Prefix Top Lists: Gaining Insights with Prefixes from Domain-based Top Lists on DNS Deployment

Johannes Naab
Technical University of Munich
naab@net.in.tum.de

Patrick Sattler
Technical University of Munich
sattler@net.in.tum.de

Jonas Jelten
Technical University of Munich
jelten@net.in.tum.de

Oliver Gasser
Technical University of Munich
gasser@net.in.tum.de

Georg Carle
Technical University of Munich
carle@net.in.tum.de

ABSTRACT

Domain-based top lists such as the Alexa Top 1M strive to portray the popularity of web domains. Even though their shortcomings (*e.g.*, instability, no aggregation, lack of weights) have been pointed out, domain-based top lists still are an important element of Internet measurement studies.

In this paper we present the concept of *prefix top lists*, which ameliorate some of the shortcomings, while providing insights into the importance of addresses of domain-based top lists. With prefix top lists we aggregate domain-based top lists into network prefixes and apply a Zipf distribution to assign weights to each prefix. In our analysis we find that different domain-based top lists provide differentiated views on Internet prefixes. In addition, we observe very small weight changes over time. We leverage prefix top lists to conduct an evaluation of the DNS to classify the deployment quality of domains. We show that popular domains adhere to name server recommendations for IPv4, but IPv6 compliance is still lacking. Finally, we provide these enhanced and more stable prefix top lists to fellow researchers which can use them to obtain more representative measurement results.

CCS CONCEPTS

• **Networks** → **Network measurement**.

KEYWORDS

Prefix Top Lists, Prefix Ranking, Internet Top Lists, Internet Measurement

ACM Reference Format:
Johannes Naab, Patrick Sattler, Jonas Jelten, Oliver Gasser, and Georg Carle. 2019. Prefix Top Lists: Gaining Insights with Prefixes from Domain-based Top Lists on DNS Deployment. In *IMC '19: ACM Internet Measurement Conference, October 21–23, 2019, Amsterdam, Netherlands.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3355369.3355598

IMC '19, October 21–23, 2019, Amsterdam, Netherlands

ACM ISBN 978-1-4503-6948-0/19/10…$15.00
https://doi.org/10.1145/3355369.3355598

1 INTRODUCTION

An important step before conducting any Internet measurements is the process of target selection. Examples of selected targets can be the complete Internet (which is not feasible in IPv6 [11, 12]), a subset of targets based on random sampling, or use of domain-based top lists to gather suitable targets. The latter method is used by many Internet measurement research papers, which leverage the Alexa Top 1M [2] and other top lists like Majestic [17] or Umbrella [8] in order to obtain a representative sample of the hosts in the Internet. Previous research has shown that top lists are notoriously unstable [21]. In addition, domain-based top lists lack any form of aggregation in terms of the underlying infrastructure or service, *i.e.*, `google.com` and `google.co.uk` are separate top list entries although providing the same service by the same company. Another issue with current top lists is the lack of weighting for each entry, *i.e.*, it is not clear how much more popular entry #11 is compared to entry #12 [3].

In this paper we propose a new technique to associate IP prefixes with a weight which represents the importance of each prefix.

As the name suggests, prefix top lists consist of network prefixes instead of domains. This allows using prefix top lists in measurement studies which go beyond the scope of domain-based top lists. We create prefix top lists by using domain-based top lists as the initial data source, mapping the domains to IP addresses and subsequently aggregating these IP addresses to prefixes. We apply a Zipf popularity distribution to each domain of the top lists to provide weights for each address, subsequently aggregating the weights for each prefix, and finally ranking the prefixes.

Prefix top lists can be leveraged in many different types of Internet measurement studies, *e.g.*, finding the top CDN networks, evaluating the Internet's infrastructure, or analyzing dependencies of important core Internet routers. In this paper we present an analysis of the DNS name server deployment using prefix top lists. To further foster the use of prefix top lists we make them available to fellow researchers at https://prefixtoplists.net.in.tum.de/.

Outline: The paper is structured as follows: In Section 2 we present our idea of creating prefix top lists and provide details on the process involved. We analyze churn and temporal stability of prefix top lists and compare them to domain-based top lists in Section 3. In Section 4 we apply prefix top lists to DNS analysis by evaluating DNS server characteristics based on their weight in prefix top lists. Section 5 lays out our key findings and discusses their implications. We compare our study to related work in Section 6 and conclude our paper in Section 7.

Rank	Domain	Weight	Top Rank	Bottom Rank
1	google.com	0.0703	1	1
2	youtube.com	0.0351	2	2
3	tmall.com	0.0226	3	4
4	baidu.com	0.0184	3	4
5	qq.com	0.0134	5	6
6	sohu.com	0.0120	5	8
7	facebook.com	0.0099	7	8
8	taobao.com	0.0092	6	9
9	login.tmall.com	0.0077	8	10
10	wikipedia.org	0.0069	9	11

Table 1: Top 10 Domains for August 1, 2019 with 7-day rolling window based on the Alexa List.

2 PREFIX TOP LIST

In this section we present our approach of prefix top lists. First, we elaborate on existing domain top lists. Then we provide details of our DNS address resolution approach. Finally, we present a top list aggregation based on IP prefixes.

2.1 Domain Top List

Existing top lists rank their entries based on some proprietary metric. The applied process and raw data sources are not publicly available. They also exhibit—often weekly—fluctuations [21].

Previous work [15] suggests merging multiple existing top lists to create more stable top lists. This can be done by averaging each top list over multiple days, creating a more stable version of the same list. As popularity in the Internet has been known to follow a Zipf distribution [1, 6, 13, 14], differently ranked input lists can be combined by aggregating their Zipf weight. The Zipf weight w for rank k in a total of N elements is calculated as follows:

$$w = \frac{1/k^s}{\sum_{n=1}^{N} 1/n^s}$$

The parameter s determines the slope of the distribution, larger values for s increase the weight for the top-ranking elements, while reducing the weight for low-ranking elements. The specific parameters which contribute to the ranking of domains in the underlying lists are not universally available. Based on previous work [15], we use a parameter of $s = 1$.

We extend this approach by not only using the Zipf distribution to assign the rank in the combined top list, but also to transfer the weight to the resulting top list for further analysis. The different input top lists are joined based on domain names, and the average weight is calculated. If a domain name does not appear in an input list, it is implicitly assigned a weight of 0.

Our Zipf-weighted input list is created by averaging the input top lists' Zipf weights over the previous week. This approach was chosen to mitigate weekly fluctuations.

This weighted list is called a domain top list. The domain top list for August 1, 2019, based on the Alexa list, is shown in Table 1. We analyze the stability of those lists in Section 3.

2.2 Address Resolution

In order to create prefix and AS based top lists, the domain names need to be resolved.

For collecting a rich DNS data set as part of this process, we deploy a custom DNS resolver. This full resolver discovers the zone setup by help of QNAME minimization (cf. RFC 7816 [7]). To get a complete picture of the DNS deployment, our full resolver queries all available authoritative name servers for each query. The name servers for each zone are discovered by (1) name in the delegation from the parent zone, (2) trustworthy (in bailiwick) glue records in the delegation, and (3) additional `NS` records in the zone apex (also called root domain or naked domain). The individual name server names are resolved by the same process. The executed queries against the authoritative name servers and the individual zone setups are saved including their metadata.

Compared to traditional resolvers, this allows for in-depth investigation of the zone setup and name servers.

All domains of the daily domain top list are resolved every day in a randomized order. For fault isolation, the resolution process is split into multiple shards. The domains are resolved over the course of 24 hours in order to avoid high load spikes on shared authoritative name servers.

The following shards are scanned and later merged into a single data set:

Umbrella t_0: the Umbrella top list for today

Alexa $t_0 \cup$ Majestic t_0: the deduplicated union of the Alexa and Majestic top list for today

Backfill: $\bigcup_{n=-1}^{-6}$ Alexa t_n, Majestic t_n, Umbrella $t_n \setminus$ (Alexa $t_0 \cup$ Majestic $t_0 \cup$ Umbrella t_0): the set of domains which were contained within the top lists over the last week but are currently no longer included in the top lists.

After the zones and domain names have been resolved, the IP addresses are extracted. If a domain points to multiple IP addresses, we extract all of them. This can occur for two reasons: An authoritative name server can return multiple IP addresses as target or different authoritative name server can return distinct IP addresses. For the rest of this work, these addresses are treated equally. Since the domain lists are only resolved from a single vantage point, this resolution process represents a local view and is susceptible to DNS-based load balancing. We discuss DNS-based load balancing and its effects on prefix top lists in Section 5.

2.3 Prefix-based Top Lists

For the generation of prefix-based top lists, different aggregation levels can be chosen: ASes, BGP announced prefixes, and normalized prefixes. Normalized prefixes are /24 prefixes (/48 for IPv6 resp.), as they represent the smallest generally propagated BGP announcements. The aggregation levels are hierarchical, with each normalized prefix belonging to one BGP announced prefix, and each BGP announced prefix belonging to one AS. For the assignment of BGP prefixes and their respective origin AS, a localized BGP dump is used. When the prefix has multiple origin ASes, we take the lower AS number from the possible paths.

For the aggregation and ranking on different levels, the weight of an input domain is transferred based on the resolved mappings. In case multiple objects are reached (i.e., a domain resolves to

AS	Rank BGP	Norm	Object	Weight	Domains	IPs
1			AS15169 – GOOGLE	0.1454	103264	15278
	1	1	172.217.18.0/24	0.0178	1039	35
	2	2	172.217.16.0/24	0.0175	1000	33
	3	3	172.217.22.0/24	0.0173	1041	42
	4		216.58.206.0/23	0.0165	973	35
		6	216.58.207.0/24	0.0151	726	21
	5	5	172.217.23.0/24	0.0164	775	23
	7	7	216.58.208.0/24	0.0154	443	14
	10		216.58.204.0/23	0.0098	547	15
		10	216.58.205.0/24	0.0097	547	15
2			AS13335 – CLOUDFLARE	0.1049	310574	91869
	28		23.227.38.0/23	0.0045	42809	13
		25	23.227.38.0/24	0.0045	42809	13
3			AS16509 – AMAZON-02	0.0651	88373	70888
	20		99.84.88.0/21	0.0057	11988	168
		18	99.84.92.0/24	0.0054	11951	132
4			AS37963 – CNNIC-ALIBABA	0.0478	7266	6733
	6		140.205.64.0/18	0.0160	6	4
		4	140.205.94.0/24	0.0159	3	2
	9		140.205.128.0/18	0.0116	12	12
		9	140.205.130.0/24	0.0113	1	1
5			AS54113 – FASTLY	0.0284	16752	887
	16		151.101.0.0/22	0.0063	4566	192
		30	151.101.1.0/24	0.0032	2895	76
6			AS14618 – AMAZON-AES	0.0248	46028	33443
	175		54.234.0.0/15	0.0008	434	295
		127	54.235.145.0/24	0.0006	2	2
7			AS23724 – CHINANET	0.0238	866	821
	14		220.181.32.0/19	0.0092	9	5
		13	220.181.38.0/24	0.0092	7	3
8			AS4837 – CHINA169	0.0150	457	451
	8		111.160.0.0/13	0.0134	3	4
		8	111.161.64.0/24	0.0134	1	2
9			AS8075 – MICROSOFT	0.0141	21685	10035
	38		13.64.0.0/11	0.0034	3941	1992
		59	13.82.28.0/24	0.0016	10	1
10			AS9808 – CMNET-GD	0.0138	81	77
	12		39.156.0.0/17	0.0096	7	3
		14	39.156.69.0/24	0.0092	6	1

Table 2: IPv4 Object Ranking for August 1, 2019 based on Alexa List. If both BGP- and Norm-Rank are provided the weight is given for the BGP prefix. The weight for the normalized prefix can differ due to varying splitting of the domain weights if multiple objects are reached.

multiple prefixes), the input weight is split between the targeted prefixes. This serves to keep the sum of weights at one (100%). An alternative would be to copy the weight to all prefixes. This variant is susceptible to manipulation: a domain can point to a multitude of prefixes and each gets the full zone weight. High ranking zones with fewer addresses would have less influence on the prefix top list. To remediate these problems, we decided to split the weight.

We generate our prefix top list by aggregating IP addresses to prefixes with the Zipf weights of a domain top list.

The rankings for the top 10 ASes, BGP prefixes, and normalized prefixes are shown in Table 2. For each object, we include the top ranked children, if applicable. The /24 norm-prefixes are weighted and ranked by summing up the weight of all domains pointing to them. Norm-prefixes are assigned to their AS by mapping them to announced BGP prefixes. The AS weight (and therefore its rank) is the sum of the BGP prefix weights. For example, the domain *tmall.com* transfers its weight of 0.0226 from Table 1 to its A-records which are in the prefixes 140.205.94.0/24 and 140.205.130.0/24.

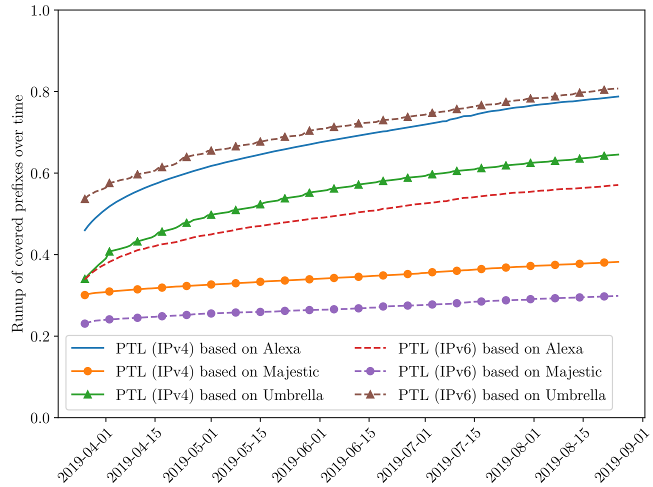

Figure 1: Coverage runup of discovered prefixes by prefix top list (PTL) per IP version over time.

The resulting rankings are our Prefix Top Lists, which present a novel view on the importance of IP prefixes in the Internet.

The ranking does not only export the weight, but also exports how many domains point into the respective object, as well as how many distinct IP addresses of those objects are referenced. Notably, some ASes have many domains pointing to them (*e.g.*, *Cloudflare*). Other ASes also have a high rank but host only few comparatively high-ranked domains (*e.g.*, *CNNIC-ALIBABA*).

3 ANALYZING PREFIX TOP LISTS

In this section we analyze prefix top lists based on BGP announced prefixes (1) by evaluating the increase caused by daily new prefixes added to prefix top lists and (2) by comparing their stability over time to regular domain-based top lists.

3.1 Daily New Prefixes

We first evaluate prefixes which have not yet been seen before for each prefix top list. This analysis shows how much data needs to be collected over time to achieve good prefix coverage.

In Figure 1 we depict the incremental coverage of generated prefix top lists for each day. We find that all prefix top lists contribute multiple ten-thousands of IPv4 prefixes and multiple thousands of IPv6 prefixes on the first day. The share of prefixes seen on the first day compared to all prefixes seen over the measurement period, differs by source. We find that Alexa provides the best coverage of IPv4 prefixes, whereas Umbrella covers most IPv6 prefixes. Interestingly, Majestic has the lowest coverage in IPv4 as well as IPv6 and sees almost no increase over time. New prefixes are still added even after five month of measurements (*e.g.*, $\approx$ 200 for Alexa IPv4 and Umbrella IPv4). These new trends are side-effects of the highly volatile nature of domain-based top lists [21], the initial source of prefix top lists.

When looking at Figure 2 we see that the new prefixes added after the first day stem with few exceptions, such as *wikipedia.org* on June 27, 2019, from low-ranked domains. The figure shows the

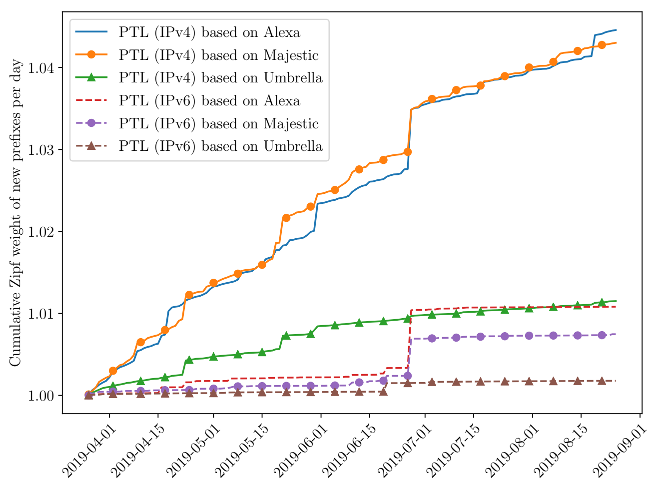

Figure 2: Cumulative Zipf weight of new prefixes per prefix top list (PTL) over time. The jump on June 27, 2019 is caused by previously unseen prefixes for *Wikipedia*.

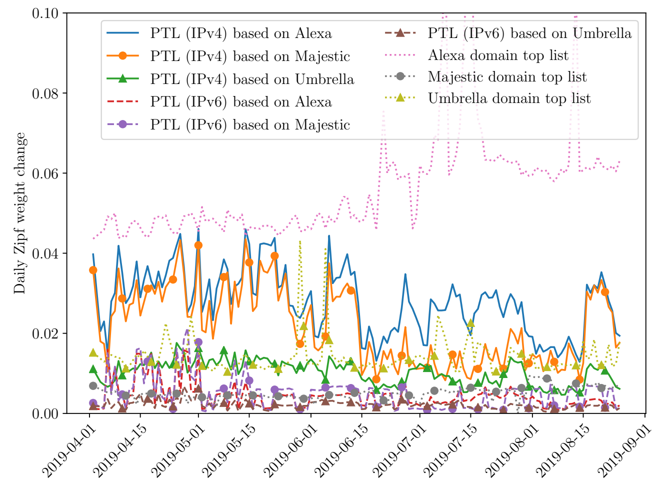

Figure 3: Daily Zipf weight changes for domain and prefix top lists (PTLs) over time.

empirical CDF of the sum of Zipf weights added by new prefixes over time, *i.e.*, it displays how important the newly added prefixes are. Newly added prefixes generally host low-ranked domains as is evident by the Zipf weight increase of less than 0.014 over more than one month of measurements and 0.045 in five months. We also find that new prefixes in the Umbrella-based prefix top lists are about four times less important in terms of aggregated Zipf weight over time compared to Alexa and Majestic.

We conclude that prefix top lists see hundreds to thousands of new prefixes being added even after multiple weeks. These prefixes, however, stem from low-ranked domains, while high-ranked prefixes are added on day one.

3.2 Temporal Stability

Next, we assess the stability of different domain and prefix top lists over time. To account for weekly patterns, which have been shown to be present in top lists [21], we aggregate domain and prefix top lists in a seven-day rolling window. We then calculate the daily Zipf weight change for each domain and prefix, respectively.

Figure 3 shows the daily Zipf weight change during our four-month measurement period. To better illustrate the magnitude of changes, we lay out the following example: Assume, on Monday the complete Zipf weight of 1.0 is given to prefix A, whereas on Tuesday the complete Zipf weight of 1.0 is given to prefix B. This constitutes a change of -1 for prefix A and a change on +1 for prefix B. In this example the daily Zipf weight change between Monday and Tuesday is 2, which is the maximum possible Zipf change per day.

As seen in Figure 3, Zipf weight changes are much more subtle in reality. None of the domain and prefix top lists sees consistent Zipf weight changes above 0.1. The least stable top list is the domain-based Alexa top list, which confirms findings in related work [21]. The most stable top list is the IPv6 prefix top list based on Umbrella. We find that IPv6 prefix top lists are especially stable. The Alexa and Majestic-based prefix top lists are comparatively volatile. This is likely due to high-weight domains performing DNS load balancing, which leads to Zipf weights being shifted to different prefixes.

To summarize, Zipf weight changes in all top lists are much smoother compared to raw domain-based top lists, with less than 0.1 of weight changes. The Alexa domain top list is the least stable, whereas IPv6 prefix top lists exhibit the greatest stability. The Alexa domain top list became even more unstable over time, especially since it often includes less than 1M domains. Starting from July 14, 2019, the list contains only about 600k to 950k domains. We assume, that this is likely due to Alexa updating its generation method and dropping the tail of domains with unreliable ranks. The lower number of domains is automatically accounted for in our Zipf weighting. The sum of all weights remains 1, therefore the importance of domains is dynamically adjusted to the number of provided domains. Nevertheless, adding or deleting up to 300k domains from one day to another results in the discovered artifacts.

4 NAME SERVER ANALYSIS

In this section we analyze the name server records collected by our DNS scanner. The evaluation is on data from August 1, 2019.

RFC 2182 requires that each zone has topologically different name server IP addresses to ensure resilience against routing issues [10]. This means the addresses should be at least in two different normalized prefixes for each used IP version. We use the results of our DNS scanner to determine if the zones of the top list domains comply with this requirement. In Figure 4 and Figure 5 we sort the domains according to their top list rank and the rank of the prefix they are in. In order to make the graphs comparable the x-axis represents the share of ranks per top list (see Table 3 for total numbers).

Figure 4 shows the domain based top list results. The relative graphs for all top lists and both IP versions are similar. Their distribution of non-compliant zones is nearly linear. Thus, compliance to the RFC requirement seems to be independent from the top list rank. We evaluate the Top 1k zones, which are not visible in the graph as they represent less than 0.1% of all zones. In this analysis we

	Alexa		Majestic		Umbrella	
	IPv4	IPv6	IPv4	IPv6	IPv4	IPv6
Domains	2.9M	1.4M	1.0M	505k	327k	206k
Non-compliant	411k	610k	121k	250k	26k	81k
Share	14.26%	43.54%	12.08%	49.56%	8.03%	39.52%

Table 3: Non-compliance of domains per top list.

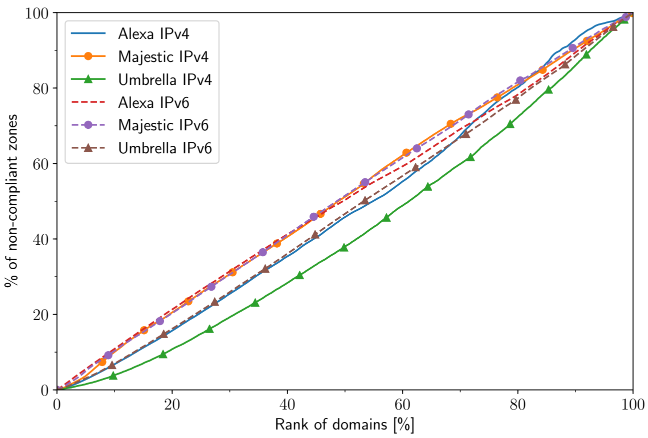

Figure 4: Distribution of non-compliant zones ordered by their rank. The x-axis is normalized to the number of domains per top list. The absolute number for the axes can be found in Table 3.

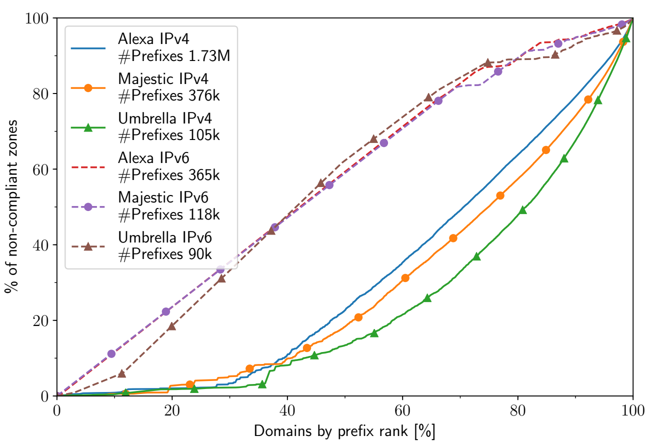

Figure 5: Non-compliant zones ranked by the prefix they are in. Similar to Figure 4 the x-axis also got normalized to the number of successful resolved domains.

find that 2.3% of the Alexa Top 1k are violating zones. The Majestic and Umbrella evaluation show similar results. For domains beyond rank 1000 we observe a steep increasing share of non-compliant domains, until the stable value becomes visible in the figure. Smaller variations in the lower ranks may be caused by clustering effects in top lists described by Rweyemamu *et al.* [20].

In total, we find 12% of all IPv4 zones not respecting the RFC requirement. This aligns with findings by Allman [4]. For domains with IPv6 name servers, about half of all entries do not comply with the RFC requirement (see Table 3). We find that even the top Alexa entries *google.com* and *youtube.com* violate the RFC requirement for IPv6. In comparison, for IPv4 the first offender is on rank 89 (*w3schools.com*) followed by rank 95 (*onlinesbi.com*) and rank 111 (*ettoday.net*).

Even though Google only hosts a small share of domains, the importance of these domains corresponds to a large Zipf weight. Cloudflare on the other side hosts more than 300k zones of the Alexa list. Our results show that Cloudflare uses a single /48 for their IPv6 name servers. United Internet, again one of the top five providers, also uses a single /48 for their IPv6 name servers. These three providers are among the most important ones in our prefix top list. The fact that even these are non-compliant indicates a general problem with the topological diversity of configured IPv6 enabled name servers.

Figure 5 shows the domains sorted by the rank of the highest-ranking normalized prefix. In comparison to Figure 4, the IPv4 prefix sorted domains provide a smaller share on the top ranks. While the domain top list ranked violations grow linearly from the start, the prefix sorted ones start with a near linear growth only after about 30%. The short and steep increase for Umbrella at about 35% is caused by the domain parking service *Bodis*. We assess the DNS availability of parked domains as not critical. Parked domains by this provider also appear in the Alexa and Majestic list but on lower ranks. This is due to Umbrella reflecting the Cisco OpenDNS query count while the other use a more complex metric.

This analysis shows that the topological diversity of configured name servers supporting IPv6 is in a worse state compared to IPv4. Nearly half of all zones have name servers in only one /48 prefix. With this analysis we also provide a metric to compare different top lists according to the deployed configurations. When evaluating the RFC requirement on the stricter BGP prefixes we found similar results as with normalized prefixes. We provide a list of prefixes with the number of non-compliant zones to interested researchers. This list can be used for prefix prioritization or comparisons.

5 DISCUSSION AND LIMITATIONS

In this section we present the key results of our prefix top list approach and discuss implications, limitations, detail our guiding ethical considerations, and elaborate on data publication.

Prefix Top List stability and global view: As described in Section 2 we provide a local view with our prefix top list. Furthermore, as shown in Section 3, the prefix top lists are influenced by fluctuations in domain-based top lists as well as DNS load balancing. To further improve our resulting prefix top list in terms of stability and global applicability, we propose to perform distributed DNS measurements and repeatedly resolve popular domains to cover a larger share of their prefixes.

DNS-based load balancing: In order to gauge the impact of geo-load balancing on the prefix top list, we performed limited measurements using commercial virtual machines in Newark and Singapore. The measurements on these vantage points include only the top 65k domains from each of the three domain-based top lists. Analyzing the Alexa-based top list, we find that the additional vantage points were able to resolve only about 90% of domains compared to our local vantage point. Further analysis is needed to understand this divergence. Within the successfully resolved domains, 7% resolve to different IPv4 addresses on at least one vantage point. When aggregated to the normalized prefix level, the smallest object in the prefix top list, only 5% of domains differ in their result.

In order to rate the impact of these differences, we recalculate the AS-, BGP- and normalized prefix ranking by integrating the additional addresses. As mentioned in Section 2.3, the base weight for a domain name is shared between the resolved prefixes. We compare the rankings by the sum of the absolute differences between the two variants (local view, vs.semi-global view), similar to Section 3.2, *i.e.*, 0 indicates no changes, while 2 stands for two completely different lists. For the AS ranking we observe a Zipf weight change of 0.0133. With respect to the BGP- and normalized prefix based top lists we observe a change of 0.2244 and 0.2266, respectively. The observed change for the normalized prefixes is only slightly larger than for the BGP prefix. This can indicate that address changes for load balancing purposes target different BGP announcements as well. The low change rate with respect to the AS ranking indicates that the load balanced domains are, however, served from within the same AS. Those are strong indicators for DNS-based geo-load balancing.

The impact of DNS geo-load balancing is clearly visible, and it needs to be taken into account when using this prefix top list. The two additional measurements of a subset of domains provide a lower bound estimation to what impact needs to be expected. To this goal, the measurements have to be scaled up to continuously cover the entire domain lists. Moreover, additional vantage points help to further investigate the impact of geo-load balancing.

Localized BGP dump: From the BGP dump we extract the origin prefix and AS number. The path is not relevant for our results. We compared the results of our BGP dump to archives of the Route Views project [18]. Only about 1% of the Top 1k IPv4 BGP prefix top list entries changes. These changes are substitutions of larger prefixes with smaller ones. In IPv6 this effect is more prevalent as the address space is larger, thus there are more possibilities for more specific prefix announcements. These larger prefixes in our top list have a minor impact on the total list. In the future, we plan to merge different BGP dumps using only the longest originating prefixes across all of them.

Web-based top lists: In order to improve the prefix top list generation, the shortcomings of the generation of proprietary top lists have to be considered. The Alexa and Majestic top lists provide a web-centered view on the Internet, which does not necessarily represent the importance of generic Internet infrastructure. A prefix top list based on the Umbrella top list, which is generated from OpenDNS statistics also provides a biased view created by OpenDNS users. The top lists do not include the *www* prefix by default. The *www* subdomain DNS records can resolve to different addresses. We plan to perform such measurements in order to quantify the impact on the prefix top lists.

DNS resilience: As shown by Allman [4] and confirmed by our measurements in Section 4, the resilience of the DNS ecosystem in terms of topologically distributed name servers warrants improvements. Especially on the IPv6 side, half of all zones with IPv6 name server addresses have less than two distributed name server prefixes—contrary to what is mandated by best current practices [10]. Even the most popular domains such as *google.com* and *youtube.com* are affected. For clients with IPv6-only resolvers, this results in a decreased resilience compared to their IPv4 peers.

Ethical considerations: For our active measurements we incorporate the proposals by Allman and Paxson [5], Partridge and Allman [19], and Dittrich *et al.* [9]. As we limit our query rate and use conforming packets, we conclude that it is unlikely that our measurements will cause problems on target systems. During our measurement period, we received no abuse emails or complaints.

Data publication: We provide our prefix top list for fellow researchers on https://prefixtoplists.net.in.tum.de/.

6 RELATED WORK

In the following section we present work in the fields of top lists and DNS analysis which is most related to ours.

Top List analysis: In 2018, Scheitle *et al.* [21] published a thorough analysis of different domain-based top lists. They analyze structure and stability of top lists where the authors find low intersections between lists and daily domain churn of up to 50%. In addition, they also show that measurement studies using top lists depend on the top lists themselves and on the dates of the used top lists. Taking these findings into account, we rank domain-based top lists using a Zipf distribution and aggregate them based on prefixes, which results in more representative and stable prefix top lists.

Le Pochat *et al.* [15] confirmed many of Scheitle *et al.*'s findings. With Tranco [16] they provide an aggregated combination of other top lists over time. In addition to their work we apply a Zipf distribution to rank domains based on popularity and aggregate prefixes to create prefix top lists.

DNS analysis: Allman [4] analyzes DNS robustness over a period of nine years. He finds a downwards trend of second level domains with fewer than two name servers in distinct /24 prefixes in his zone data analysis. In 2018 he finds that 11% of second level domains violate RFC robustness requirements. With 12% we find a similar share of domains in top lists not complying with the robustness requirements. Our data also shows that the domains on the top ranks perform better.

7 CONCLUSION

With prefix top lists we proposed a Zipf-weighted aggregation of popular and distributed services (such as CDNs) into a single metric. We showed temporal stability and found less than 0.05 weight changes per day, improving on the Alexa top list. When analyzing name server resilience, we found top domains to be better configured. Finally, IPv6 resilience still lags behind its IPv4 counterpart.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers and our shepherd Priya Mahadevan for their valuable feedback. This work was partially funded by the German Federal Ministry of Education and Research under the projects X-CHECK, grant 16KIS0530, and the German-French Academy for the Industry of the Future.

REFERENCES

[1] Lada A Adamic and Bernardo A Huberman. 2002. Zipf’s law and the Internet. *Glottometrics* 3, 1 (2002), 143–150.

[2] Alexa. May 13, 2019. Top 1M sites. https://www.alexa.com/topsites. http://s3.dualstack.us-east-1.amazonaws.com/alexa-static/top-1m.csv.zip.

[3] Alexa. May 13, 2019. What’s going on with my Alexa Rank? https://support.alexa.com/hc/en-us/articles/200449614.

[4] Mark Allman. 2018. Comments On DNS Robustness. In *Proceedings of the Internet Measurement Conference 2018*. ACM. https://doi.org/10.1145/3278532.3278541

[5] Mark Allman and Vern Paxson. 2007. Issues and Etiquette Concerning Use of Shared Measurement Data. In *Proceedings of the Internet Measurement Conference 2007*. ACM. https://doi.org/10.1145/1298306.1298327

[6] Tim Berners-Lee. 1998. The Fractal nature of the Web. http://edshare.soton.ac.uk/392/3/DesignIssues/Fractal.html.

[7] Stéphane Bortzmeyer. 2016. DNS Query Name Minimisation to Improve Privacy. RFC 7816 (Experimental). https://doi.org/10.17487/RFC7816

[8] Cisco. May 13, 2019. Umbrella Top 1M List. https://umbrella.cisco.com/blog/blog/2016/12/14/cisco-umbrella-1-million/.

[9] David Dittrich, Erin Kenneally, et al. 2012. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research. *US Department of Homeland Security* (2012).

[10] Robert Elz, Randy Bush, Scott Bradner, and Michael Patton. 1997. Selection and Operation of Secondary DNS Servers. RFC 2182 (Best Current Practice). https://doi.org/10.17487/RFC2182

[11] Oliver Gasser, Quirin Scheitle, Pawel Foremski, Qasim Lone, Maciej Korczynski, Stephen D. Strowes, Luuk Hendriks, and Georg Carle. 2018. Clusters in the Expanse: Understanding and Unbiasing IPv6 Hitlists. In *Proceedings of the Internet Measurement Conference 2018*. ACM. https://doi.org/10.1145/3278532.3278564

[12] Oliver Gasser, Quirin Scheitle, Sebastian Gebhard, and Georg Carle. 2016. Scanning the IPv6 Internet: Towards a Comprehensive Hitlist. In *Proceedings of the Traffic Monitoring and Analysis Workshop 2016*.

[13] Jeremy Kepner, Kenjiro Cho, and KC Claffy. 2019. New Phenomena in Large-Scale Internet Traffic. arXiv:cs.NI/1904.04396

[14] Serge A Krashakov, Anton B Teslyuk, and Lev N Shchur. 2006. On the universality of rank distributions of website popularity. *Computer Networks* 50, 11 (2006), 1769–1780.

[15] Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczyński, and Wouter Joosen. 2019. Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. In *Proceedings of the Network and Distributed System Security Symposium 2019*. Internet Society.

[16] Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczyński, and Wouter Joosen. May 13, 2019. Tranco List. https://tranco-list.eu/.

[17] Majestic. May 13, 2019. The Majestic Million. https://majestic.com/reports/majestic-million/.

[18] University of Oregon. 2019. Route Views Project. http://www.routeviews.org

[19] Craig Partridge and Mark Allman. 2016. Ethical Considerations in Network Measurement Papers. *Commun. ACM* (2016). https://doi.org/10.1145/2896816

[20] Walter Rweyemamu, Tobias Lauinger, Christo Wilson, William Robertson, and Engin Kirda. 2019. Clustering and the Weekend Effect: Recommendations for the Use of Top Domain Lists in Security Research. In *Proceedings of the Passive and Active Measurement Conference 2019*.

[21] Quirin Scheitle, Oliver Hohlfeld, Julien Gamba, Jonas Jelten, Torsten Zimmermann, Stephen D. Strowes, and Narseo Vallina-Rodriguez. 2018. A Long Way to the Top: Significance, Structure, and Stability of Internet Top Lists. In *Proceedings of the Internet Measurement Conference 2018*. ACM. https://doi.org/10.1145/3278532.3278574