Illuminating Router Vendor Diversity Within Providers and Along Network Paths

Taha Albakour
TU Berlin

Oliver Gasser
Max Planck Institute for Informatics

Robert Beverly
Center for Measurement and Analysis of Network Data

Georgios Smaragdakis
Delft University of Technology

ABSTRACT

The Internet architecture has facilitated a multi-party, distributed, and heterogeneous physical infrastructure where routers from different vendors connect and inter-operate via IP. Such vendor heterogeneity can have important security and policy implications. For example, a security vulnerability may be specific to a particular vendor and implementation, and thus will have a disproportionate impact on particular networks and paths if exploited. From a policy perspective, governments are now explicitly banning particular vendors—or have threatened to do so.

Despite these critical issues, the composition of router vendors across the Internet remains largely opaque. Remotely identifying router vendors is challenging due to their strict security posture, indistinguishability due to code sharing across vendors, and noise due to vendor mergers. We make progress in overcoming these challenges by developing LFP, a tool that improves the coverage, accuracy, and efficiency of router fingerprinting as compared to the current state-of-the-art. We leverage LFP to characterize the degree of router vendor homogeneity within networks and the regional distribution of vendors. We then take a path-centric view and apply LFP to better understand the potential for correlated failures and fate-sharing. Finally, we perform a case study on inter- and intra-United States data paths to explore the feasibility to make vendor-based routing policy decisions, i.e., whether it is possible to avoid a particular vendor given the current infrastructure.

CCS CONCEPTS

• **Networks** → *Network protocols*; **Network management**; • **Security and privacy** → **Network security**.

KEYWORDS

Device Fingerprinting, Network Security, Network Measurement.

ACM Reference Format:
Taha Albakour, Oliver Gasser, Robert Beverly, and Georgios Smaragdakis. 2023. Illuminating Router Vendor Diversity Within Providers and Along Network Paths. In *Proceedings of the 2023 ACM Internet Measurement Conference (IMC '23), October 24–26, 2023, Montreal, QC, Canada.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3618257.3624813

IMC '23, October 24–26, 2023, Montreal, QC, Canada

ACM ISBN 979-8-4007-0382-9/23/10.
https://doi.org/10.1145/3618257.3624813

1 INTRODUCTION

The Internet is exemplified by distributed control, varied policies, and autonomy. This inherent heterogeneity extends to the physical infrastructure. We explore one specific physical property in detail: the composition of Internet router *vendors*. The set of vendors through which data packets traverse end-to-end has both direct and nuanced implications, for instance reliability and fate-sharing, but also security and policy when a vendor is untrusted. Indeed governments have explicitly forbidden particular vendors, while others have threatened to do so [20].

Unfortunately, the distribution and composition of router vendors across the Internet remains largely opaque. Operators consider their network configurations sensitive and proprietary, and do not publicly publish information on vendors. While remote vendor and operating systems fingerprinting is common for end hosts, e.g., Nmap [35], these techniques rely on active services and responsive network stacks. However, routers are typically configured with strict security to block arbitrary requests [16, 28]. Indeed, multiple security incidents have demonstrated the value in router fingerprinting and reconnaissance in mounting attacks [4, 19, 34].

While general purpose remote fingerprinting tools have been used to great effect, e.g., for Internet-wide surveys and scanning [24, 41, 50], these tools are ill-suited to router fingerprinting. Existing techniques are frequently unable to make a reliable vendor inference and typically send a large number of packets making them (i) impractical to operate at a large scale and (ii) intrusive.

We leverage a technique based on SNMPv3 which by itself is highly accurate, but provides poor coverage. Our primary insight is to collect lightweight network and transport layer fingerprints and use this prior technique as a source of ground-truth vendor labels—this allows us to find new feature sets that uniquely identify vendors even when they do not respond to SNMPv3. In this fashion we more than double the coverage compared to the SNMPv3 technique. By tuning our technique to routers, we achieve this coverage while simultaneously improving efficiency by sending approximately two orders of magnitude fewer active probe packets as compared to existing state-of-the-art that sends up to thousands of packets per inference. Our contributions include:

- *Lightweight FingerPrinting (LFP)*, a novel lightweight, accurate, and more complete remote vendor fingerprinting methodology and tool (§3).
- Using SNMPv3 to find sets of TCP/IP stack features for LFP that uniquely identify 82% of ground-truth routers and provide 95% accuracy alone in fingerprinting major router vendors (§4).
- Applying LFP to the widely used RIPE Atlas and CAIDA Internet Topology Data Kit (ITDK) router datasets to classify 64% of the

active router IPs—more than double the coverage as compared to current state-of-the-art (§7).
- An accuracy evaluation of LFP compared to current tools and techniques showing that it is at least as good as Nmap while sending orders of magnitude fewer packets and improving coverage (§7.3).
- Inference of router vendors in more than 6,700 networks, including around 1,800 networks for which no vendor information was available in previous studies (§7.5).
- End-to-end data path-based router vendor analysis and case studies that provide valuable insights for current security and policy-based routing decisions (§6).
- LFP is publicly available, along with the the derived signatures and classification results from this study to enable reproducibility and future work [3].

2 RELATED WORK

Prior work developed passive and active techniques that leverage open ports, identifiers, and implementation-specific differences to fingerprint devices at various granularities. Most of these techniques were developed for generic hosts, while a few attempt to fingerprint routers.

Nmap: Nmap [35] is an open-source network scanning and reconnaissance tool. It performs remote OS fingerprinting by running up to sixteen tests that send ICMP, UDP, and TCP packets with different field values, flags, and options. By examining the responses, Nmap finds the best matching operating system from a database of fingerprints. The latest Nmap version (7.93) contains more than six thousand fingerprints; of these, approximately 160 and 20 correspond to Cisco and Juniper routers, respectively. Two drawbacks of Nmap are its reliance on open ports and the large amount of probe packets needed to perform fingerprinting. We compare our approach against Nmap in (§7.3.1).

Hershel: Hershel [47] is a low-overhead framework that models the problem of single-packet OS fingerprinting and develops novel approaches for tackling delay jitter, packet loss, and user modification to SYN-ACK features. Based on this theory, a classification method is developed that increases the accuracy of single-packet fingerprinting. Censys [13] includes Hershel signatures in recent raw scanning data. We compare LFP to Hershel in (§7.3.2).

Banner Grabbing: A popular technique for remote operating system fingerprinting and vendor information is "banner grabbing," whereby publicly available services leak information. For instance, the Cisco SSH server implementation returns identifying information in its response string. Internet-wide scanning and banner grabbing are performed regularly [13, 18, 19, 27, 28]. In a recent paper [26], the authors utilize banners augmented with active measurements to perform large-scale network equipment vendor classification. Similar techniques can be applied in passive measurements as well for automatic traffic classification [25]. Unfortunately, banner analysis requires an open remote service that returns this discriminating information. Routers are frequently tightly secured and unresponsive to banner queries. Moreover, banner datasets are typically proprietary or commercial, with some offering free academic licenses [13].

TCP Stack Fingerprinting: Many TCP stack variables, e.g., Window Size and Maximum Segment Size are implementation-specific [8, 33, 39]. These variables can differ between operating systems and versions. Consequently, TCP features can form a unique signature that can be used for fingerprinting. For instance, the initial TCP SYN-ACK packet provides some valuable information about a target's TCP stack characteristics such as the initial Time to Live (TTL) value, sequence number, and window size. When combined with the behavior of the re-transmission timeout of the SYN-ACK packets it was shown to serve as a fingerprinting technique to 25 different operating systems [51], and in another work this was extended to cover more than 90 OSes [31].

Sundials: Sundials [45] uses ICMP timestamps for fingerprinting purposes. Even though NTP has replaced ICMP timestamps, approximately 15% of 14.5M IP addresses in this study responded to ICMP timestamp requests. Sundials uses the variety of response behaviors as a new fingerprinting technique. However, given filtering and the relative lack of ICMP timestamp support among routers, this method has limited coverage for our fingerprinting purposes.

IPID-based Fingerprinting: The IP identifier (IPID) is a mandatory IPv4 header field used for fragmentation and reassembly. Thus, it is frequently possible to elicit an IPID value from a router via a simple ICMP echo. RFC 4413 [33] classifies IPID behavior into three classes: *(1) Sequential jump*: an incremental IPID counter that is used for all packet streams, or *(2) Random*: a pseudo-random number generator is used for the IPID value, or *(3) Sequential*: an incremental IPID counter on a per-stream basis. The IPID may also have a static value, e.g., zero. While the limited size (16 bits) of the IPID counter can be problematic, Internet researchers have utilized the IPID field for a broad range of applications. Bellovin [7] uses IPID to count NATed hosts, alias resolution tools such as MIDAR [30] and Ally [48] use monotonic IPID counters to infer aliases, and Chen et al. [15] use IPIDs to characterize end-systems. In this work, we utilize the differences in IPID value generation between router vendors across protocols for fingerprinting purposes.

TTL-based Fingerprinting: Vanaubel et al. propose a router fingerprinting technique based solely on TTL responses [50]. They send TCP, UDP, and ICMP probes toward the target, and show that the tuple of inferred initial TTL (iTTL) values from the responses can coarsely differentiate between some well-known vendors, including Cisco and Juniper. Unfortunately, the possible iTTL value range is small, and can lead to a large number of incorrect inferences, e.g., we find that Huawei has the same iTTL signature as Cisco. Nevertheless, we use the iTTL values as part of a larger feature set.

SNMPv3-based Fingerprinting: Most recently, research has shown that the adoption of the SNMPv3 protocol offers an opportunity for remote fingerprinting of network infrastructure [2] including routers. In addition to gathering detailed information about network devices, such as vendor, uptime, and the number of restarts, the reply also contains a strong, persistent identifier that allows for lightweight alias resolution and dual-stack association. We leverage this SNMPv3 technique to build a ground truth, and use it as a baseline for our proposed LFP method.

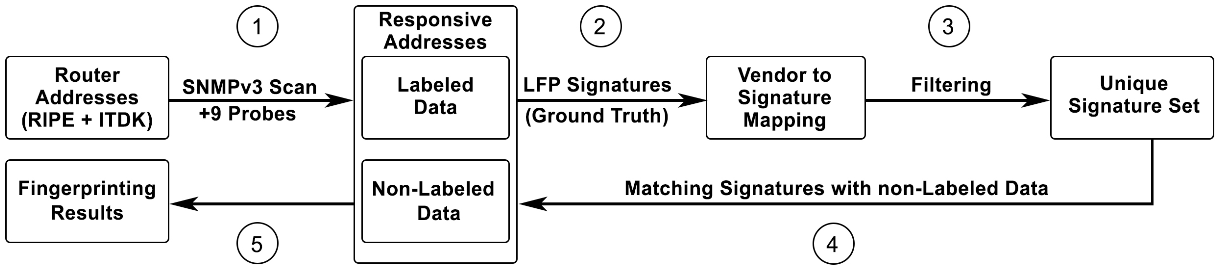

Figure 1: Data Collection Pipeline.

3 METHODOLOGY

This section presents our methodology to scalably classify routers in the wild at vendor granularity. We first give an overview, then describe our dataset, measurement probes, and the features we use for classification. Subsequently, we generate signatures based on these features and we describe how we handle classification edge cases due to ambiguity or lack of data. Finally, we elaborate on the limitations of our methodology. We refer the reader to Section 5 for the ethical principles guiding our measurements. For a pipeline of our methodology, we refer the reader to Figure 1.

3.1 Overview

Our methodology builds a model based in part on high-confidence router vendor labels and then uses that model to extend coverage and improve accuracy. Such ground-truth data can be collected using private information about the deployment of routers in a network or via information leakage using protocol scanning [13]. In our method, we utilize SNMPv3 as described in [2] which is able to accurately label around 20% of routers in the wild by sending a single unsolicited and unauthenticated request. The reply to this request contains detailed information including a router "Engine ID." This ID easily and reliably identifies the router vendor. We leverage the SNMPv3 technique and scan a set of router addresses to create labeled data and build a classification model for router fingerprinting measurement.

We expect that, typically, routers will not expose services to the public Internet. We decide to use three types of probe packets over ICMP, TCP, and UDP. ICMP has been used before to fingerprint routers, as many routers respond to ICMP packets. For TCP and UDP, we expect that routers do not expose such services to the public Internet. However, the response to packets targeting a closed port for these protocols can provide useful information towards fingerprinting the router vendor. In addition to the SNMPv3 requests, we send three single-packet probes over each of the three primary transport protocols, namely, ICMP, TCP, and UDP for a total of nine probes per IP address (Figure 1 ①). We explain the rationale to use these three protocols in Section 3.4. The feature values of the responses to our transport protocol measurements, listed in Table 1, are used to build a signature database for *Lightweight Fingerprinting (LFP)*. For IPs that are responsive to SNMPv3 requests, we extract the vendor information and used it as a label (Figure 1 ②). Note that our methodology is not dependent on the SNMPv3 to label routers. In principle, any reliable router label source can be used as input to our classification method.

3.2 Datasets

To select target router IPs, we leverage two complementary public router datasets: the RIPE Atlas traceroutes dataset [44] and the ITDK dataset [11]. We list the router datasets with dates, address counts, and AS coverage in Table 2.

RIPE Atlas Traceroutes. We extract intermediate IP hops from RIPE Atlas traceroute measurements to obtain router IPv4 addresses. We explicitly ignore the last responsive hop, if it is the same as the targeted host IP, to ensure that we only include router IPs. We utilize five snapshots of traceroute data over a ten-month period from January – November 2022. We extract between 446k to 496k router IPs from each snapshot. Further, each snapshot covers between 18.3k to 20.2k ASes. In order to increase the coverage, we utilize all five snapshots to gather signatures and evaluate their stability over time. Moreover, we find that RIPE Atlas traceroutes are relatively stable across the ten-month period, with a pairwise router IP overlap of about 88% between consecutive collections. Therefore, we utilize the most recent RIPE Atlas snapshot, i.e., RIPE-5, for our IP level analysis.

ITDK Router-Level Topologies Dataset. In addition to IP level traceroute data, we also use the router topology from CAIDA's February 2022 ITDK [11]. This complementary dataset contains router alias sets (excluding singletons) inferred via MIDAR [30] and iffinder measurements. This dataset covers fewer IP addresses and about half of the number of ASes compared to RIPE Atlas. This is expected as addresses in this dataset must respond over at least one protocol (ICMP, UDP, or TCP) which is required for alias resolution. This is also evident in our active measurement where we note a higher responsiveness for the ITDK data compared to RIPE Atlas as shown in Figure 4. The complementary nature of this dataset is underscored by a relatively low overlap of at most 26% router IPs present in any of the RIPE Atlas traceroute datasets. We use the ITDK data for gathering signatures and router level analysis.

The union of all RIPE Atlas traceroute and ITDK MIDAR datasets covers more than 970k router IP addresses in about 25k ASes. We note that our methodology is not limited to these selected datasets, but in fact other datasets containing candidate router IP addresses could be used as well. Next, we run active measurements toward targets in each of these datasets to gather features and build signatures for router fingerprinting.

3.3 Active Scanning Packets

To collect router fingerprints, we send 10 packets in total per target IP: 3 for each transport protocol and a single SNMPv3 request. We

Table 1: List of features used with possible values.

Feature	Possible Values
ICMP IPID echo	true, false
ICMP IPID counter	incremental, random, static, zero, duplicate
TCP IPID counter	incremental, random, static, zero, duplicate
UDP IPID counter	incremental, random, static, zero, duplicate
TCP UDP ICMP shared counter	true, false
TCP ICMP shared counter	true, false
UDP ICMP shared counter	true, false
TCP UDP shared counter	true, false
UDP iTTL	32, 64, 128, 255
ICMP iTTL	32, 64, 128, 255
TCP iTTL	32, 64, 128, 255
ICMP echo response size	variable
TCP response size	variable
UDP response size	variable
TCP SYN sequence number	zero, non-zero

Table 2: Overview of router address datasets: Number of unique IP addresses and Autonomous Systems. We utilize all data sources for signatures gathering. However, we use RIPE-5 for path and IP-level analysis and ITDK for router analysis.

Data Source	Date	# IPv4 addrs.	# ASes
RIPE-1	2022-01-24	494,867	20,178
RIPE-2	2022-02-24	484,930	19,989
RIPE-3	2022-06-09	496,167	20,085
RIPE-4	2022-07-04	446,629	18,304
RIPE-5	2022-11-07	476,577	18,837
ITDK	2022-02	343,312	9,922
Union		971,343	24,909

aim to reduce the impact of our scan on the target by using simple ping probes and avoiding any malformed packets. For ICMP, we send three echo requests. For each of these requests we expect an echo reply. For TCP and UDP we target port 33533, with the assumption that no services are active on this port. For TCP we send two ACK packets and one SYN packet with a non-zero acknowledgment number. We expect that all three TCP packets - both the ACK and SYN - will elicit a TCP RST response. For UDP we send three packets, each with 12 bytes of all zero payload. For each packet, we expect to receive an ICMP port unreachable response.

3.4 Feature Set

We limit our methodology to features that can be extracted mainly from the IP layer. In total, we extract 15 features from our 9 probe packets (see Table 1). We consider four groups of features:

3.4.1 IPID. We send a trio of consecutive packets and collect the IPID values from all responses. We then construct IPID sequences for each protocol. Previous work [30, 48] showed that IPID sequences exhibit distinct patterns, e.g., they can be monotonically increasing or random. These patterns can not only be used to perform IP-alias resolution as shown in previous work, but also facilitate the identification of a router's vendor. One test for device fingerprinting is checking if ICMP request and response IPID values match [5, 35]. IPID sequences differ among different protocols, but some implementations use the same sequence across all protocols. As we show in Table 1, the ICMP IPID Echo feature indicates whether Echo request and response IPID values are the same (true) or different (false). The IPID counter for any of the three protocols (ICMP, TCP, and UDP) can be characterized as incremental (which can also include wrap-around from the largest 16 bit value back to starting at zero), random, static (always the same value, other than zero), zero (always responds with an IPID of zero), or with duplicates (where exactly two responses have the same IPID value).

3.4.2 iTTL. Previous work [50] showed that different initial TTL (iTTL) values may differ between different protocols and even message types. We collect the iTTL values for each response that we receive. Typically, the iTTL value depends on the operating system or network card per vendor. Indeed, in Table 1, we show the different values, four in total, that we have collected in all our experiments (see Section 4 for details).

3.4.3 Response Size. To further diversify our features, we collect the response size for all protocols. We notice that typically, the ICMP and TCP response size often do not provide any information gain. However, the ICMP port unreachable response to a UDP request packet can differ between router vendors. This depends on whether the request packet is fully or partially quoted (and if so, how much of the original packet is quoted) in the ICMP response packet [6, 38]. As we show in Table 1, the characteristic value is variable and differs by router vendor and implementation, which allows us to make use of the response size for router fingerprinting.

3.4.4 Additional Features. RFC 793 [39] states that if a port is closed, any incoming segment except a reset triggers a reset response. If the segment has an ACK field, the reset takes its sequence number from the ACK field, otherwise, it uses a sequence number of zero. We noticed that only a few vendors are compliant with the RFC in this regard.

For the set of features and the possible feature values we refer to Table 1. We note that most of these features are only available for IPv4. Thus, in this paper we focus only on the classification of IPv4 router interfaces.

3.5 Signatures

We assemble all responses for each IP address into a feature vector based on Table 1. We use the instances of a particular feature vector that are associated with a vendor obtained from the SNMPv3 probes to create a mapping of a feature vector to a vendor. We then apply a basic filter based on occurrences threshold as described in (§4.3) At this point the feature vector is used as a *signature* for the vendor (Figure 1 ③).

Unique Signatures. If a signature is mapped only to a single vendor, then we call this a *unique signature.* In this case we have high confidence in the accuracy of the signature.

Non-Unique Signatures. When a signature is associated with *multiple* vendors, we characterize this as a *non-unique signature.*

This may happen, e.g., due to the change of the default router configuration by network operators, or simply a shared TCP/IP stack implementation between multiple vendors. As we will show in Section 4, typically there is one vendor that dominates even for non-unique signatures, or the non-unique signatures map to a family of routers that are based on the same OS or network stack. However, for the purpose of this study, we take a conservative approach and only consider *unique* signatures in our analysis.

Partial Signatures. There are also cases where a router IP responds only to a subset of the all three protocols (ICMP, TCP, and UDP). In this case, we characterize the signature as a *partial signature*. Even partial signature may prove useful to identify the vendor of a router. If the partial signature is unique for a vendor, then we call this a *partial unique signature*. If this partial signature is associated with multiple vendors, we call it a *partial non-unique signature*.

As we elaborate in detail in Section 4, it is common for a single vendor to have multiple signatures. This is to be expected as vendors often offer multiple products and versions of the same product, or it can be an artifact of acquisitions. Once the list of signatures for a given vendor has been compiled, we can match the signatures using our active measurements to infer the vendor of unlabeled routers (Figure 1 ④). With this technique we can substantially increase the coverage of routers that we can fingerprint in the wild (Figure 1 ⑤).

3.6 IPID Threshold

In order to determine whether an IPID counter is incremental or random, we investigate the returned IPID values per IP address and across all three protocols. Consequently, we sample the IPID values only for fully responsive addresses. We calculate the step values for each consecutive packet pairs and aggregate them by applying a maximum function[1]. In Figure 2 we show the distribution of maximum IPID step per IP in the responses to all three protocols. In order to distinguish random from sequential increases, we check for a knee in the distributions of Figure 2. We empirically take a conservative threshold value of 1,300 to distinguish between sequential and random IPID counters. Note that a sequential increase can be larger than '1', as concurrent traffic from that router also leads to an increase in sequential IPIDs.

We evaluate the empirical threshold by estimating the probability of misclassifying a random IPID counters as a sequential. Recall that we sent 9 packets in total and calculated 8 IPID steps by determining the difference between two consecutive IPID values. Given our threshold, the probability of a random IPID counter generating a value less than or equal to the threshold is $1301/2^{16}$ which is ≈ 0.019. For our classifier to misclassify a random counter as sequential, all eight IPID steps need to be less than or equal to the threshold, which has an extremely low probability of 0.019^8 when considering all protocols, or 0.019^2 when considering each protocol separately.

We further explore the empirical threshold in Figure 3, where we plot the distribution of the IPID difference for consecutive responses for fully responsive IPs in the RIPE-5 dataset. It is clear

[1]We obtain similar results when applying an average function instead of a maximum function. Since the maximum function is more conservative, we use it in our methodology.

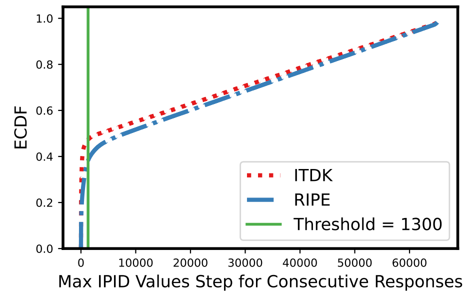

Figure 2: Maximum IPID step distribution per IP address. The vertical line shows the chosen threshold between sequential and random IPID increases.

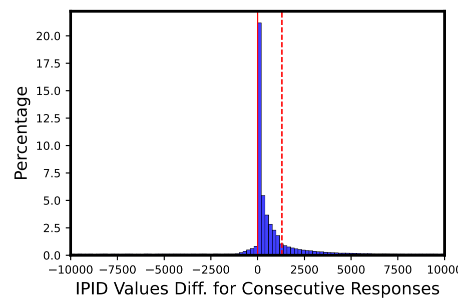

Figure 3: Distribution of IPID difference values for consecutive responses.

that around 20% of IPID differences are close to zero. Close to 90% of the IPID difference values are included by setting a threshold of 1,300, as shown with the dashed vertical line We use this threshold to differentiate between incremental values and random that are dispersed across the full range of possible IPID difference values.

Note that an effectively random IPID might by chance fall within this 1,300 threshold. Since with LFP we take the conservative approach of using the maximum IPID difference between consecutive probes, this random effect is very unlikely to occur twice in a row and thus strongly minimizes the number of false positives.

Finally, we also characterize the sharing of IPID sequences across pairs or all protocols as true of false, if this takes place or not, respectively.

3.7 Limitations

LFP improves the state-of-the-art in remote router vendor fingerprinting. However, we acknowledge that several limitations remain:

- Our classification builds on highly accurate vendor data obtained via SNMPv3 probing, however SNMPv3 coverage is not universal and imparts bias. While the SNMPv3 technique obtains correct labels for approximately 20% of the routers and 30% of the router IPs we probed, we do not generate a signature for vendors that do not implement or do not respond to SNMPv3 requests. This results in a bias toward SNMPv3 enabled routers and can negatively affect the uniqueness of generated signatures.
- As we elaborate in Section 4, a non-negligible fraction of routers do not respond to *any* remote probe. This differs across sets

of router datasets, but for these routers our technique cannot provide any insights.

- New signatures may be created as as new router models or vendors are introduced in the market. Although in Section 4 we show that over a period of ten months, and for different router datasets, the signatures we discover remain stable, over longer period of time, e.g., years, new measurements may be required to keep LFP signatures up-to-date.
- We restrict our analysis to core routers. A primary challenge to fingerprinting edge routers is the greater diversity of Customer Premise Equipment (CPE) and residential equipment, along with substantial amounts of IP churn. Although we believe that our technique can be used to fingerprint edge network equipment, we defer such an investigation to future work.
- We may misclassify random IPID response sequences as sequential. To significantly minimize the potential for erroneous inference, we we take the maximum IPID step difference among all pairwise IPID values (see Section 3.6).
- We focus our study on the IPv4 Internet. Many of the features that LFP relies on (see Table 1) are not available in IPv6 or do not provide the same discriminatory opportunities for fingerprinting. For instance, the IPv6 header does not include an IPID field unless fragmentation is induced [32], rendering all IPID-related features inapplicable for fingerprinting. Furthermore, all IPv6 implementations use the recommended initial TTL value of 64 [10]. The remaining features do not provide significant information gain to produce an accurate vendor signature.
- We limit the scope of our work and focus only on the technical aspects of remote router vendor fingerprinting that can be used to inform routing decisions. We recognize that better insight into vendors within ASes and along end-to-end paths is especially interesting given the current climate where, e.g., some countries are imposing restrictions on the use of equipment from particular vendors. In this paper, we discuss this issue in a purely factual, impartial, and non-political manner. Since we are not legal or political science scholars, we do not discuss, opine, or speculate on non-technical matters, e.g., the legal, financial, and social impact of our work.

4 ACTIVE EXPERIMENTS

We now apply our LFP methodology in active experiments to fingerprint routers in the wild. We run six measurements, one for each data source (five RIPE Atlas traceroutes, and one ITDK's MIDAR dataset, cf. Table 3). We find the five RIPE Atlas based measurements to be relatively consistent. Between 82k and 100k IPs are responsive to SNMPv3. Of those around 50k respond to all LFP probes, i.e., our labeled dataset which we use to extract vendor information. Another 58k–77k respond only to LFP probing, i.e., our dataset that we can fingerprint with the LFP technique without the IPs responding to SNMPv3. The ITDK dataset provides more SNMPv3-responsive IPs, with a similar number of LFP responses compared to RIPE Atlas traceroutes.

4.1 Responsiveness

Next, we analyze how responsive the target datasets are to LFP probes. This determines the upper limit of our coverage with LFP. Figure 4 shows the distribution of the number of responsive protocols (TCP, UDP, ICMP) per IP, comparing the ITDK and RIPE-5 datasets. Since we rely on responsiveness to create signatures, the higher the number of responsive protocols, the higher the entropy in the signature. Generally, we find that ITDK provides more responsive protocols compared to RIPE. About 50% of ITDK IPs are responsive on all three protocols, which is only 35% for RIPE. It is very encouraging, however, that we get responses from at least one protocol for 90.7% and 72.3% for ITDK and RIPE, respectively.

One other factor influencing the uniqueness of our signatures is the number of responses *per protocol*. Figures 5 and 6 show the responsiveness per protocol for RIPE-5 and ITDK, respectively. In both datasets we see that ICMP is more likely to elicit responses compared to TCP or UDP. Moreover, we see that an IP responds either to all three probe packets per protocol or to none, i.e., the line from zero to three in the plot is almost horizontal. Finally, we find that IP addresses from the ITDK dataset are generally more likely to be responsive compared to the RIPE dataset: 84.4% vs. 65.7% are responsive on all three ICMP probes for ITDK and RIPE, respectively; for TCP and UDP the difference is 63.6% in ITDK compared to 39.5% in RIPE.

4.2 Signatures

After collecting all responses from our measurements, we extract features (cf. Section 3.4) and create signatures based on our labeled SNMPv3 data (cf. Section 3.5). As can be seen in Table 3, each dataset individually contributes 34–62 unique signatures and 7–13 non-unique signatures. Unique signatures give us a high confidence when applying our LFP technique, as all labeled instances can be mapped to the same router vendor. Note, that if the same unique signature would be found with different vendors in different datasets, we count it as a non-unique signature. We find this case to be relatively rare, however, with only 2 occurrences for our five datasets. In our fingerprinting analysis, we exclude any non-unique signature and use the union of all five datasets to create a total of 89 unique signatures. We set a threshold of a minimum of 20 router samples per signature. Setting the threshold lower will only increase the covered routers by 1%, but disproportionally increases the number of signatures.

4.3 Occurrences Threshold

We perform a sensitivity analysis to understand the impact of the threshold of the minimum number of occurrences for signatures to be considered. In Figure 7, we vary the threshold (see x-axis), i.e., the minimum number of IPs with the same signature. A low threshold leads to a very high number of both unique and non-unique signatures. This is to be expected as uncommon vendors or a small number of configuration changes may lead to many different signatures. However, when we set the threshold of minimum occurrences to consider a signature to 10 or more, the number of signatures, both unique and non-unique, converges. As Figure 7 shows, choosing 10 or 20 as the threshold does not change the number of signatures substantially. A closer investigation shows that the set of signatures is also not affected. Therefore, we choose a threshold of 20 router IPs per signature in our study as it provides a good trade-off between considering a low number of popular

Table 3: Measurement overview: Responsive IPs (IPs), responsive IPs to SNMPv3 (SNMPv3), to SNMPv3 and LFP (SNMPv3 ∩ LFP), only to LFP (LFP \ SNMPv3), number of unique signatures (Unique sigs), and non-unique signatures (Non-unique sigs).

Measurement	IPs	SNMPv3	SNMPv3 ∩ LFP	LFP \ SNMPv3	Unique sigs	Non-unique sigs
RIPE-1	359,263	99,560	55,116	58,266	62	9
RIPE-2	355,709	95,600	54,933	59,400	46	8
RIPE-3	363,464	94,699	53,196	58,843	47	10
RIPE-4	323,141	82,047	48,360	72,969	49	11
RIPE-5	327,534	90,540	47,700	77,298	51	13
ITDK	311,607	113,089	58,492	53,952	34	7
Union	736,260	218,129	132,524	169,143	89	23

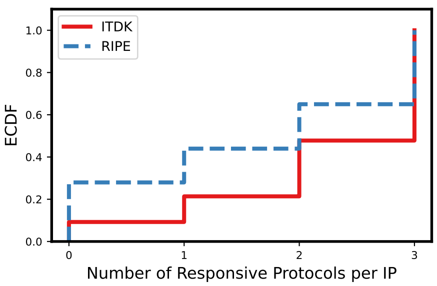

Figure 4: Responsive protocols per IP for the RIPE-5 and ITDK datasets.

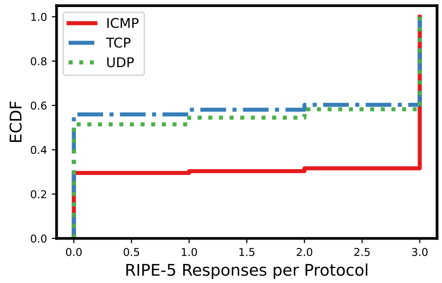

Figure 5: Responsiveness per protocol for the RIPE-5 dataset.

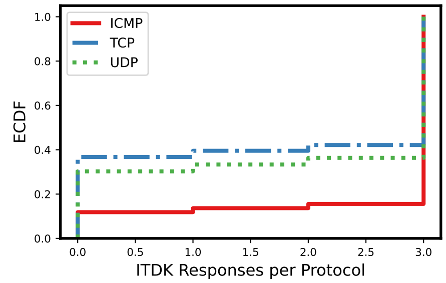

Figure 6: Responsiveness per protocol for the ITDK dataset.

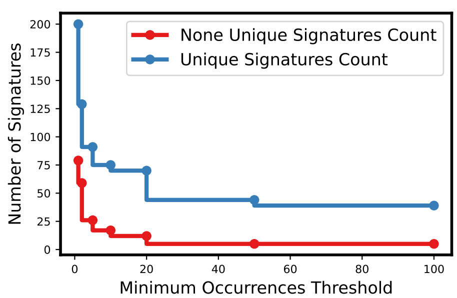

Figure 7: Sensitivity analysis: impact of setting a threshold on the number of occurrences for signatures on the number of unique and non-unique signatures.

signatures and excluding a large number of rare signatures with only a few occurrences in the hundreds of thousands of IPs in our dataset. In total, for this study, we identified 89 unique signatures and 23 non-unique signatures. We provide the full list of signatures in [3].

In addition to signatures where we get responses from all protocols, we also leverage partial signatures. Table 4 shows the partial fingerprints for different combinations of partial protocol responsiveness. We find that if we see responses from two protocols (i.e., TCP & UDP, ICMP & UDP, or ICMP & TCP), the majority of partial signatures are still unique and can therefore be leveraged by the LFP technique. Regarding single protocol signatures, the results are mixed. Most signatures are unique for TCP-UDP, ICMP-UDP and ICMP-TCP, while about half are unique for just TCP, UDP or ICMP.

Table 4: Partial signatures for different responsive protocol combinations.

Protocols	Total	Unique	Non-unique
TCP & UDP	61	43	18
ICMP & UDP	60	42	18
ICMP & TCP	51	36	15
UDP	20	12	8
ICMP	19	9	10
TCP	17	10	7

In general, utilizing unique partial signatures expands coverage by ≈ 15% while maintaining accuracy.

4.4 Mapping Signatures to Vendors

In Table 5 we show the vendor distribution based on the labeled dataset (i.e., SNMPv3-responsive addresses). To our positive surprise, more than 82% of the IPs map to a vendor with a unique signature. In total, our dataset covers 16 different vendors. We find Cisco to be the dominant router vendor for our labeled dataset with 51% of labeled router IPs with unique signatures, followed by Juniper and Huawei with 10% each.

For the major router vendors, the majority of IPs can be mapped to unique signatures, which increases our confidence in applying our technique to non-labeled data. Indeed, this is the case for 100% of Juniper, Alcatel/Nokia, and Ericsson router IPs, 98% of Cisco router IPs, and 86% of Huawei router IPs. Two notable exceptions

Table 5: Number of signatures in ground-truth dataset per router vendor. (#IPs are noted in parentheses).

Vendor	Labeled	Unique sigs	Non-unique sigs
Cisco	83,918	25 (82,020)	1 (1,898)
MikroTik	28,989	26 (9,489)	4 (19,500)
Huawei	19,869	8 (17,034)	4 (2,835)
Juniper	17,665	15 (17,665)	0 (0)
H3C	2,469	5 (358)	5 (2,111)
Alcatel/Nokia	1,111	2 (1,111)	0 (0)
Ericsson	200	1 (200)	0 (0)
Other	9,676	4 (497)	18 (9,179)

are MikroTik and H3C. For these two vendors, we might attribute a lower bound of routers. Note that both these vendors utilize UNIX-based solutions. We use the union of signatures in the following sections for router fingerprinting: network homogeneity, and end-to-end path analyses.

5 ETHICAL CONSIDERATIONS

During the design and the application of our methodology we took care to minimize any potential harm to the operation of routers and networks. First, the load of our measurements is very low. More specifically, we send ten packets, i.e., one SNMPv3 request and nine probes, three for each one of ICMP, TCP, and UDP. We do not send any malformed packet to avoid any unexpected behavior. Moreover, we coordinated with our local network administrators to ensure that our scanning efforts do not harm the local or upstream network. We follow current best practices [17, 19, 37] for active measurements and ensure that our prober IP address has a meaningful DNS PTR record. Additionally, we show information about our measurements and opt-out possibilities on a website of our scanning servers. During our active experiments, we did not receive any complaints or opt-out requests. Our work uncovers potentially sensitive security, robustness, and business information about network providers, e.g., router vendor. For this we plan to respond to any request by operators regarding their networks.

6 ROUTER VENDORS ON A PATH

In this section, we apply LFP to study the diversity of vendors along data-plane forwarding paths. Such insights are helpful as they could inform routing policy decisions by taking the equipment on a path into account. For example, if policy or law restricts a specific vendor, e.g., [20], a different path without this vendor might be selected. For this analysis, we use the most recent RIPE dataset, namely RIPE-5 (see Table 3), consisting of 7.3M traceroutes.

Figure 8 shows the ECDF of the number of hops per path in the RIPE-5 dataset. In this traceroute dataset, more than ≈ 7.1M of the paths (95%) have at least three IP hops. For our analysis, we consider only routable IPv4 addresses and we exclude any addresses that are private, or reserved. Moreover, more than 95% of the paths have a length of at most 15 hops.

Figure 9 shows the fraction of router IPs that we can map to a router vendor. We notice that for traceroutes with at least three

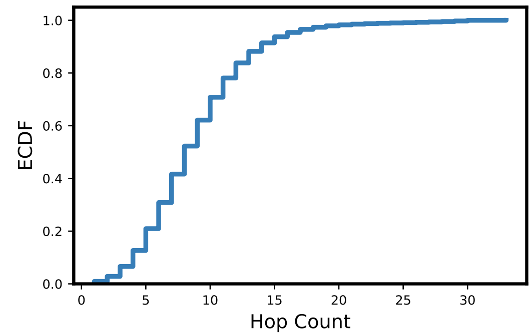

Figure 8: Path length distribution in the RIPE-5 traceroute dataset.

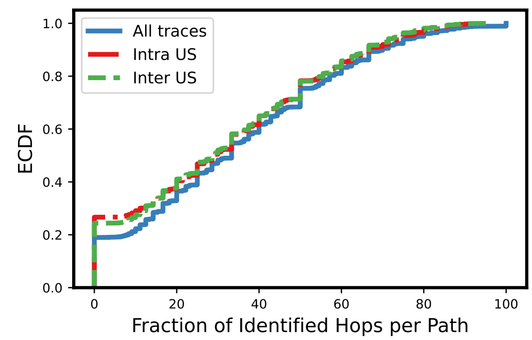

Figure 9: Identifiable routers on a path (RIPE-5).

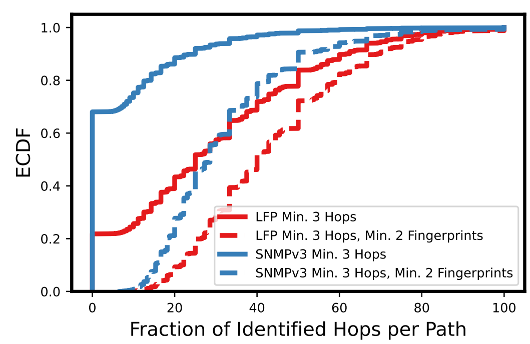

Figure 10: LFP vs. SNMPv3: Identifiable routers distribution on a path.

hops, LFP can identify at least two of the hops in 62% of the cases. This fraction increases to 82% to identify the vendor of at least one hop. This is a substantial improvement compared to the baseline with the SNMPv3 remote router vendor fingerprinting technique alone, as shown in Figure 10, where at least one vendor can be identified for only 35% of the traceroutes.

6.1 Identifying Router Vendors on a Path

First, we investigate the diversity of router vendors per path as fingerprinted with LFP. Figure 11 shows the number of unique vendors identified on paths where we can identify at least one hop; we identify around 650 unique sets of vendors. However, for around 50% of paths, LFP identifies only a single vendor. For around 40% paths, LFP identifies two vendors, and only 7% of the paths have

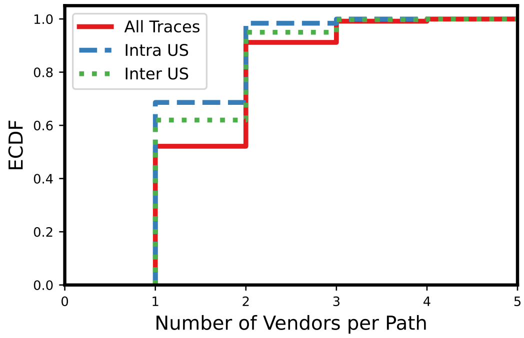

Figure 11: Router vendor diversity on a path.

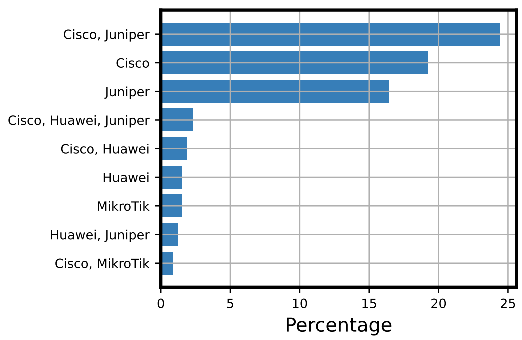

Figure 12: Top router vendor combinations for paths in the RIPE-5 dataset.

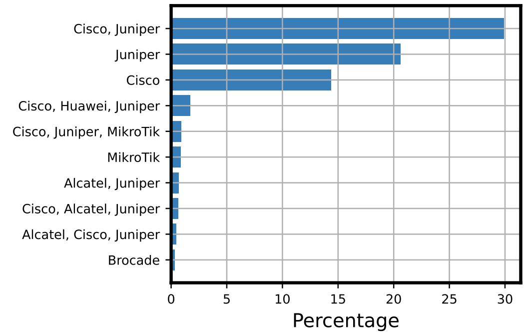

Figure 13: Top router vendor combinations for intra-US paths in the RIPE-5 dataset.

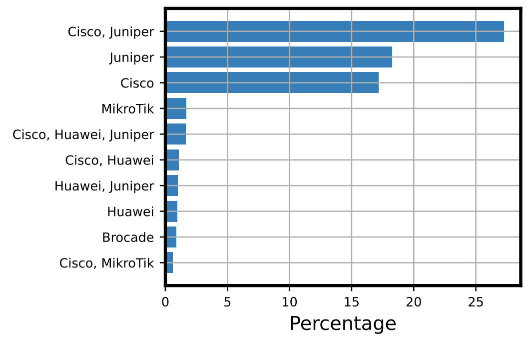

Figure 14: Top router vendor combinations on inter-US paths (source or destination US).

three distinct vendors. Four or more different router vendors are identified in fewer than 2% of the paths.

Next, we analyze the most popular combinations of router vendors on paths (without respect to their order along the path). Figure 12 shows that the top nine sets of vendors cover more than 95% of the RIPE-5 paths. The top three vendor combinations only involve Cisco and Juniper, making up almost 60% paths. Traceroute paths with all other combinations account for fewer than 3% each.

6.2 Case Study: US-related Paths

As a case study, we consider router vendor diversity specifically for the United States. There are ongoing discussions whether traffic that originates from the US, or has as a destination in the US, should be carried by "untrusted vendors" [43]. Moreover, if a vulnerability for a specific router vendor is discovered [23, 46, 49], paths with these vendors might, in theory, be avoided until a patch is developed and applied. With knowledge about vendors on a forwarding path, possible alternative paths from a source to a destination may receive preferential treatment in routing decisions by network operators. This could be facilitated with source routing techniques [40] or enforced by the upstream provider [21].

6.2.1 Intra-US Paths. First, we investigate the case that both the source and the destination of a traceroute are within the US. To geolocate the endpoints, we rely on IP address registry information. While other (more fine-grained and more accurate) geolocation techniques exist, we are primarily interested in policies and regulations that are frequently governed by the home country of the service provider, which is best reflected in the address registry. We exclude from our analysis anycast IPs [9] as they may be announced from different locations. The RIPE-5 dataset contains 395,775 traceroutes with at least three hops where both the source and the destination IP geolocate to the US. For around 60%, of them, we can identify two or more router IPs and assign them to vendors using LFP. Furthermore, we find that for more than half of intra-US traces we can identify at least a third of the router vendors on the path.

Moreover, in Figure 11, we show that in around 70% of the intra-US paths, all the IPs belong to a single vendor. The majority of the remaining 30% of intra-US paths have routers that belong to two distinct vendors, and the cases where there are three or more vendors is negligible. Indeed, our results suggest a high degree of consolidation of router vendors. Regarding the most popular set of router vendors for intra-US paths (cf. Figure 13), we see a similar picture compared to the overall dataset. Combinations of Cisco and Juniper dominate, even more so than in the overall dataset, making up more than two thirds of all intra-US paths combined. This shows that intra-US paths have low vendor diversity, consisting mostly of Cisco, Juniper, or a combination of both. Such homogeneity may be indicative of potential critical infrastructure weaknesses e.g., where all devices are affected by a vulnerability.

6.2.2 Inter-US Paths. We also investigate the case that only one of the source and the destination are in the US. In the RIPE-5 dataset, there are 3M traceroutes of least three hops where only the source or only the destination IP geolocate to the US. For around 58% of these, we can identify the vendors for two or more router IPs using LFP. For more than half of inter-US traces, we can identify the vendor of at least a third of the router IPs on the path, showing a similar distribution as intra-US as well as other paths.

Table 6: Two sample unique signatures: top for Juniper and bottom for Cisco. By changing the default value of Juniper for ICMP iTTL from 64 to 255 (values in box), the classifier misidentifies Juniper routers as Cisco.

Juniper	False	r	r	r	False	False	False	False	255	$\boxed{64}$	64	84	40	56	0
Cisco	False	r	r	r	False	False	False	False	255	$\boxed{255}$	64	84	40	56	0

Moreover, in Figure 11, we show that in around 60% of inter-US paths, all IPs belong to a single vendor. Almost all of the remaining paths contain two mappable router vendors. These observations are similar to the intra-US study, and show a high degree of vendor consolidation. Cisco and Juniper are again the most prominent vendor combinations (see Figure 14), showing a similar distribution to intra-US and overall paths. However, the results suggest that inter-US paths exhibit more heterogeneity than the intra-US paths.

6.3 Case Study: Informed Routing

Knowing the vendors across the path can inform routing policy. For example, a sender may want to avoid sending traffic through ASes dominated by hardware from vendors they do not trust. Thus, the routing policy could choose an alternative path if available. Our methodology can inform the possible alternatives and may serve as a step toward enforcing such policies. As a case study, we find vendor homogeneous ASes in the RIPE-5 dataset: ASes with at least 1k router IPs where LFP finds at least 85% of the IPs belong to a single vendor. Next, we use the CAIDA AS relationship dataset [12] to find AS paths where these vendor homogeneous ASes serve as transit ASes. Then, we consider the destinations ASes where the homogeneous transit AS appears on the path. For these destinations, we investigate if there exists an alternative path from the *same destination* but with a transit AS using a *different* vendor. Note that while our analysis utilizes the CAIDA AS relationships in order to identify policy-compliant transit ASes, such inferences may be limited by the available data and the visibility of *all* AS paths toward a given AS. We acknowledge that there may exist paths that cannot be observed from publicly available data, or that an alternative path may not be compliant in the traditional economic or valley-free routing sense.

As a demonstration of the insights possible from this analysis, we examine two networks: AS9808 and AS3786. AS9808 is a large transit provider where LFP infers Huawei to be the the dominant router vendor. We identify 25,134 AS paths where AS9808 serves as a transit provider. For 68 destination ASes, no alternative path that does not transit AS9808 is visible[2]. On the other hand, for 167 destination ASes, an alternative path via ASes that operate non-Huawei routers is available.

As a second example for a different router vendor, LFP shows that Juniper is the dominant router vendor for AS3786. We identify 1.3M AS paths, and 436 unique destinations where AS3786 appears as a transit provider. For 53 destinations there is no alternative path visible to us. Naturally, our inferences depend on our visibility into the AS, however, this result suggests that our methodology can similarly be applied to any destination when the set of paths is available.

[2]Note that not all AS paths are visible in the BGP [1, 14, 22, 36, 52], thus our analysis is limited to the visible paths only.

7 ROUTER FINGERPRINTING

With the signatures collected in our active experiments, we now apply our fingerprinting technique to the router datasets. We leverage 89 unique signatures and 78 partial unique signatures from the union dataset (cf. Tables 3 and 4). Recall that both full and partial unique signatures provide exact matches between a signature and a vendor.

7.1 IP to Vendor Mapping

We use our combined full and partial unique signatures on the latest RIPE dataset, i.e., RIPE-5, and ITDK datasets to map IP addresses to vendors. For RIPE-5, our analysis shows that our method fingerprints 56.7% of router IPs when we use unique signatures. For reference, the SNMPv3 technique fingerprints only 26% of the router IPs. LFP alone fingerprints 49%.

Figures 15 and 16 show the fingerprinting results based on responsive IPs from the RIPE-5 and ITDK datasets, respectively. We report the router IPs identified only by LFP, only by SNMPv3, and by both methods. We find that our LFP technique roughly doubles the number of fingerprintable IP addresses for both datasets. Moreover, we see that the number of fingerprintable IPs increases quite drastically for certain vendors: Juniper sees an increase of 650% and 259.3%, and Huawei 249.8% and 136.4% for RIPE-5 and ITDK, respectively. Generally, we see a more balanced router vendor distribution, with the most dominant vendor Cisco decreasing its share from ≈65% with SNMPv3 only to ≈50% for SNMPv3 + LFP. We provide an analysis for the non-unique signature precision and recall in Appendix B.

7.2 Router to Vendor Mapping

Next, we make use of the ITDK dataset not only containing IP address information, but also alias sets. We apply our signatures to all non-singleton router alias sets. First, we check if all IPs within fingerprinted alias sets report the same vendor. We find this to be the case for ≈ 99% of all alias sets, with 498 router IPs producing conflicting vendor inferences (0.65%). Second, we plot the router vendor distribution counted by alias set in Figure 17. The router distribution is similar to the IP-based distribution (cf. Figure 16), with Cisco being the dominant vendor, followed by Huawei and Juniper. Again, we can map about 96.4% more routers with the combined SNMPv3 + LFP technique, compared to SNMPv3-only.

7.3 Comparison with other Tools

To evaluate the accuracy of vendor fingerprinting by LFP and the associated bandwidth requirements, we conduct a comparison with Nmap [35] and Hershel [47]. For this, we acquire a set of addresses from Censys, which are known to reveal vendor information through service banners. Censys also provides Hershel fingerprints and OS identification where available. For each of the top six vendors found via LFP, we randomly select 500 IP addresses

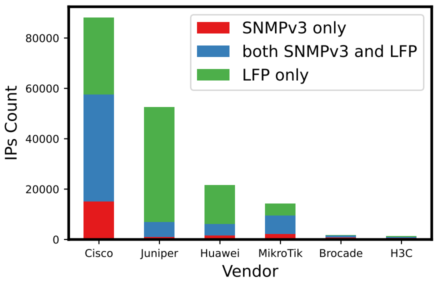

Figure 15: IPs to vendors: SNMPv3 vs. LFP for the RIPE-5 dataset.

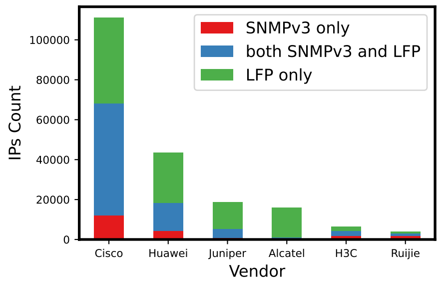

Figure 16: IPs to vendors: SNMPv3 vs. LFP for the ITDK dataset.

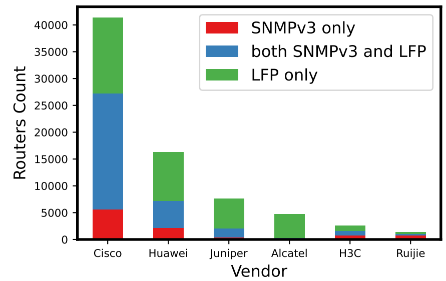

Figure 17: Routers to vendors: SNMPv3 vs. LFP for the ITDK dataset.

Table 7: Comparing coverage and accuracy of LFP and Nmap for Censys-labeled data.

	Coverage		Accuracy	
Vendor	LFP	Nmap	LFP	Nmap
Cisco	40%	10%	95%	84%
Juniper	81%	31%	99%	98%
Huawei	49%	20%	55%	50%
Ericsson	93%	6%	80%	0%
Mikrotik	83%	15%	10%	5%
Alcatel	38%	11%	48%	16%

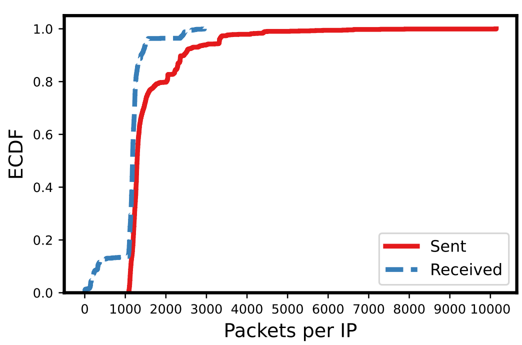

Figure 18: Sent and received packets from Nmap.

and perform tests using both LFP and Nmap. Additionally, we compare our findings with Hershel fingerprints, wherever possible.

7.3.1 Comparison with Nmap. Table 7 shows the coverage and accuracy results for LFP and Nmap for six different vendors. Coverage refers to the percentage of responsive IPs for each vendor, while accuracy refers to the percentage of correct fingerprints for the responsive IPs. Although both tools have similar accuracy, LFP has the ability to achieve substantially higher coverage.

After evaluating the coverage and accuracy of LFP and Nmap, we proceeded to analyze their respective bandwidth requirements. Specifically, LFP sends a consistent 9 packets (3 ICMP, 3 TCP, and 3 UDP) to each targeted IP address. In contrast, Nmap sends substantially more packets when attempting to fingerprint a device. Figure 18 shows the distribution of packets sent and received by Nmap using the default OS fingerprinting options. We find that Nmap sends more than 1000 packets to more than 80% of all targeted IPs. Moreover, our analysis shows that Nmap sends an average of 1,538 packets per IP and receives 1,065 packets. However, it should be noted that in certain cases Nmap may send an extremely high number of packets to a single IP address, exceeding 10 *thousand* packets. This behavior is largely influenced by the services operating on the target system. In contrast, we observe that the LFP technique has considerably lower bandwidth requirements compared to Nmap, making it a more lightweight option overall.

7.3.2 Comparison with Hershel. we also compare LFP with Hershel fingerprints. By design Hershel requires a single packet to obtain a fingerprint, which is even less bandwidth-intensive than the 9 packets sent by LFP, and much less intrusive than the multitude of packets sent by Nmap. Our analysis of the test sample shows that Hershel has an overall coverage of approximately 50%. Furthermore, we find that Hershel is only able to identify the target vendor with less than 1% accuracy for our top 3 vendors. This suggests that Hershel—while it may perform well for servers—is not a suitable tool for router fingerprinting. Additionally, we observe that Hershel often identifies Linux-based systems (such as Mikrotik) simply as Linux machines. This is due to the limited number of signatures for router vendors in the Hershel fingerprinting database.

In summary, LFP achieves a balance between coverage and accuracy while also having a low network footprint.

7.4 Family-level Fingerprinting

After discovering that many vendors have not a single but multiple signatures, we investigate whether these signatures can be linked to different router models or families. To test this hypothesis, we collect a sample of 400 Cisco IPs running SNMPv2c and query for the Sys.desc O.I.D. [42]. This provides a small ground-truth sample with fingerprinting information beyond the vendor. Next, we run LFP against these targets, and collect their signatures. The results show that the collected signatures belong to the top 13 most common Cisco signatures, which cover over 96% of labeled Cisco data. Additionally, we identify a unique signature for three different IOS families (3 XR, 3 NX, and 7 IOS signatures), which are not shared with the other versions. However, due to the limited ground-truth dataset, it is not possible to evaluate the accuracy of these results in detail, and we leave this task for future work. Overall, the sample data supports the assumption that different signatures can be linked to specific router models or families, which can lead to a more fine-grained router fingerprinting.

7.5 New Insights on Router Deployment

Using the collected router fingerprints, we next conduct a comprehensive analysis of global router vendor distribution by comparing our findings with a similar study [2]. Specifically, we utilize LFP to identify the vendor of routers and examine the global distribution of these vendors. Our analysis provides a detailed overview of the global router vendor landscape.

For our analysis, we focus on the ITDK dataset (see Table 3). Recall that this dataset has information about all the interfaces (IPs) associated with the same router via alias resolution. LFP can identify unique signature routers in 6,743 ASes, compared to 4,929 ASes with the SNMPv3 method. Thus, not only can LFP identify more than double the number of router IPs (see the previous section), but it also identifies routers in 1,814 additional ASes (+36.8%). This is a substantial contribution of LFP as it sheds light on previous blind spots in the Internet and contributes to a better estimation of the global router vendor distribution.

In Appendix A we demonstrate the efficacy of utilizing LFP to enhance router coverage in a network. Our findings reveal that LFP can identify more than twice the number of routers in large networks, thereby substantially improving coverage. Additionally, LFP provides a comprehensive analysis of router homogeneity across different points, offering a more detailed report on homogeneity.

8 DISCUSSION

Obfuscating remote router vendor fingerprinting: Our analysis shows that it is possible to hide from remote router fingerprinting. The obvious way is to drop UDP and TCP traffic, especially from non-whitelisted sources. But even if UDP and TPC traffic is not dropped, it is still possible to create rare signatures that are more difficult to be mapped to a specific vendor. It is also possible by configuring a router to confuse the classification algorithm (similar to an adversarial attack on classifiers). Some of the features are difficult to change, e.g., ICMP, TCP, or UDP IPID counters, if they can be configured at all since they might be directly implemented in the router OS. However, it is easier to change default iTTL values. In Table 6, we present two unique sample signatures for Juniper (top) and for Cisco (bottom). By changing the default value of the ICMP iTTL (see Table 1 for details) of the Juniper routers from 64 to 255, LFP would misclassify the Juniper router as a Cisco router.

Using additional sources of information for fingerprinting: Our methodology relies solely on network characteristics and active probing. Other techniques utilize other sources of information, e.g., banners, that offer good coverage [26]. Banner data analysis requires the development of heuristics. One of the benefits of using a simple rule-based approach such as LFP compared to machine learning (ML) techniques, is that it is clear why certain decisions are being taken, whereas ML techniques usually suffer from a lack of explainability. Furthermore, complex ML models in networking can suffer from deficits such as shortcut learning, spurious correlations, and vulnerability to out-of-distribution samples [29]. Future work should explore the possibility of using explainable ML models for router fingerprinting. Moreover, banners' raw data is less accessible, typically proprietary, that comes with commercial or limited academic licenses. Nevertheless, banner data analysis can complement our technique and improve fingerprinting coverage and granularity. As part of our future work, we plan to use information fusion of our data and banner data and assess the benefit of using additional information sources for router fingerprinting, especially for vendors with non-unique signature, and hopefully for finer-grained fingerprinting, e.g., model-level fingerprinting.

Non-Unique Signatures: While we only utilize unique signatures in this study, non-unique signatures can offer insights into router deployments. This is particularly relevant when a single vendor dominates a non-unique signature with thousands of instances. Additionally, utilizing non-unique signatures can increase LFP coverage to 64% in the RIPE-5 dataset. We explore the precision and recall of non-unique signatures in Appendix B and intend to investigate additional features to enhance the uniqueness of such signatures in future research.

Integrating LFP into Nmap: We also plan to investigate how the insights gained by our study can be transferred and integrated into Nmap [35]. Our analysis shows that LFP can achieve better accuracy with ten packets (including the SNMPv3 request) than the default Nmap OS detection mode, which sends up to thousands of probe packets. At least in the case of router fingerprinting, LFP has proven to be more scalable, less intrusive, and more accurate. We are already developing a Nmap variation that will replicate our experiment, and we will share it with the Nmap community to get feedback and comments. This way, we can improve our methodology and enable more researchers and engineers to use our technique.

Longitudinal analysis: As part of our future research agenda, we would like to investigate how we can use our classification methodology and our collected data to perform a large-scale longitudinal analysis of vendor changes over time, vendor changes for a network, or vendor changes per router interface IP. So far, we have collected data that spans more than six months, but the real potential of our technique will be unveiled by collecting data that spans multiple years. We plan to publicly make the tools and data available to the research community and report on our results. We also plan to investigate how geopolitical events, economic changes, security incidents, and vendor strategies may influence the distribution of routers by different vendors across different time scales and geographical regions.

9 CONCLUSION

In this paper, we have shown that only 10 packet probes per router IP are enough to accurately fingerprint up to 64% of routers in the IPv4 Internet. We developed and evaluated LFP—a lightweight fingerprinting technique that sends three probe packets for three transport protocols, namely, ICMP, TCP, and UDP. By augmenting our traces with labeled router data that relies on SNMPv3 responses, we generated around 90 unique signatures that can accurately identify all major router vendors. To our surprise, more than half of the routers replied to our probe packets. The vast majority of the responsive routers (more than 82%) can be assigned to only one vendor using our classification. Our results showed that compared to the state-of-the-art, we more than doubled the coverage of routers that we can remotely fingerprint, and more accurately inferred the router vendor compared to popular tools like Nmap. All of this is achieved with orders of magnitude less probing packets than required by Nmap. Thus, our mechanism is more scalable, less intrusive, and does not rely on external and proprietary data like banner grabs. Our classification provides valuable insights into the deployment of routers within networks and regions, and the router vendor equipment on a given path. Thus, it can be used to inform routing decisions, to assesses router deployment strategies, to analyze hardware manufacturer market share, and to help estimate the potential impact of router vulnerabilities in a network or a region. Finally, to enable further research in the area, we plan to make our LFP tool publicly available.

ACKNOWLEDGEMENTS

This work was supported in part by the European Research Council (ERC) under Starting Grant ResolutioNet (ERC-StG-679158).

REFERENCES

[1] B. Ager, N. Chatzis, A. Feldmann, N. Sarrar, S. Uhlig, and W. Willinger. 2012. Anatomy of a Large European IXP. In *ACM SIGCOMM*.
[2] Taha Albakour, Oliver Gasser, Robert Beverly, and Georgios Smaragdakis. 2021. Third Time's Not a Charm: Exploiting SNMPv3 for Router Fingerprinting. In *ACM IMC*.
[3] Taha Albakour, Oliver Gasser, Robert Beverly, and Georgios Smaragdakis. 2023. Illuminating Router Vendor Diversity Within Providers and Along Network Paths (artifacts). https://routerfingerprinting.github.io/. (2023).
[4] Manos Antonakakis, Tim April, Michael Bailey, Matthew Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J. Alex Halderman, Luca Invernizzi, Michalis Kallitsis, Deepak Kumar, Chaz Lever, Zane Ma, Joshua Mason, Damian Menscher, Chad Seaman, Nick Sullivan, Kurt Thomas, and Yi Zhou. 2017. Understanding the Mirai Botnet. In *USENIX Security Symposium*.
[5] Ofir Arkin and Fyodor Yarochkin. 2002. Xprobe v2.0: A "Fuzzy" Approach to Remote Active Operating System Fingerprinting. http://www.ouah.org/Xprobe2.pdf. (2002).
[6] Fred Baker. 1995. Requirements for IP Version 4 Routers. RFC 1812 (Proposed Standard). (Jun 1995). https://doi.org/10.17487/RFC1812 Updated by RFCs 2644, 6633.
[7] Steven M. Bellovin. 2002. A Technique for Counting NATted Hosts. In *ACM SIGCOMM Workshop on Internet measurment*.
[8] Robert Beverly. 2004. A Robust Classifier for Passive TCP/IP Fingerprinting. In *PAM*.
[9] bgp.tools. 2023. Anycast Prefixes list as bgp.tools sees it. https://github.com/bgptools/anycast-prefixes. (2023).
[10] Wesley G. Bofman and Fernando Maniego. 2019-06. FINGERPRINTING IPV4 AND IPV6 ROUTERS USING ICMP. (2019-06). https://hdl.handle.net/10945/63434
[11] CAIDA. 2021. Macroscopic Internet Topology Data Kit (ITDK). https://www.caida.org/catalog/datasets/internet-topology-data-kit/. (2021).
[12] CAIDA. 2022. AS Relationships. https://publicdata.caida.org/datasets/as-relationships/serial-1/. (2022).
[13] Censys. 2021. Censys Scanning and Data Collection. https://censys.io/. (2021).
[14] Hyunseok Chang, Ramesh Govindan, Sugih Jamin, Scott J. Shenker, and Walter Willinger. 2002. Towards capturing representative AS-level Internet topologies. In *ACM SIGMETRICS*.
[15] Weifeng Chen, Yong Huang, Bruno F. Ribeiro, Kyoungwon Suh, Honggang Zhang, Edmundo de Souza e Silva, Jim Kurose, and Don Towsley. 2005. Exploiting the IPID Field to Infer Network Path and End-System Characteristics. In *PAM*.
[16] J. Czyz, M. Luckie, M. Allman, and M. Bailey. 2016. Don't Forget to Lock the Back Door!A Characterization of IPv6 Network Security Policy. In *Network and Distributed Systems Security (NDSS)*.
[17] David Dittrich, Erin Kenneally, et al. 2012. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research. *U.S. Department of Homeland Security* (2012).
[18] Zakir Durumeric, David Adrian, Ariana Mirian, Michael Bailey, and J. Alex Halderman. 2015. A Search Engine Backed by Internet-Wide Scanning. In *ACM CCS*.
[19] Zakir Durumeric, Eric Wustrow, and J. Alex Halderman. 2013. ZMap: Fast Internet-Wide Scanning and its Security Applications. In *USENIX Security Symposium*.
[20] J. C. Gallagher. 2022. U.S. Restrictions on Huawei Technologies: National Security, Foreign Policy, and Economic Interests. (2022). https://crsreports.congress.gov/product/pdf/R/R47012.
[21] Vasileios Giotsas, Matthew Luckie, Bradley Huffaker, and kc claffy. 2014. Inferring Complex AS Relationships. In *ACM IMC*.
[22] Enrico Gregori, Alessandro Improta, Luciano Lenzini, Lorenzo Rossi, and Luca Sani. 2014. A novel methodology to address the Internet AS-level data incompleteness. *IEEE/ACM Transactions on Networking* 23, 4 (2014).
[23] Jessica Lyons Hardcastle. 2022. Patch these Juniper Networks bugs, CISA says. (2022). https://www.theregister.com/2022/07/15/cisa_critical_juniper_bugs/
[24] Nadia Heninger, Zakir Durumeric, Eric Wustrow, and J. Alex Halderman. 2012. Mining your Ps and Qs: Detection of widespreadweak keys in network devices. In *USENIX Security Symposium*.
[25] Jordan Holland, Paul Schmitt, Nick Feamster, and Prateek Mittal. 2021. New Directions in Automated Traffic Analysis. In *ACM CCS*.
[26] Jordan Holland, Ross Teixeira, Paul Schmitt, Kevin Borgolte, Jennifer Rexford, Nick Feamster, and Jonathan Mayer. 2020. Classifying Network Vendors at Internet Scale. arXiv, https://arxiv.org/abs/2006.13086. (2020).
[27] Liz Izhikevich, Renata Teixeira, and Zakir Durumeric. 2021. LZR: Identifying Unexpected Internet Services. In *USENIX Security Symposium*.
[28] Liz Izhikevich, Renata Teixeira, and Zakir Durumeric. 2022. Predicting IPv4 Services Across All Ports. In *ACM SIGCOMM*.
[29] Arthur S Jacobs, Roman Beltiukov, Walter Willinger, Ronaldo A Ferreira, Arpit Gupta, and Lisandro Z Granville. 2022. AI/ML for Network Security: The Emperor has no Clothes. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*. 1537–1551.
[30] K. Keys, Y. Hyun, M. Luckie, and k claffy. 2013. Internet-Scale IPv4 Alias Resolution with MIDAR. *IEEE/ACM Trans. Networking* 21, 2 (2013).
[31] Derek Leonard and Dmitri Loguinov. 2010. Demystifying Service Discovery: Implementing an Internet-wide Scanner. In *ACM IMC*.
[32] Matthew Luckie, Robert Beverly, William Brinkmeyer, and kc claffy. 2013. Speedtrap: Internet-Scale IPv6 Alias Resolution. In *ACM IMC*.
[33] Stephen McCann and Mark A. West. 2006. TCP/IP Field Behavior. IETF RFC 4413. (2006).
[34] Alper T. Mizrak, Stefan Savage, and Keith Marzullo. 2008. Detecting Compromised Routers via Packet Forwarding Behavior. *IEEE network* 22, 2 (2008), 34–39.
[35] Nmap. 2021. Nmap: the Network Mapper - Free Security Scanner. https://nmap.org/. (2021).
[36] Ricardo Oliveira, Dan Pei, Walter Willinger, Beichuan Zhang, and Lixia Zhang. 2009. The (in) completeness of the observed Internet AS-level structure. *IEEE/ACM Transactions on Networking* 18, 1 (2009).
[37] Craig Partridge and Mark Allman. 2016. Ethical Considerations in Network Measurement Papers. *Commun. ACM* (2016).
[38] J. Postel. 1981. Internet Control Message Protocol. RFC 792 (Internet Standard). (Sep 1981). https://doi.org/10.17487/RFC0792 Updated by RFCs 950, 4884, 6633, 6918.
[39] J. Postel. 1981. Transmission Control Protocol. (1981). https://tools.ietf.org/html/rfc793 Library Catalog: tools.ietf.org.
[40] Stefano Previdi, Clarence Filsfils, Bruno Decraene, Stephane Litkowski, Martin Horneffer, and Rob Shakir. 2016. Source Packet Routing in Networking (SPRING) Problem Statement and Requirements. IETF RFC 7855. (2016).
[41] F. Qian, A. Gerbe, Z. M. Mao, S. Sen, Oliver Spatscheck, and W. Willinger. 2009. TCP Revisited: A Fresh Look at TCP in the Wild. In *ACM IMC*.
[42] Global OID reference database. 2023. Reference record for OID 1.3.6.1.2.1.1. https://oidref.com/1.3.6.1.2.1.1. (2023).
[43] Reuters. 2021. White House vows to protect U.S. telecoms network from Huawei security threat. https://www.reuters.com/article/us-usa-huawei-security-idUSKBN29W2HV. (2021).
[44] RIPE NCC. 2021. RIPE Atlas. https://atlas.ripe.net/. (2021).
[45] Erik C. Rye and Robert Beverly. 2019. Sundials in the Shade: An Internet-wide Perspective on ICMP Timestamps. In *PAM*.
[46] Omar Santos. 2023. Threat Actors Exploiting SNMP Vulnerabilities in Cisco Routers. (2023). https://blogs.cisco.com/security/threat-actors-exploiting-snmp-vulnerabilities-in-cisco-routers
[47] Zain Shamsi, Ankur Nandwani, Derek Leonard, and Dmitri Loguinov. 2014. Hershel: Single-Packet OS Fingerprinting. In *ACM SIGMETRICS*.
[48] Neil Spring, Ratul Mahajan, and David Wetherall. 2002. Measuring ISP topologies with Rocketfuel. In *ACM SIGCOMM*.
[49] The MITRE Corporation CVE. 2022. CVE.org Search Results for routers. https://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=router. (2022).
[50] Yves Vanaubel, Jean-Jacques Pansiot, Pascal Mérindol, and Benoit Donnet. 2013. Network Fingerprinting: TTL-Based Router Signatures. In *ACM IMC*.
[51] Franck Veysset, Olivier Courtay, and Olivier Heen. 2003. Snacktime: A Perl Solution for Remote OS Fingerprinting. https://www.packetfu.com/wp/snacktime.html. (2003).
[52] Florian Wohlfart, Nikolaos Chatzis, Caglar Dabanoglu, Georg Carle, and Walter Willinger. 2018. Leveraging interconnections for performance: the serving infrastructure of a large CDN. In *ACM SIGCOMM*.

APPENDIX

A ROUTER VENDOR DISTRIBUTION: REVISITED

In Figure 19 we plot the ECDF of the percentage of identified routers per network (AS) using LFP. When we consider all the ASes, we find that for approximately 60% of the ASes, LFP identifies all the routers. In these ASes we notice a bias: for about half of the ASes there is only one router in the dataset. When we increase the minimum threshold of routers per AS to consider them in our study, we notice that for at least 75% of the ASes LFP identifies at least half of the routers in an AS. The coverage decreases for large networks, i.e., with more than 1,000 routers, which is expected as they may have more routers with closed ports or behind firewalls and other provisions.

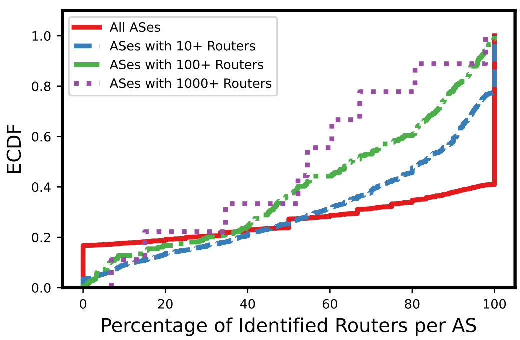

Figure 19: LFP coverage distribution per AS for different minimum thresholds of routers per AS.

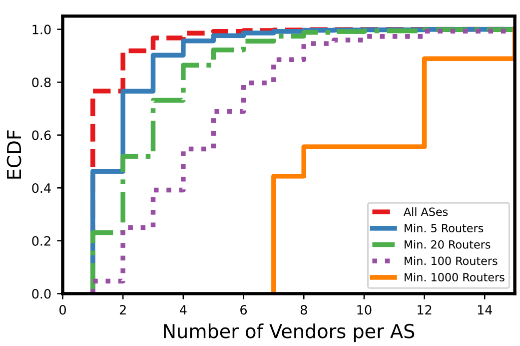

Figure 20: Assessing homogeneity of router vendors per AS.

A.1 Revisiting Homogeneity

Next, we revisit router vendor homogeneity per AS using LFP. In Figure 20 we plot the ECDF of the number of vendors per network (AS). A network is homogeneous when all the routers it hosts are from the same vendor. Our analysis shows that, indeed, when considering networks with five routers or more, around half of them have routers of one vendor and around 75% with up to two vendors. When we consider larger networks with more than 20 or 100 routers, we notice that there is only a vendor for about half and a quarter of networks, respectively. For large networks, i.e., more than 1,000 routers, LFP typically identifies multiple vendors. This is to be expected as large networks offer multiple services that may require specialized routers from various vendors. Even a few routers from different vendors can contribute to the heterogeneity in terms of router vendor per AS.

A.2 Regional Characteristics

We then study regional characteristics of deployment of router vendors and their global market share. In Figure 21 we report the number of routers we can identify with LFP per continent and vendor. The router is assigned to a location based on the headquarters location of the host network. Our analysis shows that with LFP, we can double the routers that are identified in all continents. Overall, the market is very consolidated. A small number of router manufacturers are responsible for more than 95% of the routers in a continent. We notice that in western regions like Oceania, North America (NA), and Europe, the penetration of Cisco is very high, with 81.7%, 70.3%, and 63.2%, respectively. Cisco also has around 64% of the market share in Africa. Huawei has a substantial market share in Asia and South America, with 40.6% and 36.3%, respectively. Juniper has a significant market share in North America, more than 17%.

In Europe and Asia, the additional contribution of LFP when compared with the SNMPv3-based fingerprinting is 100%, i.e., half of the routers could not be identified with the SNMPv3-based fingerprinting. In North America (NA) and South America, the additional contribution of LFP is around 87% and 76%, respectively. In North America, one of the reasons is that many Cisco routers can already be identified with the SNMPv3-based technique and routers in North America is predominantly Cisco. For South America, the reason is that Huawei already has a strong presence there, and it can be identified with SNMPv3. The highest additional contribution of LFP when compared with the SNMPv3-based fingerprinting is in the two regions with the lower number of identified routers, namely, Oceania and Africa, with 205% and 141%, respectively.

Finally, when we turn our attention to the top-13 networks in terms of the number of routers that we can identify with LFP, we notice that they are spread around the globe, see Figure 22. We also notice that the additional contribution in identifying routers with LFP compared to SNMPv3-based fingerprinting varies across the different networks. Indeed, for the top network in Asia, the increase is more than 100%, but for others, e.g., the third one also in Asia, the additional contribution of LFP is almost negligible.

Table 8: Precision and Recall: data random split (80/20)

Vendor	Recall	Precision	Total (test)
Cisco	0.99	0.99	6,754
Mikrotik	0.99	1.0	919
Juniper	0.97	0.99	789
Huawei	0.96	0.98	450
Brocade	0.64	0.72	153
H3C	0.20	0.23	123
Nokia	0.8	1	64
Ruijie	0.77	1	10
Ericsson	0.77	1	10
net-snmp	0.35	0.37	315

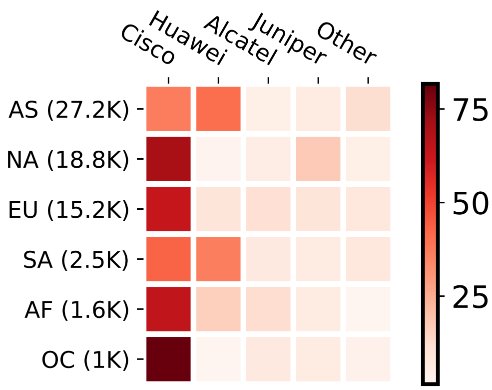

Figure 21: Router vendor popularity per continent.

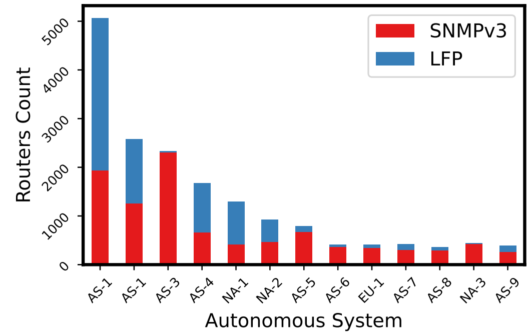

Figure 22: Additional contribution of LFP on router vendor identification in large networks.

B PRECISION AND RECALL

For the labeled RIPE-6 data with SNMPv3 information, we perform a precision and recall study. We do a 80/20 random split where we use 80% of the data for training and the other 20% for testing. The results per vendor for precision and recall are presented in Table 8. For the major vendors, namely, Cisco, Juniper, and Huawei, both precision and recall are very high, close to 1. Precision is also high for popular vendors, e.g., Nokia, Ruijiem, and Ericsson, but the recall is lower. We attribute this to the relatively low testing sample. The precision and especially recall is very low for UNIX-based vendors, e.g., net-snmp, Brocade, and H3C.