This article has been accepted for publication in IEEE Transactions on Network and Service Management. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2024.3470989

Metis: Selecting Diverse Atlas Vantage Points

Malte Tashiro, Emile Aben, Romain Fontugne

***Abstract*—The popularity of the RIPE Atlas measurement platform comes primarily from its openness and unprecedented scale. The platform provides users with over ten thousand vantage points, called probes, and is usually considered as giving a reasonably faithful view of the Internet. A good use of Atlas, however, requires a clear understanding of its limitations and bias. In this work we highlight the influence of probe locations on Atlas measurements and advocate the importance of selecting a diverse set of probes for fair measurements. We propose Metis, a data-driven probe selection method, that picks a diverse set of probes based on topological properties (e.g., round-trip time or AS-path length). Using real experiments we show that, compared to Atlas' default probe selection, Metis' probe selections collect more comprehensive measurement results in terms of geographical, topological, RIR, and industry-type coverage. Metis triples the number of probes from the underrepresented AFRINIC and LACNIC regions, and improves geographical diversity by increasing the number of unique countries included in the probe set by up to 59%. In addition, we extend Metis to identify locations on the Internet where new probes would be the most beneficial for improving Atlas' footprint. Finally, we present a website where we publish periodically updated results and provide easy integration of Metis' selections with Atlas.**

***Index Terms*—Bias, Internet measurements, Open Access, RIPE Atlas, vantage point selection**

I. Introduction

SINCE the early days of the Internet, the research community has developed tools to monitor the Internet, and the RIPE Atlas measurement platform [1] became a popular choice for Internet-wide measurements. The scale of Atlas is one of its strengths, with over ten thousand vantage points (VPs) it enables a myriad of ways to study the Internet. These studies all start by creating a new measurement, which mainly consists of specifying the type of measurement (e.g., traceroute), a target (e.g., hostname), and a set of VPs, called *probes* in the Atlas terminology. A user can only assign up to one thousand probes per measurement in order to prevent harmful use and to ensure fair resource sharing. Depending on the goal of the study, users may ask Atlas to randomly select probes. In particular, broad probe sets can be selected with the "Worldwide" area option. Random selection is handy and popular for Internet-wide [2] and regional analysis [3], [4], but due to the Atlas bias towards certain countries and autonomous systems (ASes) [5]–[8] some studies need to normalize collected data [9] or design their own VP selection procedure [10].

Manuscript received 25 October 2023; revised 1 April 2024 and 19 September 2024; accepted 23 September 2024. Date of publication ?? 2024; date of current version ?? 2024. *(Corresponding author: Malte Tashiro.)*

Malte Tashiro and Romain Fontugne are with the IIJ Research Laboratory, Chiyoda, Tokyo 102-0071, Japan (e-mail: malte@iij.ad.jp).

Emile Aben is with RIPE NCC, Amsterdam, 1012 AB, Netherlands.

Digital Object Identifier ?.?/TNSM.2024.?

Figure 1a shows the ten countries hosting the largest number of probes. Germany and the United States each represent over 13 % of all probes. The top six countries host 50 % of probes and we found 26 countries with only one probe. This uneven distribution of probes is certainly a concern when selecting probes for measurements. It also raises the question of the utility of Atlas' random probe selection, and consequently, the need for a better selection mechanism for wide scale studies. Simple mechanisms, like limiting the number of probes taken from each country or AS, could provide better geographical balance or AS diversity but are not directly addressing Atlas' inherent topological imbalance. In order to make an informed decision about probe similarities we need to inspect empirical data that embeds the probes' topological characteristics.

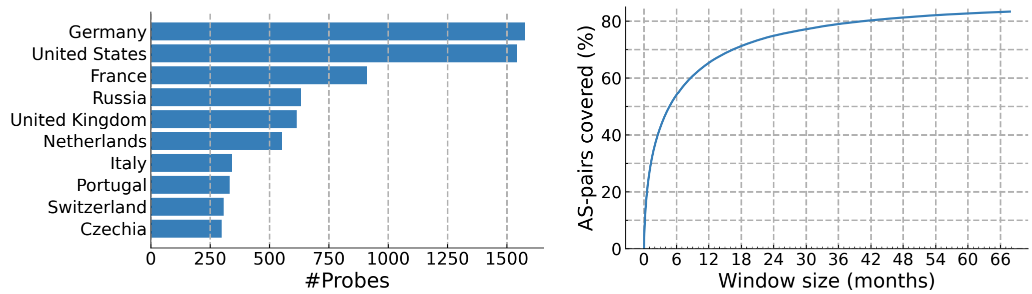
Goal In this paper we aim to assist researchers in selecting a diverse set of probes for Internet-wide studies. We argue that a random selection may lead to wrong inferences and measurement budget waste. The problem is not caused by random sampling techniques, as these are effective methods to reduce a dataset while preserving its main characteristics [11], [12]. The issue comes primarily from Atlas' deployment bias, which is still present in a random set of probes.

Overview In order to make the best use of Atlas we propose Metis, a data-driven method to select a set of diverse probes. Because the definition of probe diversity may vary from one study to another, we propose a general approach designed around a user-given distance metric. Using that metric Metis computes a distance matrix for Atlas probes and iteratively discards the probe closest to all others. This sifting process retains a set of probes that are spread out in the given space and the process can be stopped as soon as the desired number of probes is met. Computing the probe distance matrix is eased by the availability of Atlas' built-in topology measurements, which are an attempt to daily traceroute all globally routable prefixes. We leverage this dataset to obtain distances between probes at no additional measurement cost.

To evaluate the benefits of Metis we compare measurement results obtained with Atlas' default selection to results obtained with various Metis selections. For this comparison we experiment with three common distance metrics (round-trip time, AS-path length, and IP hops) and show that they all outperform Atlas' default selection in terms of geographical, topological, and regional Internet registry (RIR) coverage. Although Metis does not take probe geolocation and RIR data into account, it naturally picks probes in up to 59 % more countries and triples the number of probes from regions that are underrepresented in Atlas (i.e., AFRINIC and LACNIC).

An intuitive extension of this work is the possibility to identify locations for deploying new probes that would help to

This article has been accepted for publication in IEEE Transactions on Network and Service Management. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2024.3470989



(a) The ten countries hosting the most Atlas probes

(b) Probe-AS pair coverage for historical data.

Figure 1. (a) Atlas probes are not evenly distributed, showing a clear concentration in Germany and the United States. (b) historical topology measurement data does not suffice to reach full probe-AS coverage: the entire dataset contains traceroutes between only 83 % of probe ASes.

diversify Atlas' footprint. We also explore this idea and identify the regions and ASes that would be the most beneficial for mitigating Atlas' deployment bias.

Contributions In summary, our main contributions are:

- An updated report on RIPE Atlas deployment bias.
- Metis: a data-driven method for selecting diverse probes.
- An evaluation of Metis with real experiments and comparisons to Atlas' default probe selection.
- A deterministic method to identify candidate locations for new Atlas probe deployment.
- The release of probe selection and probe deployment results [13].

This article is an extended version of [14] which was published at TMA 2022. New contributions of the extended work include:

- Additional analysis of the probe selections which includes the industry types represented by the different sets.
- Two additional evaluation use cases targeting DNS-based load balancing and latency of anycast targets.
- An interactive website on which we provide continuous updates to the dataset as well as visualizations, historical data, and tools for easy use of the Metis selections with RIPE Atlas.
- Improved explanations and clarifications in parts of the text based on user and operator feedback, including a discussion of traceroute symmetry.

II. Methodology

Metis consists of three steps:

1) Probes are mapped to a user-defined space which is represented by a distance matrix (Section II-A).
2) The distance matrix is iteratively sifted to uncover a diverse probe set (Section II-B).
3) The sifting process stops when there is a suitable number of remaining probes (Section II-C).

A. Probe Distance Matrix

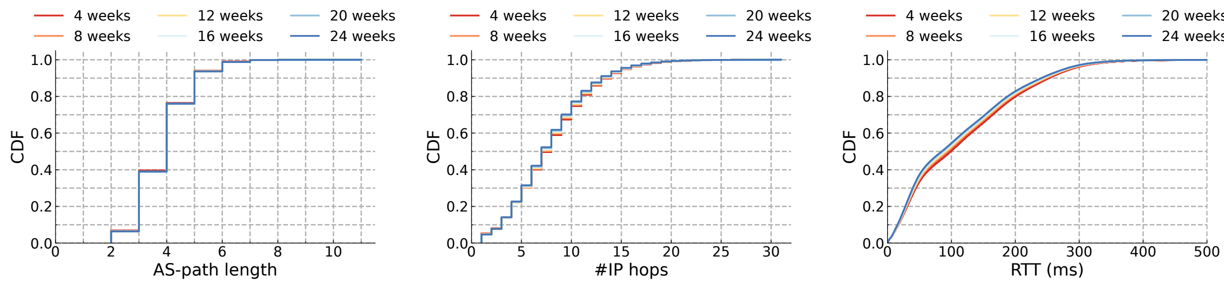
A probe distance matrix represents the distance between probes. To ease computation and data collection we group probes at AS granularity. Thus, a row or column in the distance matrix represents an AS hosting at least one probe. At the time of writing, the Atlas probes cover 3607 ASes for IPv4 and 1657 ASes for IPv6. Building a complete distance matrix for all these ASes requires over 13 million pairwise measurements. We avoid this burden by recycling data collected by Atlas' built-in topology measurements [15].

1) Topology Measurements: The goal of the topology measurements is to reveal a large fraction of Internet paths by running traceroutes from all probes to all globally reachable IP prefixes. However, a single Atlas measurement by design can only target a single destination (IP or hostname). Therefore, to assign multiple targets, Atlas uses a special hostname that resolves to a different IP address for each DNS query, which is the .1 address from a randomly selected IP prefix in a global routing table. Every probe resolves this hostname and performs a traceroute every 15 minutes, resulting in over 2.2 (1) million results for IPv4 (IPv6) per day. From this dataset we extract all traceroutes between probe ASes.

2) Distance Metric: To illustrate the flexibility of Metis and explore different notions of probe diversity, we retrieve three distances from the above traceroute data: round-trip time (RTT), number of IP hops, and AS-path length.

Extracting RTT values and the number of IP hops from traceroute is trivial. However, converting traceroutes to AS paths requires a more thorough process. We use a combination of Route Views [16] and PeeringDB [17] data to map IPs to ASes, as done in previous work [18]–[20]. If an IP address can not be mapped, the hop is ignored, which is a known shortcoming of using traceroute to infer AS paths [21]–[23], but has limited impact on our metric. In our dataset, the IP-to-AS mapping fails for about 4 % of hops. Since traceroute sends more than one packet per hop, it is possible to receive replies from different IP addresses. If these IP addresses map to different ASes, they are included as an AS set. After mapping all hops we remove duplicate ASes, keeping only the first occurrence of each AS in the path. Finally, we only include traceroutes that reached the AS of the intended target. We do not require the traceroute to reach the intended target IP, since there is no guarantee that the addresses used by the topology measurements are indeed responsive.

3) Time Window: Next we have to determine the time window required for extracting relevant topology measurement traceroutes. Since the topology measurements target the entire reachable IP space at random, only a small fraction of the



(a) AS-path length (b) IP hops (c) RTT

Figure 2. Comparing the value distributions of different time window sizes reveals that there is no conceivable difference for (a) AS-path length, and only minor differences for (b) IP hops and (c) RTT.

traceroutes ends up in target probe ASes. It is therefore necessary to choose a suitable time window that includes traceroutes between as many probe-AS pairs as possible while minimizing the risk of including topological changes that might skew computed distances.

To deal with this trade-off we inspect the probe-AS pair coverage achieved with different time windows. An AS pair (A, B) is covered if there exists a traceroute from AS A to AS B. Figure 1b shows the evolution of coverage for different window sizes. All windows start at 2021-01-01 and the window size increases in monthly increments all the way back to the beginning of the topology measurements in 2016. We assume that the distance metric between two ASes is symmetric, i.e., a traceroute in one direction suffices to satisfy coverage for both directions. However, even with this assumption only a maximum of 83 % of probe-AS pairs can be covered. This is because of both, the high number of ASes that contain only one probe, and the prevalence of ASes with few prefixes. Indeed, the randomized address mechanism make it unlikely that these ASes will target, or will be targeted by, all other probe ASes.

A four-week window results in 25 % coverage. Doubling this coverage requires a five times larger window (20 weeks), which increases the risk of including topological changes and stale data. However, we find that our metrics do not significantly benefit from a larger window and we can sacrifice coverage to minimize the impact of topological changes.

To highlight this, we use 24 weeks ($\approx$ 6 months) of traceroute data from August to December 2021 and window sizes ranging from 4 to 24 weeks in four-week increments. In addition, we verify the consistency of data over time by computing distance values for each time window shifted in one-week increments over the entire dataset. This process results in 21 shifts for the four-week window and no shifts for the 24-week window, since it already covers the entire timespan. Each shift covers a different part of the data and each window size covers a different amount (e.g., 16.7 % for the four-week window). With this comparison, we found that the start and size of the window have little influence on the resulting value distribution. A longer window provides more coverage, up to 50.72 % in case of the 24-week window, whereas the shorter four-week windows cover 25.35 % of probe-AS pairs on average. However, looking at the value distributions of the distances for all window sizes in Figure 2 reveals that there is no significant difference. Each plot in Figure 2 contains one CDF per shift (i.e., 21 CDFs for the four-week window), with different colors to separate window sizes. There is no conceivable change in terms of AS-path lengths (Figure 2a) and only a slight, but negligible, difference in case of IP hops (Figure 2b) and RTT (Figure 2c).

The set of included probe ASes is also nearly independent of the chosen window size for our considered range. The smallest overlap we observe between any four-week window and the 24-week window is still 99%, i.e., even in short windows the same probe ASes are present.

We therefore employ a four-week window for our experiments as it minimizes the risk of including topological changes while still resulting in a decent coverage and representative distance distributions.

4) Building the Matrix: We can now compute distance matrices from selected traceroutes. A distance matrix M is a $m \times m$ matrix, where m is the number of probe ASes. An entry $M_{x,y}$ represents the distance from AS x to AS y, where x and y are indices, not AS numbers. The matrix is initially filled with placeholders, indicating the absence of values. We then iterate over the data window and fill the matrix by extracting distances from traceroute results. To increase the amount of filled entries in the matrix for AS-path length and IP hops, sub-paths of each traceroute are also included. For example, a traceroute $A \rightarrow B \rightarrow C$ not only results in distances for $A \rightarrow B$ and $A \rightarrow C$, but also $B \rightarrow C$. For RTT, this is not possible as traceroute's RTT values are always bound to the source probe AS. If we obtain multiple distance values for the same probe-AS pair, we include only the lowest one in the distance matrix. Therefore, the distance matrices represent the shortest distance observed between probe ASes.

Once the entire data window has been processed, we symmetrize the matrix by mirroring the contents of cells $M_{x,y}$ and $M_{y,x}$. If both cells already contain values, the smaller value is selected. Traceroutes are generally not symmetric at the IP level. However, relaxed distance metrics like number of IP hops and AS-path length, which ignore the explicit IPs and order of hops, are less impacted by this simplification. We call these distance metrics *symmetric* if they are equidistant in both directions or *asymmetric* if one direction has a different distance value. We assess the amount of asymmetry for these metrics empirically at the end of this section.

In a final step, we remove rows and columns that have no

This article has been accepted for publication in IEEE Transactions on Network and Service Management. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2024.3470989

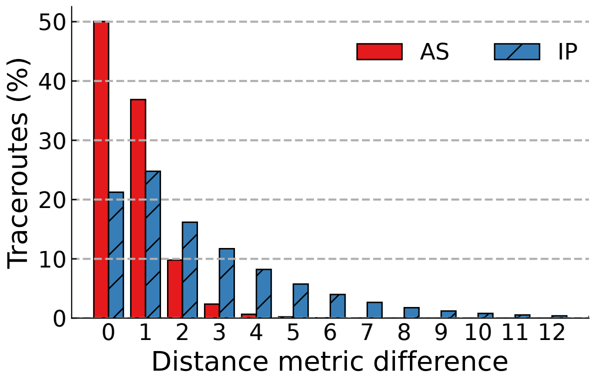


Figure 3. Distribution of distance metric differences for traceroutes between anchors. A difference larger than zero indicates asymmetry.

value. This can happen if there is no valid traceroute for a probe AS within the data window.

Using a four-week window this process starts with 3607 (1657) probe ASes for IPv4 (IPv6) and finishes with 3550 (1609) ASes, resulting in a coverage of 25.62 % (20.54 %) of AS pairs.

Distance Metric Symmetry We investigate the symmetry of traceroutes in terms of number of IP hops and AS-path length by looking at the anchor mesh measurements[1] from RIPE Atlas. These are measurements between a specific set of probes, called anchors, that perform a full mesh of traceroutes between each other every 15 minutes. This gives us a large body of bidirectional traceroutes. We analyze 24 hours of data containing measurements for 772 anchors resulting in 294k anchor pairs and 28m traceroutes. The traceroutes are processed in pairs such that a traceroute in one direction is matched with the counterpart that has the closest temporal proximity, allowing a maximum difference of up to 15 minutes. This excludes around 10k traceroutes that have no close partner.

We first inspect the symmetry in terms of number of IP hops and, as expected, it is not very high: Only 15 % of anchor pairs are symmetric during the entire data interval. The remaining anchor pairs show at least one case of asymmetry within 24 hours. We can investigate these cases further by looking at the metric difference of the traceroutes as shown in Figure 3. Around 21 % of traceroutes are symmetric, indicating that there are anchor pairs that had both a symmetric and asymmetric distance at different times in the interval. Almost 25 % show a metric difference of 1 hop hinting at a slight asymmetry. In total, 46 % of traceroutes have no or slight asymmetry, however most higher metric differences are also located at the lower end of the spectrum, with 90 % having a difference of 5 hops or fewer.

Using the more relaxed AS-path length metric results in better symmetry. A third of the anchor pairs and 50 % of all traceroutes are symmetric. In total, a majority of 87 % of traceroutes are symmetric or only have slight asymmetry (see Figure 3). In light of these results we acknowledge that assuming symmetry for the IP hops metric is not entirely accurate, but argue that the simplification works well for AS-path length, which thus should be the preferred metric.

[1]https://atlas.ripe.net/about/anchors/

B. Sifting

The next step is to select probe ASes that are spread out according to the distance matrix. We perform this filtering process by iteratively removing the probe AS closest to all others still remaining in the matrix. The iteration stops once a user-defined number of ASes is left. Closeness is defined by aggregating all distance values of an AS into a single value. However, since the computed matrices are sparse, the number of distance values may vary drastically for each AS. In preliminary experiments we found that traditional aggregators like average and median are not suitable, mainly because they may be skewed by outliers or produce incomparable aggregates for ASes with a very different number of distances.

We design our own aggregation function aiming to (1) emphasize ASes that are very close to each other but not favor ASes that are far from only a few ASes, (2) normalize distance vectors, and (3) produce single comparable values. For (1), we employ inverse distance values, thus highlighting small distances and reducing variance for large ones. Then for (2), we normalize distance vectors based on their value distribution. Finally (3), we sum up the normalized inverse distances to obtain an overall closeness score per probe AS.

Formally, let D be a vector of discrete distances from/to a probe AS (RTT values are rounded to milliseconds) and D_u the set of *unique* distance values. For each unique value $d \in D_u$ we compute s_d, the inverse distance weighted by d's relative frequency in D:

$$s_d = \frac{1}{d} \cdot \frac{c_d}{|D|} \tag{1}$$

where c_d is the number of times d occurs in D.

The final closeness score S for an AS is the sum of the weighted inverse distances:

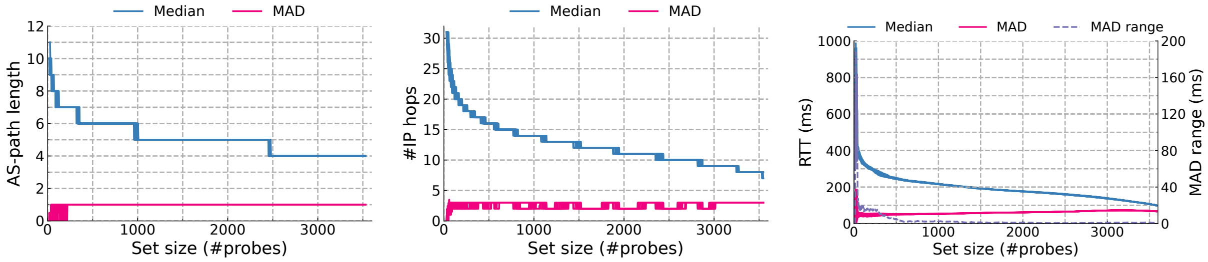
$$S = \sum_{d \in D_u} s_d \tag{2}$$

A small score implies that the AS is far away to most ASes. An AS with many small distances results in a high score. Since each iteration removes one AS from the matrix and thus shortens the length of D, the score of all remaining ASes is recalculated in each iteration.

To test the robustness of this score function we evaluate how sensitive it is to noise compared to simply computing the mean value of D. We build a distance matrix using a four-week window and the AS-path length metric and compute the removal order with both the presented function S and the mean. We then add noise in the form of outliers, in this case high AS-path length values, at random cells in the matrix and recompute the orders. To quantify the robustness of each function to the added noise, we employ the average overlap (AO; also called intersection metric [24]) to compare the noisy orders to the original one of the respective function.

AO measures the overlap between increasing proportions of two orders. Let N and R be two orders and $N_{..d}$ denote the first d elements of N, then $|N_{..d} \cap R_{..d}|/d$ is the overlap

This article has been accepted for publication in IEEE Transactions on Network and Service Management. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2024.3470989



(a) AS-path length (b) IP hops (c) RTT

Figure 4. The median and MAD evolution of all 21 data window shifts for increasing set sizes. The distance between probes decreases for all metrics with a larger set size. The MAD range of the (c) RTT for the different window shifts stabilizes for sets with more than 500 probes.

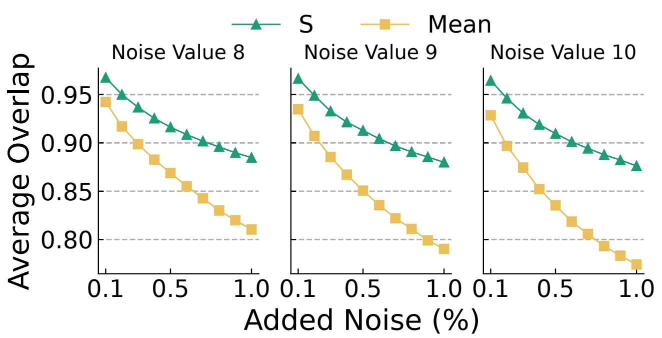


Figure 5. Impact of noise on the removal order for the presented score function (S) and the mean. The mean is more susceptible to noise resulting in a stronger degree of reordering.

of N and R for depth d. AO(N, R) is the average over all $d \in \{1...k\}$ where k is the total length of the order.

The impact of varying amounts of noise (0.1% to 1%) and levels (AS-path length value of 8 to 10) is highlighted in Figure 5. We repeat the process 10 times per amount and level of noise. Thus, each point represents the average value of the 10 repetitions, and we observe no significant spread ($max(\sigma) = 0.003$). A higher value indicates less reordering, where a value of 1 indicates equality of the compared orders. We see that S is more robust to noise as indicated by higher values. In addition, it is less impacted by the strength of the noise since the slope of the curves stays almost the same for different noise levels. We therefore conclude that the proposed score function is less impacted by unexpected outlier values than simpler approaches.

C. Choosing a Suitable Set Size

The final step is to determine a suitable stopping criterion for the above sifting procedure. We face another trade-off where a large probe set may inherit the characteristics of the total set, including its bias towards specific regions. A small set may result in probes that are spread out, but might completely miss desirable regions.

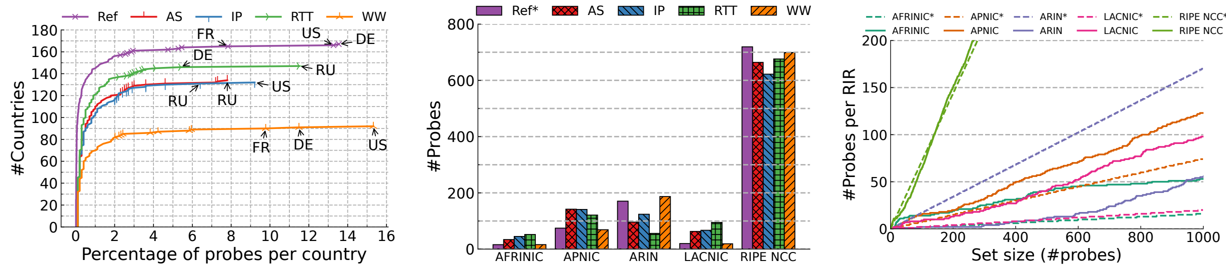
To better understand the impact of the size of the probe set, we compare the distributions of the distance matrices obtained with different probe set sizes. We use two metrics to characterize the distributions: the median and the median absolute deviation (MAD). The median indicates the center of a distribution, and the MAD quantifies its dispersion.

We compare distributions from all 21 shifted four-week data windows computed in Section II-A3 and plot the median and MAD values for different set sizes in Figure 4. The evolution of the median for all three distances confirms the effectiveness of our sifting procedure — a larger set includes more ASes that are closer together, therefore the median distance decreases. Due to the discrete nature of the AS-path length (Figure 4a) and IP hops (Figure 4b), the MAD for these metrics is rather flat and becomes unstable below 250 probes. We therefore focus on the RTT distributions (Figure 4c) to decide on a value, for which we also show the range of MAD values for the different shifts on a secondary y-axis. The MAD of the RTT stabilizes at a set size of about 500 probes and above. Therefore, for our experiments we should employ a probe set size greater than 500, and we decide to use a set size equal to the limit given by Atlas, one thousand probes, which is also a practical value for Atlas users.

III. Evaluation

We evaluate the benefits of Metis by comparing it to Atlas' "Worldwide" area selection (WW). We selected 1000 probes with each selection method (WW and Metis with the three different distance metrics) and ran traceroute measurements towards 25 targets distributed all over the world. In order to reduce dependency on the RIPE ecosystem, we choose M-Lab [25] servers as measurement targets. We pick five targets within five different regions (Africa, Asia/Oceania, Europe, North America, South America). The decision maximizes geographical distance, i.e., we use the locations of all servers within a region and select the set of five locations that offers the largest sum of pairwise distances between each other. Since multiple servers can be present at a single geographical location, we use the AS number as a tiebreaker. This process results in 25 servers from 23 distinct ASes — both the North and the South American regions contain two servers from the same AS — for IPv4. There are only two geographically distinct locations available in Africa for IPv6, resulting in 22 servers from 19 ASes — three servers in North America are in the same AS and one AS is present in both Europe and South America. Due to Atlas' daily credit limit, the measurements are spread out over four days, one day per selection method. As part of this extended version we performed two additional traceroute

This article has been accepted for publication in IEEE Transactions on Network and Service Management. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2024.3470989



(a) Percentage of IPv4 probes per country (b) Number of IPv4 probes per RIR (c) RIR distribution using the RTT selection

Figure 6. Selecting IPv4 probes with a distance metric (a) increases the number of included countries and (b) provides a better balance between RIRs. (c) the RTT selection increases the representation of AFRINIC, APNIC, and LACNIC for smaller set sizes as well, compared to the scaled Atlas distribution.

measurements targeting hostnames that use DNS-based load balancing and assessing the latency of anycast targets.

In the following sections we first inspect the selected probe sets and their characteristics, both in terms of geographical and topological diversity. Then we analyze the differences in traceroute results, highlighting the topological coverage and increased variety of the observed paths. Finally, we highlight the usefulness of Metis for other use cases, showing that more hosts are discovered in DNS and a wider variety of RTT values is seen for the anycast targets.

A. Probe Selection

We first take a look at the geographical distribution of the selected probes. Figure 6a shows the cumulative percentage of probes per country where the maximum values stand for the total number of countries represented by each probe set. The plot is essentially a CDF with an absolute y-axis, where the y-value indicates the number of countries and the corresponding x-value shows the percentage of probes located in each country. We plot markers for x-values larger than 2 % to highlight the points in the tails. For reference, we also show *Ref*, the distribution of all 11 655 active IPv4 probes. All probes combined cover 168 countries, with Germany and the United States containing 13.5 % and 13.2 % of probes respectively. *WW* covers only 93 countries, the United States account for 15.3 % of selected probes and the top three countries (United States, Germany, France) sum up to 36.6 %. In contrast, all Metis selections provide better geographical coverage, even though probe locations are not taken into account by Metis. The best-performing metric in terms of country distribution is *RTT*, which is expected since RTT can be an indicator for physical distance. The RTT selection improves the country coverage by 57.4 % to a total of 148 countries. The outlier of the RTT selection is Russia with 11.5 % of probes. The other two distance metrics improve probe distribution across countries but at the expense of a lower number of total countries. For example, using the AS-path length selection Russia accounts for 7.8 % of probes, but only 135 countries are covered in total.

A different way of comparing probe sets is by looking at their distribution across the RIRs. Figure 6b shows the number of probes per RIR for each selection. *Ref** refers to the overall Atlas distribution, proportionally scaled down to a set of 1000 probes. Since WW is a random[2] sample from all Atlas probes, the WW and Ref* distributions are almost identical. Both show a strong focus on the RIPE and ARIN regions, whereas the number of probes for AFRINIC and LACNIC are very low. In fact, AFRINIC is only represented by 16 probes and LACNIC by 19 probes in WW. Using the RTT selection, RIPE and ARIN probes are substituted by probes in other regions, effectively tripling the number of probes in these regions, and also increasing the representation of APNIC by 38 %. Despite these improvements, RIPE is still prevalent in all selections. This can be attributed to the large number of Atlas probes in the RIPE region. But a uniform distribution of probes over the different RIRs is also not expected. As 39 % of active ASes are indeed assigned in the RIPE region [26] we still expect probe sets representing a global view of the Internet to have a large proportion of RIPE probes.

To further investigate how the RIR distribution evolves for smaller set sizes, Figure 6c illustrates the RIR distribution of the RTT selection for different set sizes. The dashed lines show the proportionally scaled Atlas distribution for each RIR, and the solid lines show the distributions provided by the RTT selection. Even for smaller set sizes, the RIPE region is prevalent. While the Atlas distribution strongly favors ARIN, and almost ignores AFRINIC and LACNIC (two dashed lines close to the x-axis), the RTT selection keeps the representation of AFRINIC, APNIC, and LACNIC close together up until a set size of 500, before AFRINIC reaches a plateau and the share of ARIN probes increases. This supports the fact that Metis better balances between regions as it favors probe diversity as opposed to Atlas' inherent probe distribution.

In addition to the geographical distribution we also inspect the topological distribution of the probes by looking at the number of probe ASes contained in each set of selected probes. Metis chooses one probe per AS by design, so we focus on the WW selection. The Atlas selection contains probes from 569 unique ASes out of which 20.6 % have more than one probe. The top ten ASes have ten or even more probes. While it may be reasonable to select multiple probes in some large ASes, this is another challenge that raises new questions, such as, how many and which probes to select from which AS.

The reduced topological diversity of the WW selection

[2]We got feedback from Atlas operators that the load of a probe factors into the selection, making it not truly random.

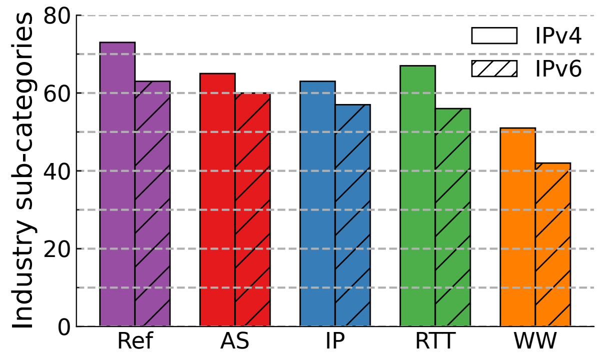


Figure 7. Number of industry sub-categories represented by the selections.

is also apparent when inspecting the usage types of ASes included in the respective sets. Ziv et al. [27] provide a research dataset that maps public ASes to industry types using data from business intelligence databases, website classifiers, and a machine learning algorithm. The assigned types are divided into two layers: The first layer contains broader categories (e.g., "Computer and Information Technology"), whereas the second layer provides more specific sub-categories (e.g., "Internet Service Provider (ISP)"). A single AS can be mapped to up to three industry types and not all ASes have sub-category assignments. There is a total of 17 layer-1 categories that are all present in every selection (including the set of all probes). Comparing the sub-categories, however, reveals that sets selected by Metis are more diverse. There are 85 sub-categories, out of which 73 are present in the Ref dataset. Figure 7 shows the number of sub-categories represented by the different selections. There is a clear distinction between Metis selections (63 to 67 sub-categories) and the default WW selection (51 sub-categories), showing a wider range of AS types in general. These results again confirm the diversity of ASes that are selected by Metis. Although Metis does not take the industry types of ASes into account when selecting probes, selecting ASes that are far apart in the Internet topology helps to diversify the type of ASes that are selected.

B. Measurement Results

We now look at the results of the three example use cases in detail and discuss their characteristics.

1) Traceroute: First, we compare traceroute results collected using the four different probe selections. We look at the topology covered by the paths found in traceroutes and then, based on a few representative results, discuss changes in the observed AS-path lengths, IP hops, and RTT.

We quantify the topological coverage of collected traceroutes by counting the unique number of ASes and IPs in the traceroute results. For fairness and given the lower number of unique probe ASes for WW, this analysis excludes the probe ASes and only focuses on other ASes on the path. Figure 8c shows the average number of unique ASes per selection over all regions for IPv4 on the left. As expected, using the AS-path length as distance metric maximizes the number of visible ASes, hence producing an average increase of 3.84 × compared to WW. The highest increase is in the African region by 4.03 ×, or 542 ASes more than WW, for a total of 729 (up from 187) unique ASes visible on the paths. Interestingly, the IP-hops and RTT selections also provide a substantial increase of 3.1 × and 2.8 × over WW.

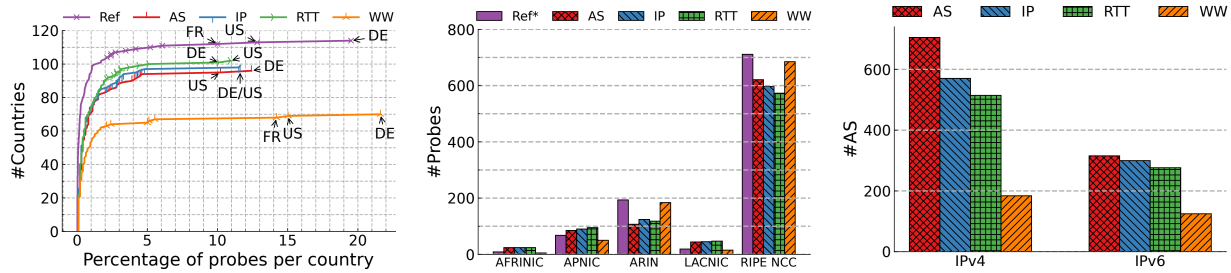
The improved topological coverage is also apparent in the number of unique IPs seen on the paths. Although the differences are less pronounced compared to the number of ASes, using the IP-hops selection produces an average increase of 23.22 % over WW. The maximum increase is seen in Europe with 24.97 %, or 1319 IPs more than WW, for a total of 6586 (up from 5269) unique IPs visible. Again, the other distance metrics are also providing substantial coverage improvements, namely, 18.37 % and 14.66 % for the AS-path length and RTT selections respectively.

Next, we analyze the distributions of AS-path length, IP hops, and RTT of traceroutes towards the M-Lab servers. Due to space constraints we present only results for servers in the North American region (Figure 9) but the same conclusions are drawn for other regions. Figure 9a depicts the AS-path length distribution for all traceroutes towards the five targets in North America. Most notably, all Metis selections increase the overall AS-path lengths. The high probe concentration of WW in the North American region results in very short paths. We observe only 2 ASes for 34 % of paths (i.e., the probe and target AS) and 75 % have length 3 or less. All Metis selections manage to reduce the fraction of very short paths. The AS-path length selection results in only 5 % of AS-path with length equal to 2. In addition, the AS-path length selection increases the average path length to 4.5, which is almost equal to the global average of 4.4 as observed in BGP data [26]. In contrast, the WW selection provides an average AS-path length of 3.5, falling one hop short. For all regions, the observed average AS-path length of the Metis selections is consistently closer to the global average than the WW selection, which is an evidence that paths collected with Metis better resemble the global Internet.

For the RTT distribution (Figure 9c), comparing the median values, WW is the lowest with 133 ms, the RTT selection is the highest with 159 ms, closely followed by IP hops and the AS-path length selections 152 ms and 151 ms. Overall, the Metis selections produce up to 10 % less low RTTs (less than 100 ms) which is expected given that RTT is tightly related to geographical distances, that intercontinental RTTs are usually over 100 ms [7], [28], and that these experiments focus on worldwide views. The IP-hops selection slightly increases the number of observed IP hops (Figure 9b), as expected, although there is no significant change for this metric. In summary, using specific distance metrics enables a better coverage of the topology, as seen by increased AS- and IP-path lengths, and a better geographical coverage as shown by the RTT increase as well as probes' country and RIR distribution. In addition, the flexibility of Metis allows the user to explore different distance metrics that best suit their use case.

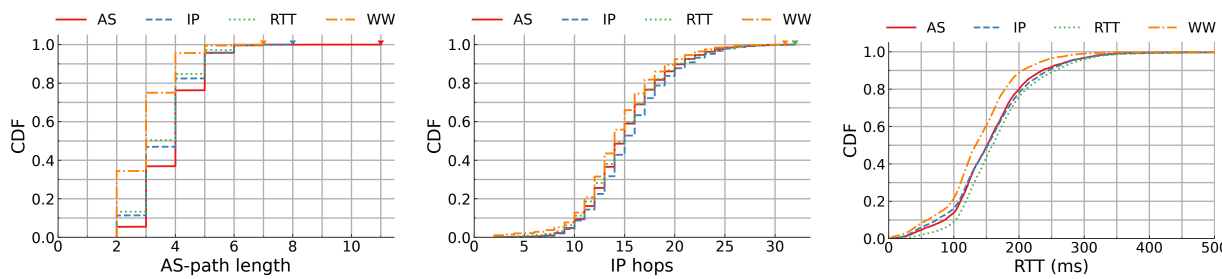
2) DNS: The second measurement targets hostnames that use DNS-based load balancing, a practice that assigns multiple A records (domain to IP mappings) to a single domain name. We perform traceroute measurements and use the local DNS

This article has been accepted for publication in IEEE Transactions on Network and Service Management. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2024.3470989



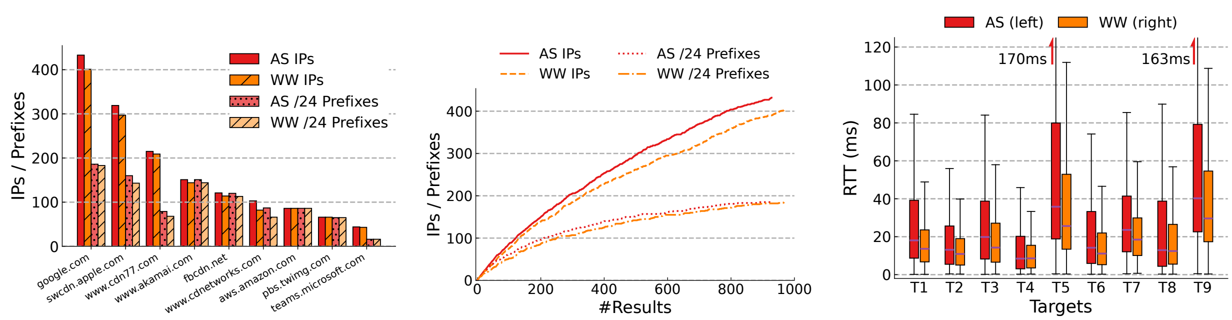
(a) Percentage of IPv6 probes per country (b) Number of IPv6 probes per RIR (c) Average number of unique path ASes

Figure 8. All distance selections (a) increase the number of countries represented by IPv6 probes and (b) better balance RIR distribution compared to the default Atlas selection. Metis' selections increase the topology covered by the traceroutes as indicated by (c) the number of unique ASes visible on the paths.



(a) AS-path length (b) IP hops (c) RTT

Figure 9. IPv4 measurement results for North America show that distance selections can (a) reduce the topological clustering between probes and targets and increase distance in terms of (b) IP hops and (c) RTT.



(a) Resolved IPs and /24 prefixes (b) Resolved IPs and /24 prefixes for different set sizes; google.com (c) RTT distribution for anycast targets

Figure 10. Metis discovers (a) more hosts and prefixes for DNS-based load balanced services and (b) the benefit increases with larger set sizes. For the latency evaluation of anycast targets (c) Metis provides more diverse results, preventing overly optimistic conclusions. The key for (c) is shown in Table I.

Table I
KEY FOR ANYCAST TARGETS OF FIGURE 10C.

#	Organization	ASN	IP
T1	Automattic (WordPress)	2635	192.0.78.23
T2	Google	15169	216.239.32.117
T3	OpenDNS	36692	146.112.41.2
T4	Cloudflare	13335	172.64.151.101
T5	Akamai	21342	184.85.248.128
T6	Fastly	54113	151.101.130.137
T7	Microsoft	58068	13.107.6.156
T8	Packet Clearing House	42	194.0.17.1
T9	OVHcloud	16276	198.27.92.14

resolver of the probes to resolve nine diverse domain names that use DNS-based load balancing.

Figure 10a shows the number of resolved IPs and /24 prefixes for the AS-path length and WW selections. Metis achieves equal or better variety in all cases, discovering up to 32 IPs (google.com) and 21 prefixes (www.cdnetworks.com) more than WW. The higher prefix variety in particular highlights the diversity of the Metis selection, since prefixes are usually distributed in the topology.

To assess the benefit of Metis for smaller set sizes, we plot the evolution of resolved IPs and prefixes for google.com in Figure 10b. Since some probes fail the DNS resolution, which is a technical problem not in our control, we plot the number of valid results. Metis results are selected in order of diversity, whereas WW results are selected in the order in which they were served by the Atlas API. While Metis resolves more IPs and prefixes even for smaller set sizes, the benefit only increases with more probes. The maximum difference reaches 50 IPs at 670 results and 16 prefixes at 312 results.

3) Anycast: The last measurement assesses the latency of anycast targets. With anycast, a single IP is shared between multiple hosts and the destination of a packet is the host topologically closest to the sender. We selected nine IPs that (1) are part of an anycast prefix according to [29], (2) host a popular domain name[3] that is present in the Cisco Umbrella Popularity List [31], and (3) are not in the same AS as other selected IPs. The resulting targets are shown in Table I.

Figure 10c shows the RTT distributions when targeted by the AS-path length and WW selections. Each box represents the interquartile range (IQR), the whiskers span $1.5 \times$IQR, and the purple line indicates the median. Metis measurements show both a wider RTT spread in terms of IQR and variability outside the upper and lower quartiles. In addition, the median of the distributions is noticeably higher for seven targets, indicating that a performance analysis based on WW results alone might lead to an overly optimistic conclusion.

IV. IPv6 Measurements

We also evaluated Metis with IPv6 probes. The total number of IPv6 probes in Atlas is 5502, spread over 1657 ASes and 115 countries. We, again, conducted experiments with selections of 1000 probes. The distribution of probes per country for each selection are analogous to IPv4 (Figure 8a). All Metis selections result in more diverse sets of countries compared to WW, for instance, RTT improves the number of covered countries by 45 % to a total of 103 (up from 71). Furthermore, WW picks over 50 % of probes in only three countries (Germany, United States, France), whereas the RTT selection manages to spread the same fraction of probes over eleven countries.

The RIR distribution is also similar to IPv4, with only subtle differences (Figure 8b). First, the number of probes for AFRINIC and LACNIC is even less, both for the WW and Metis selections. For the default Atlas selection, AFRINIC is only represented by five probes, LACNIC by 15. The RTT selection is able to improve that to 24 probes for AFRINIC and 47 probes for LACNIC. While these numbers might seem low, they already contain 50 % of all AFRINIC IPv6 probes (48 probes) and 45 % of LACNIC IPv6 probes (104 probes). Another difference to IPv4 is that IPv6 favors more ARIN probes, although RIPE is still prevalent in the sets with roughly two thirds of probes.

The smaller probe pool also has an influence on the number of unique source ASes covered by the WW set. Only 358 distinct ASes are selected, out of which 23.7 % contain more than one probe. In addition, the first ten ASes represent over 40 % of probes.

Finally, there is also a decrease in the number of industry sub-categories represented by all selections (see hatched bars in Figure 7). The total number of sub-categories present in the Ref dataset is reduced to 63 from 73. The WW selection follows this decline by dropping to 42 from 51, whereas Metis selections suffer a smaller impact. Indeed, the AS selection is almost capable of the same representation as the entire Ref probe set, only missing three sub-categories and achieving a total representation of 60 sub-categories.

Looking at the unique ASes visible on the paths in Figure 8c, we observe that all Metis selections are able to at least double the number of ASes compared to WW.

The traceroute results are overall more homogeneous — possibly attributed to the fact that the selection of 1000 probes represents a large share of all available probes. Indeed, since Metis selects one probe per AS and the number of available IPv6 probe ASes is even smaller, the Metis selections share 62.5 % of probe ASes. As a consequence, there are almost no discernible differences between the Metis selections in terms of RTT, and all of them provide a slightly increased median of 275 ms to 278 ms, compared to 254 ms for WW. In terms of AS-path length, all Metis selections feature less short paths and come closer to the average path length observed in global routing [26].

V. RIPE Atlas Probe Placement

The above experiments demonstrate that Metis identifies distant sets of Atlas probes. In this section, we apply the same method to address a related problem, the identification of ASes on the Internet that are distant from Atlas probes. The goal here is different, we are aiming at revealing locations where Atlas probes are needed, thus providing help for a better probe deployment [32].

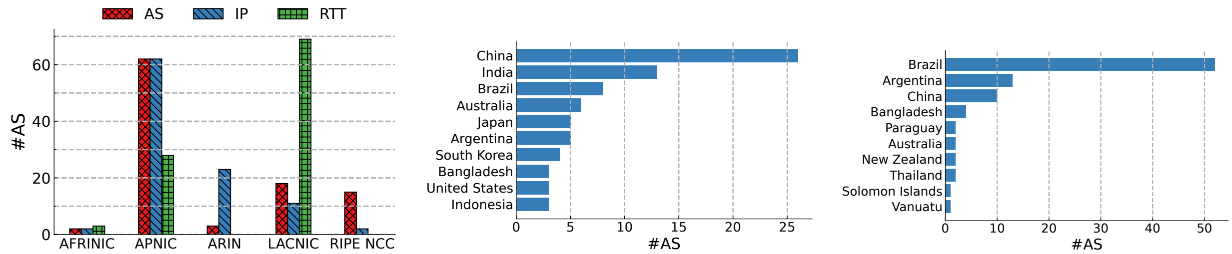
This application requires some adjustments to the distance matrices. First, we compute distance matrices from all results of the topology measurements, including traceroutes towards non-probe ASes. Next, we keep only the discrete distance vectors of non-probe ASes. Each vector represents the distances from multiple probe ASes to the specific non-probe AS. Finally, in order to make relevant recommendations we only consider ASes that were reached by at least 50 % of the probe ASes. The sifting procedure is unchanged but we stop it after the first iteration, then report ASes with a low score.

For this analysis we also tuned the time window differently. Due to the random selection of targets in the topology measurements, ASes with many announced IP prefixes are more likely to appear in the distance matrices. Consequently, for a four-week window the distance matrices contain mainly large non-probe ASes with a median number of prefixes equal to 216. In contrast, using a longer time window of 24 weeks we collect traceroutes from more ASes with fewer prefixes, hence the median number of prefixes for non-probe ASes is reduced to 48. Therefore, in this section we employ a 24-weeks time window, which results in a total of 67 420 non-probe ASes targeted by traceroutes from which 2990 remain after applying the minimum threshold of 1804 distance values (50 % of the 3607 probe ASes). Computed scores are available at [13].

The following analysis focuses on the 100 ASes with the lowest S score (i.e., farther from Atlas), hereafter referred to as recommendations. The distribution of recommendations in terms of RIRs and countries is shown in Figure 11. There is

[3]Domain-IP mapping provided by OpenINTEL [30].

This article has been accepted for publication in IEEE Transactions on Network and Service Management. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2024.3470989



(a) RIR distribution (b) Country distribution; AS-path length (c) Country distribution; RTT

Figure 11. The distribution of the top 100 recommendations in terms of (a) RIR distribution show a clear tendency towards APNIC and LACNIC. The country distribution for the (b) AS-path length selection reveals that the recommendations are mainly in China and India, whereas the (c) RTT selection recommends placement in Brazil.

a noticeable difference in recommendations depending on the distance metric. Especially AS-path length (Figures 11a and 11b) and RTT (Figures 11a and 11c) recommend very different regions. These differences confirm that AS-path length and RTT do not necessarily correlate: An AS may be reached by a long AS path, but with a small RTT. The recommendations in terms of AS-path length are mostly focused on the APNIC region with 62 ASes, followed by LACNIC (18) and RIPE (15). Many of the APNIC ASes are located in China (26) followed by India (13) highlighting the topological distance of these countries from existing probes. The average AS-path length to the first 100 recommendations is 5.6 which is much larger than the average observed in global routing [26].

Using RTT as the distance metric shifts the focus to the LACNIC region (Figure 11a). 69 recommendations are located in LACNIC, followed by 28 in APNIC. Notably, there are no ASes from neither ARIN nor RIPE in the top 100 recommendations. This is somewhat expected as the existing Atlas probes already have a strong presence in these regions and RTT values are related to geographical distances. Further inspection of the LACNIC recommendations shows that the majority of ASes are located in Brazil (52), followed by Argentina (13). Therefore, the South American region, and Brazil in particular, offers the potential of increasing Atlas' RTT diversity. The top 100 recommendations show an average RTT of 244 ms, with a maximum of 319 ms.

Finally, there is a noticeable lack of recommendations in the AFRINIC region. We attribute this to the fact that there are a lot less active ASes in the AFRINIC region (less than 1600 at time of writing, compared to, e.g., almost 11 000 for LACNIC [26]), resulting in a smaller footprint in the topology measurements. However, even without our recommendations the RIPE NCC is actively trying to improve visibility in underrepresented regions via community outreach and ambassador programs. [33]

VI. Community Contributions

In our endeavor to help the networking community to make the best use of the RIPE Atlas measurement platform, we implemented a web application that provides Metis results at

https://metis.ihr.live

This application consists of two webpages, one for selecting a diverse set of Atlas probes, and one for finding ASes where Atlas probe deployment would be the most beneficial. Hence, a user can get Metis results in just a few clicks. And for better integration with existing tools one can also query the data via a REST API (https://ihr.iijlab.net/ihr/en-us/api). This API can also be used to retrieve historical data.

The data is updated every Sunday/Monday night at midnight UTC and is based on the last four weeks of measurement data.

A. Probe selection

The probe selection page provides a simple interface to select a probe set based on the latest topology measurement results. On this page the user can select the number of requested probes, a distance metric, and the IP version to retrieve Metis results.

The results are given as a list of ASes instead of probes, so that Atlas can dynamically select probes that are active and available at the time the user creates a measurement. To help users integrate Metis results into their tools, we provide the API link that can generate the displayed results and an Atlas API specification that the users can utilize when creating new Atlas measurements.

B. Probe deployment

The probe deployment page shows deployment recommendations for the RIPE Atlas measurement platform as described in Section V. These recommendations consist of the top 100 ASes that are the most distant to the existing Atlas probe ASes based on the presented metrics. The web interface displays these results by country and RIR so that one can easily identify regions that need the most attention. Additional recommendations can be retrieved via the API. These results are also updated weekly.

VII. Related Work

The uneven probe distribution of Atlas has been mentioned multiple times in the literature [5]–[7]. Notably, in 2015 18 % of Atlas probes were hosted by only ten ASes [5]. More recently a comparison with a larger measurement platform reveals the lack of Atlas probe in countries of South America, Africa, and Asia [7] which corroborate with our findings.

Sermpezis et al. [8] quantify the multi-dimensional bias in different internet measurement platforms, including RIPE Atlas. They confirm formally what we only show empirically: that the probe distribution of RIPE Atlas contains bias both in the geographical and the topological dimension. In addition, they show that the WW selection method yields a set of probes that has significantly higher bias, even when compared to a random sample. However, they do not propose any method for a better selection.

Aware of these limitations different approaches have been used to work with large-scale Atlas data. Numerous studies derive Internet characteristics by selecting probes per region [3], [4], [34], or using all Atlas probes worldwide [35], [36] and acknowledging for potential bias. A few other studies are carefully selecting probes to achieve probe diversity while maintaining a broad geographic coverage [10], or normalizing data collected with Atlas to address its bias [9]. All these studies could benefit from our probe-selection framework.

Similar to our work, a probe selection based on paths similarities is proposed in [37]. The probe similarity metric of [37] is however much more rigid than our approach using a user-given distance metric since it is based solely on IP paths similarities. In addition, the flexibility of our approach allows us to plan probe deployment which is not possible with the probe similarity metric.

VIII. Conclusion

This paper demonstrates the geographical and topological disparities of Atlas' default probe selection. To mitigate this we develop a flexible distance-based probe selection, Metis, and show its benefits for worldwide probe selection. Overall, Metis enables a better use of Atlas and may become even more valuable in the future as Atlas users have to pick a limited number of probes from an ever-growing set of probes. Although the presented experiments focus on worldwide selections, the proposed approach is also applicable to smaller regions, for example, selecting a set of probes from a single continent or country. But in this case users should adjust the number of selected probes accordingly to the number of available probe in the region.

In addition, Metis provides recommendations for the deployment of new Atlas probes which has rarely been addressed in the literature. We hope this work can initiate more research efforts in this direction hence improving the development of Atlas and related measurement platforms.

We presented the original work at RIPE 85 [38] to gather feedback from both the RIPE community and Atlas operators. It was well received and deemed useful, however there is no official integration planned at this point in time. Since the presentation we also received feedback from operators in the LACNIC region that use our recommendations as one input in their decision process for new probe deployment.

References

[1] RIPE NCC Staff, "RIPE Atlas: A Global Internet Measurement Network," *The Internet Protocol Journal*, vol. 18, no. 3, pp. 2–26, Sep. 2015.

[2] J. Blendin, F. Bendfeldt, I. Poese, B. Koldehofe, and O. Hohlfeld, "Dissecting Apple's Meta-CDN during an iOS Update," in *Internet Measurement Conference (IMC)*, 2018, pp. 408–414.

[3] M. Candela, E. Gregori, V. Luconi, and A. Vecchio, "Using RIPE Atlas for Geolocating IP Infrastructure," *IEEE Access*, vol. 7, pp. 48 816–48 829, 2019.

[4] R. Fanou, P. Francois, and E. Aben, "On the Diversity of Interdomain Routing in Africa," in *Passive and Active Measurement Conference (PAM)*, 2015, pp. 41–54.

[5] V. Bajpai, S. J. Eravuchira, and J. Schönwälder, "Lessons Learned from using the RIPE Atlas Platform for Measurement Research," *ACM SIGCOMM Computer Communication Review*, vol. 45, no. 3, pp. 35–42, Jul. 2015.

[6] V. Bajpai, S. J. Eravuchira, J. Schönwälder, R. Kisteleki, and E. Aben, "Vantage Point Selection for IPv6 Measurements: Benefits and Limitations of RIPE Atlas Tags," in *IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, 2017, pp. 37–44.

[7] T. K. Dang, N. Mohan, L. Corneo, A. Zavodovski, J. Ott, and J. Kangasharju, "Cloudy with a Chance of Short RTTs: Analyzing Cloud Connectivity in the Internet," in *Internet Measurement Conference (IMC)*, 2021, pp. 62–79.

[8] P. Sermpezis, L. Prehn, S. Kostoglou, M. Flores, A. Vakali, and E. Aben, "Bias in Internet Measurement Platforms," in *Network Traffic Measurement and Analysis Conference (TMA)*, 2023, pp. 1–10.

[9] R. Singh, A. Dunna, and P. Gill, "Characterizing the Deployment and Performance of Multi-CDNs," in *Internet Measurement Conference (IMC)*, 2018, pp. 168–174.

[10] L. Davisson, J. Jakovleski, N. Ngo, C. Pham, and J. Sommers, "Reassessing the Constancy of End-to-End Internet Latency," in *Network Traffic Measurement and Analysis Conference (TMA)*, 2021.

[11] A. Pescapé, D. Rossi, D. Tammaro, and S. Valenti, "On the impact of sampling on traffic monitoring and analysis," in *International Teletraffic Congress (ITC)*, 2010, pp. 1–8.

[12] N. Duffield, "Sampling for Passive Internet Measurement: A Review," *Statistical Science*, vol. 19, no. 3, pp. 472–498, Aug. 2004.

[13] Internet Health Report - Metis, 2022. [Online]. Available: https://ihr.iijlab.net/ihr/en-us/metis

[14] M. Appel, E. Aben, and R. Fontugne, "Metis: Better Atlas Vantage Point Selection for Everyone," in *Network Traffic Measurement and Analysis Conference (TMA)*, 2022.

[15] E. Aben, "Measuring More Internet with RIPE Atlas," Jan. 2016. [Online]. Available: https://labs.ripe.net/author/emileaben/measuring-more-internet-with-ripe-atlas/

[16] "University of Oregon Route Views Project." [Online]. Available: http://www.routeviews.org/

[17] "PeeringDB." [Online]. Available: https://www.peeringdb.com/

[18] H. Chang, S. Jamin, and W. Willinger, "Inferring AS-level Internet Topology from Router-Level Path Traces," in *Scalability and Traffic Control in IP Networks*, vol. 4526, 2001.

[19] Y. Hyun, A. Broido, and kc claffy, "Traceroute and BGP AS Path Incongruities," 2003.

[20] G. Nomikos and X. Dimitropoulos, "traIXroute: Detecting IXPs in traceroute paths," in *Passive and Active Measurements Conference (PAM)*, 2016, pp. 346–358.

[21] Y. Hyun, A. Broido, and kc claffy, "On Third-party Addresses in Traceroute Paths," in *Passive and Active Network Measurement Workshop (PAM)*, 2003.

[22] Z. M. Mao, J. Rexford, J. Wang, and R. H. Katz, "Towards an Accurate AS-Level Traceroute Tool," in *Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM)*, 2003, pp. 365–378.

[23] Y. Zhang, R. Oliveira, Y. Wang, S. Su, B. Zhang, J. Bi, H. Zhang, and L. Zhang, "A Framework to Quantify the Pitfalls of Using Traceroute in AS-Level Topology Measurement," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 9, pp. 1822–1836, Sep. 2011.

[24] R. Fagin, R. Kumar, and D. Sivakumar, "Comparing top k lists," *SIAM Journal on Discrete Mathematics*, vol. 17, no. 1, pp. 134–160, 2003.

[25] "M-Lab." [Online]. Available: https://www.measurementlab.net/

[26] P. Smith, "BGP Routing Table Analysis." [Online]. Available: https://thyme.apnic.net/

[27] M. Ziv, L. Izhikevich, K. Ruth, K. Izhikevich, and Z. Durumeric, "ASdb: A System for Classifying Owners of Autonomous Systems," in *Internet Measurement Conference (IMC)*, 2021, pp. 703–719.

[28] R. Fontugne, J. Mazel, and K. Fukuda, "An Empirical Mixture Model for Large-Scale RTT Measurements," in *IEEE Conference on Computer Communications (INFOCOM)*, 2015, pp. 2470–2478.

[29] BGP.Tools, "Anycast Prefixes list as bgp.tools sees it." [Online]. Available: https://github.com/bgptools/anycast-prefixes

[30] R. van Rijswijk-Deij, M. Jonker, A. Sperotto, and A. Pras, "A High-Performance, Scalable Infrastructure for Large-Scale Active DNS Measurements," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 6, pp. 1877–1888, Jun. 2016.

[31] Cisco Umbrella, "Umbrella Popularity List." [Online]. Available: https://s3-us-west-1.amazonaws.com/umbrella-static/index.html

[32] E. Aben, "Route Collection at the RIPE NCC - Where are we and where should we go?" Oct. 2020. [Online]. Available: https://labs.ripe.net/author/emileaben/route-collection-at-the-ripe-ncc-where-are-we-and-where-should-we-go/
[33] L. Hestina, "Help Build a Bigger, Better RIPE Atlas," Oct. 2021. [Online]. Available: https://labs.ripe.net/author/lia_hestina/help-build-a-bigger-better-ripe-atlas/
[34] P. Gigis, V. Kotronis, E. Aben, S. D. Strowes, and X. Dimitropoulos, "Characterizing User-to-User Connectivity with RIPE Atlas," in *Applied Networking Research Workshop (ANRW)*, 2017, pp. 4–6.
[35] R. de Oliveira Schmidt, J. Heidemann, and J. H. Kuipers, "Anycast Latency: How Many Sites Are Enough?" in *Passive and Active Measurement Conference (PAM)*, 2017, pp. 188–200.
[36] R. Fontugne, A. Shah, and K. Cho, "Persistent Last-mile Congestion: Not so Uncommon," in *Internet Measurement Conference (IMC)*, 2020, pp. 420–427.
[37] T. Holterbach, E. Aben, C. Pelsser, R. Bush, and L. Vanbever, "Measurement Vantage Point Selection Using A Similarity Metric," in *Applied Networking Research Workshop (ANRW)*, 2017, pp. 1–3.
[38] M. Tashiro, "Metis: Better Atlas Vantage Point Selection for Everyone," RIPE 85 MAT-WG Session, Oct. 2022. [Online]. Available: https://ripe85.ripe.net/archives/video/938/

Malte Tashiro received the M.Sc. degree in computer science from Saarland University, Saarbrücken, Germany in 2020.

He is a researcher at the IIJ Research Laboratory, Tokyo, Japan and is pursuing his Ph.D. degree at the Graduate University for Advanced Studies. His research interests include network measurement and Internet topology analysis.

Emile Aben received the M.Sc. degree in chemistry from the Radboud University, Nijmegen, Netherlands.

He is a Senior Research Engineer at the RIPE NCC, Amsterdam, Netherlands since 2009. Before that he worked at the Center for Applied Internet Data Analysis, San Diego, CA, USA. His research interests include network measurement, technology changes (IPv6 and RPKI) and network security.

Romain Fontugne received the Ph.D. degree in computer science from the Graduate University for Advanced Studies, Tokyo, Japan in 2011.

He is the deputy director of IIJ Research Laboratory, Tokyo, Japan and an active member of the WIDE Project. His research interests include network measurement, networked data analytics, and network security.