Towards Understanding Middlebox Deployments in Dutch ASes: Impact of IP Sampling Size

Bulut Ulukapi
University of Twente

Anna Sperotto
University of Twente

Ralph Holz
University of Münster

SYNOPSIS

Middleboxes are pervasive network elements that modify, delay, or drop network packets. Understanding their deployment within Autonomous Systems (ASes) enhances transparency, secure operations, and informed routing. However, detecting middleboxes at scale remains challenging when scanning network prefixes, due to the limited coverage of address space per prefix. In this preliminary study, part of a broader investigation into middlebox deployment in Dutch ASes, we use Yarrpbox to examine how the number of sampled IPs per prefix affects detection. By varying per-prefix sampling, we evaluate the resulting number of detected middleboxes, their network locations, and reachability. Results show larger samples yield more detections and higher responsiveness, with diminishing returns beyond 600 IPs per prefix. Moreover, our initial analysis reveals 56% of middleboxes are located in hosting providers.

CCS CONCEPTS

• **Networks** → **Middle boxes / network appliances**; **Network measurement**; *Network management*.

KEYWORDS

middlebox detection and analysis, internet measurement

ACM Reference Format:
Bulut Ulukapi, Anna Sperotto, and Ralph Holz. 2025. Towards Understanding Middlebox Deployments in Dutch ASes: Impact of IP Sampling Size. In *Applied Networking Research Workshop (ANRW '25), July 22, 2025, Madrid, Spain.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3744200.3744765

1 INTRODUCTION

One critical aspect of Autonomous System (AS) infrastructure is the deployment of middleboxes, which significantly influence traffic management within and across networks. These devices go beyond standard IP forwarding [1] by performing tasks such as packet rewriting, filtering, traffic shaping, and inspection. While they enhance security and performance, they can also cause protocol interference, reduce transparency, and complicate network management. Despite their importance, middlebox deployment patterns, particularly across AS types, remain underexplored.

A key methodological challenge in middlebox detection is the selection of IP addresses sampled within network prefixes, as the sampling strategy directly affects detection coverage. Although large-scale studies [2, 4, 5] advanced our understanding of middlebox prevalence, recent studies [5] relied on fixed sampling rates , typically one IP address for each /24 prefix, leaving unanswered the question of how varying the IP sampling size might influence middlebox detection accuracy and coverage while scanning prefixes.

In this preliminary study, drawn from a broader ongoing effort to investigate and categorize middlebox functions and deployments in Dutch ASes, we study how the number of sampled IPs per prefix influences middlebox detection effectiveness using active measurements with Yarrpbox. Our experiments are scoped to Dutch ASes and to Yarrpbox-based active probing. Additionally, to provide an initial view on the distribution of middleboxes across AS types, we analyze high-confidence middleboxes using IPinfo dataset [6], which provides IP-to-company mappings. Our analysis reveals that approximately 56% of these middleboxes are located within hosting provider networks. However, given that IPinfo provides only coarse-grained categories, we highlight the need for more accurate and tailored AS classifications—including categories such as critical infrastructure, governmental and financial institutions—to enable more reliable comparisons of middlebox deployments across network types.

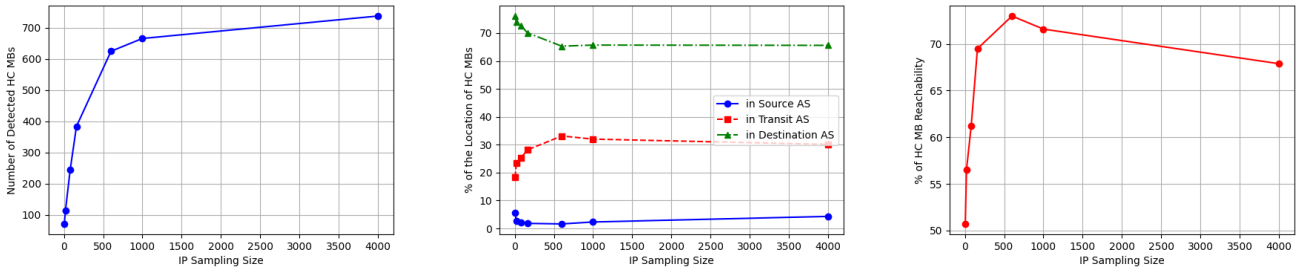
2 APPROACH

Data Collection: We used the RIPE API [7] to collect 1584 ASNs registered in the Netherlands and their associated 14579 IP prefixes. To evaluate the impact of IP coverage on middlebox detection, we systematically varied the number of sampled IPs per prefix.

To select IPs for probing, we first used ZMap [3] to scan port 80 across the collected prefixes and identify responsive hosts. We then sampled IPs from the responsive set within

Applied Networking Research Workshop (ANRW '25), July 22, 2025, Madrid, Spain, 2025.

ACM ISBN 979-8-4007-2009-3/2025/07
https://doi.org/10.1145/3744200.3744765



(a) On detected high-confidence middlebox number

(b) Location of detected high-confidence middleboxes.

(c) Reachability of detected high-confidence middleboxes.

Figure 1: IP sampling size vs. middlebox detection.

each prefix. If too few responsive IPs were found, we supplemented the sample with randomly chosen unique IPs from the same prefix.

We tested seven sampling sizes per prefix: 5, 20, 80, 160, 600, 1000, and 4000 IPs. This resulted in $\approx 73k$, $\approx 292k$, $\approx 1.2M$, $\approx 2.4M$, $\approx 5.1M$, $\approx 6.5M$, and $\approx 10.3M$, total target IP addresses, respectively. We should note that the maximum sampling size applied to each prefix was bounded by its address space, with smaller prefixes receiving proportionally fewer IPs.

Middlebox Detection: We used Yarrpbox [5] to probe the sampled IP addresses. By default, it probes one IP per /24 prefix when given prefixes as input, but coverage can be increased by supplying a pre-generated list of IPs instead. All scans were run from a single host at our university in Europe. We conducted preliminary experiments with scan rates of 500, 1000, 2000, 5000, and 10,000 packets per second. Detection counts ranged between 55 and 71, with 5000 packets per second offering the highest distinct middlebox coverage while maintaining response reliability and avoiding rate-limiting or filtering. We therefore adopted this scan rate for the remainder of the study. Our scans were performed on 9 April 2025, utilizing TCP SYN probes directed at port 80.

AS Usage Types: Following the identification of high-confidence middleboxes, we mapped them to their respective Autonomous Systems (ASes) using ASN data from RIPE. To classify the AS usage types, we used IPinfo[6] data, which provides four categories: ISP, hosting, business, and education.

3 RESULTS AND DISCUSSION

Effect of IP Sampling Size On General Detection Statistics: Our results show a correlation between the size of IP sampling per prefix and the effectiveness of middlebox detection. As the number of sampled IPs increases (Fig. 1a) , both the total number of detected middleboxes and the number of high-confidence detections increase, though after the sample size of 600, the curve seems to stabilize. The middleboxes we discuss are high confidence middleboxes which are classified based on their "location error" [5]. The reachability of middleboxes (Fig. 1c), which is their responsiveness to ping , starts to decrease at sampling sizes greater than 600. Additionally, the number of detected middleboxes in transit ASes increases with larger IP sample sizes (Fig. 1b). This may suggest either that a broader set of IP probes results in interactions with a wider range of network paths, or that, although yarrpbox employs randomized, stateless probing to enhance coverage and minimize measurement bias, specific design aspects or probe patterns may still inadvertently introduce bias or overrepresent certain middleboxes. Explicit analysis of path diversity is needed to clarify this. Overall, larger sampling sizes improve middlebox detection coverage and visibility while targeting prefixes with yarrpbox.

Effect of IP Sampling Size On Detected Modifications: Across all sampling sizes, the most frequently detected middlebox-induced modifications to TCP headers were adding *NOP*, removal of the *Multipath Capable* flag, and modifying the *Urgent Pointer* and/or *Receiver Window*. These modifications alter packet headers in ways that can interfere with protocol operation by preventing the use of Multipath TCP, suppress features the endpoints negotiated, and potentially impact performance. As the sampling size increased, we also observed a rise in Sequence Number modifications. This may occur because some middleboxes affect only specific routing paths within a prefix[2], and thus small samples may miss these path-specific middlebox behaviors.

Investigating AS Usage Types: An initial analysis shows that 56.2% of detected middleboxes are in hosting provider networks, followed by ISPs (36.3%), business (4.8%), and education ASes (2.7%). These results suggest that hosting and ISP networks dominate middlebox deployments, likely for performance and security, but finer AS distinctions are needed to better understand their roles in specialized environments.

ACKNOWLEDGMENTS

We thank IPInfo for providing us with academic licence. This work is supported by the research project 'CATRIN' (NWA.1215.18.003) as part of the Dutch Research Council's (NWO) National Research Agenda (NWA) and partly financed by the Province of Gelderland and Centre for Safety & Digitalisation.

REFERENCES

[1] Brian Carpenter and Scott Brim. 2002. *Middleboxes: Taxonomy and issues.* Technical Report.

[2] Gregory Detal, Benjamin Hesmans, Olivier Bonaventure, Yves Vanaubel, and Benoit Donnet. 2013. Revealing middlebox interference with tracebox. In *Proceedings of the 2013 conference on Internet measurement conference.* 1–8.

[3] Zakir Durumeric, Eric Wustrow, and J Alex Halderman. 2013. ZMap: Fast Internet-wide scanning and its security applications. In *22nd USENIX Security Symposium.*

[4] Korian Edeline and Benoit Donnet. 2019. A bottom-up investigation of the transport-layer ossification. In *2019 Network Traffic Measurement and Analysis Conference (TMA).* IEEE, 169–176.

[5] Fahad Hilal and Oliver Gasser. 2023. Yarrpbox: Detecting Middleboxes at Internet-Scale. *Proc. ACM Netw.* 1, CoNEXT1, Article 4 (jul 2023), 23 pages. https://doi.org/10.1145/3595290

[6] IPinfo. 2025. IP to Company Database. https://ipinfo.io/products/ip-company-database

[7] RIPE Network Coordination Centre. 2025. RIPEstat Data API Documentation. https://stat.ripe.net/docs/02.data-api/ Accessed: 2025-04-09.