A Replication Study of DSAV Measurement

Christopher Kitras and Bryson Schiel
Department of Electrical and Computer Engineering
Brigham Young University
Provo, UT, United States
{chkitras,schielb}@byu.edu

ABSTRACT

Destination-side source address validation (DSAV) plays an important role in the modern internet. Despite this, many autonomous systems have a history of this service being disabled, as was indicated in a 2020 measurement study on DSAV. This paper acts as a replication study for that work and seeks to measure any change in DSAV utilization. While little has changed overall in global DSAV utilization between the two studies, several outliers are discovered and discussed.

1 INTRODUCTION

Researchers continue to play a crucial part in helping to identify vulnerabilities in the internet. One such vulnerability that researchers have focused on is attackers sending packets to a network with an IP address that originates from within that network (see Figure 1). In the case of a network space that has a NAT or some other gateway device, the network depends on destination-side source address validation (DSAV) to keep the network protected. DSAV verifies that the incoming packet is from a valid source, and if an external packet's source is spoofed to be from an internal address, then DSAV should reject it. Where DSAV fails, however, the packet gets forwarded to its target within the network (see Figure 2).

This attack and its measurement were the focus of a 2020 study performed by Deccio *et al.* [2]. In this study, the researchers transmitted DNS queries to recursive DNS resolvers. These packets had spoofed source addresses that made them appear to have originated within an internal network that is accessible to the DNS server. The DNS query itself was directed to an authoritative DNS server that the researchers controlled; they would log the queries coming in and see if any matched up the queries they had sent to the recursive resolver.

In the event that a query to the authoritative server matched up with a query from the experiment, Deccio *et al.* could determine that the recursive resolver that query went to was not protected by DSAV.

This paper will perform much of the same testing and determine if any major changes can be noted in the results.

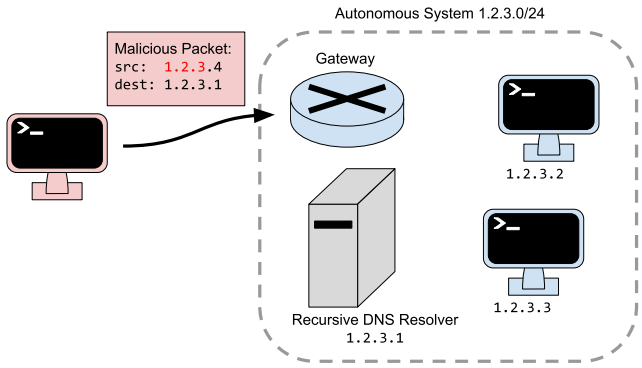

Figure 1: An attacker from outside a network sends a packet addressed as if from within the system. If the gateway device lets this packet in, other devices might assume they are receiving messages from within the network. This can then form a basis for DOS and other attacks.

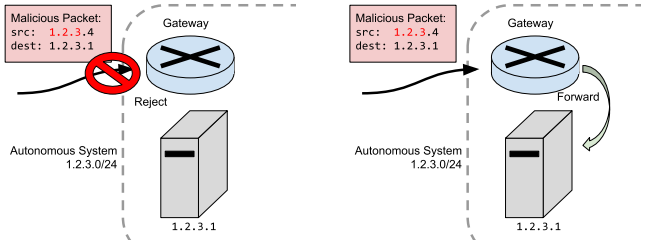

Figure 2: (a) shows successful DSAV, where the gateway rejects an external packet with an internal source address, while (b) shows a DSAV failure.

2 METHODOLOGY

In this study, we aim to closely follow the methodology laid out in Deccio *et al.*, but some changes were made to simplify the scale of the research.

2.1 Changes from Deccio *et al.*

The most notable change in our methodology comes down to the source of data that we use. Our resource to identify DNS

recursive resolvers comes from DNS logs for the byu.edu authoritative DSN servers, whereas Deccio *et al.* used a Day in the Life (DITL) list for their research. This gives us a narrower slice of the internet to query, but we still have access to hundreds of thousands of servers and get a good representation from much fo the world's internet in our research. We also limit our queries to use IPv4 networking addresses, while Deccio *et al.* used both IPv4 and IPv6; this is due to network administration limitations at the time of our measurements.

Additionally, while Deccio *et al.* sent approximately 100 queries to each DNS server, our work limits it to 25, as this still gives a good sense of the networks without using too many resources for a replication study.

Lastly, we use the Python library Scapy [1] to spoof our packets. This provides a very simple interface for modifying address contents, and it provides an update to the methodology as performed by Deccio *et al.*

2.2 Measurements

We analyze DNS logs for the byu.edu authoritative server to identify IP addresses of recursive DNS servers that we can test against. From this list, we specifically look for servers with IPv4 addresses that query other IPv4 address resources ("A"). We do this as our sending infrastructure is limited to IPv4 resources for this study. From this specification, and using DNS logs from August 26 to September 15, 2024, we gather a total of 444,781 IPv4 addresses for DNS servers.

For each IP address, we send up to 25 packets with spoofed source addresses, with the expectation that these will be considered "within" the server's network. As we send these spoofed packets, we would hope that the gateway to the DNS server performs DSAV and drops the packet (see Figure 2a). Of course, there are many other reasons why a DNS server would refuse to service a query, so we cannot assume that all dropped packets are caused by a successful DSAV application.

Alternatively, any instance where a query is sent to our authoritative server, we know that DSAV has failed (see Figure 2b), and the recursive DNS server has received a packet that should have only come from its internal network. These failures are the main focus of our analysis.

For the 25 packets we send, they are composed of the following:

- *Different prefix*: 21 packets from other-prefix addresses associated with the AS.
- *Same prefix*: an IP address form the *same* /24 prefix as the DNS server.
- *Unique local*: 192.168.0.10
- *Destination-as-source*: the target IP address itself.
- *Loopback*: 127.0.0.1

This is the exact same list of spoofed source addresses used in Deccio *et al.* with the previously noted exception of using fewer *different prefix* addresses.

2.3 Ethics

To minimize the impact of the various DNS servers, we limit our transmission rate to 700 messages per second. Additionally, any HTTP lookup to the server our queries reference goes to a webpage hosted by the lab that owned the DNS server. This webpage, `https://imaal.byu.edu/`, informs the searcher that the work on the server is related to internet measurement. There is also contact information in case the server administrator wishes the experiment to exclude their server from the measurement.

3 ANALYSIS

For our analysis, we start by characterizing several of the countries we observe in our experimentation; this is covered in Subsection 3.1. We then compare our results to those discovered by Deccio *et al.* in Subsection 3.2

3.1 Vulnerability by Country

As part of our analysis, we characterize each country by its vulnerability. In Figure 3, we sort the countries that proved vulnerable in our study from least-to-most vulnerable (if a country has a vulnerability of 0%, it can be found on the left of the graph, meaning that we have a non-zero amount of ASes that we target, but zero of them sent DNS queries in response; if all ASes that we target sent DNS responses, then the country is on the right with a vulnerability of 100%).

A few things can be noted from this graph:

3.1.1 Endpoints. There is very little data for the endpoints at the extremes of the graph, where the vulnerability is roughly either 0% or 100%. Looking at the countries involved, many of them fall in the developing world, and they have very few ASes we can query (between 1-4). This is far from a statistically significant portion of the networking resources of those countries, and so little can be determined from these endpoint data locations.

3.1.2 (Most of) The Top 10 Countries. Looking through the top 10 countries from Table 1, they almost all fall within the range of 30-70% reachable. The United States accounts for a massive portion of the queried ASes, which makes sense as it is a center for internet research and activity, and since our test sample draws from the DNS server of a US university. Outside of that spike, however, most of the top 10 countries account for fairly small portions of our AS space (10% of the US AS count). There are, however, two major outliers on either end of the spectrum.

Country	ASes Total	ASes Reachable	ASes Old Rate	IP Targets Total	IP Targets Reachable	IP Targets Old Rate
United States	4,340	1,746 (40.23%)	↓ 28%	248,300	7,155 (2.88%)	↑ 3.2%
Brazil	2,613	1,741 (66.62%)	↓ 59%	8,316	2,888 (34.72%)	↓ 4.8%
Russia	1,161	80 (6.89%)	↑ 59%	4,713	119 (2.52%)	↑ 11.6%
Germany	606	236 (38.94%)	↓ 36%	19,846	1100 (5.54%)	↓ 3.8%
India	584	403 (69.00%)	↓ 41%	11,651	1258 (10.79%)	↑ 11.6%
Indonesia	507	382 (75.34%)	–	6,006	834 (13.88%)	–
Great Britain	489	218 (44.58%)	↓ 33%	22,831	814 (3.56%)	↑ 4.5%
Poland	472	312 (66.10%)	↓ 52%	2,068	612 (29.59%)	↓ 6.0%
Canada	445	205 (46.06%)	↓ 36%	12,110	615 (5.07%)	↓ 2.8%
Ukraine	386	256 (66.32%)	↓ 63%	1,033	460 (44.53%)	↓ 15.4%

Table 1: DSAV results for the 10 countries associated with the largest fraction of ASes (IPv4 only) in our set of target IP addresses compared with values from the original paper. The red down arrow signifies a lower original rate and the green uparrow, a higher original rate.

Country	ASes Total	ASes Reachable	IP Targets Total	IP Targets Reachable
French Guiana	4	3 (75.00%)	19	18 (94.74%)
Virgin Islands	3	2 (66.67%)	6	5 (83.33%)
Tajikistan	13	11 (84.62%)	35	29 (82.86%)
Mali	3	3 (100.00%)	10	8 (80.00%)
Niger	5	4 (80.00%)	10	8 (80.00%)
Solomon Islands	3	3 (100.00%)	5	4 (80.00%)
Anguilla	4	3 (75.00%)	5	4 (80.00%)
Antigua and Barbuda	2	2 (100.00%)	5	4 (80.00%)
Guinea	7	6 (85.71%)	12	9 (75.00%)
Eswatini	4	3 (75.00%)	8	6 (75.00%)

Table 2: DSAV results for the 10 countries associated with the largest non-100% percentage of IP addresses (IPv4 only) reachable by spoofed-source packets.

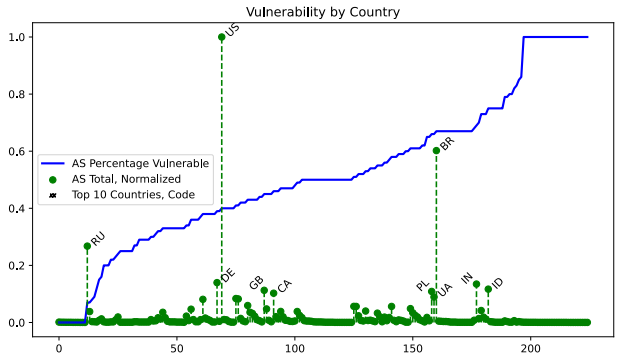

Figure 3: Characterizing the countries that we test. The blue line indicates how vulnerable a country is. The green dots represent the actual amount of ASes we test against, normalized to 1 (the US got the most traffic, so its spike goes to 1.0; every other dot is compared to the amount the US received). The dots of the top ten countries from Table 1 are labeled with their 2-letter country code.

3.1.3 Outliers: Russia and Brazil. Two countries stand out among the top 10 from Table 1 when you observe their quantity of ASes compared to their vulnerability to DSAV failure in our measurements. These are Russia and Brazil.

Russia: When observing Russia's ASes and their vulnerability, we see that they have a very low reachability rate, only about 6.9%. Despite this, they account for over a quarter of the ASes relative to what the US provides. Russia stands out among our top 10 as the most secure country while also being the third-most active.

Brazil: Similar to Russia, Brazil accounts for a major portion of the ASes we query, over half the amount in the US. However, it falls on the other end of the vulnerability spectrum compared to Russia—Brazil is in the top half of vulnerable countries, with an overall reachability of 66.6%. This indicates a very active, very vulnerable country for DSAV failure, presenting a large attack space for a single country.

These two countries provide data for outliers in typical DSAV adoption rates as a function of internet traffic, Russia being a positive outlier, and Brazil a negative one. By better identifying and measuring these types of cross-sections in data, we can more clearly characterize countries and their attack space for techniques such as DSAV failure.

3.2 Comparison to Deccio *et al.*

Although the methodology varies by only dataset and frequency of spoofed addresses querying a given target, we

	Category-Inclusive (one-or-more)				Category-Exclusive (only)			
Source Category	**Addresses**	**Old Addresses**	**ASNs**	**Old ASNs**	**Addresses**	**Old Addresses**	**ASNs**	**Old ASNs**
All Queried	444,781	11,204,889	17,487	53,922	444,781	11,204,889	17,487	53,922
All Reachable	28,931 (6.5%)	↓ 4.6%	9,007 (51.5%)	↓ 49%	28,931 (6.5%)	↓ 4.6%	9,007 (51.5%)	↓ 49%
Other Prefix	26,324 (90.9%)	↓ 78%	8,238 (91.5%)	↑ 97%	7,444 (25.7%)	↑ 33%	1,381 (15.3%)	↓ 6.9%
Same Prefix	21,439 (74.1%)	↓ 63%	8,288 (92.0%)	↓ 91%	2,565 (8.8%)	↑ 17%	1,258 (13.9%)	↓ 1.4%
Private	214 (.74%)	↑ 3.4%	90 (0.9%)	↑ 12%	20 (.06%)	↑ 0.5%	13 (.14%)	↑ 0.4%
Dst-as-Src	0 (0.0%)	↑ 17%	0 (0.0%)	↑ 47%	0 (0.0%)	↑ 2.6%	0 (0.0%)	↑ 0.8%
Loopback	0 (0.0%)	↑ 0%	0 (0.0%)	↑ 0%	0 (0.0%)	↑ 0%	0 (0.0%)	↑ 0%

Table 3: Number of IP addresses or ASNs for which at least one spoofed-source reached its target ("Category-Inclusive") or for which a spoofed-source category was the *only* one to reached its target ("Category-Exclusive"). Percentages in the "All Reachable" row represent the fraction of *all* targets queried, whereas other percentages in other rows represent the fraction of *reachable* targets. All these values are compared with values from the original paper. The red down arrow signifies a lower original rate and the green uparrow, a higher original rate.

still noticed some substantial differences between the original work and our results. We do not believe these results invalidate the previous findings but rather bolster the original claims and show the evolution of the topology of DSAV adoption within the last 4 years.

3.2.1 Top 10 Countries with Most Represented ASes. Comparing Table 1 to its corresponding figure in [2], we note that the top 10 countries remain mostly the same, albeit with a different order in the bottom six and the replacement of Australia with Indonesia. This indicates that our dataset reflects almost the same distribution of represented countries as that of the data from DITL logs. While we do not assert that the quality of the dataset is any better than that found in [2], we do believe it to be comparable in the sense that it does not overrepresent and particular country's data anymore than the original work.

We note that in the Top 10, the majority of the countries have seen a decrease in DSAV-compliant hosts. The only country that improved in its DSAV adoption rate across the ASes was Russia, going from a shocking 59% to just 6.89% of ASes that were open, and the IP address susceptibility rate dropped from 11.6% to 2.52%. The only other countries that saw marginal improvement in observing DSAV implementation were the United States, India, and Great Britain regarding individual IP address susceptibility.

3.2.2 Top 10 Countries with non 100% Susceptible IP Pool. It is interesting to note that in the original work, almost all the countries that had the most vulnerable IP addresses in their available pools were not to be found on our lists. Instead, Table 4 shows that the only country that carries over from the top vulnerable IP pool is Eswatini. All of the other countries both in the top 12 vulnerable IP pools by country and the following top 10 non-100 countries with highly vulnerable IP pools as seen in Table 2 are mainly island nations or are located in Africa. Most of these countries also have a very small representation in the total amount of IP addresses that responded, insinuating that the poorer in infrastructure the country is, the less it takes higher-level security concepts like DSAV into account.

3.2.3 Susceptibility Topology Across All Addresses and ASes. In [2], the authors created a group of different spoofed IP addresses that could potentially fool a DNS server into thinking the incoming packet was coming from within its network. The groups of spoofed IP addresses are the ones listed in Section 2. While looking at the percent of all IP addresses and hosts reachable, we observe a slight increase compared to the original study. This is a little disappointing since [2] provides specific source-spoofing techniques that concerned network administrators of these ASes could design against. However, upon closer analysis, we do notice an interesting trend in exactly which techniques see an increase in effectiveness while others have been completely nullified.

Taking the target address and generating a random IP address with the same prefix or even a different prefix within the same IP blocks owned by the AS are still the most effective methods to successfully deliver a spoofed IP address. However, we see a dramatic dip in the effectiveness of using an IP address from the reserved private IPv4 address space. Even more surprising is the absolute lack of success in using the target address as the source address. What used to have almost a 50% success rate in the old ASNs has dropped to 0% across ASNs and IP addresses entirely. Finally, we see that using the loopback address is ineffective with no queries getting through to any of our IPv4 targets.

4 DISCUSSION

Looking through our data and comparing it with Deccio *et al.*, we notice several similarities in the major contributors to the

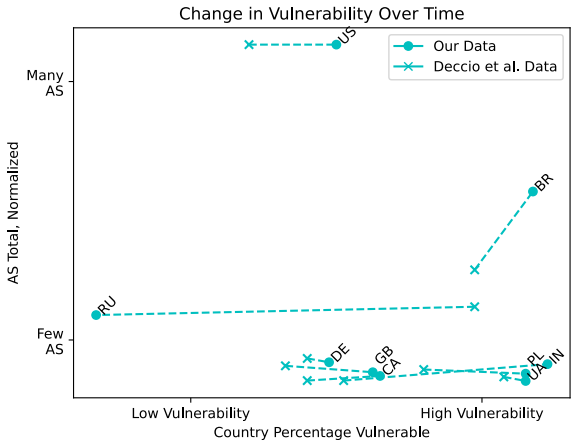

Figure 4: A measurement of change over time, using the same metrics as Figure 3. Note that all countries have increased in vulnerability with the dramatic exception of Russia, which has seen a major reduction. Brazil has also seen a major increase in ASes relative to the United States.

data. For example, nine of the top ten source countries have stayed the same. At the same time, there are notable shifts within these countries. All but one Russia have increased in vulnerability, and Brazil has dramatically risen in resource provision. This adjustment over time can be seen in Figure 4.

It is of course possible that the origin for these two outliers in the data comes from the dataset itself; we use a single DNS server's logs rather than something more authoritative like a DITL list. However, this likelihood seems mitigated by the close correlation between our top ten and Deccio *et al.*'s top ten ignoring these two countries.

5 CONCLUSION

Even though not much time has passed since the original study for this paper, we have observed that little has changed by and large in the effort to implement DSAV across the internet. Notwithstanding, Table 3 shows that even though the percentage of leakages has risen in the last 4 years, the type of leakages tend are now more concentrated into spoofing packets with believable prefixes instead of relying on operating system misconfiguration. Future work in DSAV may look into easy-to-implement methods that can detect when traffic coming from outside of the AS has addresses that are found within that same AS.

REFERENCES

[1] Philippe Biondi and contributors. 2024. Scapy: A Python tool for packet crafting and manipulation. https://scapy.net/ Version 2.5.0, accessed December 9, 2024.

[2] Casey Deccio, Alden Hilton, Michael Briggs, Trevin Avery, and Robert Richardson. 2020. Behind Closed Doors: A Network Tale of Spoofing, Intrusion, and False DNS Security. In *Proceedings of the ACM Internet Measurement Conference* (Virtual Event, USA) *(IMC '20)*. Association for Computing Machinery, New York, NY, USA, 65–77. https://doi.org/10.1145/3419394.3423649

A TOP 12 COUNTRIES WITH 100% EXPOSED IP ADDRESSES

	ASes		IP Targets	
Country	**Total**	**Reachable**	**Total**	**Reachable**
Sudan	3	3 (100.00%)	5	5 (100.00%)
American Samoa	2	2 (100.00%)	5	5 (100.00%)
Vanuatu	2	2 (100.00%)	4	4 (100.00%)
Palau	2	2 (100.00%)	4	4 (100.00%)
Caribbean Netherlands	2	2 (100.00%)	3	3 (100.00%)
Guadeloupe	2	2 (100.00%)	4	4 (100.00%)
Sint Maarten	2	2 (100.00%)	4	4 (100.00%)
Cook Islands	1	1 (100.00%)	3	3 (100.00%)
Saint Martin	1	1 (100.00%)	2	2 (100.00%)
Saint Lucia	1	1 (100.00%)	3	3 (100.00%)
Turkmenistan	2	2 (100.00%)	2	2 (100.00%)
Guernsey	1	1 (100.00%)	1	1 (100.00%)

Table 4: DSAV results for the 12 omitted countries associated with the largest 100% percentage of IP addresses (IPv4 only) reachable by spoofed-source packets.