS	NS
ng1	64439	14929	38319	7781	2688
ng2	63847	15292	37017	8012	2743
ng3	62805	14661	37161	7841	2397
ng4	61283	14755	35488	7785	2471

TABLE III

NUMBER OF QUERIES AT LOCAL (VALIDATION FAIL)

scenarios. We observe approximately 7000 A record queries for all the settings. This is because the TTL values of A record (600s) are fixed in the experiment. The query interval is less than the TTL of A record, thus the caching effect is limited as expected. However, it is difficult to infer the exact numbers of queries due to several reasons, e.g., caching timing, caching ignorance, and query duplication, different from simulation approaches.

The numbers of DNSKEY queries are varied with the TTL setting value of the query. We confirm that the shorter TTL settings make the number of queries larger; The shorter TTL setting (600s for ok1 and ok3) generate almost twice queries as the longer TTL setting (7200s for ok2 and ok4). However, it is noted that the ratio of the two values of TTL is 12 times larger. This result suggested that longer TTLs of DNSKEY do not help linearly decreasing the number of queries so much. The numbers of NS queries also follow the similar trend.

The numbers of DS queries are almost similar among the setting due to the fixed value of the DS TTL. However, the numbers of them are about 30-40% of the numbers of A queries though the TTL settings are more than 10 times different. This is likely due to the caching effect at DNS resolvers; we would expect one or two DS queries per probe, considering the TTL (7200s) value and the measurement period (8100s). DS records are sometimes appeared in the additional section of the reply of other queries (e.g., A, NS records). However, this result suggests that many validators issues explicit DS queries instead of trusting such additional information. Thus, observing DS queries is still a good way to track DNSSEC validators as shown in the past literature [11].

Next, we investigate the number of queries at the local authoritative server in the DNSSEC validation failed scenarios. We observe a larger number of queries in the failure case as listed in Table III. Type A queries are almost twice to those in the successful validation scenarios. The reason of this is that once the validation is failed, the validator do not cache such failed queries. However, again, it is not easy to estimate the exact number of queries from the real experiment data due to complicated caching mechanisms in DNS. We further confirm the increase in the number of other query types. One notable point is that the number of DNSKEY queries is much larger (over 7x) than that of the successful case. This DNSKEY's trend is more highlighted than that of DS queries. This result suggests that counting the number of queries, especially DNSKEY queries, would be a hint to detect the DNSSEC validation failure.

takeaway: the number of queries increase during DNS validation failure. In particular, the change of the number of DNSKEY is significant compared with that of DS queries. In addition, we confirm that shorter TTLs do not increase DNSSEC related queries so much.

B. Observing at vantage point

Next, we focus on the number of appeared queries at the vantage point (the authoritative servers of ".jp").

name	All	A	DKEY	DS	NS
ok1	1456	1305	61	87	2
ok2	1816	1705	14	91	1
ok3	1415	1268	39	101	0
ok4	1285	1171	12	99	0

TABLE IV
NUMBER OF QUERIES AT JP-DNS (VALIDATION SUCCESS)

name	All	A	DKEY	DS	NS
ng1	2342	2081	136	114	1
ng2	1958	1777	71	105	1
ng3	1594	1372	105	110	1
ng4	1468	1279	81	104	1

TABLE V
NUMBER OF QUERIES AT JP-DNS (VALIDATION FAIL)

Table IV lists the number of DNS queries observed at the vantage point (JP-DNS) for the successful validation scenarios. Comparing to the local authority's results (Table II), the numbers of the queries is quite smaller than those at the local authoritative server due to caching effect at caching resolvers. For example, once NS record of dnslab.jp is resolved, following queries for dnslab.jp (e.g., A www.dnsok1.dnslab.jp) do not arrive at the JP-DNS. Thus, at the ccTLD level, most of related queries from one resolver are disappeared. The attenuation of A, DNSKEY, DS and NS queries are 19%, 1%, 3%, and 0.3%, respectively.

Finally, we show the numbers of DNS queries appeared at the vantage point during a DNSSEC failure scenario. An increase in the number of A queries is due to negative cache (NCACHE: 900s) different from the successful scenario. An interesting point is that the number of DS queries does not increase much but we observed much more DNSKEY queries; 1.2x for DS and 2.2x - 6.8x for DNSKEY. Compared with the local authority's results, the change of DNSKEY still has impact on the validation failure detection.

takeaway: the number of observable queries at the vantage point is much less than those at the local authority because of caching effects. Nonetheless, the change of DNSKEY queries is a good hint to characterize the DNSSEC validation failure.

V. DISCUSSION

The ultimate goal of our work is to detect DNSSEC validation failures from changes of DNS queries at the ccTLD authority (JP-DNS). In our experiment, we observed a limited number of queries at the vantage point due to caching effects. However, our results show that the number of DNSKEY queries changed significantly in the successful and failed validation scenarios. Here, we mainly discuss two issues: query features for distinguishing DNSSEC validators, and TTL settings for mitigating validation failure effect.

We review query characteristics of successful DNSSEC validators, failed validators, and non-validators (i.e., query with DO bit but never validate the reply) from our results. To identify DNSSEC validator and non-validator a prior work [11] demonstrated that the ratio of DS queries to A queries is significantly higher in DNSSEC validators, because non-validators do not explicitly send DS queries. On the other

type	#A	#DS	#DNSKEY
non-validator	n	0	0
successful validator	n	m	l
failed validator	$> n$	$> m$	$> l$

TABLE VI
QUERY FEATURES OF DNSSEC VALIDATORS

hand, failed validators still send DS and DNSKEY queries for trying to validate the requested zone, as well as more A queries (because failed queries are not cached). Note also that simply checking changes of A/AAAA queries could be enough to detect DNSSEC validation failure from our results, but it is likely true that changes of A queries are more sensitive than those of DNSSEC related queries because of non-validator's queries. i.e., the change by the failure is hidden by other legitimate A/AAAA queries. We list the query features for three types of DNSSEC validators in Table VI. Here, symbols n , m , and l indicate the approximate number of queries for A, DS, and DNSKEY respectively ($n \gg m \approx l$), depending on TTL values. In summary, we can classify non-validators, failed validators, and successful validators from passively collected DNS query traces.

Next, we discuss the TTL settings of DNSSEC related records. There is no consensus on the TTL value of each record. In general, longer TTL is preferred for stable servers and shorter is for dynamic contents. In terms of the DNSSEC validation, once DNSSEC settings are fail at an authoritative server, the effect of the failure is distributed in DNSSEC validators (caching resolvers) and last until their cache expiration. Shorter TTL is beneficial to minimize the caching effect due to the failure, though the number of queries increases. Our results show that the increase of the queries is not linearly proportional to the decrease in the TTLs (e.g., the number of A records does not increase twice if the TTL of A record is half). The impact of the short TTL is not significant in the number of queries. Thus, setting shorter TTL would be beneficial for mitigating the impact of DNSSEC validation failure.

We observed notable changes of DNS queries in successful and failed DNSSEC validation scenarios. However, still it is not clear how to identify such changes in a huge number of queries aggregated at a vantage point. We will apply time series based anomaly detection methods to DNS query data at JP-DNS. The main challenge of the future work will be to quantify the detection accuracy of DNSSEC validation failure when we inject queries in failed validation to passively collected DNS traces.

VI. CONCLUSION

DNSSEC is a promising technique to enhance the security of DNS, by authenticating a chain of trusts in the DNS hierarchy. However, the deployment of DNSSEC is still on the way. One of the reasons of this under deployment is due to not enough reliability of DNSSEC validations in real world operations, especially DNSSEC validation failures due to operation misses.

In this paper, we aim at quickly and reliably detecting such DNSSEC validation failures from passively collected DNS queries at TLD-level DNS authoritative servers. In particular, we focus on changes of DNS query patterns at the server before and after DNSSEC validation failures. Our large-scale active and passive controlled measurements with RIPE Atlas probes and authoritative servers demonstrate that increases in DNSKEY queries are a promising candidate of metrics to detect the failures from passively collected query data. Furthermore, we showed that the increase in the number of queries is limited even for short TTL values of DNSSEC queries, thus shorter TTL is beneficial for mitigating caching effect for DNSSEC validation failures.

ACKNOWLEDGEMENT

We are grateful to thank Satoru Kobayashi for their helpful discussion. We also thank the anonymous reviewers for their helpful comments. Kensuke Fukuda's work is partially supported by JSPS KAKENHI (18H03237).

REFERENCES

- [1] B. Ager, H. Dreger, and A. Feldmann. Predicting the dnssec overhead using dns traces. In *Annual Conference on Information Sciences and Systems*, pages 1484–1489, Princeton, NJ, 2006.
- [2] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. DNS Security Introduction and Requirements. RFC4033, Mar 2005.
- [3] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. Protocol Modifications for the DNS Security Extensions. RFC4035, Mar 2005.
- [4] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. Resource Records for the DNS Security Extensions. RFC4034, Mar 2005.
- [5] J. Bau and J. C. Mitchell. A security evaluation of dnssec with nsec3. In *NDSS*, page 18, San Diego, CA, 2010.
- [6] T. Chung and et al. A longitudinal, end-to-end view of the dnssec ecosystem. In *Usenix Security'17*, pages 1307–1322, Vancouver, Canada, 2017.
- [7] T. Chung, R. van Rijswijk-Deij, D. Choffnes, D. Levin, B. M. Maggs, A. Misllove, and C. Wilson. Understanding the role of registrars in dnssec deployment. In *IMC'17*, pages 369–383, London, UK, 2017.
- [8] CZ.NIC Labs. Dnssec/tlsa validator. <https://www.dnssec-validator.cz/>.
- [9] C. Deccio, J. Sedayao, K. Kant, and P. Mohapatra. Quantifying and improving DNSSEC availability. In *ICCCN'11*, Maui, HI, 2011.
- [10] DNS-OARC. DSC. <https://www.dns-oarc.net/tools/dsc/>.
- [11] K. Fukuda, S. Sato, and T. Mitamura. A technique for counting dnssec validators. In *INFOCOM2013*, pages 80–84, Turin, Italy, Apr 2013.
- [12] A. Herzberg and H. Shulman. Dnssec: Security and availability challenges. In *CNS'13*, pages 365–366, National Harbor, MD, 2013.
- [13] G. Huston and G. Michaelson. Measuring dnssec performance. The ISP Column, 2013.
- [14] IANIX. Major DNSSEC outages and validation failures. <https://ianix.com/pub/dnssec-outages.html>, 2020.
- [15] W. Lian, E. Rescorla, H. Shacham, and S. Savage. Measuring the practical impact of dnssec deployment. In *USENIX Security*, pages 573–587, Washington, D.C, 2013.
- [16] D. Migault, C. Girard, and M. Laurent. A performance view on dnssec migration. In *CNSM'10*, pages 469–474, Niagara Falls, Canada, 2010.
- [17] G. C. M. Moura, J. Heidemann, R. de O. Schmidt, and W. Hardaker. Cache me if you can: Effects of dns time-to-live. In *IMC'19*, pages 101–115, Amsterdam, Netherlands, 2019.
- [18] Ólafur Gusmundsson and S. D. Crocker. Observing DNSSEC validation in the wild. In *Workshop on Securing and Trusting Internet Names (SATIN2011)*, page 8, Teddington, UK, Apr 2011.
- [19] E. Osterweil, M. Ryan, D. Massey, and L. Zhang. Quantifying the Operational Status of DNSSEC Deployment. In *IMC2008*, pages 231–241, Vouliagmeni, Greece, Oct 2008.
- [20] RIPE. Ripe atlas. <https://atlas.ripe.net/>.
- [21] K. Schomp, T. Callahan, M. Rabinovich, and M. Allman. On measuring the client-side DNS infrastructure. In *ACM IMC'13*, pages 77–90, Oct. 2013.

- [22] R. van Rijswijk-Deij, A. Sperotto, and A. Pras. Dnssec and its potential for ddos attacks: a comprehensive measurement study. In *IMC'14*, pages 449–460, Vancouver, Canada, 2014.