UNIVERSITY OF CALIFORNIA SAN DIEGO

Evaluating the Effectiveness of BGP Origin Validation Mechanisms Using Public Data Sources

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Computer Science

by

Chongyang “Ben” Du

Committee in charge:

Professor Kimberly Claffy, Co-Chair
Professor Alex C. Snoeren, Co-Chair
Professor Margaret E. Roberts
Professor Stefan Savage
Professor Cecilia Testart

2025

The Dissertation of Chongyang “Ben” Du is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2025

DEDICATION

This dissertation is dedicated to my four families: my parents, the Hoffmaster family, the Claffy & Dunnigan family, and the Izhikevich family. Each of you has profoundly shaped who I am, and I would not have reached this milestone without your support.

EPIGRAPH

"It's not possible."
"No. It's necessary."

— *Interstellar* (2014)

TABLE OF CONTENTS

Dissertation Approval Page iii

Dedication iv

Epigraph v

Table of Contents vi

List of Figures ix

List of Tables xi

Acknowledgements xii

Vita xv

Abstract of the Dissertation xvii

Introduction 1

Chapter 1 BGP Origin Validation Background and Relevant Public Datasets 6
1.1 Border Gateway Protocol 6
1.2 BGP Hijacking 9
1.3 Defenses against BGP hijacking 11
1.4 Internet Routing Registry 12
1.4.1 IRR database objects 13
1.4.2 Reliability issues in the IRR 13
1.5 Resource Public Key Infrastructure 14
1.5.1 RPKI weaknesses 16
1.6 Mutually Agreed Norms for Routing Security 18
1.6.1 MANRS limitations 19
1.7 Public Measurement Datasets and their Limitations 20
1.7.1 BGP datasets 20
1.7.2 Limited visibility in BGP data 21
1.7.3 IRR datasets 22
1.7.4 IRR incompleteness 22
1.7.5 RPKI datasets 22
1.7.6 RPKI ROV deployment uncertainty 23
1.7.7 Other datasets 24

Chapter 2 Reliability of Established Trust in the IRR 26
2.1 Overview 27
2.2 Related Work 29
2.3 Methodology 30

2.3.1 Inter-IRR consistency . . . 30
2.3.2 Consistency between IRR and RPKI . . . 32
2.3.3 Classification of ASes registered in IRR . . . 34
2.3.4 Irregular route object inference . . . 35
2.4 IRR Characteristics . . . 36
2.4.1 Inconsistency across IRR databases . . . 36
2.4.2 Overlap with BGP announcements . . . 36
2.5 Inconsistency between IRR and RPKI . . . 38
2.5.1 IRR consistency with RPKI . . . 39
2.5.2 Historical consistency between IRR and RPKI . . . 39
2.5.3 Causes of prefix length inconsistency . . . 44
2.5.4 Analysis of ASN inconsistency . . . 45
2.6 ASes behind IRR records inconsistent with RPKI . . . 47
2.7 Irregular Route objects . . . 49
2.7.1 RADB analysis . . . 49
2.7.2 ALTDB analysis . . . 51
2.8 Summary . . . 51

Chapter 3 Inadvertent Delegation of Trust to Malicious Actors . . . 53
3.1 Overview . . . 54
3.2 Related Work . . . 55
3.3 Definition of IP Leasing . . . 56
3.3.1 RIR address policy . . . 56
3.3.2 IP leasing business model . . . 57
3.4 Dataset . . . 59
3.5 Methodology . . . 59
3.5.1 Processing RIR WHOIS databases . . . 59
3.5.2 Inferring leased address space . . . 61
3.5.3 Curating evaluation dataset . . . 63
3.5.4 Limitations . . . 63
3.6 Results . . . 64
3.6.1 IP address leasing ecosystem . . . 64
3.6.2 Evaluating inference with brokered leases . . . 66
3.6.3 Potential abuse of leased prefixes . . . 68
3.6.4 Defense against abuse . . . 69
3.7 Summary . . . 70

Chapter 4 Adoption of Origin Validation Practices by MANRS Networks . . . 72
4.1 Overview . . . 73
4.2 Methodology . . . 75
4.2.1 MANRS participation . . . 76
4.2.2 MANRS conformance with Action 4 and Action 1 . . . 77
4.2.3 MANRS impact . . . 79
4.3 MANRS Ecosystem Characteristics . . . 80

4.4 MANRS Action 4: Prefix Origination Behavior 82
4.4.1 RPKI prefix validity 82
4.4.2 IRR prefix validity 85
4.4.3 AS level conformance to Action 4 85
4.4.4 Analyzing unconformant ASes 86
4.4.5 Conformance stability 88
4.4.6 Presence in RPKI of MANRS ASes 89
4.5 MANRS Action 1: Route Filtering Behavior 90
4.5.1 Propagation of RPKI-Invalid prefixes 90
4.5.2 Propagation of IRR-Invalid prefixes 91
4.5.3 AS level conformance to Action 1 92
4.5.4 MANRS RPKI filtering effectiveness 93
4.6 Discussion 94
4.7 Summary 95

Chapter 5 Future Partial Deployment Strategies of BGP Origin Validation 96
5.1 Overview 97
5.2 ROV Simulation Methodology 98
5.2.1 Metric to quantify ROV deployment impact 98
5.2.2 Simulation assumptions 99
5.2.3 Simulation construction 100
5.2.4 Approximating currently ROV-deploying ASes 102
5.2.5 AS rankings for selecting future ROV-deploying ASes 102
5.2.6 Validating simulation with in-the-wild RPKI-Invalid prefixes 104
5.3 Simulation Results 105
5.3.1 Simulation validation 105
5.3.2 Current ROV deployment does not sufficiently reduce RPKI-invalids propagation 107
5.3.3 Difference in effectiveness in using different AS Ranking metrics 109
5.3.4 U.S. FCC-Listed service providers offer limited marginal benefit 112
5.4 Limitations 114
5.5 Summary 115

Chapter 6 Conclusion 117

Bibliography 121

LIST OF FIGURES

Figure 1.1. Sub-prefix hijack scenario where AS66 attempts to hijack 1.1.0.0/24 from legitimate origin AS1. To prevent propagation, AS3 must perform Route Origin Validation (ROV) and there must be a ROA authorizing only AS1 to originate 1.1.0.0/16. 16

Figure 1.2. MANRS has 4 programs and each program includes a set of mandatory and recommended actions. We focus on the ISP (Network Operator) and CDN/Cloud Provider categories, and the two actions in each category related to route registration and route filtering. 18

Figure 1.3. Example topology where AS5 originates an RPKI-invalid prefix 7.0.0.0/24, and a BGP collector at AS1 attempts to observe it. Whether AS1 sees the announcement depends on the ROV deployment status of AS2–AS4. 23

Figure 2.1. Structure of the chapter. Methodology sections are in blue and results sections are in green. 28

Figure 2.2. Number of v4 and v6 IRR records in RADB, RIPE, APNIC, and AFRINIC IRR databases. 33

Figure 2.3. Inconsistency between IRR databases. For example, when comparing RIPE IRR to ARIN IRR, 104 route objects in RIPE have corresponding route objects in ARIN with the same prefix, and 60% of those have no matching origin ASes. 37

Figure 2.4. In May 2023, 13 of 17 IRR databases had more RPKI-consistent route objects than RPKI-inconsistent ones. 39

Figure 2.5. RADB IPv6 records were more consistent with RPKI than IPv4 records. . 40

Figure 2.6. In IRR databases maintained by the RIRs, the number of consistent records was at least 10 times the number of inconsistent records within the same RIR. 42

Figure 2.7. BGP sometimes agrees with IRR records but not RPKI, showing that some IRR records in the Inconsistent Length category may actually be correct. . 44

Figure 2.8. Sometimes inconsistent ASN records are registered more recently than their corresponding consistent records. 46

Figure 3.1. An entity (lessee) can follow this two-step pipeline to establish Internet presence without having their own IP addresses or even network infrastructure. 58

Figure 3.2. Inference diagram of an example leased prefix. The bold orange rectangle marks the leased prefix. The leased prefix's BGP origin (AS15169) is related to neither the ASN assigned to its address provider (AS8851) nor the BGP origin of the portable parent prefix (also AS8851). 62

Figure 3.3. The RPKI and BGP behaviors of an IPXO-leased prefix suggests its periods of lease. 69

Figure 4.1. Structure of the chapter. Methodology sections are in blue and results sections are in green. 74

Figure 4.2. MANRS participants have grown significantly between 2015 and 2022, especially after 2019. 80

Figure 4.3. MANRS ASes and share of routed IPv4 address space over time. 81

Figure 4.4. Prefixes originated by MANRS and non-MANRS ASes (shared legend). Large MANRS ASes were more likely to originate RPKI Valid prefixes than IRR Valid prefixes in May 2022. 83

Figure 4.5. Percentage of RPKI covered address space for MANRS and non-MANRS networks. RPKI covered address space of MANRS networks grew faster between 2019 and 2022. 89

Figure 4.6. RPKI and IRR Invalid prefixes propagated by MANRS and non-MANRS networks in May 2022. 90

Figure 4.7. Unconformant prefixes propagated by MANRS and non-MANRS networks by network type. The median large MANRS AS propagated only 2.5% MANRS-unconformant prefixes. 92

Figure 4.8. Prefix preference for MANRS transit by RPKI status (0 means no preference). RPKI Invalid BGP announcements were less likely to propagate through MANRS networks than non-MANRS networks. 94

Figure 5.1. Example AS-level topology and the prefixes originated by each AS. 101

Figure 5.2. Comparison of deployment strategies using AS-level metrics. 109

Figure 5.3. Comparison of deployment strategies using IP-level metrics. 110

Figure 5.4. Impact of adding FCC-listed significant service providers and other top ranked U.S. ASes as ROV-deploying ASes. The median polluter could reach 9,617 U.S.-based ASes with only RoVista ASes deploying ROV, compared to 9,561 when the additional 16 FCC SSPs were included. 113

LIST OF TABLES

Table 1.1. Glossary of Terms Related to BGP Origin Validation 7

Table 2.1. Sizes of IRR databases grew between November 2021 and May 2023. ARIN-NA and RIPE-NA are ARIN-nonauth and RIPE-nonauth, respectively. ARIN-nonauth, OPENFACE, and RGNET databases were retired before May 2023. 30

Table 2.2. RPKI coverage of IPv4 and IPv6 address space expanded almost 10 fold. . . 33

Table 2.3. IRR overlap with BGP. The middle column shows the number of route objects present between November 2021 and May 2023. The right column shows the percentage of route objects that had the same prefix and origin AS in BGP over the same 1.5-year period. 38

Table 2.4. Classification of ASes that registered IPv4 records in RADB. MANRS ASes were more consistent than other ASes. 48

Table 2.5. Classification of ASes that registered IPv4 records in RIPE IRR. Most ASes were in the RIPE region and were highly consistent. 48

Table 2.6. Number of unique prefixes in each step of filtering potentially irregular IRR objects in RADB. 49

Table 3.1. Number of prefixes in each group per RIR according to the inference process (§3.5.2). The 47k inferred leased prefixes were 4.1% of all routed prefixes in April 2024. 64

Table 3.2. Top 3 IP holders by # of inferred leases, April 2024. 65

Table 3.3. Confusion matrix: we evaluated our results against 14,856 validated prefixes. 67

Table 4.1. Number of ASes originating RPKI Invalid prefixes by size category. In May 2022, MANRS ASes of all size categories were more likely to originate RPKI Invalid announcements compared to non-MANRS ASes. 84

Table 4.2. Non-conformant prefix origins from the MANRS networks we contacted. . 87

Table 4.3. Action 1 (filtering) conformance. ASes with no customers are conformant by default and included in right side columns. Results for transit ASes are shown in left side columns. Only 23% of small MANRS ASes provided transit and 97.1% of them were conformant. 93

Table 5.1. Mapping of significant service providers listed by the FCC to their primary AS numbers. 112

ACKNOWLEDGEMENTS

I would first like to acknowledge my outstanding advisors, KC Claffy and Alex Snoeren, for their invaluable guidance in both my research and career. I am also grateful to Amogh Dhamdhere, with whom I worked as an undergraduate research assistant, for teaching me foundational research skills. I thank my co-authors, Cecilia Testart and Romain Fontugne, for their collaboration and for staying awake with me during the long nights before the 5am paper submission deadlines.

I would like to thank Professor Liz Izhikevich for her invaluable advice on writing papers, statements, and proposals, as well as for helping me practice conference presentations and seminar talks. I look forward to continuing our work together at UCLA.

I am also thankful to Stefan Savage, whose engaging undergraduate computer security course (CSE 127) sparked my initial interest in network security research, and to Geoff Voelker, who forwarded KC's email to the CSE mailing list and led me to joining her lab. I also appreciate his generosity in later assigning me an office space in 3140, where I spent a summer "squatting", and giving me the invaluable opportunity to *network* with others.

I would like to acknowledge Dr. Audrey Arinello-Randall. She has been an amazing cohort who not only supported my research besides my advisors but also dragged me into trying all kinds of sports. I would also like to formally thank her for sponsoring the catering for my defense.

I am also grateful to Victor Vianu. I had the honor to serve as a teaching assistant for his undergraduate database class (CSE 132A) since 2018. Working alongside him and engaging with students was both rewarding and fun.

I would also like to thank Professor Emma Benser, who greatly helped me prepare for my first course as an instructor and made lecture material more engaging. I also appreciated the time we spent together watching videos on how to prepare for the academic job market.

I would like to thank Alex Yen for accompanying me to countless cafés and boba shops, which helped me make it through the writing of this dissertation.

I would like to thank Niharika Bhaskar, whom I had the privilege of advising as an undergraduate researcher in my lab and later came to know as a friend. I am grateful to her for keeping me updated on what the younger generation is up to these days.

I wish to especially thank Katherine Izhikevich, who has been my distinguished research consultant, event co-organizer, tour guide, taxi driver, delivery person, artist, connoisseur, xxxxxx xxxx, xxxx xxxxx, pastry chef, botanist, champion, and best friend. I would not have been able to make CSE as fun and chaotic without her support.

Finally, I would like to thank all my friends and colleagues, including but not limited to Dr. Amanda Tomlinson, Dr. Zesen Zhang, Dr. Gautam Akiwate, Dr. Keegan Ryan, Dr. Stewart Grant, Dr. Anil Yelam, Dr. Ariana Mirian, Dr. Enze Liu, Dr. Nishant Bhaskar, Sumanth Rao, Alisha Ukani, Luoxi Meng, Max Gao, Shivani Hariprasad, Elisa Luo, Alex Bellon, Eric Mugnier, Paul Chung, Ye Shu, Seoyoung Kweon, Yibo Wei, and Ximi.

Chapter 2, in part, is a reprint of the material as it appears in *Proceedings of Passive and Active Measurement Conference 2022.* Ben Du, Gautam Akiwate, Thomas Krenc, Cecilia Testart, Alexander Marder, Bradley Huffaker, Alex C. Snoeren, and kc claffy. The dissertation author was the primary investigator and author of this paper. This chapter also includes material as it appears in *Proceedings of the Internet Measurement Conference 2023.* Ben Du, Katherine Izhikevich, Sumanth Rao, Gautam Akiwate, Cecilia Testart, Alex C. Snoeren, and kc claffy. The dissertation author was the primary investigator and author of this paper.

Chapter 3 in part, is a reprint of the material as it appears in *Proceedings of the Internet Measurement Conference 2024.* Ben Du, Romain Fontugne, Cecilia Testart, Alex C. Snoeren, and kc claffy. The dissertation author was the primary investigator and author of this paper.

Chapter 4 in part, is a reprint of the material as it appears in *Proceedings of the Internet Measurement Conference 2022.* Ben Du, Cecilia Testart, Romain Fontugne, Gautam Akiwate, Alex C. Snoeren, and kc claffy. The dissertation author was the primary investigator and author of this paper.

Chapter 5 in part, is a reprint of the material as it was submitted to *International Confer-*

ence on emerging Networking EXperiments and Technologies (CoNEXT) 2025 Ben Du, Shivani Hariprasad, Romain Fontugne, Cecilia Testart, Alex C. Snoeren, and kc claffy. The dissertation author was the primary investigator and author of this paper.

VITA

2018	Bachelor of Science, University of California San Diego
2020	Master of Science, University of California San Diego
2025	Doctor of Philosophy, University of California San Diego

PUBLICATIONS

"RIPE IPmap Active Geolocation: Mechanism and Evaluation." Ben Du, Massimo Candela, Bradley Huffaker, Alex C. Snoeren, and kc claffy. ACM SIGCOMM Computer Communication Review (CCR), vol. 50, no. 2, pp. 3–10. May 2020.

"IRR Hygiene in the RPKI Era." Ben Du, Gautam Akiwate, Thomas Krenc, Cecilia Testart, Alexander Marder, Bradley Huffaker, Alex C. Snoeren, and kc claffy. Passive and Active Measurement (PAM) Conference, Virtual, March 2022.

"Stop, DROP, and ROA: Effectiveness of Defenses through the Lens of DROP." Leo Oliver, Gautam Akiwate, Matthew Luckie, Ben Du, and kc claffy. ACM Internet Measurement Conference (IMC), Nice, France, October 2022.

"Mind Your MANRS: Measuring the MANRS Ecosystem." Ben Du, Cecilia Testart, Romain Fontugne, Gautam Akiwate, Alex C. Snoeren, and kc claffy. ACM Internet Measurement Conference (IMC), Nice, France, October 2022.

"Poster: Taking the Low Road: How RPKI Invalids Propagate." Ben Du, Cecilia Testart, Romain Fontugne, Alex C. Snoeren, and kc claffy. ACM SIGCOMM Conference. New York, NY. September 2023.

"IRRegularities in the Internet Routing Registry." Ben Du, Katherine Izhikevich, Sumanth Rao, Gautam Akiwate, Cecilia Testart, Alex C. Snoeren, and kc claffy. ACM Internet Measurement Conference (IMC), Montréal, Canada, October 2023.

"Sublet Your Subnet: Inferring IP Leasing in the Wild." Ben Du, Romain Fontugne, Cecilia Testart, Alex C. Snoeren, and kc claffy. ACM Internet Measurement Conference (IMC), Madrid, Spain, November 2024.

"From Scarcity to Opportunity: Examining Abuse of the IPv4 Leasing Market." Bernhard Degen, Ben Du, Ricky K. P. Mok, Raffaele Sommese, Mattijs Jonker, Roland van Rijswijk-Deij, and kc claffy. Network Traffic Measurement and Analysis Conference (TMA), Copenhagen, Denmark, June 2025.

"Locating and Enumerating Anycast: A Comparison of Two Approaches". Remi Hendriks, Tim Betzer, Ben Du, Raffaele Sommese, Mattijs Jonker, and Roland van Rijswijk-Deij. Applied Networking Research Workshop (ANRW), Madrid, Spain, July 2025.

ABSTRACT OF THE DISSERTATION

Evaluating the Effectiveness of BGP Origin Validation Mechanisms Using Public Data Sources

by

Chongyang “Ben” Du

Doctor of Philosophy in Computer Science

University of California San Diego, 2025

Professor Kimberly Claffy, Co-Chair
Professor Alex C. Snoeren, Co-Chair

A key vulnerability of the Border Gateway Protocol (BGP) is its lack of origin validation, which allows attackers to falsely advertise IP address space they do not legitimately control, exposing victims to significant risks including service disruption, data theft, and financial loss. To mitigate such threats, the Internet community has developed origin validation mechanisms that operate in parallel with BGP to serve as an external source of trust. These systems rely on authoritative databases where networks voluntarily register their route origins and use the published information to verify BGP messages they receive. However, the effectiveness of these systems remains uncertain, as evolving usage and operational constraints may undermine the

protection they could provide for the routing system. Evaluating such effectiveness is challenging, especially without access to how authoritative databases are operated or incorporated into routing operations by networks. Nonetheless, the public nature of these databases and the external observability of BGP behavior provide a useful basis for assessing the deployment and limitations of these mechanisms.

This dissertation demonstrates that it is possible to evaluate the effectiveness of BGP origin validation mechanisms using public measurement data. I present four studies that use such data to examine key aspects of BGP origin validation mechanisms. First, I assess the integrity of authoritative databases by identifying outdated and attacker-falsified records. Second, I analyze how adversaries have exploited the emerging IP address leasing market in ways that undermine the effectiveness of origin validation. Third, I evaluate the adoption of origin validation mechanisms among networks participating in the MANRS initiative. Finally, I simulate partial deployment strategies to identify high-priority networks for targeted BGP origin validation deployment. These analyses reveal weaknesses in how trust is established, delegated, and adopted in the systems network operators use to validate BGP announcements. My findings offer the networking community actionable insights to improve the reliability and deployment of these origin validation mechanisms, and provide stakeholders with guidance to prioritize the most impactful networks for advancing routing security. Together, these studies present a repeatable methodology for continuously evaluating the effectiveness of origin validation mechanisms and monitoring their evolution over time.

Introduction

The Internet connects billions of users worldwide through a global routing system operated by a diverse set of organizations. These organizations, including Internet service providers (ISPs), cloud providers, educational institutions, and other companies, independently manage their own networks, known as Autonomous Systems (ASes). ASes exchange reachability information using the Border Gateway Protocol (BGP), which serves as the backbone of the interdomain routing system. However, BGP lacks a built-in mechanism to verify the authenticity of propagated routing information [124]. BGP assumes by default that the routing information received from neighboring routers is correct and, in turn, propagates that information to the receiving networks' neighbors.

This lack of validation gives attackers the opportunity to conduct BGP hijacking: an AS can falsely announce that it hosts a block of IP addresses (a prefix) that it does not have the authority to use (*a.k.a.* origin hijack), or claim to be on the path to a prefix, even though it has no legitimate route to that prefix (*a.k.a.* path hijack). As a result, attackers can impersonate other networks and intercept sensitive traffic, leaving the global routing system exposed to both accidental and intentional compromises. Over the past decade, malicious BGP hijacks have disrupted major Internet services [125], facilitated spam campaigns [148, 186], and resulted in substantial financial losses [112, 118, 161]. Meanwhile, accidental misconfigurations continue to cause widespread outages and instability [113, 170].

In response to this long-standing threat, several IETF-standardized protocols and research efforts have proposed solutions to prevent both origin and path hijacks [86, 87, 100, 143, 189]. While these approaches offer strong security guarantees, they introduce significant computational

and administrative overhead, as each router must perform additional cryptographic operations for every hop in the path. This complexity has hindered their real-world deployment [33].

Given the challenges of securing the entire path in BGP announcements, the Internet community shifted its focus to a narrower but more tractable goal: verifying whether the origin AS — the network that originates a prefix — is authorized to do so. Rather than modifying BGP itself, this form of validation relies on external systems to check whether an AS has legitimate control over the IP space it claims to originate, reducing the computational burden on routers. These systems follow a common pattern: first, an authoritative database must exist to record the legitimate origin AS for each IP prefix; second, networks must validate BGP announcements by checking them against this database and rejecting messages that do not match.

Over the years, the Internet community has utilized two primary frameworks to support this model. In the early 1990s, operators introduced the Internet Routing Registry (IRR) to facilitate the sharing of routing policy information [4]. Although the IRR was not originally designed for security purposes, operators used it as a source of information for origin validation. In contrast, the Resource Public Key Infrastructure (RPKI), developed in 2012, was explicitly designed to cryptographically authenticate the origin AS of BGP announcements against trusted sources [99]. In this dissertation, we refer to both systems as *BGP origin validation mechanisms*.

However, a critical question remains: how effective are these trust-based BGP origin validation mechanisms? Several factors can undermine their effectiveness. First, the reliability of the underlying databases is not guaranteed. For example, if authoritative databases contain inaccurate or outdated information, networks cannot use them to accurately filter incorrect BGP announcements. Worse, if attackers are able to inject malicious data into these databases, they can compromise both the data integrity and the security mechanisms that depend on them, facilitating successful BGP hijacks. Even when they cannot directly tamper with authoritative databases, attackers may exploit emerging business ecosystems, such as the IP address leasing market, to obtain legitimate authorizations and undermine the effectiveness of origin validation. Moreover, the adoption of these mechanisms is still evolving. If not enough networks actively

deploy these mechanisms, large portions of the routing system will still be vulnerable to BGP hijacking.

Evaluating the reliability and deployment of these mechanisms is challenging, especially without insider knowledge of how IRR and RPKI are operated by different database providers or incorporated into routing operations by individual networks. Likewise, the opaque nature of the IP address leasing market makes it difficult for independent researchers, who are not involved in the brokerage or transaction processes, to observe how malicious actors might exploit it. Despite these challenges, because trust-based BGP origin validation mechanisms are designed to share information publicly (e.g. IRR and RPKI records) and because routing behavior can be observed through BGP routing tables, public data sources offer a valuable opportunity to evaluate the real-world effectiveness and limitations of these mechanisms.

In this dissertation, I demonstrate that it is possible to use public measurement data to evaluate the effectiveness of trust-based BGP origin validation mechanisms. I empirically assess how trust is established, adopted, and delegated in these systems. Specifically, I use public IRR, RPKI, and BGP data in combination to: (1) identify inaccuracies and attacker manipulation in authoritative databases; (2) investigate how attackers can maliciously acquire origin authorization by exploiting the IP address leasing market. (3) assess the real-world deployment of BGP origin validation mechanisms; and (4) analyze partial deployment strategies to determine those that offer the greatest marginal benefit to the routing system.

I first evaluate the reliability of the IRR as a source of trusted routing information by assessing the accuracy of its published records. While network operators have used both IRR and RPKI as authoritative databases to obtain validated routing information, the IRR is significantly older and lacks consistent maintenance and validation standards, making it more prone to inaccuracies. In contrast, RPKI is a newer system that relies on cryptographic attestation of routing authorizations. By comparing records in the largest IRR database against corresponding entries in RPKI, I find that 62% of overlapping records were in conflict as of October 2021, highlighting the widespread presence of likely outdated or inaccurate information in the IRR.

Beyond general inaccuracies, I also identify thousands of suspicious IRR entries, many of which were created by networks previously involved in BGP hijacking, providing strong evidence of attacker-falsified records, which could diminish the effectiveness of IRR.

In contrast to the IRR, RPKI ensures that only legitimate parties can authorize ASes to originate their prefixes using cryptographic attestations. However, if these parties are unaware of who they are granting access to, they may unintentionally authorize malicious actors to use the address space for illicit purposes. To understand how this authorization process may be abused in the real world, I examine how attackers exploit RPKI's delegated trust model by leveraging emerging business practices, such as the IP address leasing ecosystem, to maliciously obtain authorization. I quantified the prevalence of IP address leasing and found that 4.1% of all advertised IPv4 prefixes were leased in April 2024. Notably, I found that leased prefixes were eight times more likely to have a valid ROA authorizing an AS known for abusive or malicious activity, indicating that attackers can undermine the effectiveness of RPKI by exploiting the IP leasing market.

Despite the above limitations, IRR and RPKI remain the primary systems that operators use for BGP origin validation. However, these systems cannot be effective unless networks actively deploy them. To promote broader adoption, the Mutually Agreed Norms for Routing Security (MANRS) initiative was launched, with participating networks committing to a set of recommended origin validation practices. To understand how well these commitments translate into practice, I assess the adoption of trust-based BGP security mechanisms by analyzing their deployment among networks participating in the Mutually Agreed Norms for Routing Security (MANRS) initiative. As part of their commitment, MANRS participants are expected to incorporate IRR or RPKI into their routing operations. However, MANRS does not enforce strict compliance with routing security requirements, allowing potential non-compliant participants to weaken its overall effectiveness. I find that, as of May 2022, over 17% of MANRS participants had not properly adopted routing security mechanisms, exposing a lack of accountability even among those formally committed organizations.

Despite MANRS's efforts to promote BGP origin validation practices, deployment remains limited. This lack of progress has drawn attention from government agencies seeking to strengthen routing infrastructure security through targeted regulations. To help policymakers and stakeholders make empirically informed decisions about improving routing security, I analyze future deployment strategies for RPKI using empirical data. I develop a routing topology analysis technique that builds on inferred AS relationships from public data to simulate and evaluate partial RPKI deployment strategies. Based on public routing data collected as of April 2025, I find that the set of significant services providers currently targeted by the U.S. Federal Communications Commission (FCC) for routing security adoption provides only limited benefit to U.S networks and the global Internet. My analysis also shows that RPKI deployment is still in its early stages, where targeting large and well-connected networks brings substantial and initially increasing marginal benefits, though these gains begin to level off as deployment expands further.

The remainder of this dissertation is organized as follows. Chapter 1 provides background on how BGP operates, the vulnerabilities that make it susceptible to hijacking, and the trust-based defenses developed to mitigate these threats. Chapter 2 presents my analysis of inconsistencies between IRR and RPKI records and identifies suspicious IRR entries likely falsified by attackers. Chapter 3 investigates the IP address leasing ecosystem and its implications for routing security, particularly how attackers could abuse it to acquire delegated trust. Chapter 4 introduces a measurement framework to evaluate the routing behavior of MANRS participants, characterizing the networks involved and assessing their adherence to routing security commitments. Chapter 5 describes my simulation-based evaluation of future RPKI deployment strategies and their effectiveness in reducing the propagation of incorrect or malicious routing information. Finally, Chapter 6 summarizes the dissertation, highlights key takeaways and actionable insights for network operators and other stakeholders, and outlines directions for future research.

Chapter 1

BGP Origin Validation Background and Relevant Public Datasets

This chapter provides background on BGP origin validation mechanisms and introduces the public data sources that support the analyses in this dissertation. I first review how BGP functions and how it is susceptible to hijacking. I then describe the Internet Routing Registry (IRR) and the Resource Public Key Infrastructure (RPKI), which we collectively refer to as *BGP origin validation mechanisms* throughout this dissertation. I also introduce the Mutually Agreed Norms on Routing Security (MANRS) initiative and explain its role in promoting origin validation. Finally, I present the best publicly available data sources I use to evaluate the effectiveness of BGP origin validation mechanisms.

1.1 Border Gateway Protocol

The Border Gateway Protocol (BGP)[149] is the de facto interdomain routing protocol that enables networks across the Internet, known as Autonomous Systems (ASes), to exchange reachability information. Each AS is an independent network under the control of an administrative entity, such as an Internet Service Provider (ISP), cloud provider, enterprise, or university. BGP allows each AS to advertise a set of destination IP addresses in the form of IP prefixes, enabling other networks to route traffic toward those addresses. This exchange of reachability information forms the foundation of the global Internet routing system.

Table 1.1. Glossary of Terms Related to BGP Origin Validation

Term	Definition
IP address holder	An organization or entity that has been allocated or assigned a block of IP addresses by a Regional Internet Registry (RIR).
IP prefix	A block of contiguous IP addresses, typically represented in CIDR notation (e.g., `192.0.2.0/24`).
Autonomous System (AS)	A collection of networks and routers under the control of a single organization that presents a unified routing policy to the Internet.
Regional Internet Registry (RIR)	One of five organizations (e.g., ARIN, RIPE NCC, APNIC, LACNIC, AFRINIC) responsible for managing IP address and AS number allocations within specific geographic regions.
Border Gateway Protocol (BGP)	The de facto interdomain routing protocol used to exchange routing information between Autonomous Systems (ASes) on the Internet.
Origin AS	The Autonomous System (AS) that is listed as the origin in a BGP route announcement, indicating where the route originated.
BGP announcement / route / advertisement	A BGP message that advertises reachability to an IP prefix, including the origin AS and AS path.
BGP origin hijacking	A malicious or accidental event in which a network falsely announces IP prefixes it does not legitimately control, causing traffic misdirection or disruption.
BGP origin validation	The process of verifying that the origin AS in a BGP announcement is authorized to advertise the associated IP prefix.
Internet Routing Registry (IRR)	A distributed database system where networks publish their routing policies and prefix-origin information, often used to generate BGP filtering rules.
Resource Public Key Infrastructure (RPKI)	A cryptographic framework that allows IP address holders to authorize specific ASes to originate their prefixes in BGP.
Route Origin Authorization (ROA)	A digitally signed object in RPKI that authorizes a specific AS to originate a particular IP prefix (or set of prefixes) in BGP.
Validated ROA Payload (VRP)	A data structure derived from validated ROAs. It contains the IP prefixes, authorized origin ASes, and maximum prefix lengths used for Route Origin Validation (ROV).
Route Origin Validation (ROV)	The process by which BGP routers use VRPs to validate incoming BGP announcements and assign them a status: Valid, Invalid, or Not Found.
Relying Party (RP)	Software that retrieves and cryptographically validates RPKI data (including ROAs) and compiles the resulting VRPs for use in ROV.

BGP operates as a path-vector protocol, meaning that each route advertisement includes the IP prefix being announced and the sequence of ASes (the AS path) through which the announcement has propagated. This path information serves two functions: (1) it prevents routing loops by allowing ASes to reject routes that already contain their own AS number, and (2) it enables ASes to make route selection decisions based on local routing policy, such as preferring routes with the shortest path or avoiding certain transit providers. While shortest AS path is a common route preference policy, other attributes, such as local preference, may influence the decision. After selecting the best path, the AS exports it to a subset of its neighbors according to its export policy.

When forwarding traffic, routers apply the longest-prefix match rule: if multiple prefixes cover a destination IP address, the route advertisement with the most specific prefix (i.e., the longest prefix length) is selected. This rule ensures traffic is directed to the most precise destination, which is essential because the more specific prefix may belong to a different AS than its covering prefix.

Because multiple ASes can independently announce the same prefix, the same prefix may appear with multiple origin ASes in the global routing table. This phenomenon is known as a multi-origin AS (MOAS) conflict. MOAS events can occur for legitimate reasons, such as when a large organization operates multiple ASes and announces the same prefix from different geographic locations (*e.g.,* for redundancy or traffic engineering). However, MOAS can also result from misconfigurations or malicious activity, such as origin hijacking.

For example, suppose we are an AS that peers with Google and we receive the BGP route from Google shown below. UCSD (AS7377, the *origin AS*) advertises the prefix 137.110.0.0/16 to its neighbor CENIC (AS2152), which then propagates the route to Lumen (AS3356), and then to Google (AS15169), and eventually to our AS. As a result, we learn that to reach any IP address within 137.110.0.0/16, we should forward traffic to AS15169.

```
137.110.0.0/16:  AS15169, AS3356, AS2152, AS7377
```

BGP lacks built-in mechanisms to verify the correctness of BGP advertisements received from neighbors. As a result, attackers can forge or tamper with various parts of a BGP advertisement, such as the IP prefix or any parts of the AS path, without detection.

1.2 BGP Hijacking

Over the past decade, malicious BGP hijacks have disrupted major Internet services [125, 185], enabled large-scale spam campaigns [148, 186], and caused substantial financial losses [35, 112, 118, 161]. At the same time, accidental misconfigurations continue to trigger widespread outages and routing instability [113, 170].

BGP hijacking occurs when an attacker modifies any parts of a BGP announcement to redirect Internet traffic without authorization. Sermpezis et al.[165] classified BGP hijacks to the following types based on the attribute modified by the attacker.

- Classification by modified IP prefix:

  (1) **Exact prefix hijack:** The attacker AS originates a routing entry for the prefix that is exactly the same as what the legitimate AS announced. Such an announcement will cause a multi-origin conflict (MOAS) for the originated prefix, meaning that the aforementioned prefix will have more than one origin AS. This type of hijack typically affects victim networks topologically close to the attacker because they are more likely to receive and prefer the attacker's announcement due to its shorter AS path.

  (2) **Sub-prefix hijack:** The attacker AS originates a prefix that is more specific than that originated by the legitimate prefix holder. An example of this hijack is the 2008 YouTube hijack event, YouTube originally announced its prefix 208.65.152.0/22. When Pakistan Telecom hijacked YouTube, it announced a path for YouTube's more-specific prefix 208.65.153.0/24 [125]. Since BGP routing follows the longest-prefix

matching rule, more-specific prefixes will always be preferred and passed along to adjacent ASes. Therefore sub-prefix hijacking can affect the entire Internet, and all traffic toward the hijacked prefix will be routed to the hijacker, and thus sub-prefix hijacking is more powerful than exact prefix hijacking [165].

- Classification by modified AS Path:

  (1) **Origin hijack:** The attacker originates a prefix it is not authorized to announce (a.k.a. hijacks the prefix). The origin AS in this BGP announcement is the attacker's AS.

  (2) **Path hijack:** An attacker may manipulate the AS path while keeping the legitimate route origin unchanged [73]. The attacker AS can place itself on any hop along the AS path, acting as a man-in-the-middle. Unlike an origin hijack, a path hijack can be more stealthy, as the attacker can continue forwarding traffic to the legitimate destination, allowing interception or surveillance without the victim noticing any immediate connectivity issues.

Overall, attackers can exploit BGP hijacking to achieve several malicious goals, including disrupting Internet services, eavesdropping on sensitive communications, or stealing cryptocurrency.

(1) **Denial of Service (DoS):** Attackers can render a prefix unreachable to large portions of the Internet, also known as *blackholing* [135]. This is typically done by falsely originating the victim's prefix (origin hijack) and dropping all traffic sent to it.

(2) **Impersonation:** By announcing a prefix they do not own (origin hijack), attackers can assume the identity of the legitimate owner to conduct illicit activities. For example, they may launch phishing attacks by intercepting traffic intended for the victim and serving a spoofed version of their website [77]. BGP hijacks are also commonly used in spam campaigns. Spammers operating from a small set of IP addresses risk being detected

and blocklisted – networks that receive spam from those addresses may report them to public blocklists, causing other networks to reject future email originating from the same address space. To evade such blocklisting, spammers need a large supply of fresh, unused IP addresses. One tactic is to briefly hijack unannounced prefixes, send spam during the short window before detection, and then withdraw the hijacked routes to avoid long-term consequences [148].

(3) **Interception:** In a path hijack, adversaries can redirect traffic to themselves and then forward it to the legitimate destination, performing a man-in-the-middle attack. This enables packet inspection and potential theft of sensitive information, often without detection [14]. Interception attacks have targeted blockchain infrastructure. For instance, attackers have hijacked Bitcoin-related prefixes to delay block propagation, wasting mining effort and enabling double-spending attacks by disrupting network consensus [118].

1.3 Defenses against BGP hijacking

In response to the long-standing threat of BGP hijacking, several IETF-standardized protocols and research efforts have proposed solutions to prevent both origin and path hijacks [86, 87, 100, 143, 189]. These proposals extend BGP with cryptographic mechanisms to validate the entire AS path, aiming to ensure the authenticity and integrity of routing announcements. However, such validation requires an additional cryptographic operation at each hop in the path, introducing significant computational and administrative overhead. This level of complexity has hindered the widespread adoption of these proposed protocols over the past decade [33]. As securing the entire AS path by modifying BGP saw little progress, network operators shifted their focus to securing only the origin AS using verification mechanisms external to BGP. This led to the promotion of the Internet Routing Registry (IRR), a non-cryptographic system not originally designed for origin validation but widely adopted by network operators for that purpose, and the Resource Public Key Infrastructure (RPKI), which was specifically developed

to cryptographically attest to the validity of the origin AS for a given BGP prefix. In this dissertation, we refer to both IRR and RPKI as *BGP origin validation mechanisms*.

1.4 Internet Routing Registry

The IRR, first introduced in 1995 as a routing policy storage system [17], was designed to facilitate sharing of routing policies across networks. Operators use information in the IRR for several purposes, such as building inbound BGP filters and validating BGP origins. Transit providers [11, 110], cloud providers [65, 167] and Internet Exchange Points (IXPs)[8, 39] often require peers and customers to register their IP prefixes and AS numbers in IRR databases. These IRR records are used to construct prefix filters that determine which routes are accepted from each neighbor. By relying on IRR data, operators can automate the generation of prefix lists for different customers, help debug configuration errors, and filter out unauthorized or misconfigured BGP announcements [172].

Some networks allow their customers and peers to register in any IRR database, while others support only a subset like the ones they retrieve information from [110, 138].

The IRR comprises a distributed collection of independently operated databases. As of July 2025, there were 18 databases still in operation [120]. Two of the largest and oldest IRR databases are RADB [119], operated by Merit Network, and the RIPE IRR [152], operated by the RIPE NCC. Over the past few decades, newer IRR databases have emerged, operated by commercial companies, Regional Internet Registries (RIRs), and Local Internet Registries (LIRs). Despite differences in ownership and policies, the core purpose of all IRR databases has remained the same: publishing routing information that can be used by other networks to construct BGP route filters. Each IRR database is managed independently under different policies and registration processes. These databases generally fall into two categories:

- **Authoritative IRR databases:** The five Regional Internet Registries (RIRs) – RIPE, ARIN, APNIC, LACNIC, and AFRINIC – operate authoritative IRR databases. The RIRs

validate registered routing information against their own address ownership records to ensure correctness [154]. Each RIR's authoritative IRR database includes only address space that the RIR manages within its designated service region.

- **Non-authoritative IRR databases:** Managed by commercial networks (e.g. Lumen, NTT), non-profits (e.g. Merit), or other entities, these databases typically have less stringent validation compared to authoritative IRR databases or in some cases, no validation at all [85]. Some RIRs also offer non-authoritative versions of their IRR databases, either to provide coverage for address space not officially managed by the RIR [153] or to accommodate registrants who have not signed a formal service agreement [190].

1.4.1 IRR database objects

Networks can use the Routing Policy Specification Language (RPSL) [4] to register and retrieve routing policy information from IRR databases. RPSL formats objects as lists of attribute-value pairs. The `route` and `route6` objects are particular significance to BGP origin validation. The example `route` object below shows that `AS7377` intends to announce `137.110.0.0/16`, and that UCSD is the maintainer of this object.

```
route:  137.110.0.0/16
origin:  AS7377
mnt-by:  UCSD
```

1.4.2 Reliability issues in the IRR

Despite being the most widely deployed system used by network operators for BGP origin validation, IRR have several weaknesses. First, the IRR depends on voluntary (although sometimes contractually required) contribution of routing information by network operators. Unfortunately, some Internet Service Providers (ISPs) are reluctant to share their routing policies to avoid leaking sensitive business information [50], and thus avoid using the IRR at all. Second,

information in some the IRR databases is not strictly — and sometimes not at all — validated by the operator of the IRR database to check whether the registrant is authorized to use the specified prefix. As such, the accuracy of IRR information is not guaranteed [88]. Third, there is a lack of coordination among different IRR databases. If a network registers routing information in multiple databases but only updates one of them, inconsistencies can arise across the IRR system [95]. As a result, operators who query multiple databases may retrieve conflicting entries and be unable to determine which one most accurately reflects the registrant's current routing intent. We study these IRR weaknesses in detail in Chapter 2.

We are not aware of any effort to resolve inconsistencies across IRR databases, nor a standardized procedure to validate whether a registered AS in the IRR has the authorization to announce its registered prefixes. While attempts to create variations of the IRR by adding validation mechanisms have been proposed, none have ever gained traction [62, 91]. After years of debate, a significant set of stakeholders including many ISPs agreed on an alternative to improving routing security known as the Resource Public Key Infrastructure (RPKI).

1.5 Resource Public Key Infrastructure

The Resource Public Key Infrastructure (RPKI) [99] was introduced in 2012 to allow networks to cryptographically validate the origin AS in BGP advertisements. Unlike certain IRR databases, which allow users to register routing information with limited checks on whether they are actually authorized to use the specified IP addresses, RPKI provides stronger guarantees. It binds IP addresses and AS numbers to public keys through cryptographic certificates issued by the RIRs only to legitimate address space holders. Each of the five RIRs operates as the root of trust, a.k.a. *trust anchor*, for its corresponding service region. The IP address space managed by each RIR is mutually disjoint, ensuring there is no duplication and thus no conflicting information across registries. RPKI requires two complementary actions from different parties to be effective:

ROA registration: IP address holders need to obtain certificates from their RIRs or ISPs

to cryptographically sign Route Origin Authorization (ROA) objects, which authorize specific ASes to originate their IP prefixes in BGP. There are currently two RPKI certificate hosting models: *hosted RPKI* and *delegated RPKI*. In hosted RPKI, The RIRs host Certificate Authority (CA) certificates and sign Route Origin Authorizations (ROAs) for the IP address space and AS numbers of their registered members. In delegated RPKI, RIR members such as ISPs can host their own CA certificates to sign ROAs for their own or their customers' address space.

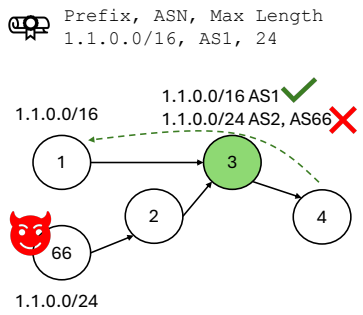
Once ROAs are registered by network operators, other networks can use this information to perform BGP origin validation and filter out route announcements that do not match the authorized origin AS, a process called Route Origin Validation (ROV) [122].

Route Origin Validation: An ROV-deploying network uses Relying Party (RP) software to download ROAs and relevant certificates from the RPKI repositories, check the validity of the certificate chain, and generate a list of validated ROA payloads (VRPs). The most important fields of a VRP are IP prefix, ASN, and max length. Inside a VRP, `IP Prefix` specifies the IPv4 or IPv6 address space held by the network. `ASN` specifies the AS number used to announce the `IP Prefix` in BGP. `Max Length` specifies the length of the most specific subprefix of the `IP Prefix` authorized in BGP by the IP address holder. The example VRP below shows that `AS7377` is authorized to announce in BGP `137.110.0.0/16` and any subprefix whose length does not exceed `20`.

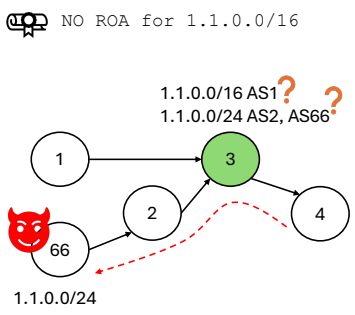
```
(IP Prefix, ASN, Max Length)
(137.110.0.0/16, AS7377, 20)
```

The VRPs are then communicated to the BGP router via the RPKI-to-router (RTR) protocol [21], who can then validate BGP announcements against covering VRPs, i.e., those whose IP prefix contains the prefix in the BGP announcement. The resulting *RPKI status* of a BGP announcement is one of the following:

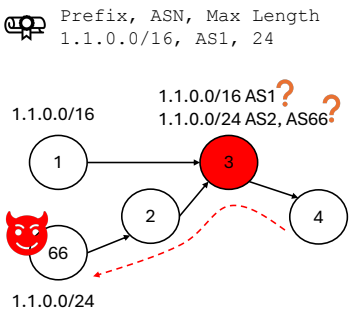
- *Not found*: The BGP prefix has no covering VRP.



(a) AS4 is protected from the hijack if 1.1.0.0/16 is covered by a ROA and AS3 enforces ROV.



(b) AS4 is vulnerable if no ROA exists for the hijacked prefix, even if AS3 performs ROV.



(c) AS4 is also vulnerable if AS3 does not perform ROV, even if a valid ROA exists for the prefix.

Figure 1.1. Sub-prefix hijack scenario where AS66 attempts to hijack 1.1.0.0/24 from legitimate origin AS1. To prevent propagation, AS3 must perform Route Origin Validation (ROV) and there must be a ROA authorizing only AS1 to originate 1.1.0.0/16.

- *Invalid ASN*: No `ASN` in any covering VRP matches the BGP origin AS.
- *Invalid prefix length*: The BGP origin AS is valid but the announced prefix is more specific than allowed by the covering VRP's `max length`. (Sometimes combined with *invalid ASN*.)
- *Valid*: the BGP prefix and origin AS agree with those in a covering VRP.

BGP routers should reject announcements that invalid (either invalid ASN or invalid prefix length) or adjust their preference to the route accordingly [122, 134]. In this process, the cryptographic validation of RPKI certificates is performed by the RP software independent from the router. As a result, this process does not impose significant computational overhead on the router itself, making it feasible for many routers to support BGP origin validation [31].

1.5.1 RPKI weaknesses

Since only address space holders can create Route Origin Authorizations (ROAs) or delegate ROA-signing authority to others, RPKI ensures that registered records more accurately reflect legitimate control, resulting in better data quality than the IRR. However, attackers may

still obtain authorization if address space holders are unaware of whom they are delegating to. We examine this potential abuse scenario in Chapter 3.

Although the RPKI deployment process is standardized and modern network equipment supports Route Origin Validation (ROV), additional challenges arise because of the configuration and operation of relying party software[93]. Some operators are concerned that RPKI centralizes too much control in the hands of the RIRs [37, 71, 168], and these concerns are compounded by legal risks [168, 191] and business uncertainties [187]. These additional cost and complexity of using these databases has hindered operator participation [95, 173, 181, 182]. Moreover, deploying ROV does not protect a network's own address space from being hijacked by others; it only filters invalid announcements from others. As a result, the incentives for operators to invest in RPKI-based countermeasures are not well aligned with the benefits, limiting broader adoption [33].

Furthermore, RPKI requires coordinated actions from multiple parties to effectively prevent a BGP hijack from propagating on the Internet: (1) the address space owner must create a ROA for the hijacked prefix, and (2) *other* networks must enforce ROV to filter out invalid announcements. If the hijacked prefix is not covered by any ROA, an ROV-deploying network cannot validate its legitimacy and will propagate it by default, resulting in traffic being misdirected to the attacker (Figure 1.1b). Without ROAs, ROV provides no filtering benefit, which diminishes its utility and can discourage deployment. Conversely, if the hijacked prefix is covered by a ROA but no networks along the path enforce ROV, the hijack can still succeed (Figure 1.1c), which may disincentivize address holders from registering ROAs.

To address this lack of incentive to implement RPKI, security-aware industry members launched the Mutually Agreed Norms on Routing Security (MANRS) initiative to encourage companies to implement routing security best practices including registering ROAs and validating their customers' BGP announcements.

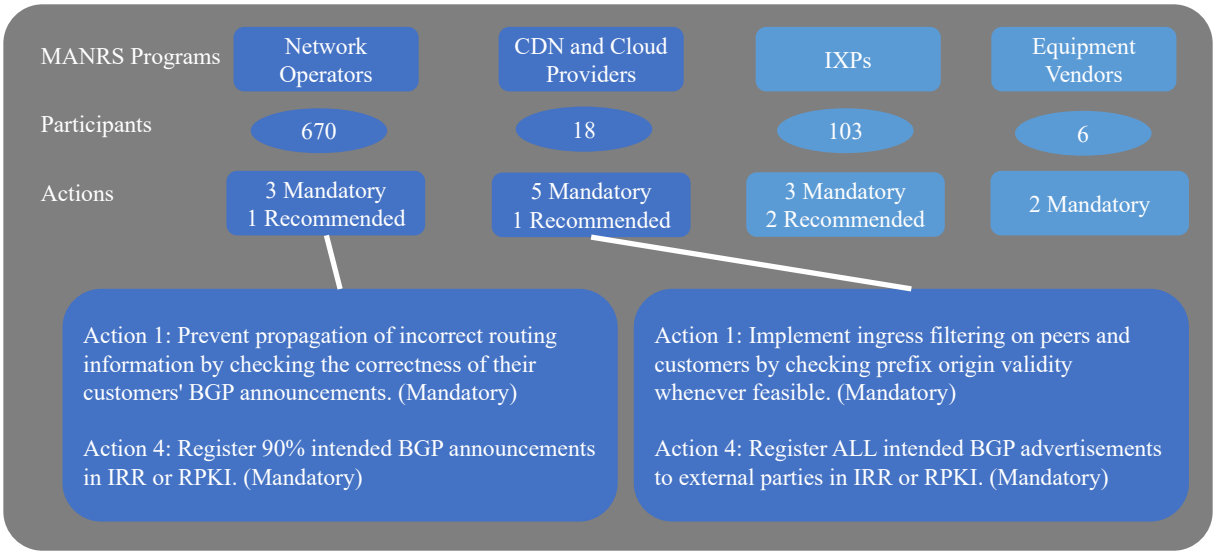


Figure 1.2. MANRS has 4 programs and each program includes a set of mandatory and recommended actions. We focus on the ISP (Network Operator) and CDN/Cloud Provider categories, and the two actions in each category related to route registration and route filtering.

1.6 Mutually Agreed Norms for Routing Security

The Mutually Agreed Norms for Routing Security (MANRS) initiative launched in 2014 and has had steady growth in membership ever since [158]. MANRS currently provides four programs with different security action sets [116] intended for different types of network operators (Figure 1.2): Internet service (transit) providers; content distribution networks (CDNs), including cloud providers; Internet exchange points (IXPs); and equipment vendors. We focus on the first two categories in this dissertation and summarize their action sets below.

MANRS for Network Operators (ISP)

- *Action 1*: Prevent propagation of incorrect routing information by checking the validity of the ISP customers' BGP announcements.
- *Action 2 (optional)*: Filter outbound traffic with spoofed source IP and run the CAIDA Spoofer software [108] to prevent DDoS attack traffic from being originated from the participant's network.
- *Action 3*: Maintain up-to-date network contact information in IRR databases or PeeringDB,

a database where IXP networks sharing interconnection policies and contact information.

- *Action 4*: Register intended BGP announcements in IRR or RPKI. Using RPKI is recommended.

MANRS for CDN and cloud providers

- *Action 1*: Implement ingress filtering on peers and customers by checking the validity of BGP announcements using IRR/RPKI.
- *Action 2*: Same as Action 2 for ISPs, but mandatory for CDNs.
- *Action 3*: Same as Action 3 for ISPs.
- *Action 4*: Register all intended BGP advertisements to external parties in IRR or RPKI.
- *Action 5*: Encourage peers to adopt MANRS.
- *Action 6 (optional)*: Provide monitoring tools to peers.

1.6.1 MANRS limitations

MANRS has not taken on rigorous enforcement of these best practices. MANRS provides some aggregated statistics from external sources [117] but declines to publicly detail non-conformance. (Instead, MANRS provides individual operators with private monthly reports regarding conformance.) Previous work has found inconsistent adherence of MANRS members to specific recommendations, e.g., the Source Address Validation action (Recommended Action #2) [108], but we are not aware of an independent study of MANRS members' overall conformance with MANRS actions. Lack of empirical assessment of conformance frustrates a collective understanding of how effective MANRS is at achieving its objectives. I use public data sources to assess the conformance level of MANRS networks in Chapter 4.

This lack of understanding of the MANRS's network's behaviors and still slow progress in the uptake in RPKI deployment recently prompted the U.S. Federal Communications Commission (FCC) to launch a Notice of Inquiry to ascertain whether there was a role for government intervention to improve routing security in 2023 [36] and proposed reporting requirements for

service providers in 2024 [51]. The White House and the U.S. National Institute of Standard and Technology (NIST) recommended routing security practices for organizations [140, 183]. In Chapter 5, I analyze future deployment strategies for BGP origin validation mechanism to help policymakers and stakeholders make empirically informed decisions about improving routing security.

1.7 Public Measurement Datasets and their Limitations

This section describes the datasets I use in the dissertation to evaluate the effectiveness of trust-based BGP origin validation mechanisms from multiple perspectives. Because BGP is a globally distributed protocol operated by thousands of autonomous networks, collecting global routing tables is infeasible for independent researchers, as one would need to operate a BGP-speaking network and establish peering relationships with a large and diverse set of other networks. Such an operational and logistical barrier is out of reach for most independent researchers. Instead, I rely on publicly available datasets collected by existing large-scale measurement infrastructures. These datasets provide valuable visibility into BGP routing behavior. However, they are best-effort in nature and subject to several limitations, including incomplete vantage point coverage and limited temporal granularity. Despite these trade-offs, these public datasets remain the best available tools to conduct meaningful empirical analysis without requiring access to internal network operations.

1.7.1 BGP datasets

The RouteViews [139] and RIPE RIS [127] projects are the two primary efforts that collect and publish BGP data for researchers. Both projects operate a globally distributed network of collectors, often deployed at Internet Exchange Points (IXPs), which establish BGP sessions with volunteer networks (peers). Through these peers, the collectors receive and archive BGP update and withdrawal messages, as well as periodic routing table snapshots, providing a comprehensive record of Internet routing activity. Several datasets derived from these archives

are used throughout this dissertation.

Prefix2AS dataset. CAIDA collects the daily snapshots of routing tables of from the `routeviews2` route collector and extract the announced IP prefix and origin AS. This data is published daily as the Routeviews Prefix to AS mappings Dataset [25]. Below is an example entry in the Prefix2AS dataset:

```
(Prefix, ASN, Prefix Length)
(137.110.0.0, AS7377, 16)
```

AS Relationship dataset. Luckie et al. [109] introduced a widely used dataset of AS business relationships, inferred from BGP data collected by RouteViews and RIPE RIS. Their methodology leverages the observation of AS paths in BGP routing tables to classify the links between ASes as either customer-provider or peering relationships. The dataset is periodically updated and made publicly available, providing one of the most comprehensive views of inter-AS connectivity on the Internet. In this dissertation, we use the AS links from this dataset to construct a global Internet topology graph.

1.7.2 Limited visibility in BGP data

Current BGP research relies on primarily two projects that collect and publish BGP data [127, 139]. Each project operates collectors that collects BGP routes from ASes that peers with at least one of those collectors (a.k.a vantage points). However, only 1% of ASes on the Internet peer with those collectors [6], creating an incomplete view of the routing system [5–7, 166]. A constructed AS-level topology of the Internet rely on public BGP data can miss peering connections [141]. Additionally, In BGP, an AS only exports a set of filtered routes to its neighbors, and thus a collector peer receives only a subset of all routes from its neighbors [149], hiding potential links between ASes. Additionally, since AS relationship data is derived from BGP data, it inherits these limitations [109, 145].

1.7.3 IRR datasets

IRR archives. Merit Network, the operator of the largest IRR database, maintains a list of current IRR databases in operation [120]. The listed IRR providers publish daily snapshots of their IRR databases on their corresponding FTP or HTTP servers.

1.7.4 IRR incompleteness

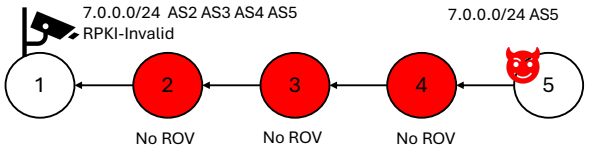
The historical data for IRR database is incomplete as RADB is the only database provider that provides historical snapshots. Although RADB mirrors entries from other IRR databases, its public archives only include records that were directly registered in RADB itself.

Also, daily IRR snapshots lacks the temporal resolution needed to capture short-lived or rapidly changing entries. IRR records can change hourly as networks update their configurations in response to operational needs.

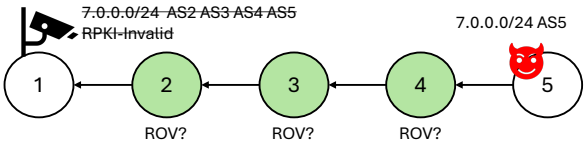
1.7.5 RPKI datasets

RPKI archives. The RIPE NCC publishes daily validated ROA payloads from all five RPKI trust anchors (APNIC, ARIN, RIPE NCC, AFRINIC, LACNIC) since, even after the retirement of their RPKI Validator[156, 157].

Researchers have used these archives to study the evolution of RPKI over the past decade. In 2017, Wählisch et al. [187] found that large CDNs and website hosting providers were reluctant to register ROAs in RPKI, limiting its adoption. By 2019, Chung et al. [31] observed a significant increase in IPv4 address space covered by ROAs since 2012, along with a decline in misconfigurations over time, indicating positive trends in RPKI ROA registration. In 2024, Testart *et al.* [182] found that small networks required additional support such as operational training to adopt ROAs, while the complexity of address delegation can discourage full ROA adoption in larger networks. I use RPKI archives to evaluate the accuracy of IRR databases in Chapter 2, identify maliciously obtained RPKI authorizations in Chapter 3, and assess MANRS networks' adoption of RPKI in Chapter 4.



(a) If AS1 observes the RPKI-invalid prefix, then all of AS2, AS3, and AS4 have propagated the invalid announcement and **none** of them has deployed ROV.



(b) If AS1 does not observe the prefix, then at least one of AS2, AS3, or AS4 filters it via ROV, but we cannot pinpoint which one(s).

Figure 1.3. Example topology where AS5 originates an RPKI-invalid prefix 7.0.0.0/24, and a BGP collector at AS1 attempts to observe it. Whether AS1 sees the announcement depends on the ROV deployment status of AS2–AS4.

RoVista dataset. In 2023, Li et al. [102] developed RoVista, an IP-ID side channel technique to check whether IP packets were able to reach RPKI-invalid destinations and inferred networks that either deployed or fully protected by ROV. They published the resulting ROV-deploying networks as a dataset [178]. ROV inference techniques have significantly advanced over time. RoVista represents the most comprehensive effort to date and is the only project that publishes a public dataset. Earlier approaches, such as Reuter et al. [150] in 2017, originated both RPKI-valid and RPKI-invalid prefixes from the PEERING testbed and checked which prefixes reached BGP route collectors, and found that three ASes deployed ROV. In 2021, Rodday et al. [159] extended the above methodology by conducting traceroutes from RIPE Atlas probes to the controlled prefixes. Other studies leveraged RPKI-invalid prefixes in the wild to infer whether an AS deployed ROV: in 2017, Gilad et al. [57] found 3 ASes deployed ROV because they only propagated non-RPKI-invalid announcements from neighbors. Testart et al. [181] found increasing ROV deployment among large transit providers (which directly peer with Routeviews collectors) from 2018 to 2020. I use this dataset to assess the effectiveness of partial RPKI deployment strategies in Chapter 5.

1.7.6 RPKI ROV deployment uncertainty

Inferring whether a specific AS currently deploys Route Origin Validation (ROV) remains an open challenge. Existing techniques [57, 102, 150, 159] can only determine whether an AS

is *protected* by ROV, not whether it actively performs validation, due to limited visibility from available BGP vantage points. Figure 1.3 shows an example: if an RPKI-invalid prefix is observed, then none of the ASes along the path between the vantage point and the originator of the RPKI-invalid prefix deploys ROV. On the contrary, if any of the ASes along the paths deploys ROV, the vantage point will not observe the RPKI-invalid prefix, making it impossible to pinpoint which AS deploys ROV without additional vantage points.

1.7.7 Other datasets

We introduce several additional public datasets that are used throughout this dissertation to complement the BGP data collected by RouteViews and RIPE RIS. These datasets provide essential contextual information about the organizations that operate Autonomous Systems, their routing security commitments, and the topological importance of ASes in the global Internet.

AS2Org dataset. The CAIDA AS-to-Organization dataset [24] maps Autonomous Systems to their registered organizations and countries. This dataset is derived from IP address and ASN allocation records published by the Regional Internet Registries (RIRs) and augmented with heuristics to merge multiple ASNs belonging to the same organization. We use this dataset to map each AS to the responsible organization and to identify sibling ASes that are operated by the same entity.

MANRS dataset. The Mutually Agreed Norms for Routing Security (MANRS) initiative maintains public lists of its participating Internet Service Providers (ISPs) and Content Delivery Networks (CDNs) [116]. These lists include information such as organization names, AS numbers, geographic service areas, and the specific routing security actions each participant has committed to implement. We use this dataset to identify MANRS ASes and audit their conformance with their stated commitments in Chapter 4.

Internet Health Report and Internet Yellow Pages. The Internet Health Report (IHR) [84] and the Internet Yellow Pages (IYP) [53] are projects developed by IIJ (Internet Initiative Japan) Research Labs that aggregate a wide variety of data sources, including BGP,

IRR, RPKI, and IP allocation datasets, to provide a comprehensive view of Internet routing and security. We use two of their key modules in this dissertation: 1) the IHR-ROV module, which reports BGP prefixes along with their IRR and RPKI validation states, allowing us to study origin validation deployment at scale. 2) The AS Hegemony module, which provides the AS Hegemony metric—a measure of how central an AS is in terms of routing dependency. AS Hegemony quantifies how often an AS appears on BGP paths to other networks, both globally and within individual countries, and serves as an indicator of the topological importance of an AS.

Chapter 2

Reliability of Established Trust in the IRR

Although there is a more recent and secure alternative to the IRR, the Resource Public Key Infrastructure (RPKI), many networks have not yet deployed RPKI-based filtering due to its cost and complexity [31, 46, 48, 57, 187]. As a result, IRR-based filtering remains the most popular route-filtering metric, even for networks participating in the Mutually Agreed Norms for Routing Security (MANRS) routing security initiative [47, 116]. Unfortunately, IRR information may be unreliable due to its lack of validation of registered information and lack of coordination among different database operators (§1.4). This unreliability may limit the ability of networks to construct correct BGP route filters and, thus, diminish the effectiveness of IRR-based BGP origin validation.

In this chapter, I show that it is possible to examine the reliability of IRR databases using public data sources, which I leverage in the following ways. First, I analyze the inconsistencies between different IRR databases because IRR providers regularly publish database snapshots. Second, I examine the accuracy of IRR databases by comparing the information within against RPKI information, which is published by the RIPE NCC, one of the five RIRs. Third, I investigate the ASes associated with inconsistent records and infer root causes by examining the organizations that operate them using publicly available AS-to-organization and AS relationship data. Finally, I combine these analyses with public BGP data to identify suspicious IRR records that are likely registered by malicious actors.

2.1 Overview

Information in an IRR database may be inaccurate due to *improper hygiene*, where IP address space holders may fail to update origin information after changes in routing policy or prefix ownership, since there is no penalty for neglecting to do so [95]. The problem is exacerbated by the existence of many separate IRR databases from which network operators can choose. As of May 2023, there were 18 active IRR databases operated by different types of organizations (e.g., commercial, non-profit, and Internet registries). Due to the diversity of IRR databases and a lack of coordination among them, information can be inconsistent: networks may register routing information in multiple IRR databases but only maintain a subset of them [95, 179]. Moreover, networks that migrate to RPKI may stop maintaining their IRR records, increasing discrepancies between IRR and RPKI and widening the gap between routing decisions based on the two systems.

In addition to operator mismanagement, attackers may also exploit IRR vulnerabilities. Different IRR databases have different procedures to authenticate database users and validate information registered in the IRR — or sometimes no validation at all, and as a result, the quality of IRR databases varies. Because many large networks permit their customers or peers to register route objects in most IRR databases [8, 11, 39, 65, 110, 167], attackers register false records in a single database, particularly one with weak or nonexistent verification procedures. They can then falsely originate the prefix in BGP, and upstream providers, believing the IRR information to be legitimate, accept and propagate the hijacked prefix to the rest of the Internet. Recent news articles and blog entries [63, 64, 184] have reported increasingly more complicated methods to falsify IRR records that have caused significant financial loss to victims.

UCSD was directly affected by a notable case of IRR abuse involving RADB, the largest IRR database. In late 2020, UCSD received an abuse report indicating that AS207427 (`GoHosted.eu`) had hijacked three UCSD prefixes in BGP for approximately 45 days, spanning the end of 2020 and the beginning of 2021. A postmortem investigation revealed that the attacker

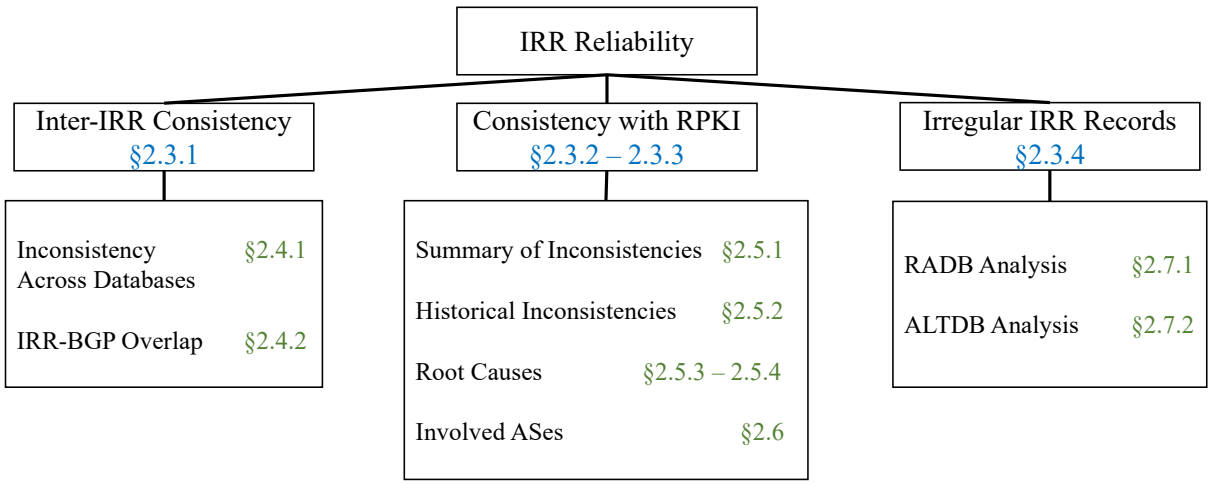


Figure 2.1. Structure of the chapter. Methodology sections are in blue and results sections are in green.

had registered `route` objects in RADB listing those UCSD prefixes with AS207427 as the origin AS. The upstream provider of AS207427 accepted and propagated the announcements to the rest of the Internet, as they appeared valid according to RADB records. RADB later removed the false `route` objects after being contacted by UCSD.

Our goal in this chapter is to systematically assess the reliability of the IRR system in light of the vulnerabilities described above. Figure 2.1 outlines the structure of our analysis. Given the lack of coordination among IRR databases, which can lead to conflicting records, we begin by quantifying inconsistencies across databases. Next, because inaccurate or outdated IRR information can undermine operators' ability to construct correct routing filters, we assess the accuracy of IRR records. Since there is no definitive ground truth for routing intent, we use RPKI as a reference point. RPKI records are cryptographically signed, can only be created by legitimate address holders, and include expiration dates, which makes them a more trustworthy source for identifying inconsistencies. To better understand the causes of IRR inaccuracies, we examine maintenance practices across ASes that registered inconsistent records and infer common causes of inaccuracy, such as complex business relationships and accidental misconfigurations. Finally, to understand how attackers have exploited these weaknesses, we identify IRR records that were likely falsified by malicious actors.

2.2 Related Work

Several existing studies have examined the accuracy of IRR data. Previous works have studied the consistency between IRR and BGP and found them to be largely consistent with each other. In 2013, Khan et al. [88] compared the prefixes registered in 14 IRR databases with BGP announcements and found 65% of IRR route objects exactly matched the prefix origin information in BGP. In 2004, Siganos and Faloutsos [171] compared network business relationships extracted from the IRR with those inferred from BGP data using the Gao-Rexford model [56], and found that 83% of the routing policies registered in the IRR were consistent with the relationships observed in BGP. In 2007, Siganos and Faloutsos [172] expanded the study to compare the consistency of BGP prefix-origins with authoritative IRRs. They concluded that RIPE was the best-maintained registry, with 73% of announced RIPE prefixes matching an existing registry record. However, their method matched maintainers of IRR `route` objects to maintainers of RIR WHOIS database records, which only works for IRR databases that are tightly coupled with their corresponding address ownership database (RIPE and APNIC databases at the time of their study).

In 2008, Sriram et al. [176] enhanced the algorithm by Siganos and Faloutsos to analyze the consistency between stored routing information and address ownership information across all authoritative IRRs and RADB. They found APNIC to be the most consistent and RADB the least. However, since RADB contains little address ownership information, such a comparison cannot effectively evaluate RADB's consistency.

To date, there has been no systematic effort to resolve inconsistencies across IRR databases, nor a standardized procedure to validate whether an AS registered in the IRR is actually authorized to announce its associated prefixes. These gaps highlight the need for an alternative approach to evaluate the reliability of RADB and other IRR databases.

Table 2.1. Sizes of IRR databases grew between November 2021 and May 2023. ARIN-NA and RIPE-NA are ARIN-nonauth and RIPE-nonauth, respectively. ARIN-nonauth, OPENFACE, and RGNET databases were retired before May 2023.

IRR	2021		2023	
	# Routes	% Addr Sp	# Routes	% Addr Sp
RADB	1,349,854	57.97	1,429,972	50.23
APNIC	608,319	7.97	654,677	8.35
RIPE	369,546	19.43	398,798	19.57
NTTCOM	451,143	10.25	380,938	9.84
AFRINIC	95,236	1.96	102,282	2.02
LEVEL3	91,563	8.95	77,939	7.89
ARIN	51,678	3.42	70,905	4.58
WCGDB	62,852	11.52	57,636	11.26
RIPE-NA	54,744	2.93	52,827	2.81
ALTDB	18,326	1.55	23,146	1.57
TC	8,353	0.12	18,010	0.17
JPIRR	11,540	4.14	12,932	4.30
LACNIC	5,789	0.74	11,074	1.02
IDNIC	4,594	0.21	5,721	0.21
BBOI	928	0.06	831	0.06
PANIX	40	0.00	40	0.00
NESTEGG	4	0.00	4	0.00
ARIN-NA	63,560	3.60	0	0.00
CANARIE	1,422	0.16	0	0.00
RGNET	43	0.00	0	0.00
OPENFACE	17	0.00	0	0.00

2.3 Methodology

In this section, we describe our process for analyzing inconsistencies across IRR databases, identifying mismatches between IRR and RPKI, investigating the underlying causes of these inconsistencies, and inferring IRR records that were likely registered by malicious actors.

2.3.1 Inter-IRR consistency

We use daily snapshots of all IRR databases we collected between November 2021 and May 2023 to conduct this analysis. In November 2021, we were able to access 21 IRR databases from 17 FTP servers, but only 18 databases in May 2023. Three IRR providers (`ARIN-NONAUTH`,

OPENFACE, RGNET) retired their databases during our data collection period and their listings have been removed [120]. CANARIE stopped responding to FTP requests before May 2023 but was still listed as active on irr.net. Table 2.1 describes the size of each IRR database in November 2021 and May 2023.

To analyze the consistency between IRR databases, we compare the route objects between every pair of IRR databases IRR^A and IRR^B as follows: Assuming we have a route object R^A in IRR^A consisting of prefix P^A and origin AS^A, we classify the route object in the following steps:

(1) Find in IRR^B a list of route objects $R_1^B \cdots R_n^B$ with prefixes $P_1^B \cdots P_n^B$ such that every prefix P_i^B is the same as P^A.

(2) If there does not exist such a list $R_1^B \cdots R_n^B$, then we classify R^A as **no overlap**.

(3) If there exists such a list $R_1^B \cdots R_n^B$ and AS^A equals any AS_i^B corresponding to R_i^B, then we consider R^A to be **consistent** with respect to IRR^B.

(4) If AS^A does not equal any AS_i^B from R_i^B, then we use the CAIDA as2org and AS Relationship dataset to check for potential sibling, customer-provider, or peering relationship between AS^A and AS_i^B. If such a relationship is found, then we also consider R^A to be **consistent** with respect to IRR^B.

(5) If none of the above criteria are satisfied, then we classify R^A to be **inconsistent**.

We then calculate the consistency of IRR^A with respect to IRR^B as the percentage of **consistent** objects all overlapping route objects.

BGP overlap. For each IRR database, we count the number of route objects with the exact same prefix and origin AS in BGP between November 2021 and May 2023 and calculate its percentage over all route objects to identify actively used IRR records.

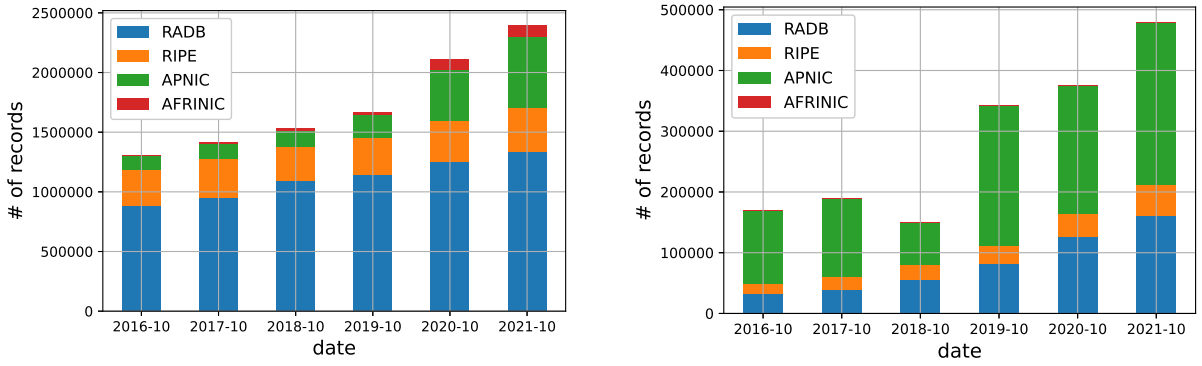
2.3.2 Consistency between IRR and RPKI

Since there is no definitive ground truth for routing intent data, it is difficult to directly verify the correctness of IRR entries. To address this, we approximate ground truth using RPKI records. RPKI is the only other global database of routing authorizations, and unlike the IRR, its records are cryptographically signed and can only be created by legitimate IP address holders. In addition, RPKI records have expiration dates, which reduce the likelihood of outdated or stale information persisting in the system. These properties make RPKI a more accurate and trustworthy source as an authoritative database, and therefore a reasonable reference point for identifying inaccuracies in IRR data (§1.5).

We first use the IRR datasets from the previous section to analyze RPKI inconsistencies for all IRR databases, and then perform a extended historical analysis of four IRR databases using archived snapshots obtained from one IRR provider and from CAIDA to include both IPv4 and IPv6 records. We use monthly snapshots of the RADB database from August 2016 to October 2021, which we refer to as the *RADB IRR dataset* in this chapter. In addition, we obtain quarterly snapshots of IRR records from RIPE, APNIC, and AFRINIC via their WHOIS databases, archived by CAIDA. These are referred to as the *RIPE*, *APNIC*, and *AFRINIC IRR datasets*, respectively. Collectively, we refer to these four sources as the *IRR datasets* and extract all `route` and `route6` objects for our analysis. Figure 2.2 provides an overview of these sources, including their growth over time. As of October 2021, RADB contained the most IPv4 route entries, while APNIC had the most for IPv6.

For comparison, we also use RPKI archive snapshots from the same period (August 2016 to October 2021), referred to as the *RPKI dataset*. Table 2.2 summarizes the number of validated ROA payloads (VRPs) for both IPv4 and IPv6. We use prefix origin pairs found in the extracted `route` and `route6` objects and classify them according to their consistency with records in the *RPKI dataset*.

To classify IRR records by their consistency with ROAs, we define the following four



(a) RADB has the most v4 records.

(b) APNIC has the most v6 records.

Figure 2.2. Number of v4 and v6 IRR records in RADB, RIPE, APNIC, and AFRINIC IRR databases.

Table 2.2. RPKI coverage of IPv4 and IPv6 address space expanded almost 10 fold.

Date	IPv4			IPv6		
	ROAs	Prefixes	ASNs	ROAs	Prefixes	ASNs
2016–10	23k	22k	3,874	3.5k	3.3k	1,911
2017–10	38k	35k	5,067	6.1k	5.3k	2,519
2018–10	50k	46k	6,465	9.0k	8.1k	3,370
2019–10	92k	84k	10,232	15k	14k	5,274
2020–10	160k	143k	16,276	26k	24k	8,651
2021–10	205k	185k	21,265	40k	37k	10,878

classes: Records that show full consistency in prefix and origin AS are classified **consistent**. When the ASN in an IRR record does not equal that of the ROA, the IRR record is classified as **inconsistent ASN**. When the ASNs are the same, but the prefix length differs, the IRR record is classified as **inconsistent length**. Otherwise, if there is no corresponding prefix in the ROA, the record is classified as **not in RPKI**. We apply the following algorithm, a modified version of Route Origin Validation [76] and similar to the IRR Explorer [133], to the *IRR datasets*.

(1) We choose snapshots of the same date between August 2016 and October 2021 and from the *IRR datasets* and the *RPKI dataset*.

(2) We denote each `route` or `route6` object as R_x. We denote the list of ROAs in the RPKI dataset as *ROALIST*

(3) For each record R_x, we denote the prefix as P_x and origin AS as AS_x.

(4) We look for exact matching prefixes or covering prefixes of P_x in *ROALIST*. The resulting list of candidate ROAs are denoted L_{ROA}

(5) If L_{ROA} is empty, then we put R_x in **not in RPKI**.

(6) For each candidate ROA, C_{ROA}, in L_{ROA}, we put C_{ROA} in a list, M_{ROA}, if the origin AS in C_{ROA} equals AS_x.

(7) If M_{ROA} is empty, then we classify R_x as **inconsistent ASN**.

(8) For each C_{ROA} in M_{ROA}, we put C_{ROA} in a final list, V_{ROA}, if the prefix length of P_x does not exceed MaxLength field in C_{ROA}.

(9) If V_{ROA} is empty, we classify R_x as **inconsistent length**, otherwise as **consistent**.

Using the example from §1, if the *RPKI dataset* contains ROA `(137.110.0.0/16, AS7377, 20)`, a record `(137.110.0.0/24, AS195)` from the *RADB IRR dataset* falls into **inconsistent ASN** while `(137.110.0.0/24, AS7377)` falls into **inconsistent length**.

2.3.3 Classification of ASes registered in IRR

To study geographical trends in how different ASes conform to IRR record maintenance practices, we discard all IRR records in the **Not in RPKI** category and classify the remaining ASes into three groups based on their IRR records:

1) An AS is **Entirely consistent (EC)** if all its IRR records are classified as **consistent**.

2) An AS is **Entirely inconsistent (EI)** if all associated records are classified as either **inconsistent ASN** or **inconsistent length**.

3) An AS is classified as **Mixed** if it is associated with both **consistent** and **inconsistent** IRR records.

2.3.4 Irregular route object inference

We outline the steps to infer route objects that may have been registered for malicious purposes.

Mismatching origin AS with authoritative IRRs. We consider the information in authoritative IRR databases to be more trustworthy than other IRRs (§ 1.4), so we classify a route object as irregular if there is a mismatch with its corresponding object in the authoritative IRR. We compare the route objects using the steps defined in § 2.3.1 where IRR^A is any non-authoritative IRR database and IRR^B is the combined 5 authoritative IRR databases. To address the possibility of networks creating ad-hoc registration of route objects containing more-specific prefixes in different IRRs for purposes such as traffic engineering, we change § 2.3.1 step (1) to the following: P_i^B is a covering prefix of P^A.

Matching IRR objects to BGP announcements. An adversary forging IRR records to circumvent filtering [142] would announce in BGP the prefix from their falsified route object to achieve their goal (e.g. origin hijacking). Therefore, for the route objects we classify as **inconsistent** above, we check whether their prefixes appeared in a BGP announcement during the same time period. We classify the **inconsistent** route objects into three categories:

(1) Route objects whose prefixes associated with the same set of ASNs in both the IRR dataset and BGP dataset are considered **full overlap**.

(2) Route objects whose prefixes associated with different sets of ASNs in IRR and BGP with partially overlapping ASNs are considered **partial overlap**. For example, if the following two IRR route objects $(P, AS1)$, $(P, AS2)$ corresponds to BGP announcements $(P, AS2)$, $(P, AS3)$, then $(P, AS2)$ is considered partial overlap. We classify route objects in this category as *irregular*.

(3) Route objects whose prefixes associated with disjoint sets of ASNs in the IRR and BGP are considered **no overlap**.

Validating irregular route objects. If a route object obtained from the steps above has a matching RPKI record in our *RPKI dataset*, we remove it from our irregular route object list. We also look for ASes from our irregular route objects in the *serial hijacker dataset* to identify irregular objects likely registered by serial hijackers.

2.4 IRR Characteristics

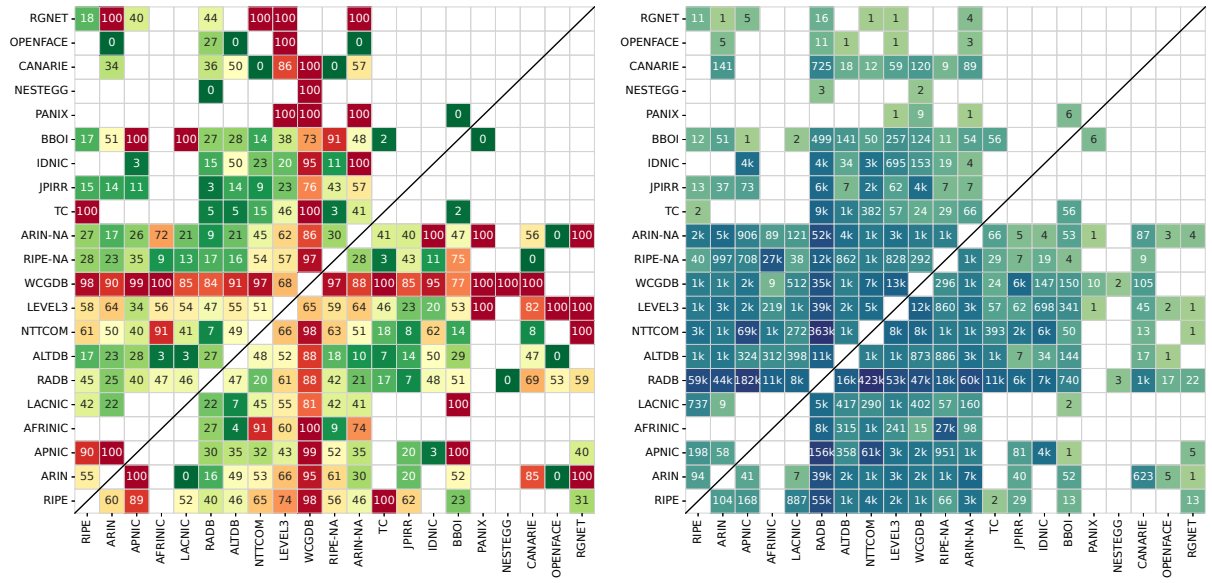
To understand the characteristics of the IRR databases, we analyze the inter-IRR consistency and calculate the overlap between BGP and IRR databases.

2.4.1 Inconsistency across IRR databases

Figure 2.3a shows the percentage of route objects with the same prefix but different origin ASes between pairs of IRRs. We find that most IRR databases have mismatching route objects with one other, consistent with persistent neglect by IRR users and thus an increasing number of outdated entries [88]. We also notice instances where a company registered route objects in multiple IRR databases, but only updated the records in one IRR database, causing inter-IRR inconsistency. Most surprising are the mismatching records between pairs of authoritative IRR databases (Figure 2.3b, lower left 5×5 grid), since each RIR only allows registration of route objects containing address blocks managed by that RIR, which do not overlap with each other. We find that those mismatching route objects correspond to address space that was transferred across RIRs, and the address owner from the previous RIR did not remove the outdated object.

2.4.2 Overlap with BGP announcements

We provide an updated analysis on IRR consistency with BGP since the previous study in 2013. We calculate the presence of route objects in BGP over the span of November 2021 and May 2023. Table 2 shows that ALTDB has more overlap with BGP, compared to RADB. Khan et al. [88] showed that in 2013 RADB had 65% overlap with BGP, significantly higher than the 29% we see in 2023. In contrast, the overlap between ALTDB and BGP increased from 50%



(a) Fraction of inconsistent route objects in the IRR on the Y-axis with respect to the IRR on the X-axis. The denominators are in Fig. 2.3b.

(b) Number of route objects from the IRRs on the Y-axis that have overlapping route objects in the IRR on the X-axis.

Figure 2.3. Inconsistency between IRR databases. For example, when comparing RIPE IRR to ARIN IRR, 104 route objects in RIPE have corresponding route objects in ARIN with the same prefix, and 60% of those have no matching origin ASes.

in 2013 [88] to 62% in 2023. This suggests that information in ALTDB was more current than RADB, consistent with the fact that ALTDB had a much higher RPKI consistency than RADB (99% vs 61% for route objects with a covering RPKI ROA).

We further study the inconsistencies between authoritative IRRs and BGP and find 6,163 (1.3% RIPE), 1,291 (1.5% ARIN), 2,804 (0.4% APNIC), 1,983 (1.9% AFRINIC), and 367 (2.7% LACNIC) route objects were inconsistent with BGP announcements that lasted more than 60 days. Although those long-lived inconsistencies may suggest inaccuracy of IRR records, it is also possible that those route objects did not cause operational harm as networks may have used those route objects along with more robust IRR filters (*e.g.,* AS-SET filtering [116]).

Table 2.3. IRR overlap with BGP. The middle column shows the number of route objects present between November 2021 and May 2023. The right column shows the percentage of route objects that had the same prefix and origin AS in BGP over the same 1.5-year period.

IRR	# Route Objects	% Route Objects in BGP
RADB	1,542,724	28.81% (444,479/1,542,724)
APNIC	681,879	17.82% (121,480/681,879)
RIPE	447,089	59.27% (265,002/447,089)
NTTCOM	465,555	14.85% (69,146/465,555)
AFRINIC	104,823	20.76% (21,759/104,823)
LEVEL3	92,673	24.23% (22,452/92,673)
ARIN	83,476	61.65% (51,467/83,476)
WCGDB	62,852	5.59% (3,514/62,852)
RIPE-NA	54,755	27.92% (15,289/54,755)
ALTDB	25,704	62.41% (16,043/25,704)
TC	19,366	77.2% (14,950/19,366)
JPIRR	13,504	67.49% (9,114/13,504)
LACNIC	13,486	72.67% (9,800/13,486)
IDNIC	5,912	67.12% (3,968/5,912)
BBOI	951	51.95% (494/951)
PANIX	40	15.0% (6/40)
NESTEGG	4	75.0% (3/4)
ARIN-NA	64,957	18.73% (12,169/64,957)
CANARIE	1,460	58.42% (853/1,460)
RGNET	44	47.73% (21/44)
OPENFACE	17	41.18% (7/17)

2.5 Inconsistency between IRR and RPKI

In this section, we present the results of our analysis of the consistency between IRR databases and RPKI. We begin by quantifying inconsistencies between IRR and RPKI across all 18 IRR databases to assess the overall accuracy of IRR and highlight the varying quality of different IRR providers. We then conduct a historical analysis of a subset of IRR databases to examine how and why inconsistencies arise over time, offering insight into the operational practices that contribute to outdated or incorrect records.

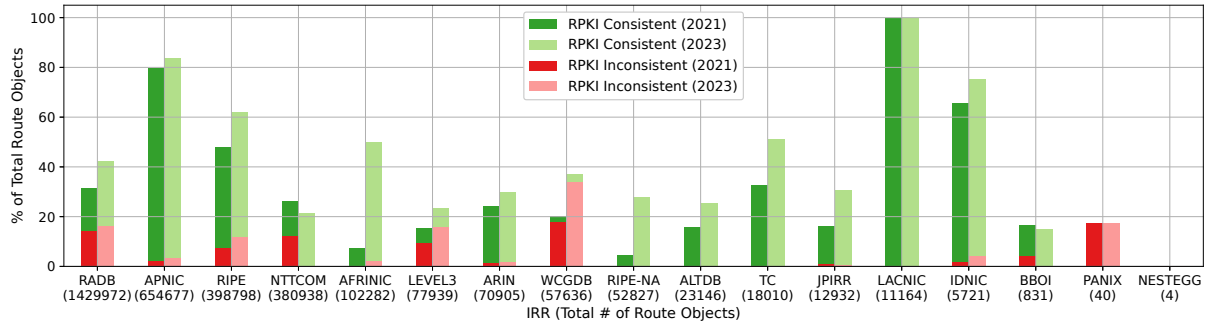


Figure 2.4. In May 2023, 13 of 17 IRR databases had more RPKI-consistent route objects than RPKI-inconsistent ones.

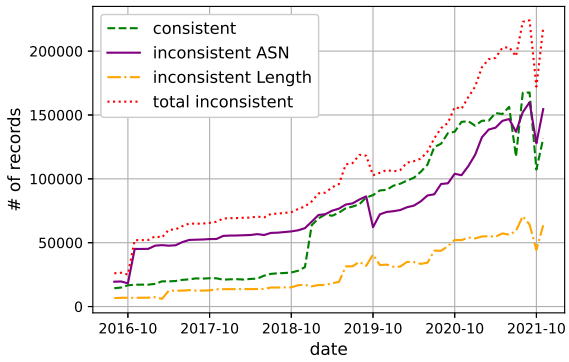
2.5.1 IRR consistency with RPKI

We find 351,404 ROAs (320,005 prefixes) in May 2023, where 120,220 new ROAs (111,340 new prefixes) were created after November 2021, showing significant growth in RPKI registration. Figure 2.4 shows the percentage of route objects that were RPKI consistent (green) and RPKI inconsistent (red) in November 2021 and May 2023. Since we are able to compare more route objects to RPKI in 2023, we discover that most IRRs had increased percentages of both RPKI consistent and RPKI inconsistent records and a decreased percentage of records not in RPKI. Some IRRs, like NTTCOM and BBOI, improved their record maintenance practices over the past 2 years by removing records with inconsistent objects.

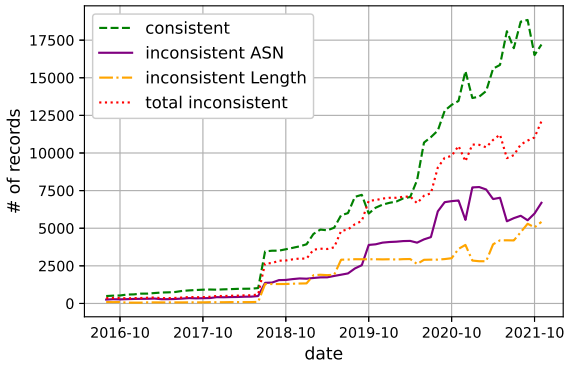
We also find that 4 IRR databases (LACNIC, BBOI, TC, NTTCOM) were 100% consistent with RPKI, likely due to a policy to reject route objects that are RPKI inconsistent [137]. Conversely, we find no RPKI-consistent records in PANIX and NESTEGG, and we recommend operators not to use them in route filtering.

2.5.2 Historical consistency between IRR and RPKI

RADB IRR. We begin by investigating the records in the *RADB IRR dataset*. From a total of 1.33M IRR records in October 2021, we find a corresponding ROA for 279,402 (21%) of the IPv4 prefix origin pairs. We find that 107,882 (or 38%) pairs were **consistent** with the ROA, while 127,099 (46%) showed an **inconsistent ASN** and the remaining 44,421 pairs (16%)



(a) RADB IPv4 inconsistencies. The uptick of inconsistent records in Oct. 2016 was due to Verisign registering RPKI records. Verisign later deleted the RADB records in Sept. 2019.



(b) RADB IPv6 inconsistencies. The uptick around Jun. 2018 was due to a large Thai ISP registering RPKI records and thus giving us visibility of an increased number of inconsistent IRR records.

Figure 2.5. RADB IPv6 records were more consistent with RPKI than IPv4 records.

exhibited an **inconsistent length**.

However, the fractions vary across our 5-year observation window (Figure 2.5a) We observe an increase in the fraction of prefixes with a corresponding ROA, starting with 882,220 (5%) in 2016, peaking at 30% in August 2021 and dropping to 1,335,602 (20%) in October 2021. We attribute this increase to accelerated adoption of RPKI during that time period (Table 2.2). The total number of **consistent** records increased by around 1000%, from 14,359 in October 2016 to its peak of 167,370 in August 2021. Also, the number of **inconsistent** records increased by around 850%, from 26,057 in October 2016 to 222,982 in August 2021.

We notice some outstanding events in Figure 2.5a: In October 2016, there was a sudden increase of 26,647 **inconsistent ASN** records. Those RADB records were registered under AS26415 by Verisign, with a description of `verisign customer route`. Customers of Verisign registered their prefixes under their own AS numbers (27 total) in RPKI, which caused this inconsistency. Later, in September 2019, Verisign deleted 26,682 records from RADB in an effort to clean up their records in RADB.

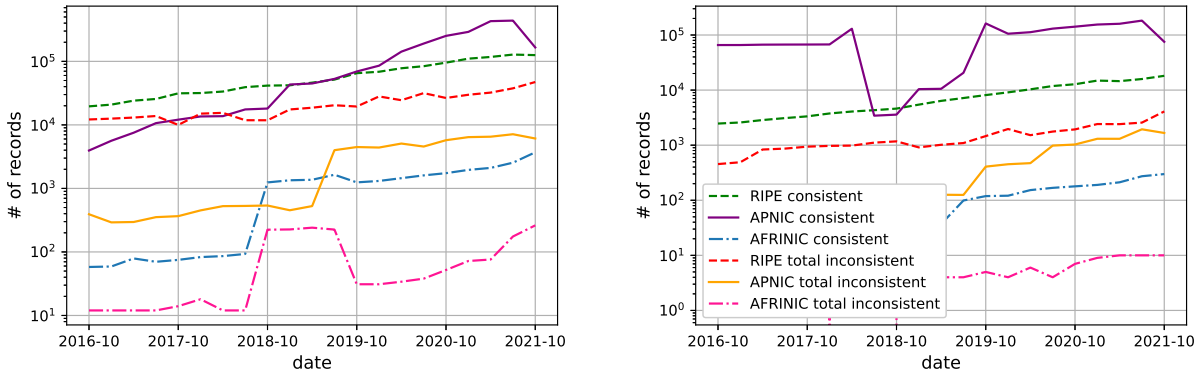
Figure 2.5a also shows an increase of 34,430 **consistent** records in January 2019. Those RADB records were registered by 10 TWNIC ASes. We speculate that this event was the outcome

of TWNIC obtaining a delegated RPKI CA certificate from APNIC in late 2018 [94]. Later, from July 2021 to October 2021, fluctuations in the green line were caused by TWNIC RPKI records disappearing and reappearing from our *RPKI dataset*, possibly due to instability caused by the migration of RPKI Relying Party (RP) software from the RIPE NCC RPKI Validator to Routinator [156].

Next, we look at the consistency of RADB IPv6 records. Figure 2.5b shows that for around two years until May 2018, there were few IPv6 records in any category, constituting fewer than 5% of all prefix origin pairs in the *IRR RADB data set*. After that, the fraction of IPv6 prefixes with a corresponding ROA increased to more than 10% and peaked at 20%, which indicates a steady adoption of RPKI for IPv6 prefixes. Of 27,540 prefixes with a matching ROA in October 2021, around 16,506 (60%) were **consistent**, 5,977 (22%) **inconsistent ASN**, and 5,057 (18%) **inconsistent length**. Interestingly, the number of inconsistent records stabilized after October 2020, while consistent records continued to increase. This contrasts with our IPv4 observations, where the inconsistency kept growing. Also, on November 25, 2019 the RIPE NCC announced the complete exhaustion of IPv4 address space [155], after which the growth of **consistent** records increased as a potential outcome of the greater incentive to deploy IPv6 operationally.

The sudden increase of records in all three categories in June 2018 can be attributed to 2,411 **consistent** records of 7 APNIC ASes, 827 **inconsistent ASN** records of 2 (a subset of the 7) ASes, and 1,180 **inconsistent length** records of 2 ASes (AS45430 and AS133481) operated by Advanced Wireless Network Company Limited (AWN), which is a large Thai ISP. Later in September 2019, AS45430 registered 2 prefixes in RPKI, causing 1,313 IRR records that belonged to AS133481 to become **Inconsistent**. Using CAIDA's AS Rank Dataset [22], we find that both ASes belonged to AWN, and AS45430 was the provider of AS133481. This is an example of the cause of some **Inconsistent ASN** cases in §2.5.4, where the customer AS failed to remove their IRR record after the provider AS reclaimed its address space.

RIPE IRR. In October 2021 in the *RIPE IRR dataset*, there were 125,424 **consistent**,



(a) RIR IRR IPv4 (legend same as Fig. 2.6b)

(b) RIR IRR IPv6

Figure 2.6. In IRR databases maintained by the RIRs, the number of consistent records was at least 10 times the number of inconsistent records within the same RIR.

30,764 **inconsistent ASN**, and 16,543 **inconsistent length** IPv4 records (Figure 2.6a). For IPv6, the three categories had 18,154, 3,374, and 721 records, respectively (Figure 2.6b). In the RIPE IRR, the prefix origin consistency for IPv4 and IPv6 records were similar. The consistent records have grown steadily while there have been a minimal increase in the number of records in the inconsistent categories, showing improved IRR hygiene over time.

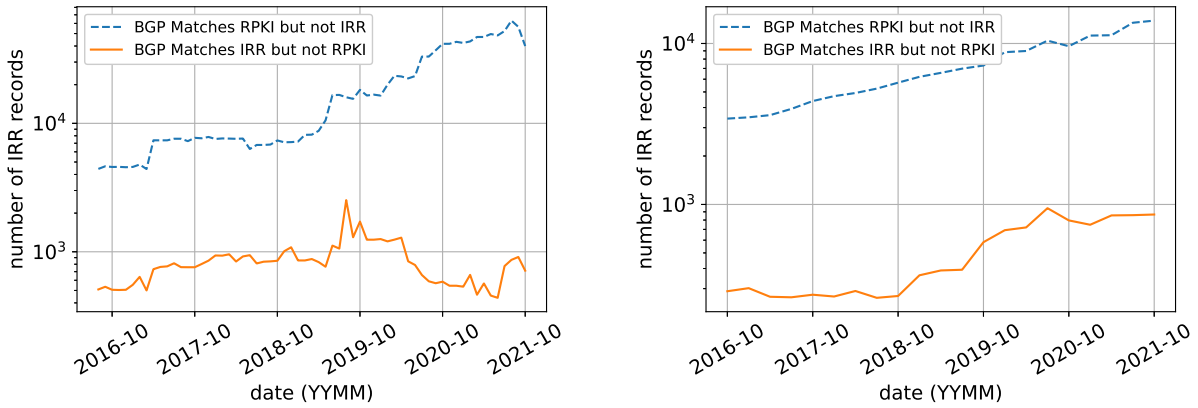
Comparing the RADB and the RIPE IRR, we confirm prevailing knowledge that RIPE IRR had better-maintained records. As of October 2021, 8.3% of total records in *RADB IRR dataset* were **consistent** compared to 34.4% of total records in the *RIPE IRR dataset*. RIPE IRR's better hygiene may be due to RIPE NCC's authorization: To create a `route` object, a validation process first checks if the maintainer (`mntner`) has the authority to announce the IP prefix by either looking for parent maintainer information or referencing IP address space ownership information [152]. Our *RIPE IRR dataset* include only route objects in the authoritative RIPE IRR, which contains only prefixes managed by the RIPE NCC.

In the *RADB IRR dataset*, 28.3% of records had matching ROAs in the *RPKI dataset*, and the corresponding fraction of RPKI-matching records for the *RIPE IRR dataset* was 45.2%. This shows that a larger fractions of networks registered in the RIPE IRR have also registered in the RPKI.

APNIC IRR. In July 2021 in the *APNIC IRR dataset*, there were 438,143 **consistent**, 3,426 **inconsistent ASN**, and 3,702 **inconsistent length** IPv4 records. For IPv6, the three categories had 182,563, 1,014, and 928 records, respectively. We find that the APNIC IRR had the highest consistency with RPKI compared to all other IRRs. The sudden drop in the solid purple line in Figure 2.6b shows the number of **consistent** records decreased by 128,056 in July 2017. This decrease was caused by AS10091 (StarHub, Singapore) and AS45224 (Lanka Bell Limited, Sri Lanka) removing the entirety of their IRR records (65,491 and 62,555 records) from the APNIC IRR. Figure 2.6b also shows that there were significantly more IPv6 records in the APNIC IRR than the other RIR IRRs. Dhamdhere et al. [155] found that APNIC made a big push towards IPv6 deployment because it was the first geographical region to experience IPv4 exhaustion. We speculate that the IPv6 push caused such a high presence of IPv6 records in the APNIC IRR.

AFRINIC IRR. In October 2021 in the *AFRINIC IRR dataset*, there were 3,702 **consistent**, 180 **inconsistent ASN**, and 82 **inconsistent length** IPv4 records. For IPv6, the three categories had 299, 5, and 5 records, respectively. Although the number of consistent records exceeded that of inconsistent records, the AFRINIC IRR overall contained fewer records than other IRRs. Figure 2.6a shows an increase of both **consistent** and **inconsistent** records in October 2018 (dash-dotted blue and pink lines). We find this event was caused by AS30844 (Liquid Telecom, UK) registering its prefixes in RPKI and caused 1,082 **consistent** and 212 **inconsistent ASN** IRR records. We speculate that AFRINIC had the lowest number of IRR records compared to other RIRs because AFRINIC had only launched its IRR in 2013, and had used the RIPE IRR before then [2].

Overall, the IRR databases operated by RIRs showed higher consistency with RPKI compared to RADB, as a result of the RIR's ability to regulate the creation process of `route` objects with address ownership validation. In the following sections, we conduct inconsistency analysis only on RADB and RIPE IRR records due to the low inconsistency of APNIC IRR



(a) RADB IPv4

(b) RIPE IRR IPv4

Figure 2.7. BGP sometimes agrees with IRR records but not RPKI, showing that some IRR records in the Inconsistent Length category may actually be correct.

records and scarcity of AFRINIC IRR records.

2.5.3 Causes of prefix length inconsistency

We further analyze the **inconsistent length** category. Networks that registered such records were one step away from good hygiene. Those networks successfully kept the origin ASes of their prefixes consistent in IRR and RPKI, but registered too-specific prefixes in the IRR. We speculate that this phenomenon could be caused by RPKI misconfiguration instead of bad IRR hygiene. To find out whether the networks registered inaccurate IRR records or incorrectly used the RPKI `Max Length`, we compare the **inconsistent length** records to their corresponding BGP announcement and RPKI ROAs. For each **inconsistent length** IRR record, we look for its exact or covering prefix in BGP. If the BGP prefix origin pair was the same as that in the IRR record, we label it *BGP matches IRR but not RPKI*, and this indicates that the IRR record was correct and the network likely misconfigured RPKI. If the BGP prefix was less specific than the IRR record prefix length and the RPKI ROA `Max Length`, we label it *BGP matches RPKI but not IRR*, and this indicates that the IRR record was inaccurate.

Figure 2.7a shows that as of October 2021, out of 44,421 **inconsistent length** IPv4 records in RADB, 713 (1.6%) agreed with BGP (*BGP matches IRR but not RPKI*) and 39,968

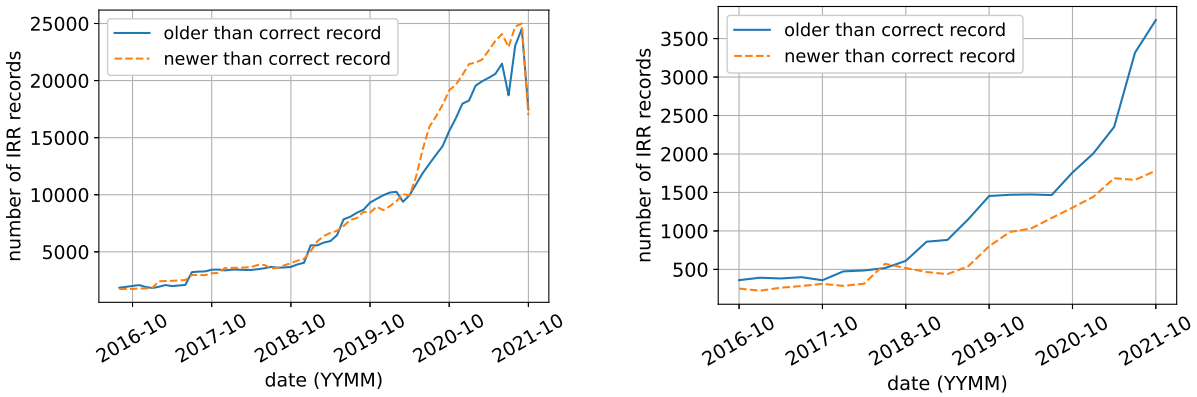
(90.0%) disagreed with BGP (*BGP matches RPKI but not IRR*). Most prefix-length inconsistency in the IRR was caused by networks registering too-specific prefixes in the IRR, while in BGP and RPKI, they aggregated the prefixes. In far fewer cases, networks used the correct prefixes in IRR and announced them in BGP, but registered less-specific prefixes in RPKI and failed to set the proper *Max Length* attribute. Figure 2.7b shows a similar situation for IPv4 records in the RIPE IRR. Of 16,543 **inconsistent length** records, 866 (5.2%) were *BGP matches IRR but not RPKI* and 13,872 (83.9%) were *BGP matches RPKI but not IRR*. We found similar distributions for IPv6 records in both RADB and the RIPE IRR (not shown due to space constraints).

To summarize, prefix length inconsistency comes from two types of mistakes with different operational impacts. The first type of mistake is having incorrect IRR entries that do not correctly reflect their prefix owners' routing intentions in BGP due to mismatching prefix lengths. If the upstream provider of the prefix owner requires the owner to register its exact BGP announcements in an IRR database, the prefix owner's current BGP announcements may be dropped (e.g. the upstream provider sees (`137.110.0.0/24, AS7377`) in IRR but sees (`137.110.0.0/16, AS7377`) in BGP).

The second type of mistake is misconfiguring the RPKI Max Length field. A too-small Max Length value in RPKI will almost immediately cause disruption to the prefix owner if its upstream provider uses RPKI filtering. The prefix owner's BGP announcement will be marked *RPKI invalid* and the upstream provider will drop the prefix owner's BGP announcement. Some operators reported that the Max Length feature of RPKI can cause confusion especially for RPKI newcomers because no similar feature exists in the IRR [121]. Gilad et al. [59] stated that the use of Max Length brings more harm than good to routing security, and operators have drafted proposals to discourage the use of the RPKI Max Length attribute [58].

2.5.4 Analysis of ASN inconsistency

We further analyze the causes of **inconsistent ASN** records. Networks may stop maintaining their IRR records after initial registration, inducing outdated information [88]. To study



(a) RADB IPv4

(b) RIPE IRR IPv4

Figure 2.8. Sometimes inconsistent ASN records are registered more recently than their corresponding consistent records.

such registration practices, we compared records within the same IRR dataset. First, we define *conflicting* records to be any two IRR records with the same IP prefix but different origin ASes. Then, we take each **inconsistent ASN** record and look for *conflicting* records in the same IRR dataset. We call the corresponding *conflicting* record a **correct** record if it is categorized as **consistent** (§2.5.2). We compare the registration dates of the **inconsistent ASN** record and the **correct** record.

We find 34,174 (26.9%) [5,524 (18.0%)] **correct** records for 127,099 [30,764] **inconsistent ASN** RADB [RIPE] IPv4 records in the October 2021 snapshot. Figure 2.8a shows that in RADB, 17,319 (13.6%) records were older than their **correct** records and 16,855 (13.2%) were newer. Figure 2.8b shows that in RIPE IRR, 3,741 (12.1%) records were older than their **correct** records and 1783 (5.8%) were newer.

This result contradicts the intuition that an **inconsistent ASN** record should be older than its **correct** counterpart, because the inaccurate IRR records are likely stale [88]. To explore this surprising phenomenon, we use the October 2021 snapshots of the *IRR datasets* and use the CAIDA AS Relationship dataset [23] from October 2021 to examine the relationship between the ASes that registered those records. We retrieve the AS relationships for 10,382 (30.4%) out of 34,174 **correct** records for RADB and 3,203 (58.0%) out of 5,524 for RIPE IRR. In RADB,

we find that out of 5,468 **inconsistent ASN** records with older **correct** records, 4,464 (81.6%) **correct** records were registered by providers of the networks that registered the **inconsistent ASN** records. Similarly in RIPE IRR, out of 839 **inconsistent ASN** records, 563 (67.1%) fell into the same situation. Such a high percentage of provider-customer relationships suggests that the providers first registered their prefixes in an IRR database, and then assigned address space to their customers. Those customers also registered their assigned prefixes in the same IRR database but failed to delete the IRR records after their providers revoked the address space. We found anecdotal evidence that some large ISPs such as Advanced Wireless Network Company Limited (AWN) in Thailand (§2.5.2) had customers who failed to delete outdated IRR records.

To prevent this inconsistency, providers could proactively require their customers to remove their IRR entries upon reclamation of address space to promote good IRR maintenance. Alternatively, if providers still intend to allow their customers to use such address space, the providers could register additional ROAs in RPKI under their customers' ASes.

2.6 ASes behind IRR records inconsistent with RPKI

To examine geographical trends in IRR record maintenance across ASes, we take the October 2021 snapshots for the *RADB IRR dataset* and the *RIPE IRR dataset* and classified the ASes according to §2.3.3. We use the October 2021 snapshots of the *as2org dataset* [24] and the *MANRS dataset* [116] to classify the ASes by their RIR and whether they were MANRS participants. We labeled the ASes in RIPE, ARIN, APNIC, LACNIC, and AFRINIC regions as EU, NA, AP, SA, and SA respectively. Table 2.4 shows that LACNIC (SA) ASes had the best IRR hygiene among all RIRs as of October 2021. Although there were more **entirely consistent (EC)** ASes than **entirely inconsistent (EI)** ASes, the number of **consistent** records was lower than **inconsistent ASN** and **inconsistent Length** records combined. This discrepancy is because the **entirely consistent (EI)** ASes registered fewer IPv4 records in RADB. We also use the *CAIDA AS Rank* [22] dataset to look at the AS *customer cone* size distribution in each category,

Table 2.4. Classification of ASes that registered IPv4 records in RADB. MANRS ASes were more consistent than other ASes. The classes are Entirely Consistent (EC), Entirely Inconsistent (EI), or Mixed.

AS Class	Record	General AS Count (9,675)						MANRS ASes
		Total	EU	NA	AP	SA	AF	(326)
EC	54,488	4,375 (45.2%)	562	1,084	1,184	1,452	67	173 (53.1%)
EI	367,795	3,513 (36.3%)	415	1,366	1,075	271	50	43 (13.2%)
Mixed	489,172	1,787 (18.5%)	271	428	739	324	16	110 (33.7%)

Table 2.5. Classification of ASes that registered IPv4 records in RIPE IRR. Most ASes were in the RIPE region and were highly consistent.

AS Class	Record	General AS Count (13,109)						MANRS ASes
		Total	EU	NA	AP	SA	AF	(220)
EC	75,589	9,339 (71.2%)	9,039	175	85	31	6	140 (63.6%)
EI	18,613	1,478 (11.3%)	1,309	86	63	7	4	7 (3.2%)
Mixed	144,284	2,292 (17.5)	2,193	70	26	0	3	73 (33.2%)

but find no correlation between AS size and AS registration practice in RADB. Table 2.5 shows that the authoritative RIPE IRR had few users outside of the RIPE service region, and the ASes had good IRR hygiene as a result of the validation requirement of the RIPE Database.

As of October 2021, of 787 MANRS ASes, 326 appeared in the *RADB IRR dataset* and had corresponding ROAs in the *RPKI dataset*. The MANRS ASes had better IRR hygiene than ASes in RADB, because the fraction of **entirely consistent (EC)** ASes was higher (53.1% vs. 45.2%). However, fewer MANRS ASes registered in the RIPE IRR compared to RADB so the fraction of **entirely consistent (EC)** ASes was below that of RIPE ASes (63.6% vs. 71.2%). Note that MANRS **only** requires their participants to register in either IRR or RPKI. MANRS ASes are not required to keep their records consistent between IRR and RPKI, and any MANRS AS that appears in Table 2.4 and Table 2.5 registered in both IRR and RPKI, which is more than required by MANRS.

Table 2.6. Number of unique prefixes in each step of filtering potentially irregular IRR objects in RADB. An RADB prefix is *consistent* with authoritative IRR if the origin ASes in both databases are related. An *inconsistent* RADB prefix partially overlaps with BGP if there is a common origin AS between RADB and BGP. At the end of the workflow, 23,353 out of 1,218,946 (0.2%) prefixes were inconsistent with authoritative IRR and partially overlapped with BGP. We consider the 34,199 route objects containing those prefixes *irregular*.

RADB	1,218,946 Total Prefixes	20.4% (249,725/1,218,946) Appear in Auth IRR	39.8% (99,323/249,725) consistent
			60.2% (150,402/249,725) **inconsistent**
		39.2% (59,024/150,402) Appear in BGP *and* inconsistent	54.7% (32,289/59,024) no overlap
			5.7% (3,382/59,024) full overlap
			39.6% (23,353/59,024) **partial overlap**

2.7 Irregular Route objects

We now examine RADB and ALTDB, two IRR databases where irregular or falsified `route` objects have been previously reported.

2.7.1 RADB analysis

We apply our workflow to filter irregular route objects in RADB. We start with 1,218,946 unique prefixes from route objects in RADB between November 2021 and May 2023. We find 196,664 unique prefixes that had a mismatching origin AS with one or more of the five authoritative IRRs (RIPE, ARIN, APNIC, AFRINIC, LACNIC). Of these prefixes, we find the mismatching ASes for 46,262 prefixes had a sibling or customer-provider relationship with each other. We remove those prefixes, leaving 150,402 prefixes that were inconsistent with authoritative IRR databases (Table 2.6 line 2).

Comparing those inconsistent prefixes to prefix origins announced in BGP, we find 3,382 prefixes in *full overlap*, 23,353 in *partial overlap*, and 32,289 prefixes in *no overlap* (Table 2.6 line 3-5). We focus on the *partial overlapped* prefixes because they had multi-origin AS conflicts (MOAS) in BGP, which has been a metric used to identify BGP hijacking [180]. We are able to match 23,353 *partial overlapped* prefixes to 34,199 prefix origins in BGP announcements

(Table 2.6 line 5). We find some prefixes belonged to different route objects with the same origin ASes but different maintainers, suggesting some networks had multiple maintainer accounts in RADB (*e.g.,* ipxo.com). We classify those 34,199 prefix origins as *irregular* and further analyzed them.

Validation. We first validate the *irregular* route objects using automated steps (§2.3.4). We perform Route Origin Validation (ROV) [122] on the *irregular* route object using the *RPKI dataset* (§2.3.2).

We find that of the 34,199 *irregular* route objects, 20,523 were consistent, 4,082 had an mismatching ASN, 144 had a prefix that was too specific, and 9,450 had no matching ROA in RPKI.

We then compare the BGP behavior of route objects with different RPKI statuses and identify cases where the RPKI-inconsistent route object was announced in BGP for over a year. Investigating those cases (*e.g.,* 24.157.32.0/19, AS54120), we find that the mismatching RPKI records were created recently relative to the time of our study, possibly as a result of the network adopting RPKI and correctly creating all required records.

Of the 13,676 *irregular* route objects that were RPKI-inconsistent/unknown, we remove the ones whose AS appeared in the RPKI-consistent route objects, leaving 6,373 irregular route objects (315 had matching BGP announcements that lasted < 30 days). Network operators who use IRR-based filtering should carefully consider those irregular route objects.

We also compare our list of 34,199 (Table 2.6) route objects with the list of serial hijackers from Testart et al. [180] and find 5,581 route objects registered by 168 serial hijacker ASes. We find one of those ASes (AS35916) to be a small US-based ISP with 10 customers according to CAIDA's AS Rank [22]. The other serial hijacker AS (AS9009) was a European hosting provider with more than 100 customers, which was also known to be exploited by attackers to abuse the DNS system [3]. However, networks may have registered both irregular and benign route objects, which can complicate the inference of suspicious route objects.

Source of false inference: IP leasing companies. During manual inspection, we find

30.4% (10,408 / 34,199) irregular route objects registered by ipxo.com, an IP leasing company. This presents a challenge to automating inference of irregular objects. IPXO registered 738 ASes under different maintainers in RADB, none of which had a sibling or customer-provider/peering relationship [24]. Those ASes displayed short-lived BGP activity, with announcement duration spanning from 10 minutes to more than 500 days. Prehn et al. [146] explored the use of different datasets to infer leasing relationships, but found limited coverage because IP leasing companies have no obligation to report their activities to the RIRs.

2.7.2 ALTDB analysis

Similar to RADB, we apply our workflow to ALTDB and identified 1,206 unique prefixes that were inconsistent with the authoritative IRR databases. Of these, we find 918 *full overlap* with BGP, 5 *partial overlap*, and 12 had no overlap. We map the 5 *partial overlap* prefixes to 11 prefix origins in BGP. Of the 11 prefix origins, one was RPKI-inconsistent with mismatching ASN and 10 were not found in RPKI. We manually inspect those 11 prefixes and find 5 highly suspicious cases. One had a prefix registered by AS58202, a Georgian network with no customers, providers, or peers. They announced the prefix, which was part of a larger prefix owned by Sprint, in BGP for only 14 hours. The other 4 prefixes were part of Verizon's address space, but were announced by unrelated ASes for less than 1 day. We also identify a benign case where the route object was registered by Akamai who could have originated the prefix on their business customer's behalf.

2.8 Summary

This chapter examines the reliability of the Internet Routing Registry (IRR) as a source of trust in BGP security mechanisms. By leveraging the increasing availability of RPKI data as a reference, I quantify inconsistencies between IRR and RPKI records and analyzed IRR maintenance practices across ASes. Comparing RADB and RPKI, I find 61.4% of IPv4 and 40.1% of IPv6 records (that appeared in both databases) are inconsistent. In contrast, the RIPE

IRR had only 27.4% IPv4 and 18.4% IPv6 inconsistent records. I uncovered some causes of inconsistency: complicated customer-provider relationships among ASes in the IRR, and possible misconfiguration in the RPKI. Furthermore, I expand the scope to all IRR databases and developed a methodology to detect potentially attacker-falsified entries without external ground truth, identifying over 5,000 irregular objects linked to known hijackers. These results highlight the operational complexity and security risks of relying on outdated IRR data, emphasize the need for improved cross-IRR coordination, and support the ongoing shift toward more robust mechanisms like RPKI.

These findings demonstrate that public data sources can be effectively used to evaluate the accuracy of IRR databases and identify IRR records that are likely falsified by attackers, which are factors that undermine the effectiveness of BGP origin validation mechanisms. The Internet operators community is aware of the unreliability of the IRR and has been promoting the adoption of RPKI as a more secure alternative. Although RPKI uses cryptographic attestations to ensure that only legitimate IP address holders can authorize ASes to originate their prefixes, its delegated trust model still introduces potential vulnerabilities. If IP address holders are unaware of who they are granting access to, they may unintentionally authorize malicious actors to use the address space for illicit purposes. In the next chapter, we examine whether attackers have exploited this trust model to subvert RPKI-based protections.

Chapter 2, in part, is a reprint of material as it appears in *Proceedings of Passive and Active Measurement Conference 2022.* Ben Du, Gautam Akiwate, Thomas Krenc, Cecilia Testart, Alexander Marder, Bradley Huffaker, Alex C. Snoeren, and kc claffy. The dissertation author was the primary investigator and author of this paper. This chapter also includes material as it appears in *Proceedings of the Internet Measurement Conference 2023.* Ben Du, Katherine Izhikevich, Sumanth Rao, Gautam Akiwate, Cecilia Testart, Alex C. Snoeren, and kc claffy. The dissertation author was the primary investigator and author of this paper.

Chapter 3

Inadvertent Delegation of Trust to Malicious Actors

In the previous chapter, I found that inaccuracies and attacker-falsified records can undermine the effectiveness of the IRR. To address such weaknesses, the operator community has promoted the adoption of RPKI as a more secure alternative. RPKI uses cryptographic attestations to ensure that only legitimate IP address holders can authorize ASes to originate their prefixes, making it far more difficult for attackers to register false records. However, its delegated trust model still carries a potential vulnerability: if an IP address holder is unaware of the ultimate party being granted authorization, they may unintentionally enable malicious use of their address space. The IPv4 leasing market provides one such opportunity for abuse. Since the exhaustion of unallocated IPv4 space, IPv4 addresses have acquired monetary value, leading to a secondary market for leasing address space. Anecdotal reports suggest that attackers have exploited this market to obtain RPKI routing authorizations from leasing providers.

In this chapter, I show that it is possible to use public data sources to systematically infer RPKI authorizations obtained by malicious actors through the IP address leasing market in the following two steps. First, leased prefixes are typically routed in BGP to allow services hosted behind them to be accessible to the broader Internet, and some leasing transactions also appear in publicly accessible IP address allocation databases (WHOIS) maintained by the RIRs. I leverage BGP and WHOIS data to characterize the structure and participants of the leasing ecosystem.

Second, the Internet community maintains blocklists of ASes involved in malicious activity, which network operators commonly use to filter spam and other unwanted traffic. I combine this information with RPKI data to assess the prevalence of cases in which adversaries exploit the leasing market to obtain RPKI authorizations.

3.1 Overview

The Internet numbers resource ecosystem has become more active since the exhaustion of unallocated IPv4 address space: the Regional Internet Registries (RIRs) have reported growing numbers of in-region and out-of-region IPv4 transfers over the past few decades [105, 106, 151]. Major cloud companies have purchased IPv4 address space from various organizations [15, 144, 169] to support their growing customer base and market share in the hosting and content distribution industry [27, 192], creating more demand in the IPv4 address market. This led to the rise of IP address brokerage businesses [144, 146] which facilitate buying, selling, and leasing of IP address space, providing a convenient platform for organizations to monetize their unused address space and for those who need address space to easily obtain it.

Recent shifts in business dynamics have contributed to the increasing demand for IP leasing. Some cloud companies introduced an IPv4 surcharge to offset the cost of obtaining IPv4 addresses [16], but allow customers to avoid these surcharges by using the cloud companies' Bring Your Own IP (BYOIP) services [16, 20, 66, 67, 98]. This pricing asymmetry has created a budget-friendly operational pipeline: customers can first lease IP address space from a brokerage company and then bring it to cloud hosting providers [81]. For example, commercial virtual private network (VPN) providers often lease different IP address blocks on a short-term basis to bypass the geo-filters of streaming services [89, 90, 107].

The legality of leasing has been debated since IANA's exhaustion of IPv4 in 2011 [103]. While most RIRs have taken a neutral position, AFRINIC explicitly prohibits leasing [123]. In 2023, ARIN clarified that leasing is permitted for address space an organization already holds

but cannot be used to justify acquiring additional allocations [177].

Leasing and the resulting IP address circulation have broader operational and security implications. It can create geolocation inconsistencies [38, 43, 72, 83, 104]: prefixes on the IPXO marketplace have been found to geolocate to multiple continents simultaneously [82]. It can also contribute to inaccuracies in IRR databases [45] and non-compliance to best routing security practices [48, 102, 182]. More seriously, attackers have used leasing platforms in combination with BYOIP services to obtain legitimate RPKI authorizations and appear legitimate in the routing system, and then conduct malicious activities which caused some of the leased IP address blocks to be blacklisted, creating operational challenges for the legitimate parties involved in the leasing process [40, 44, 54, 68]. Because the leasing market operates with limited transparency, it is often unclear who is leasing which resources and for what purpose without direct access to brokerage transactions.

Our goal in this chapter is to use public data to characterize the IP leasing ecosystem and assess its security implications, particularly in relation to RPKI-based origin validation. Our main contributions are:

(1) We develop a methodology that uses public WHOIS data to infer leased IP address space and evaluate its accuracy using a reference dataset curated from records of registered IP brokerage companies (§3.6.1, §3.6.2).

(2) We show that, compared to non-leased address space, our inferred leased address space is eight times more likely to be RPKI-authorized to ASes that have been repeatedly involved in malicious activities, showing that attackers are using the IP leasing market to undermine the effectiveness of RPKI-based origin validation mechanisms (§3.6.3).

3.2 Related Work

Prior work has studied the IPv4 transfer ecosystem and its abuse. In 2013 and 2017, Livadariu *et al.* [105, 106] studied routing activity of RIR-reported transfers and found 85%

transferred IP prefixes were advertised in BGP, suggesting that most transfers were due to legitimate operational need for IP address space. In 2020, Giotsas *et al.* [60] found that ASes that advertised transferred address space in BGP were more likely to engage in malicious activities compared to ASes that had not. While these studies focused on the IP transfer market, none examined abuse within the IP leasing market.

In 2023, Beverly [18] compared the geographical information of IP addresses registered in the RIRs to the physical locations in which the address space was being used. They found that 3.4% of 50K randomly selected IP prefixes were operating outside of the region of their allocating RIRs. Prehn *et al.* [146] provided the first quantification of the prevalence of IP leasing in the RIPE region in 2020 and found a growing number of leasing providers' websites between 2019 and 2020. They also found that the leasing ecosystem was still maturing at that time as the market price for leasing IPv4 addresses varied significantly. Given that the IP leasing market has continued to evolve, a broader, up-to-date analysis is needed to understand its current scale, structure, and potential for abuse.

3.3 Definition of IP Leasing

ARIN defines IP leasing as follows: an organization is *leasing IP addresses* when it provides its customer with IP addresses but does not provide the customer connectivity to the Internet [177]. To infer leased IP address space, we need to understand relevant RIR policies on IP resource distribution, the routing system, and the IP leasing business model.

3.3.1 RIR address policy

The RIPE NCC, ARIN, APNIC, AFRINIC, and LACNIC are the five RIRs in the world. They are each allocated IP addresses from the Internet Assigned Numbers Authority (IANA) and further distribute IP addresses to organizations within their designated geographic region. According to the RIR address policies, there are three broad categories of IPv4 address space.

(1) **Portable:** Portable address space is typically directly distributed to organizations by the RIRs. Organizations holding portable address space can choose any Internet connectivity provider, as the RIRs themselves do not offer connectivity services. Therefore, portable address space is generally not considered *leased*. Different RIRs mark portable address space with different statuses such as `Allocated PA` (Provider Aggregatable) or `Assigned PI` (Provider Independent) (RIPE [129], AFRINIC [1]), `(Direct) Allocation/Assignment` (ARIN [13, 188], LACNIC [97]), or `Allocated/Assigned Portable` (APNIC [10, 160]).

(2) **Non-portable:** Non-portable address space is typically sub-allocated or assigned by holders of portable address space (address providers) to other organizations. Organizations that hold non-portable address space are expected to use their address provider's Internet connectivity. If they do not, the non-portable address space is considered *leased* [177]. Non-portable space is labeled as `Sub-allocated PA` or `Assigned PA` (RIPE, APNIC), `Reallocation` or `Reassignment` (ARIN, LACNIC), or `Allocated Non-portable` or `Assigned Non-portable` (APNIC).

(3) **Legacy:** Legacy address space was directly distributed to organizations by IANA before the RIRs were established [130]. Since this space is not managed by the RIRs, it has no defined portability, so we do not apply the definition IP leasing to it. However, legacy address holders can choose to register their address space with RIRs to access certain services and privileges [131]. Once registered, the address space will no longer be legacy and will have defined portability.

In our study, we follow the above definitions to design our methodology and identify non-portable *leased* address space (§3.5.2).

3.3.2 IP leasing business model

We define the following terms to describe different business parties involved in the leasing process.

- **IP holder:** This entity's IP address space is portable.

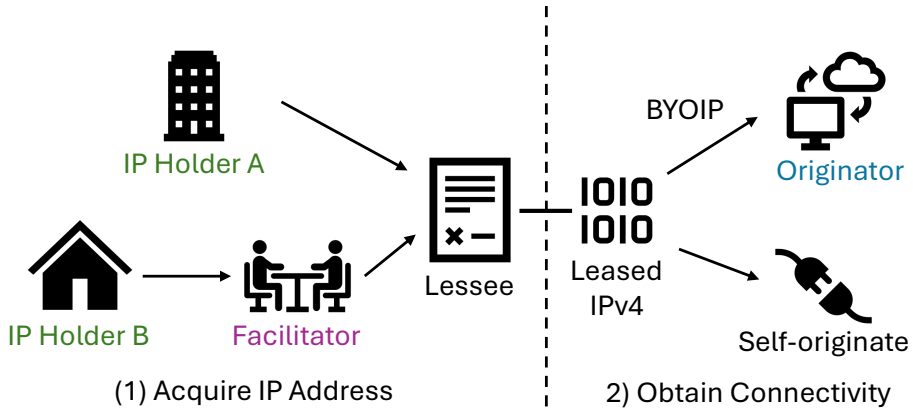


Figure 3.1. An entity (lessee) can follow this two-step pipeline to establish Internet presence without having their own IP addresses or even network infrastructure.

- **Lessee:** The entity that pays to temporarily acquire address space from an IP holder.
- **Facilitator:** An entity who facilitates the transaction between an IP holder and lessee. The IP leasing process does not always require a facilitator.
- **Originator:** The entity who advertises (leased) IP address space in BGP, *a.k.a.* the entity behind the origin AS of the advertised prefix. The originator may be the same entity as a lessee if the lessee has its own routing infrastructure.

Figure 3.1 illustrates the business model of IP leasing. An individual or organization plays the role of a lessee when it leases IP address space from an IP holder directly or through a facilitator (e.g., an IP broker). Some IP brokerage companies have multiple leasing business models: they can act as an IP holder to lease their own IP address directly to customers, or act as a facilitator between the customer and a different IP holder. After acquiring IP address space through leasing, the lessee can either provide their own Internet connectivity by originating the address space in BGP themselves or use BYOIP services to have cloud providers originate their newly leased address space.

3.4 Dataset

We use the prefix2as dataset and RPKI archives (§1.7) snapshots between April 1 and April 15, 2024 to obtain leased prefixes originated in BGP and RPKI records created for leased prefixes. We also use the following additional datasets to infer leased IP address space.

RIR WHOIS Database. We collected the April 1, 2024 WHOIS database snapshots from RIPE, ARIN, APNIC, AFRINIC, and LACNIC as our primary dataset for inferring leased IP address space.

Registered IP Brokers. ARIN and APNIC currently provide a list of registered brokers. In April 2024, ARIN listed 9 organizations as "qualified facilitators" [12] and APNIC listed 38 organizations as "registered brokers" [9]. The RIPE NCC used to provide a list of "recognized brokers" until December 2023 but decommissioned the list thereafter [128]. We used the Internet Archive to retrieve the December 2023 snapshot of RIPE NCC's recognized brokers page [126] and found 115 organizations. Overall, we construct a list of 162 organizations considered to be registered brokers.

Spamhaus. We downloaded the Spamhaus ASN-DROP lists monthly from February through May 2024. The Spamhaus ASN-DROP list contains ASes used by attackers for malicious activities such as spam operations and botnet control [175].

3.5 Methodology

In this section, we introduce our workflow to infer leased IPv4 address space based on the definition of IP leasing (§3.3.1) and our steps to curate an evaluation dataset to validate our results.

3.5.1 Processing RIR WHOIS databases

We first parse the RIR WHOIS database records to obtain the necessary parts for our inference. Figure 3.2 shows an example of processed WHOIS records relevant to a leased

prefix; we explain our workflow (circled numbers in Figure 3.2) and refer to the shapes of each component below.

① **Parse relevant WHOIS objects.** For our study, we need WHOIS database objects containing address blocks, Autonomous System (AS) numbers, and organizations. For RIPE, APNIC, and AFRINIC, those objects are `inetnum`, `aut-num`, and `organisation` respectively. We also obtain the maintainers of `inetnum` objects from their `mnt-by` fields for our inference evaluation (§3.5.3). For ARIN, those objects are `NetHandle`, `ASHandle`, and `OrgID`. LACNIC does not store organizations independently but includes the `owner` field in other objects.

② **Construct address allocation tree.** For each RIR WHOIS database, we parse all objects containing IANA allocated (non-legacy) IPv4 address blocks and convert the address-range notation into CIDR-prefix notation. We remove all hyper-specific prefixes longer than /24 as they are mostly for internal infrastructure use [164]. We input the prefixes into a prefix tree and find the root nodes and leaf nodes (Figure 3.2 rectangles). The root nodes represent portable prefixes directly allocated to the IP holders by the RIRs, and the leaf nodes contain non-portable prefixes sub-allocated or assigned by IP holders to other organizations (§3.3.1). We do not focus on the intermediate nodes in the tree, as they represent intermediate sub-allocations and are not critical to our inference process.

③ **Assign AS numbers.** For each root node, we first extract the `org` or `OrgID` fields and then find in the WHOIS databases all `aut-num` or `ASHandle` objects (ellipse) containing the same fields.

④ **Find the BGP origins.** For each leaf node, we search in our BGP dataset for its exact-matching prefix and origin AS (trapezoids). We repeat the process for the root nodes, and if an exact-matching prefix does not exist, we then search for its least-specific covering prefix and origin AS (trapezoids). This matching process accounts for cases where the organization holding consecutive portable address blocks choose to aggregate their corresponding prefixes in BGP.

3.5.2 Inferring leased address space

After completing steps ① through ④, we compare the components of the tree ⑤ to infer leased address space. We match the text colors (green, purple, and blue) in Figure 3.2 to the colors of their corresponding business parties in Figure 3.1. The inference process assigns each leaf-node prefix to one of the following groups, and further classifies the prefix as leased or not within each group:

- **Group 1:** If neither the leaf node nor its root node has a BGP origin, we consider it *unused* and thus not leased.
- **Group 2:** If only the root node has a BGP origin but not the leaf node, the leaf node prefix has likely been aggregated into its parent prefix in the root node in BGP. We call the leaf node an *aggregated customer* and thus not leased, e.g., the gray rectangle in Figure 3.2.
- **Group 3:** If the leaf node has a BGP origin but the root node does not, we classify it into group 3. If a customer-provider or sibling relationship [24, 109] exists between the leaf node's BGP origin (bottom trapezoid) and the RIR-assigned ASes associated with root nodes (top ellipse), we consider the leaf node to be an *ISP customer*, otherwise ***leased*** (bold orange rectangle). The different business parties are shown in Figure 3.2: the root node organization is the IP holder (green), the leaf node maintainer is the facilitator (purple), and the leaf node BGP origin is the originator (blue).
- **Group 4:** If both the root node and leaf node have BGP origins, we classify the leaf node into group 4. If the leaf node's BGP origin AS has a relationship to either the root node's assigned AS or the BGP origin AS, the leaf node prefix is likely a customer delegation of the root node prefix [92]. In this case, we consider the leaf node a *delegated customer*; otherwise we consider it ***leased***.

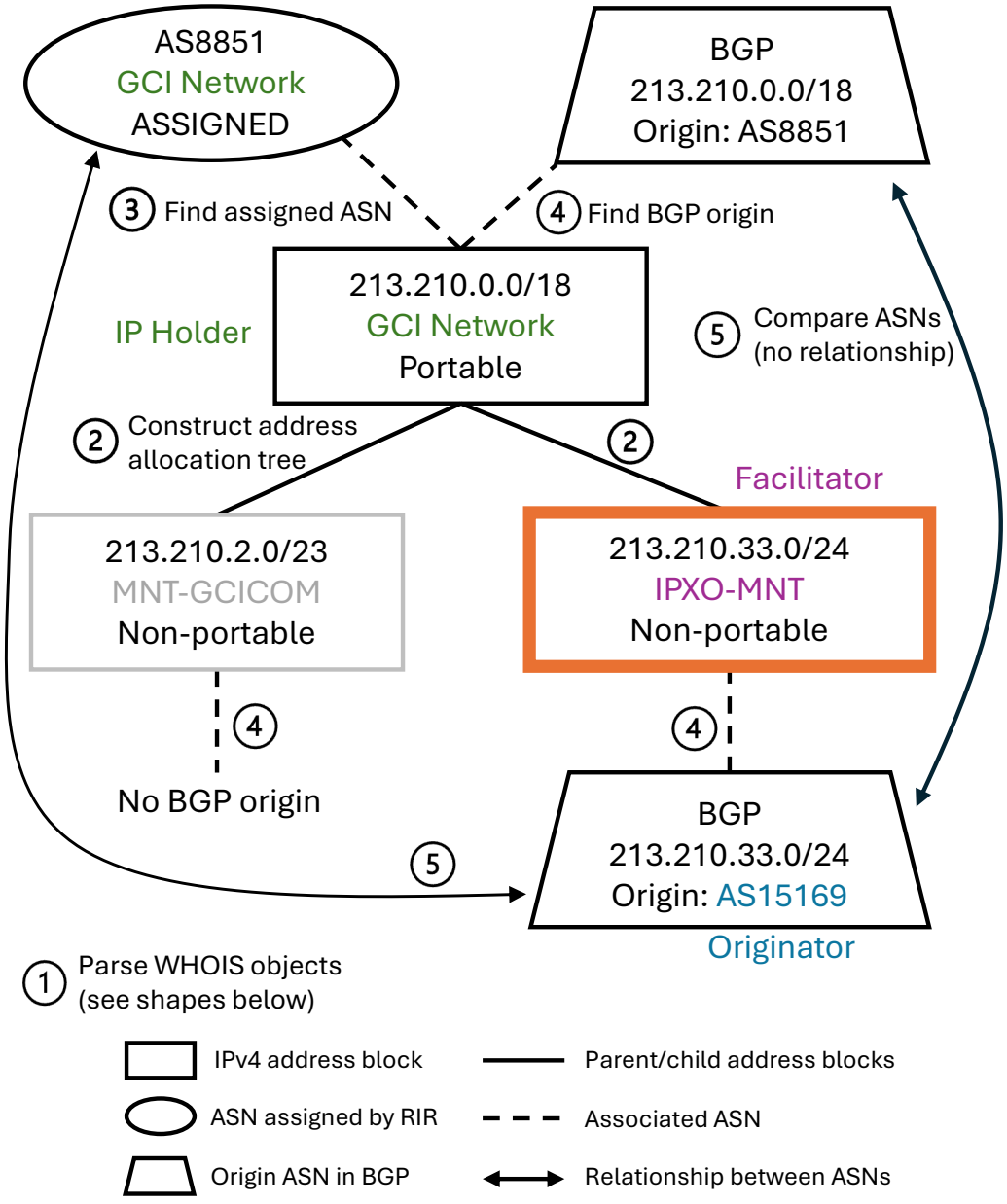


Figure 3.2. Inference diagram of an example leased prefix. The bold orange rectangle marks the inferred leased prefix. The leased prefix's BGP origin (AS15169) is related to neither the ASN assigned (AS8851) to its address provider (GCI Network) nor the BGP origin of the portable parent prefix (also AS8851). The colored text is matched to that of Figure 3.1 to show the different business parties involved in the lease.

3.5.3 Curating evaluation dataset

We manually curate a reference dataset to evaluate our inference methodology. To the best of our knowledge, no existing ground-truth dataset for leased IP address space exists. The closest available data are the IP leasing facilitators listed by the RIRs, who may act as maintainers for leased prefixes they broker. To identify the brokers' leased prefixes, we first map the company names of registered brokers to the `organisation`/`OrgID` objects in the corresponding RIR database and obtain their maintainer fields (e.g., `mnt-by`, `mnt-ref`). We then extract address blocks with matching maintainers from the WHOIS databases. Since not all address blocks managed by a broker are leased, we exclude non-leased address blocks by manually looking through their registration information. We use the remaining prefixes as our positive labels. Additionally, we manually construct a list of non-leased prefixes from five residential ISPs across three RIRs as our negative labels.

3.5.4 Limitations

We identify the following limitations of our methodology that may affect its accuracy:

Legacy IP Addresses. We do not consider the leasing of legacy address space since it is not managed by the RIRs and has undefined portability (§3.3.1). Future work could expand our methodology to support leasing inference for legacy address space.

Uncaptured AS Relationships. Complex business relationships, such as mergers, acquisitions, and international subsidiaries, that are not captured can also result in incorrect lease inferences.

Multi-homing. Some prefixes we classify as *group 4 leased* (§3.6.1) may be using both their IP holder's and originator's connectivity, which we cannot confirm with our datasets. Reactive measurements from suitable vantage points are required for further analysis.

Routing Attacks. We may mis-classify routing attacks (*e.g.,* prefix squatting [162]) as IP leasing: If a squatter happens to originate a prefix in our **unused** category, our methodology

Table 3.1. Number of prefixes in each group per RIR according to the inference process (§3.5.2). The 47k inferred leased prefixes were 4.1% of all routed prefixes in April 2024.

	Inference Group	RIPE	ARIN	APNIC	AFRINIC	LACNIC	All RIRs
1	Unused	63,670	43,011	25,437	28,936	27,551	188,605
2	Aggregated Cust.	204,337	98,316	21,515	1,741	11,950	337,859
3	ISP Customer	31,484	10,302	7,725	777	2,250	52,538
	Leased	**26,774**	**6,697**	**3,275**	**2,172**	**627**	**39,545**
4	Delegated Cust.	27,610	22,927	8,291	1,236	1,294	61,358
	Leased	**1,872**	**5,633**	**150**	**63**	**55**	**7,773**
# of Leased Prefixes / Total		**28,646 /** 355,747	**12,330 /** 186,886	**3,425 /** 70,192	**2,235 /** 45,330	**682 /** 47,861	**47,318 /** 705,016

would consider the squatted prefix as leased.

3.6 Results

In this section, we infer the leased prefixes in April 2024 for each RIR region, compare with prior work, evaluate our inference, and analyze the leasing ecosystem for each RIR.

3.6.1 IP address leasing ecosystem

Table 3.1 shows the inferred number of leased prefixes for each RIR. We infer 47,318 leased prefixes across all RIRs, which is 4.1% of all 1,146,921 observed IPv4 prefixes in BGP and 0.9% of the routed IPv4 address space. We find that RIPE had the most leased prefixes compared to the other RIRs. This finding is consistent with the description of RIPE's IP address market as the most active: RIPE had the most reported IPv4 transfers [146], the most available IPv4 blocks for sale [78], and the most available IPv4 blocks for lease [82] according to IP brokers' websites. Table 3.1 summarizes these statistics for all regions. We explain in detail our inference results for the RIPE region below.

RIPE. Overall we infer 28,646 RIPE prefixes are leased. We obtain 355,747 leaf nodes containing non-portable prefixes for RIPE (§3.3.1). Table 3.1 shows the number of prefixes in each category: we find 63,670 prefixes in group 1 *unused* that are not originated in BGP. Those

Table 3.2. Top 3 IP holders by # of inferred leases, April 2024.

RIR	Organization	Count
RIPE	Resilans AB	1,106
	Cyber Assets FZCO	941
	Russian Scientific-Research Institute	675
ARIN	EGIHosting	1,418
	PSINet, Inc.	1,233
	Ace Data Centers, Inc.	533
APNIC	Orient Express LDI Limited	145
	Capitalonline Data Service (HK)	135
	Aceville PTE.LTD.	96
AFRINIC	Cloud Innovation Ltd	2,014
	ATI - Agence Tunisienne Internet	38
	Nile Online	32
LACNIC	Radiografica Costarricense	114
	Impsat Fiber Networks Inc	88
	Newcom Limited	25

prefixes may be available for lease (*e.g.,* available on brokers' marketplaces), but we do not consider them *leased* since they were not active. We classify 204,337 prefixes into group 2 *aggregated customer*. Those prefixes were likely assigned by ISPs to customers who could not provide connectivity for themselves and then aggregated to the less-specific root node prefix in BGP. We find 31,484 prefixes in group 3 *ISP customer* that were likely ISP customer prefixes that were not aggregated to a covering prefix in the routing system. We find 26,774 prefixes in group 3 ***leased***. Some of those prefixes we may have falsely inferred to be leased due to unobserved AS relationship or uncaptured company structure (§3.5.4). We find 27,610 prefixes in group 4 *delegated customers* that were likely prefixes delegated to customers by ISPs or parent organizations [92]. We find 1,872 prefixes in group 4 ***leased***. Those prefixes could be redundantly connected to the Internet via their provider and a second upstream AS for resilience (multi-homing), and thus we could have falsely inferred them as leased prefixes (§3.5.4).

Top IP holders in each RIR. Table 3.2 lists the three IP holders that we infer to have leased out the most prefixes. In RIPE, Resilans, a Swedish IP management company, leased 806

prefixes within Sweden. Cyber Assets FZCO, a U.A.E. company, leased prefixes to 44 countries, including 332 prefixes to the U.S. and 110 to Brazil. In ARIN, the top IP holder, EGI Hosting, is a U.S. company that leased 1,418 prefixes to 18 countries, where 651 prefixes were leased in Cyprus and 195 in Panama. The second largest IP holder, PSINet, was an ISP that ceased operation in 2002 and was acquired by Cogent. Some of their prefixes were reallocated to IP leasing companies, but the majority are misclassified as leased due to our *as2org dataset* not capturing the acquisition (§3.5.4). For AFRINIC, the top IP holder, Cloud Innovations, leased 2,014 prefixes, far exceeding the second top IP holder with 38 leased prefixes. This finding is consistent with previous reports and studies of Cloud Innovations' pervasive leasing practices in AFRINIC [18, 123]. Cloud Innovations also served as the largest facilitator in AFRINIC, illustrating a case in which an IP holder facilitates its own leasing operations (§3.3.2).

3.6.2 Evaluating inference with brokered leases

We evaluate the accuracy of our inference using a manually curated dataset of prefixes leased through IP broker companies. This comparison shows that public datasets provide sufficient accuracy and coverage for large-scale leasing inference, making them a reliable basis for our study.

Of the 115 registered brokers in RIPE, we directly map 46 to their RIPE WHOIS database entries and manually match 39 due to inconsistencies such as variations in legal entity suffixes (e.g., LTD vs. L.T.D.), abbreviations, and fictitious business names. The remaining 30 brokers could not be matched as they were not present in the RIPE database. We successfully map the 85 matching organizations to 107 maintainer handles, covering 11,076 prefixes. After manually filtering out 1,621 prefixes that were likely not leased (e.g., brokers that also served as ISPs providing connectivity for their customers), we identify 9,455 actively leased prefixes.

For ARIN, we find 24 prefixes managed by 2 out of 9 brokers and manually remove 1 prefix that was not leased. In the case of APNIC, we are unable to match any registered brokers to address blocks because its database did not provide maintainer fields for organization objects.

Table 3.3. Confusion matrix: we evaluated our results against 14,856 validated prefixes.

	Inferred Lease	Inferred Non-lease	
Actual Lease	7,735 (TP)	1,743 (FN)	**Recall 0.82**
Actual Non-lease	121 (FP)	5,257 (TN)	Specificity 0.98
	Precision 0.98	NPV 0.75	**Accuracy 0.88**

Overall, our evaluation dataset includes 9,478 actual leased prefixes (positive labels). We also collect 5,378 actual non-leased prefixes (negative labels) managed by 5 ISPs (AT&T, Comcast, Orange, Vodafone, IIJ), which were originated in BGP by their respective ASNs. We reach out to IIJ and confirm that the IIJ prefixes we collect are indeed used for providing connectivity to their customers. In total, our evaluation dataset contains 14,856 prefixes.

Table 3.3 is a confusion matrix that shows that our inference methodology achieved 98% precision, 82% recall, and 88% accuracy when evaluated against the reference dataset. Upon further analysis, we find that the 1,605 false negatives we classify as group 1 *unused*. Since our methodology relies on the BGP origin of the leased prefix, it cannot detect leased prefixes that are not yet originated in BGP (inactive leases). The remaining 138 false negatives were legacy address blocks registered in the RIPE region, which are outside of the RIPE's portability definitions and thus not captured by our methodology. Additionally, we examine the 121 false positives and found that 110 were associated with Vodafone, which registered 17 `organisation` objects for different subsidiaries. These subsidiaries used different AS numbers, and their relationships were not captured by our AS Relationships dataset (§3.5.4).

Comparison with Prior Work. We conduct a preliminary comparison between our methodology and that of Prehm *et al.* [146]. Prehn *et al.* classified address blocks as leased if their maintainers differed from their parent blocks (*e.g.,* Figure 3.2). However, comparing maintainers can lead to false positives, such as classifying some customer prefixes as leased if the customer uses their own maintainer instead of the provider's (*e.g.,* leaf-node prefixes that have

different maintainers from the root-node prefixes in our *Aggregated Customer*, *ISP Customer*, or *Delegated Customer* groups). Conversely, there may be address blocks with the same maintainer that are actually leased, such as when an IP holder directly leases prefixes to lessees. On the contrary, their methodology can detect inactive leases that ours would falsely classify as *Unused*.

3.6.3 Potential abuse of leased prefixes

To assess the potential abuse of leased prefixes, we check whether leased prefixes are more likely to be originated by AS numbers on the Spamhaus ASN-DROP list compared to non-leased prefixes. In our analysis for April 2024 leased prefixes, we obtain 47,318 unique leased prefixes (Table 3.1), of which 1.1% (541) were originated by AS numbers on the Spamhaus ASN-DROP list. In contrast, we find that only 0.2% (2,224) of the 1,100,025 non-leased prefixes in our BGP dataset were originated by blocklisted AS numbers. In other words, leased prefixes were approximately five times more likely to be advertised by an AS considered abusive by Spamhaus than non-leased prefixes, and Numerous blog and forum posts discussed such abuse [68, 132].

We identify a total of 9,217 ASes originating leased prefixes and compare them against a list of 957 inferred serial BGP hijackers [180]. We find that 2.9% (269/9,217) of all originators were also serial hijackers. Those 269 serial hijacker ASes originated 13.3% (6,312) of all 47,318 inferred leased prefixes. Of the 1,100,025 non-leased prefixes, 3.1% (33,503/1,100,025) were originated by a serial hijacker AS. This shows that leased prefixes were more likely to be originated by serial hijacker ASes compared to non-leased prefixes.

Undermining RPKI Protection. We also look for ROAs associated with leased prefixes and blocklisted AS numbers in our RPKI dataset. We find 31,156 ROAs for the 47,318 leased prefixes, 1.6% (498/31,156) of which contained ASes on the blocklist. In contrast, the percentage of non-leased prefixes with ROAs containing blocklisted ASes was similar to the fraction of BGP advertisements from those same ASes, namely 0.2% (1,260/506,629). This suggests that leased prefixes are even more likely to have a ROA authorizing an abusive AS to use them. We then look into how ROAs were signed for leased prefixes and discover that some facilitators either

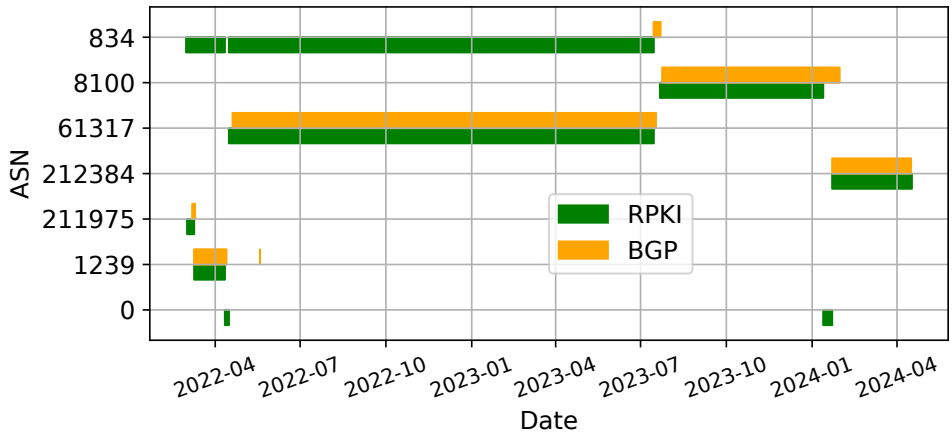


Figure 3.3. The RPKI and BGP behaviors of an IPXO-leased prefix suggests its periods of lease.

require the IP holder to create proper ROAs for the leased prefixes, or offer RPKI management services where the IP holder can delegate RPKI signing capability to the facilitator [80].

The above findings show that the IP leasing market could enable attackers to undermine routing security mechanisms such as RPKI. For example, previously spammers that hijacked another organization's address space for spam campaigns [148] might have had their unauthorized BGP announcements filtered by the RPKI. However, by leasing IP addresses, they can now bypass this obstacle.

3.6.4 Defense against abuse

We manually browse through several IP leasing companies' websites and find that one company, IPXO, provided extensive description on abuse policy [79]. Operators have also discussed their experiences with IPXO's strict abuse handling process on forums [54]. This suggests that IP leasing facilitators can step in to help mitigate and prevent abuse of leased prefixes. We also find that IPXO ranks among the top three facilitators in each of the RIPE, ARIN, and APNIC regions, indicating that large leasing providers have the potential to reduce abuse across a substantial portion of leased prefixes.

We also find that although RPKI might provide routing authorizations for abusers, it can also be leveraged to prevent abuse. Oliver *et al.* [142] showed that operators have created ROAs

with AS0 for their blocklisted prefixes, allowing them to be filtered in BGP and leading attackers to cease activity. We observe AS0 usage when analyzing the historical routing activity of an IPXO prefix. Figure 3.3 shows that IPXO used AS0 in between different leases of that prefix, likely for marking the end of a lease or abuse-related purposes such as removing the prefix from blocklists [142].

3.7 Summary

This chapter investigates the scale and security implications of the growing IPv4 address leasing market. Despite its increasing popularity, the leasing ecosystem has remained largely opaque and unexplored. We find that 4.1% (47,318 out of 1,146,921) of prefixes advertised in BGP in April 2024 were likely leased, and our inference method achieves 98% precision. Moreover, leased prefixes were eight times more likely to be RPKI-authorized to blocklisted ASes — specifically, repeat offenders of spamming — than non-leased prefixes. These findings show that malicious actors have obtained RPKI authorization through leasing arrangements, exploiting the RPKI system's delegated trust model.

These results show that public data sources can be used to accurately characterize the opaque IP leasing ecosystem and show that attackers can exploit such ecosystem to undermine BGP origin validation mechanism relying on RPKI. Although IRR and RPKI have their weaknesses, they remain the primary systems for BGP origin validation. However, neither can effectively limit the propagation of BGP origin hijacks or misconfigured routes unless networks actively deploy them. The network operators community, especially through the Mutually Agreed Norms on Routing Security (MANRS) initiative, has been promoting broader adoption of BGP origin validation mechanisms. In the next chapter, we analyze the deployment of IRR and RPKI among networks participating in the MANRS initiative and evaluate the initiative's effectiveness in improving routing security.

Chapter 3 in part, is a reprint of material as it appears in *Proceedings of the Internet*

Measurement Conference 2024. Ben Du, Romain Fontugne, Cecilia Testart, Alex C. Snoeren, and kc claffy. The dissertation author was the primary investigator and author of this paper.

Chapter 4

Adoption of Origin Validation Practices by MANRS Networks

In the previous chapters, I examined weaknesses in the established trust of the Internet Routing Registry (IRR), showing that poor maintenance of IRR databases can lead to inconsistent or outdated records. I also found that attackers may exploit the delegated trust model of RPKI to maliciously obtain authorization. These limitations can undermine the effectiveness of BGP origin validation mechanisms reliant on IRR and RPKI. Nevertheless, IRR and RPKI remain the primary systems that operators use for BGP origin validation, and many large networks require customers to register routing information in one or both systems prior to establishing peering relationships [8, 28, 65]. To further promote best practices in routing security, network operators and industry stakeholders launched the Mutually Agreed Norms for Routing Security (MANRS) initiative. The launch of MANRS raises the following questions for the Internet community: to what extent have MANRS participants adopted BGP origin validation mechanisms, and has their participation improved the effectiveness of these mechanisms?

In this chapter, I show that it is possible to assess the adoption of BGP origin validation mechanisms for networks participating in the Mutually Agreed Norms in Routing Security (MANRS) initiative using public measurement data, which I leverage in three key dimensions. First, we can directly observe whether networks have voluntarily registered routing information in authoritative databases such as the IRR and RPKI, as their prefix and AS data are explicitly

recorded. Second, we can indirectly infer whether a network has failed to filter invalid BGP messages by detecting the propagation of BGP announcements that conflict with those authoritative records. Finally, we can make a broad assessment of the effectiveness of MANRS networks as a whole in reducing the propagation of BGP announcements with incorrect/misconfigured origins.

4.1 Overview

In 2014, a group of security-conscious network operators launched the Mutually Agreed Norms for Routing Security (MANRS) initiative to promote the adoption of BGP origin validation mechanisms, and ultimately improve the overall security and resilience of the global routing system [116, 158]. First supported by the Internet Society (ISOC) and now the Global Cyber Alliance, MANRS advocates for a set of security best practices—which they call *actions*—to prevent the propagation of incorrect routing information, restrict forwarding of traffic from spoofed IP addresses, and facilitate coordination between network operators.

However, MANRS has not taken on rigorous enforcement of these best practices. ISOC provides some aggregated statistics on their participants' conformances to these best practices [117] but does not publicly detail non-conformance. Previous work has found inconsistent adherence of MANRS members to specific recommendations, e.g., the Source Address Validation action (Recommended Action #2) [108], but we are not aware of an independent study of MANRS members' overall conformance with MANRS actions. Lack of empirical assessment of conformance prevents a collective understanding of how effective MANRS is at achieving its objectives. Indeed, this lack of understanding recently prompted U.S. regulators to launch a Notice of Inquiry to ascertain whether there was a role for government intervention to improve routing security [36].

Our goal in this chapter is to evaluate the state of the MANRS initiative as of 2022 in terms of its participants, their routing behavior and the broader impact of MANRS on the routing ecosystem, and discuss potential improvements. Figure 4.1 outlines the analysis to answer the

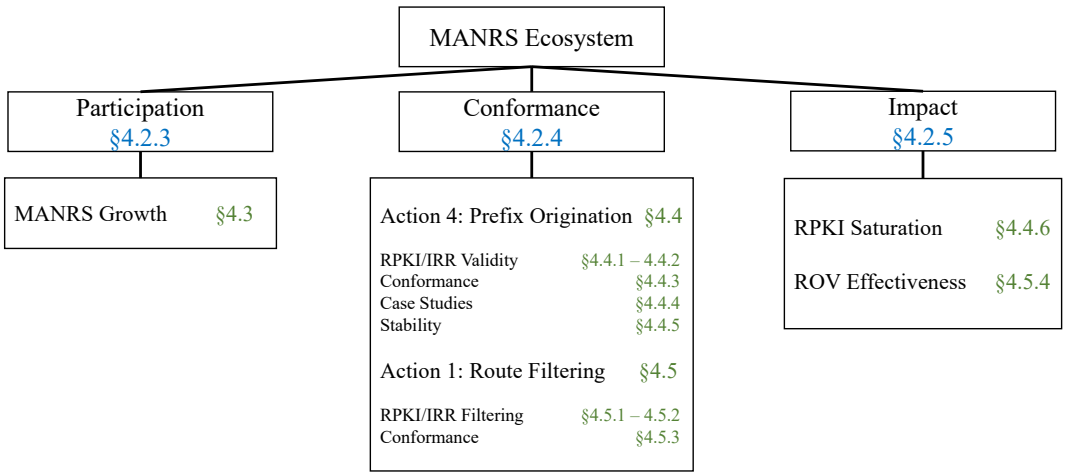


Figure 4.1. Structure of the chapter. Methodology sections are in blue and results sections are in green.

following research questions

RQ1: Who is part of MANRS? To understand where MANRS has gained traction, we look at the geographical distribution of ASes in MANRS and their service regions. We use customer-cone size, size of originated address space, and size of address space covered by RPKI objects of ASes to further characterize MANRS participants and their significance in each RIR.

RQ2: Are MANRS networks conformant with MANRS actions? We seek to assess how conformant MANRS members are to Action 1 and Action 4 of the MANRS ISP and CDN programs (§1), which relate to the origination and propagation of incorrect routing information. These actions are not new or unique to MANRS; they were proposed to improve routing security many years ago but their implementation has struggled [122]. Currently, MANRS does not enforce the implementation of these actions despite being required for members. Therefore, we want to evaluate the extent to which MANRS members actually employ these actions.

We cannot reasonably expect all MANRS members to be continually and entirely conformant to MANRS actions. Operating a large network, maintaining objects in routing registries, and coordinating with customers are all laborious and error-prone tasks, even with the help of automated tools [26]. In addition, operators have reported technical caveats that prevented them from complete ROV deployment [136], and many commercial ISPs are not filtering customers'

announcements [19, 74]. As a result, operators need to constantly monitor their resources to ensure that their—and their customers'—advertisements match IRR and RPKI entries, and may have to communicate with customers to create or update ROAs.

MANRS Action 1 for CDNs, which relates to the propagation of incorrect routing announcements from either MANRS members or their customers, allows propagated announcements to be more specific than the records registered in IRR. Such a mismatch in prefix lengths is generally due to prefix de-aggregation in the context of traffic engineering practices [32]. Thus, we treat BGP announcements with IRR status of *invalid prefix length* as conformant to MANRS Actions 1 and 4.

RQ3: What is the impact of MANRS on the broader routing ecosystem? We want to gauge the effect MANRS has in securing the routing system. First, we evaluate if MANRS networks indeed have better security practices than non-MANRS networks. Specifically, we assess AS behavior with respect to registering RPKI ROAs for its prefixes and performing route origin validation In addition, we measure if MANRS members reduce the number of invalid prefixes transiting their networks.

4.2 Methodology

We evaluate the MANRS ecosystem along three key dimensions: participation, conformance, and impact. We also describe our AS classification methodology. Our study draws on several public data sources, including *prefix2as* snapshots from 2015 to 2022, the *as2org dataset* from April 2022, *RPKI archives* from 2014 to 2022, and *IRR archives* collected between November 2021 and May 2022.

We capture MANRS participants in the ISP and CDN programs as of May 1, 2022, and refer to them as the *MANRS ISP dataset* and *MANRS CDN dataset*, respectively. The Internet Society also kindly provided us with the dates when each network joined its respective MANRS program; we refer to this as the *historical MANRS dataset*.

Additionally, we process the *Internet Health Report* (IHR) to extract IRR and RPKI validation status for observed BGP prefix-origin pairs (*IHR prefix-origin dataset*) and for ASes appearing in paths to those prefixes (*IHR transit dataset*).

To evaluate the routing behavior of MANRS participants and compare it to that of non-participants, we rely on the following RPKI and IRR validation statuses, as well as additional classification metrics:

RPKI validity We use the RPKI validity statuses defined in §1.5. For RPKI, a prefix origin is *Valid* if there is at least one VRP with prefix, ASN, and max length attributes all matching the route. If all VRPs covering the route have an ASN different from the route's origin AS then the route is *Invalid*. If at least one VRP has a matching ASN but the Max Length attribute is not covering the route, then the route is classified as *Invalid Length*. A route is *Not Found* if no VRP covers it.

IRR validity For IRR, we apply the same classification method as RPKI, but since there is no standardized max length attribute in IRR, we consider the prefix length as the max length value for IRR entries.

AS Customer Degree. The routing complexity of a network increases with its customer degree, especially when additional route filtering is deployed. To perform a fair comparison of conformance between ASes of similar routing complexity, we classify ASes into three sizes by their number of AS-level customers inferred by CAIDA's AS Rank [23], using the thresholds defined by Dhamdhere et al. [41]:

- **Small networks**: $CustomerDegree \leq 2$
- **Medium networks**: $2 < CustomerDegree \leq 180$
- **Large networks**: $CustomerDegree > 180$

4.2.1 MANRS participation

To understand the MANRS ecosystem, we begin by examining key characteristics of participating networks, including their global footprint and the extent to which they engage with

existing routing security mechanisms.

Geographical Distribution We use the *as2org dataset* to find the countries of headquarters of the MANRS organizations (ISPs and CDNs) and the RIRs that allocated their corresponding ASes, and match the information in our *historical MANRS dataset* to obtain a distribution of MANRS ASes by RIR over time.

Routing Table Presence We use the *prefix2as dataset* and the *historical MANRS dataset* to find the MANRS ASes in the BGP table and calculate the fraction of IPv4 address space announced by them per year.

Registration Completeness We use the *as2org dataset* to identify all ASes allocated to each MANRS member organization, including those not registered in MANRS. We manually check the correctness of this dataset in our case studies. We then calculate the fraction of ASes in the MANRS participants list over all ASes of each organization. We also use the *prefix2as dataset* to calculate the fraction of IP address space originated by registered MANRS ASes over total address space originated by all ASes of each organization.

4.2.2 MANRS conformance with Action 4 and Action 1

In the *IHR prefix origin dataset* and the *IHR transit dataset*, we define a prefix-origin pair to be *MANRS-conformant* if its RPKI status is Valid, or its IRR status is Valid or Invalid length (given IRR does not have a max len attribute), and *MANRS-unconformant* if it is RPKI Invalid or (RPKI NotFound, IRR Invalid).

Prefix Origination Behavior We use the *IHR prefix origin dataset* to obtain the prefixes originated by each AS and their RPKI and IRR statuses. We calculate the RPKI origination validity ($OG_{RPKIvalid}$) for each AS using Formula 4.1:

$$OG_{RPKIvalid} = \frac{\#\ of\ RPKI\ Valid\ prefixes}{total\ \#\ of\ originated\ prefixes} \times 100\% \tag{4.1}$$

We then calculate the IRR origination validity ($OG_{IRRvalid}$) for each AS using Formula

4.2:

$$OG_{IRRvalid} = \frac{\#\ of\ IRR\ Valid\ prefixes}{total\ \#\ of\ originated\ prefixes} \times 100\% \tag{4.2}$$

We then calculate the MANRS Action 4 Conformance ($OG_{conformant}$) for each AS using Formula 4.3:

$$OG_{conformant} = \frac{\#\ of\ MANRS\text{-}conformant\ prefixes}{total\ \#\ of\ originated\ prefixes} \times 100\% \tag{4.3}$$

Route Filtering Behavior It is challenging to measure ROV deployment at scale, and as of 2022, no previous work has done so with high confidence and sufficient validation (§1). Instead of measuring ROV deployment, we analyze the prevalence of RPKI Invalid announcements that propagated through each AS. We then use the *IHR transit dataset* to obtain the BGP announcements propagated by each AS and their RPKI statuses and calculate the RPKI propagation invalidity ($PG_{RPKIinv}$) for each AS using Formula 4.4. We discuss the limitations of our methodology in §3.5.4.

$$PG_{RPKIinv} = \frac{\#\ of\ RPKI\ Invalid\ +\ RPKI\ Invalid\ Length\ prefixes}{total\ \#\ of\ propagated\ prefixes} \times 100\% \tag{4.4}$$

We calculate the IRR propagation invalidity (PG_{IRRinv}) for each AS using Formula 4.5.

$$PG_{IRRinv} = \frac{\#\ of\ IRR\ Invalid\ prefixes}{total\ \#\ of\ propagated\ prefixes} \times 100\% \tag{4.5}$$

We identify the direct customers of each AS using the *AS Relationship dataset* and focus on BGP announcements from customers. We calculate the MANRS Action 1 Unconformance (PG_{unc}) for each AS using Formula 4.6.

$$PG_{unc} = \frac{\#\ of\ MANRS\text{-}unconformant\ prefixes}{total\ \#\ of\ propagated\ customer\ prefixes} \times 100\% \tag{4.6}$$

4.2.3 MANRS impact

To study the impact of MANRS networks, we quantify their level of completion in RPKI registration and overall ROV effectiveness.

RPKI Coverage To conduct historical analysis on RPKI registration, we use the annual snapshots of the *prefix2as dataset* from 2015 to 2022 and *RPKI dataset* snapshots with matching dates, and use the Route Origin Validation (ROV) algorithm [122] to calculate the RPKI status of all BGP announcements for each snapshot. We then calculate the RPKI coverage, denoted $RCov_m$, for MANRS ASes defined by Equation 4.7.

$$RCov_m = \frac{ROA\ covered\ MANRS\ Routed\ Address\ Space}{Total\ MANRS\ Routed\ Address\ Space} \times 100\% \tag{4.7}$$

We also calculate the RPKI coverage for non-MANRS ASes ($RCov_n$) defined by Equation 4.8.

$$RCov_n = \frac{ROA\ covered\ non\text{-}MANRS\ Routed\ Address\ Space}{Total\ non\text{-}MANRS\ Routed\ Address\ Space} \times 100\% \tag{4.8}$$

Overall ROV Effectiveness To study the overall ROV deployment of MANRS networks collectively, we analyze whether RPKI invalid (and invalid prefix length) BGP announcements are more likely to propagate through MANRS networks than non-MANRS networks. Intuitively, if MANRS networks have more effective ROV deployment than non-MANRS networks, RPKI invalid announcements should be less likely to propagate through the former.

Using the *IHR transit dataset*, for each prefix origin pair PO_k, we find its transit ASes and denote the hegemony score (§1.7.7) of each MANRS AS H_{ASi}^{MANRS} and non-MANRS AS H_{ASj}^{XMANRS}. The list of transit AS hegemony scores is denoted below:

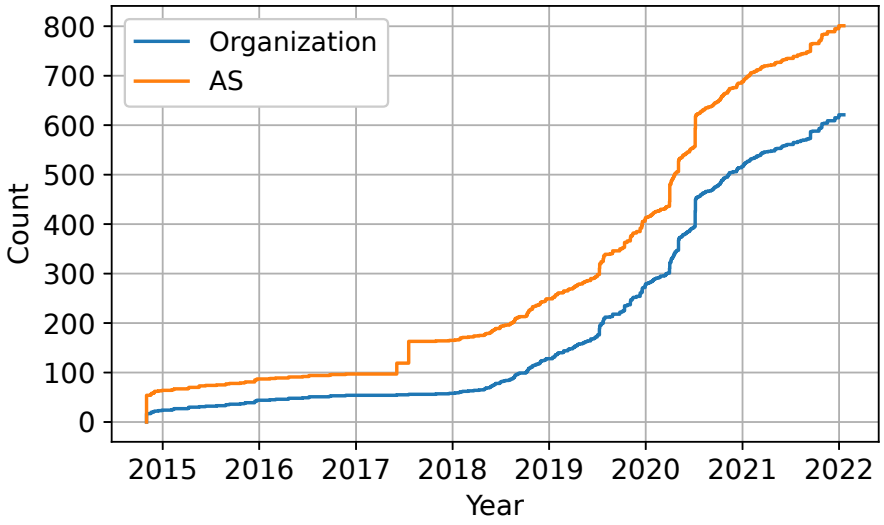


Figure 4.2. MANRS participants have grown significantly between 2015 and 2022, especially after 2019.

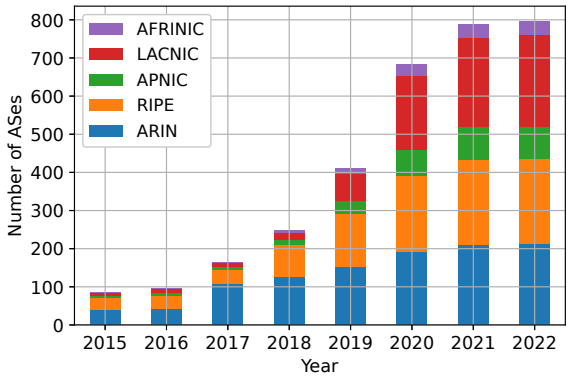
$$\{PO_k : H_{AS1}^{MANRS} \cdots H_{ASm}^{MANRS}, H_{AS1}^{XMANRS} \cdots H_{ASn}^{XMANRS}\}$$

Each AS hegemony value represents how likely an AS is to be on the paths towards PO_k. Equation 4.9 defines a *MANRS preference score* PS_k^{MANRS} for PO_k, which represents how much more (less) likely PO_k is to go through MANRS transit ASes. PO_k is more likely to propagate through MANRS networks if PS_k^{MANRS} has a value greater than 0. We compare the distribution of *MANRS preference scores* for RPKI Invalid, Valid, and NotFound prefix origin pairs.

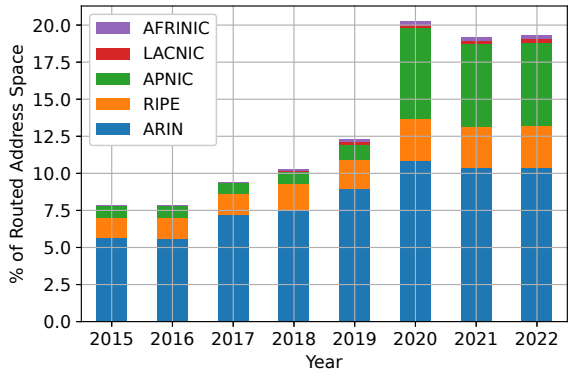
$$PS_k^{MANRS} = \sum_{i=1}^{m} AS_i^{MANRS} - \sum_{j=1}^{n} AS_j^{XMANRS} \tag{4.9}$$

4.3 MANRS Ecosystem Characteristics

Figure 4.2 shows the growth of MANRS between 2015 and 2022 in terms of both organizations and ASes. We analyze this growth by geographic region, in terms of member ASes (Figure 4.3a) and address space originated by those ASes (Figure 4.3b).



(a) MANRS ASes over time. Brazil (in LACNIC region) added 90 small ASes in 2020 due to local outreach efforts.



(b) Percentage of MANRS routed IPv4 address space. MANRS ASes in the ARIN region announce the most address space.

Figure 4.3. MANRS ASes and share of routed IPv4 address space over time. The excessive jump in LACNIC ASes was caused by a strong outreach effort by NIC.br in Brazil (184 of the 239 LACNIC ASes are registered in Brazil), but those ASes contributed little additional (0.24%) routed IPv4 address space. The large jump in address space (b) was caused by China Telecom (AS4134, 4.0% of routed v4 address space) joining MANRS. Most large networks (transit providers, CDNs) are from the ARIN region.

Some anomalies in Figures 4.3a and 4.3b merit further explanation. The significant jump in LACNIC ASes was caused by a strong outreach effort by Brazil's national Internet registry (NIC.br) [174], but those ASes contributed little additional IPv4 address space. In contrast, the large jump in APNIC address space was due mostly to a single ISP – China Telecom – joining MANRS in 2020. Also in 2020, the increase in ARIN routed IPv4 address space was caused by the introduction of the MANRS CDN and Cloud Provider Program, where Amazon (AS16509) represented 1.3% of routed IPv4 address space. In 2021, Level3 (AS3356, now Lumen Technologies) and China Telecom (AS4134) announced fewer prefixes than in 2020, causing a drop in routed address space for those regions (Figure 4.3b).

Many ISPs own more than one AS number, and MANRS allows each organization to specify which of their AS numbers will become MANRS members and thus subject to the MANRS requirements. We use the *as2org dataset*, the *MANRS ISP dataset*, and the *MANRS CDN dataset* to identify sibling ASes (i.e., owned by the same organization) of registered MANRS ASes. We then use the *IHR prefix origin dataset* and estimate that, as of May 2022, 463

(70%) organizations registered all of their AS numbers in MANRS, and 543 (82%) announced IPv4 address space *only* through registered ASes. We find that, as of May 2022, 117 MANRS organizations announced some of their IP address space from ASes that were not MANRS members, and, 8 of these 117 only announced their IP address space from non-MANRS ASes. This means that for those organizations, MANRS conformance could not accurately reflect their routing security practices because some of their originated prefixes were not under MANRS scrutiny. On the contrary, we find 80 organizations that did not register all of their ASes in MANRS but only announced IP address space from the MANRS ASes, indicating that they did not register their quiescent ASes in MANRS. Some examples include cloud hosting companies that only provided services to their customers through their main AS.

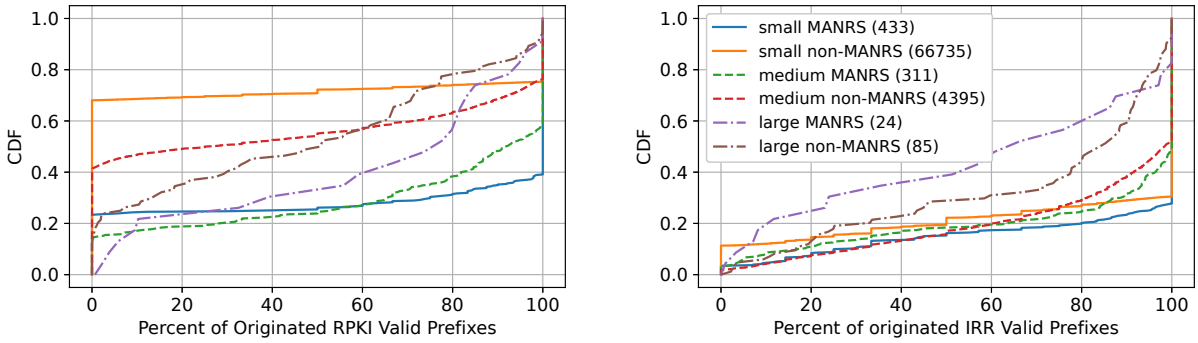
4.4 MANRS Action 4: Prefix Origination Behavior

MANRS Action 4 in the ISP and CDN programs requires participating ASes to originate prefixes that are either RPKI Valid or IRR Valid (§1.6). In this section, we quantify the validity of prefixes originated by MANRS ASes, compare to that of non-MANRS ASes, and separately check the conformance of ISPs and CDNs to their corresponding action. The MANRS ISP program states that its members must originate at least 90% IRR/RPKI Valid prefixes, while the MANRS CDN program requires 100%.

4.4.1 RPKI prefix validity

Figure 4.4 depicts the distribution of the overall percentage of valid prefixes originated by small, medium and large MANRS and non-MANRS ASes. Figure 4.4a shows that for small networks, RPKI validity distribution was bimodal: in May 2022, prefixes originated by most ASes were either entirely valid or entirely invalid. The bimodal distribution was due to small networks generally originating few prefixes - the 75th percentile of small networks originated only 5 prefixes, and thus could more easily maintain RPKI registration.

Small MANRS ASes were about 2.5 times more likely to register ROAs for their prefixes.



(a) MANRS ASes were more likely than non-MANRS ASes of the same size to originate RPKI Valid prefixes.

(b) Large MANRS ASes were less likely to originate IRR Valid prefixes than large non-MANRS ASes.

Figure 4.4. Prefixes originated by MANRS and non-MANRS ASes (shared legend). Large MANRS ASes were more likely to originate RPKI Valid prefixes than IRR Valid prefixes in May 2022.

Indeed, in May 2022, of 433 small MANRS ASes that originated prefixes, 264 (60.1%) originated only RPKI Valid prefixes while 102 (23.6%) originated only non-RPKI-Valid prefixes. In contrast, of 66,735 small non-MANRS ASes, 45419 (68.1%) ASes originated only RPKI Invalid/NotFound prefixes while 16,492 (24.7%) ASes originated only RPKI Valid prefixes.

The RPKI validity difference between medium MANRS and non-MANRS ASes was similar to that of the small networks, with medium MANRS ASes almost twice as likely to originate only RPKI Valid prefixes. In May 2022, of 311 and 4,395 medium MANRS and non-MANRS ASes, 129 (41.5%) and 1,044 (23.8%) originated only RPKI Valid prefixes, while 46 (14.8%) and 1,818 (41.4%) originated only RPKI Invalid/NotFound prefixes. This also shows a positive impact of MANRS for medium networks.

The validity distribution for large networks was less polarized than that of small and medium networks. However, while all large MANRS ASes originated some RPKI Valid prefixes, we find that 10 (11.8%) non-MANRS ASes did not register ROAs for any prefix they announced and thus originated only RPKI NotFound prefixes. Additionally, another large network (AS23947, an Indonesian ISP) did not originate any RPKI Valid prefix and originated 2 RPKI Invalid prefixes. We further study this case below. On the other end, 3 (12.5%) out of 24 large MANRS ASes and

Table 4.1. Number of ASes originating RPKI Invalid prefixes by size category. In May 2022, MANRS ASes of all size categories were more likely to originate RPKI Invalid announcements compared to non-MANRS ASes.

Size Category	Group	# ASes Orig. RPKI-Inv. (%)	Total ASes	# Prefixes
Small	MANRS	0 (0.0%)	433	0
	Non-MANRS	489 (0.7%)	66,735	1,097
Medium	MANRS	9 (2.8%)	311	14
	Non-MANRS	198 (4.5%)	4,395	1,401
Large	MANRS	5 (20.8%)	24	13
	Non-MANRS	28 (32.9%)	85	724

5 (5.9%) out of 85 large non-MANRS ASes originated only RPKI Valid prefixes.

We further study the ASes that originated RPKI Invalid prefixes, since ROV-deploying networks will drop those routes while allowing RPKI NotFound announcements to pass. Table 4.1 shows that, in May 2022, no small MANRS AS originated RPKI Invalid prefixes while 489 (0.7%) small non-MANRS ASes originated 1,097 RPKI Invalid prefixes. 9 (2.8%) medium MANRS ASes originated 14 RPKI Invalid prefixes and 198 (4.5%) medium non-MANRS ASes originated 1,401 RPKI Invalid prefixes. We also find that 5 (20.8%) large MANRS ASes originated 13 RPKI Invalid prefixes, while 28 (32.9%) of large non-MANRS ASes originated 724 RPKI Invalid prefixes, showing that MANRS networks were less likely to originate RPKI Invalid announcements.

We note that ASes may originate RPKI Invalid prefixes for legitimate reasons, such as routing misconfigurations. We further analyze the Indonesian ISP (AS23947) that originated 2 RPKI Invalid prefixes and found that the prefixes were registered under `AS0` in RPKI. Since both prefixes were registered under `AS23947` in RADB and have been announced consistently since May 2019, we speculate that this ISP misconfigured its ROA.

4.4.2 IRR prefix validity

Figure 4.4b shows that the median large MANRS network originated 63.5% IRR Valid prefixes of all prefixes the network originated, which was lower than the 84.0% median for the non-MANRS counterpart. We speculate this difference is due to (1) IRR being more widely adopted than RPKI and (2) networks that adopt RPKI (especially MANRS networks with high RPKI registration rates) leaving IRR records unmaintained, causing BGP announcements to become IRR Invalid and creating inconsistency between IRR and RPKI records [42].

In contrast, MANRS and non-MANRS ASes in the small and medium classes had similar likelihood in originating IRR Valid prefixes. In May 2022, the percentage of networks that only originated IRR Valid prefixes was 72.3% for small MANRS, 70.0% for small non-MANRS, 52.1% for medium MANRS, and 48.0% for medium non-MANRS.

We also find that non-MANRS networks were more likely to register only in IRR but not RPKI compared to MANRS networks. In May 2022, 23.6% (102/433) small MANRS ASes versus 65.4% (43,659/66,735) small non-MANRS ASes only registered in IRR. Similarly, 14.8% (46/311) medium MANRS compared to 41.0% (1,803/4,395) non-MANRS ASes, and 0% (0/24) large MANRS ASes compared to 11.8% (10/85) large non-MANRS ASes, registered only in IRR. Registering only in IRR is less optimal than registering in RPKI, since IRR may contain inaccurate records due to looser validation standards compared to RPKI [42]. The lower fractions of IRR-only MANRS ASes suggests MANRS participants were more likely to adopt RPKI than their non-MANRS counterparts in 2022.

4.4.3 AS level conformance to Action 4

After looking at the prefix validity levels of ASes, we look at how conformant were MANRS ISPs and CDN program participants to required validity levels. We first look at how conformant MANRS CDN participants were to the Action 4 requirements in the MANRS CDN program, which requires CDNs to originate all prefixes as either RPKI or IRR Valid. We discover

one MANRS CDN AS did not originate any prefix according to our *prefix2as* dataset. In the May 1st snapshot of the *IHR prefix origin dataset*, we find that the AS belonged to a company that registered two ASes in MANRS but was only using one AS to originate address space in BGP (similar to other cases mentioned in §4.3). We consider that AS to be trivially conformant. We find that 17 out of 20 CDN ASes were conformant, i.e., 100% of the prefixes they originated were either RPKI Valid or IRR Valid. The other three ASes still originated more than 98% of their prefixes as either IRR or RPKI Valid. These three ASes were among the top 5 MANRS CDN ASes who originated the most prefixes, two of which originated more than 3,500 prefixes in BGP. For ASes that originate so many BGP announcements, it is reasonable that a small fraction of them were neither RPKI Valid nor IRR Valid due to their complicated relationship with business customers. We provide case studies on those three CDNs (§4.4.4).

We then analyze the conformance of MANRS ISP participants to the MANRS ISP program Action 4 requirements, which requires ISPs to have over 90% of the prefixes they originate in BGP to be either IRR or RPKI valid. In the same *IHR prefix origin dataset* snapshot from May 2022, we find 95 ASes that did not originate any prefix and considered them trivially conformant. We then find that only 39 out of 754 ASes were unconformant, where fewer than 90% of the prefixes they originated were IRR or RPKI valid. The 39 ASes belonged to 15 organizations total: 24 ASes were from one large ISP (ISP1), 2 ASes belonged to ISP2, and the remaining 13 organizations each had one AS. The percentages of IRR/RPKI Valid prefixes originated by those 39 ASes ranged from 0% to 89%. We manually inspect some of those ASes with no IRR/RPKI Valid prefixes and find they were stub ASes of large networks who originated fewer than 3 prefixes. We reached out to some of those non-conformant networks and discuss their behavior next.

4.4.4 Analyzing unconformant ASes

We manually analyze 3 unconformant ISPs and 3 unconformant CDNs to shed light on the reasons for their unconformance.

Table 4.2. Non-conformant prefix origins from the MANRS networks we contacted. CDN2 had no Invalid prefixes and only one RPKI-NotFound prefix. Prefixes in Sibling/C-P were originated by a sibling or customer AS, and no relationship was found for prefixes in the Unrelated column. The RPKI Invalid and IRR Invalid columns each sums the 2 subsequent columns.

	RPKI Invalid (NotFound)	Sibling/C-P	Unrelated	IRR Invalid & No RPKI	Sibling/C-P	Unrelated
CDN1	3	3 (100%)	0	48	38 (79%)	10 (21%)
CDN2	(1)	0	1 (100%)	0	0	0
CDN3	0	0	0	5	5 (100%)	0
ISP1	1	0	1 (100%)	302	154 (51%)	148 (49%)
ISP2	8	6 (75%)	2 (25%)	272	152 (56%)	120 (44%)
ISP3	1	1 (100%)	0	486	359 (74%)	127 (26%)

We conduct case studies for all three unconformant CDNs and the three largest unconformant ISPs. Table 4.2 provides an overview of the prefix-origins that were not MANRS-conformant. The RPKI Invalid and IRR Invalid columns contain the number of prefix-origins with those statuses (prefix-origins in the IRR Invalid column were all RPKI NotFound), which combined is a subset of the total non-MANRS-conformant prefixes of the network (excluding prefix-origins that were both RPKI NotFound and IRR NotFound). The Sibling/C-P columns denote the number of prefix-origins whose BGP origin AS and mismatching RPKI/IRR origin AS belonged to the same organization (i.e., siblings according to the *as2org dataset*) or had a customer-provider relationship (C-P, according to the *AS relationships dataset*). The Unrelated columns denote the number of remaining prefix-origins whose BGP origin ASes had no relationship with the corresponding IRR/RPKI origin ASes.

Most prefix-origins that were not conformant were IRR Invalid instead of RPKI Invalid. Since RPKI Invalid prefix-origins suffer more visibility reduction in the global routing table (due to ROV filtering) than RPKI NotFound or IRR Invalid ones, networks who originated IRR Invalid prefixes due to temporary traffic engineering efforts or misconfigurations are less likely to have their services negatively impacted. In addition, in almost all cases, the majority of the RPKI Invalid/IRR Invalid prefix-origins were under the Sibling/C-P category, meaning the lack

of compliance was likely due to internal misconfigurations or business practices, and thus could be easily corrected.

All six organizations in our case study were partial MANRS participants according to our analysis using the *as2org dataset*. We also manually check the AS numbers belonging to those organizations and contact their network operators to confirm that they did not register all AS numbers in MANRS. ISP2 mentioned that they intended to register all AS numbers in MANRS but might have missed some newer AS numbers. CDN1 listed only one AS number in MANRS, with 98.7% MANRS-conformant prefix-origins, but had 11 out of 12 unlisted ASes with 100% conformant prefixes. Similarly for CDN2, CDN3, and ISP1, 75%, 100%, and 100% of their unlisted ASes had 100% conformant prefixes.

4.4.5 Conformance stability

We next look at MANRS conformance over time. We take 12 weekly snapshots from the *IHR prefix origin dataset* between February 1st, 2022 and May 1st, 2022 (1 snapshot was removed due to missing data) to analyze the conformance stability of MANRS ASes. We find that 17 out of 20 CDNs stayed in conformance for all 12 snapshots, and the 3 CDNs stayed unconformant for 12 weeks. For MANRS ISPs, we found 46 ASes were not conformant in some of the 12 snapshots, where 35 ASes were unconformant consistently across all snapshots. The remaining 11 ASes that were unconformant for fewer than 12 snapshots belonged to 10 organizations. We also find one AS had fluctuating conformance status where its MANRS-conformant prefix-origins dropped below 90% in early February and again in late March. Overall, most ASes were stable in their conformance and consistently stayed either conformant or unconformant.

We look at the 3 CDNs to study prefix-level conformance stability. For CDN1, we find that between February and May 2022, it stopped announcing 80 prefixes and announced 141 new prefixes, while the active 3,822 prefixes remained conformant. Similarly for CDN2, no prefix changed conformance status over the 3-month period, and for CDN3, only 2 out of 902 prefixes changed conformance status. Overall, prefixes were also likely to have stable conformance status,

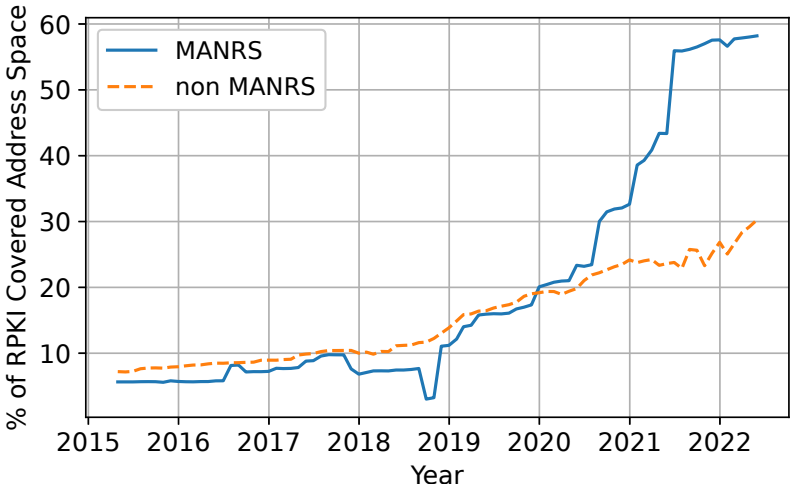


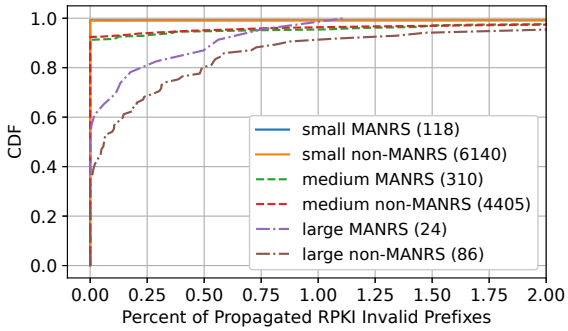
Figure 4.5. Percentage of RPKI covered address space for MANRS and non-MANRS networks. RPKI covered address space of MANRS networks grew faster between 2019 and 2022.

possibly due to infrequent changes in RPKI and IRR registration.

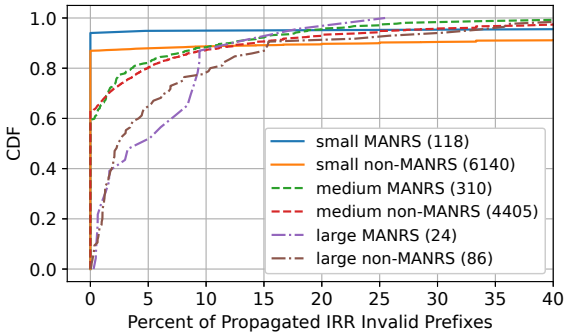
4.4.6 Presence in RPKI of MANRS ASes

We find that in May 2022, 64.8% of the routed IPv4 address space was not covered by RPKI VRPs, while only 5.3% was not covered by IRR route objects. This shows that the RPKI has significant room for growth. We find that MANRS ASes had higher RPKI coverage than non-MANRS networks.

Figure 4.5 shows that in May 2022, MANRS networks had reached an RPKI coverage level of 58.2% (*i.e.,* that fraction of their collective address space had been RPKI signed), greater than 30.2% for non-MANRS networks. Although there is still room for MANRS RPKI coverage to grow, it is unlikely to reach 100% due to the barriers in RPKI registration of legacy IPv4 address space [49]. We discover that the significant increase of MANRS RPKI coverage after 2020 was due to the introduction of the MANRS CDN program, where large CDNs and cloud providers such as Amazon (AS16509) and Cloudflare (AS13335) registered more than 1,700 prefixes in RPKI. This change correlates with the increase of ARIN MANRS address space shown in Figure 4.3b. Similarly, 95.0% of MANRS address space was covered by IRR and



(a) Percent of RPKI Invalid prefixes propagated by MANRS and non-MANRS networks by size. Large MANRS networks propagated limited RPKI Invalid prefixes, no more than 1.1% of their total.



(b) Percent of IRR Invalid prefixes propagated by MANRS and non-MANRS networks by size. Large MANRS networks propagated fewer than 25% IRR Invalid prefixes.

Figure 4.6. RPKI and IRR Invalid prefixes propagated by MANRS and non-MANRS networks in May 2022.

97.6% was MANRS-conformant; 84.6% of non-MANRS address space was covered by IRR and 87.2% was MANRS-conformant.

4.5 MANRS Action 1: Route Filtering Behavior

In addition to requirements on prefix origination, MANRS requires networks to ensure the correctness of customer announcements. In this section, we quantify the RPKI/IRR Invalid prefixes propagated through MANRS networks and originated by their customers, compare to that of the non-MANRS networks, and analyze conformance to the Action 1 requirements.

4.5.1 Propagation of RPKI-Invalid prefixes

Figure 4.6a shows the distribution of the percentage of valid prefixes forwarded by networks for small, medium and large, MANRS and non-MANRS networks. It shows a distinct difference between large MANRS and non-MANRS networks. 11 (45.9%) of 24 large MANRS networks propagated no RPKI Invalid prefixes compared to 31 (36.0%) of 86 for large non-MANRS networks. The large MANRS network that propagated the most RPKI-Invalid announcements did so for 1.1% of its address space, compared to 6.4% for the top large

non-MANRS network.

In contrast, there was no such difference between MANRS and non-MANRS ASes for small and medium networks. Indeed, small networks propagated almost no RPKI Invalid announcements. Of 118 small MANRS ASes that propagated some prefix, 117 (99.2%) propagated no RPKI Invalid prefixes while 1 (0.8%) propagated only 1 RPKI Invalid prefix. Of 6,140 small non-MANRS networks, 6,083 (99.1%) ASes propagated no RPKI Invalid prefixes. This similarity was because small networks are mostly edge ASes with almost no customers and therefore propagate very few prefixes in general: the 75th percentile of small networks propagated only 5 prefixes. Since RPKI Invalid announcements occupied less than 1% of the routing table, it is very unlikely for small networks to encounter RPKI Invalid announcements.

Similarly, medium MANRS and non-MANRS networks were almost indistinguishable with respect to the percentage of propagated RPKI Invalid announcements. Of 310 and 4,405 medium MANRS and non-MANRS networks, 283 (91.3%) and 4,072 (92.4%), respectively, propagated no RPKI Invalid prefixes.

4.5.2 Propagation of IRR-Invalid prefixes

Figure 4.6b shows that in our May 2022 data, small MANRS networks were less likely to propagate IRR Invalid prefixes than small non-MANRS networks. However, medium MANRS and non-MANRS networks had similar likelihood of propagating IRR Invalid prefixes. In May 2022, the percentage of networks that only originated IRR Valid prefixes was 94.1% for small MANRS, 85.5% for small non-MANRS, 59.4% for medium MANRS, and 63.0% for medium non-MANRS. Large MANRS networks propagated at most 25.5% IRR Invalid announcements of all their received announcements, which was lower than the 74.5% of their non-MANRS counterpart. We speculate this difference was due to inconsistent IRR adoption and filtering among ISPs [179], and especially among non-MANRS networks without standard filtering procedures.

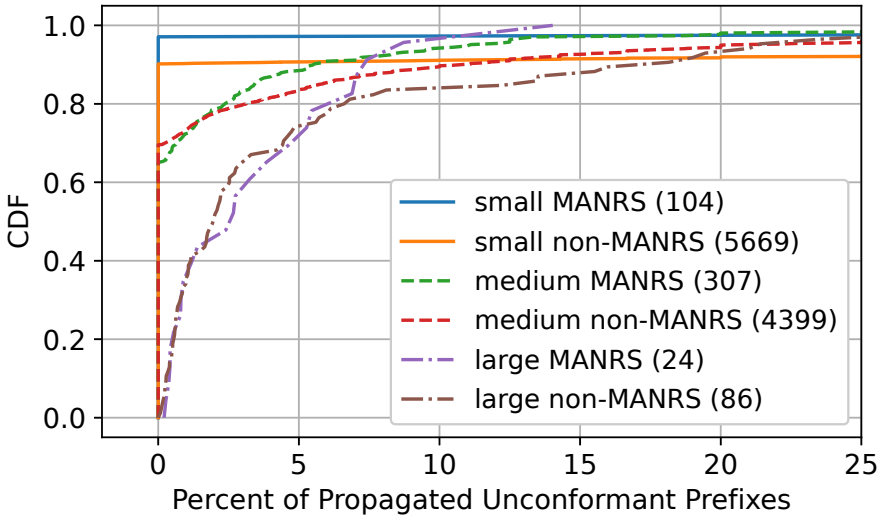


Figure 4.7. Unconformant prefixes propagated by MANRS and non-MANRS networks by network type. The median large MANRS AS propagated only 2.5% MANRS-unconformant prefixes.

4.5.3 AS level conformance to Action 1

To assess networks' level of conformance to the filtering requirement, we combine members of MANRS ISPs and CDNs programs and calculated their fractions of propagated MANRS-unconformant announcements (§4.2.2) from their *direct customers*. Since the MANRS ISP program Action 1 does not provide a threshold, and the MANRS CDN program Action 1 only describes a validation flow in its recommended additional requirements, a MANRS AS is *fully* Action 1 conformant if *none* of the announcements they propagate are MANRS-unconformant. MANRS ASes that did not propagate any announcements are considered trivially conformant

Figure 4.7 shows that in May 2022, all types of MANRS ASes were more likely to be Action 1 conformant than their non-MANRS counterpart. All large MANRS ASes propagated fewer than 15% MANRS-unconformant prefixes, while the highest percentage of MANRS-unconformant prefixes propagated by large non-MANRS ASes was 41.4%. The long tails for small and medium non-MANRS ASes both end at 100%, where we found one medium non-MANRS AS propagated more than 800 RPKI or IRR Invalid prefixes. To emphasize, those prefixes were received from the direct customers of the ASes, hence the fewer ASes per category

Table 4.3. Action 1 (filtering) conformance. ASes with no customers are conformant by default and included in right side columns. Results for transit ASes are shown in left side columns. Only 23% of small MANRS ASes provided transit and 97.1% of them were conformant.

	Transit Conformant	Total Transit	Total Conformant	Total MANRS
Small	101 (97.1%)	104	448 (99.3%)	451
Medium	200 (65.1%)	307	212 (66.4%)	319
Large	0 (0%)	24	0 (0%)	24

(Figure 4.7 legend vs. Figure 4.6b).

Table 4.3 lists the number of fully conformant MANRS ASes. The Total Transit column shows the number of MANRS ASes that actually propagated some announcement. We consider the remaining MANRS ASes that we did not observe to have propagated any announcements trivially conformant. We obtain the Total Conformant values by adding the number of trivially conformant ASes to the number of Transit Conformant ASes. We find that 347 of 451 small MANRS ASes did not provide transit to any prefixes at all. Overall, of MANRS ASes across all size categories were more than 93% conformant to Action 1.

4.5.4 MANRS RPKI filtering effectiveness

To measure whether RPKI Invalid announcements were more likely to propagate through MANRS transit networks, we first estimate how likely BGP announcements were to propagate through MANRS networks by calculating the *MANRS preference scores* for RPKI Valid and RPKI NotFound announcements. Figure 4.8 depicts their distribution and shows that 34% of RPKI Valid prefix origin pairs preferred to transit via MANRS networks (*MANRS preference scores* greater than 0). Similarly for RPKI NotFound announcements, 36% preferred MANRS transit networks. The similarity between these *MANRS preference scores* is expected as networks propagate RPKI Valid and NotFound in the same way, even if they implement ROV. In contrast, only 14% of RPKI Invalid prefix origin pairs preferred MANRS networks, suggesting that MANRS networks collectively were more effective in filtering RPKI Invalid announcements

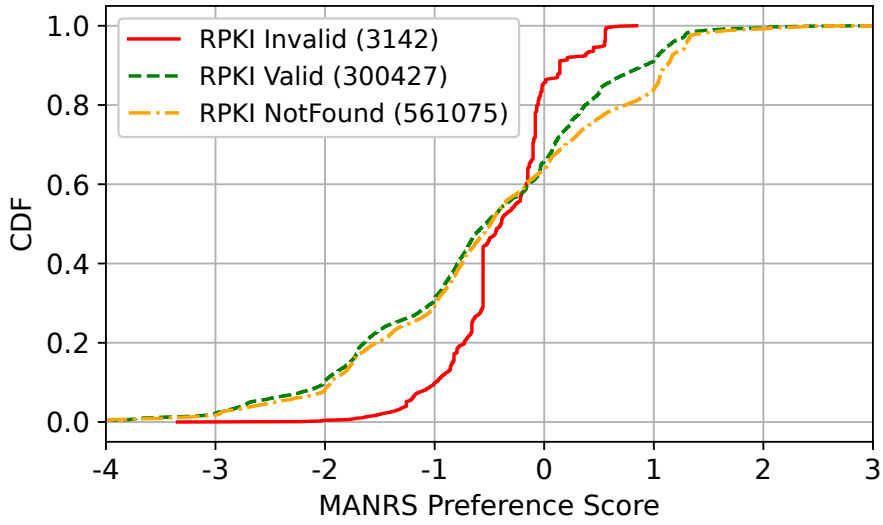


Figure 4.8. Prefix preference for MANRS transit by RPKI status (0 means no preference). RPKI Invalid BGP announcements were less likely to propagate through MANRS networks than non-MANRS networks.

than non-MANRS networks, and therefore suggesting better routing security.

4.6 Discussion

We surveyed six MANRS-participating network operators whose prefixes were found to be unconformant with IRR or RPKI validation policies. We contacted each operator to share our findings and received replies from five of them. Two operators identified the root cause of the unconformant announcements and confirmed that they had corrected the issue, which we were able to independently verify at the time of writing. The remaining three networks reported that they were using our findings to investigate the problem internally.

We also asked these operators for feedback on the monthly MANRS conformance reports. Three networks indicated that they read the reports but found the information insufficiently actionable. The other two operators were not aware of the reports, suggesting that we may have reached someone other than the designated MANRS point of contact.

In follow-up discussions, two operators noted that complying with MANRS requirements was challenging due to complex business relationships and legacy infrastructure. These responses

point to potential areas where MANRS support and reporting could be improved to better serve diverse operational contexts.

4.7 Summary

This chapter provides a first look into the MANRS networks' conformance to routing security practices. We investigated the geographical and address space distribution of MANRS members, and found the presence of very large networks as part of MANRS and a large proportion of MANRS members in Brazil. Our analysis of IRR and RPKI revealed that MANRS members were more likely to register and maintain routing objects than non-MANRS members. The analysis on route filtering showed that MANRS members were likely to have better routing practices than non-MANRS members. We inferred that, as of May 2022, the vast majority of, but not all, MANRS members were conformant with MANRS actions, and found that conformance for very large networks is difficult to achieve. We also found that MANRS networks collectively were more effective in filtering out RPKI Invalid announcements than non-MANRS networks.

These findings demonstrate that public data sources can be effectively used to independently assess the conformance of MANRS participants to their commitment to deploying origin validation practices, and that the initiative has been effective in reducing the propagation of BGP announcements with incorrect origins through MANRS ASes. While this study focuses on MANRS participants, these public data sources can be also be used to evaluate partial RPKI deployment strategies across the broader Internet. Recently, U.S. government agencies proposed regulations to accelerate partial RPKI deployment and to maximize its impact. We focus on these emerging policy efforts in the next chapter.

Chapter 4 in part, is a reprint of the material as it appears in *Proceedings of the Internet Measurement Conference 2022*. Ben Du, Cecilia Testart, Romain Fontugne, Gautam Akiwate, Alex C. Snoeren, and kc claffy. The dissertation author was the primary investigator and author of this paper.

Chapter 5

Future Partial Deployment Strategies of BGP Origin Validation

Although MANRS participants have made a positive impact on improving routing security, the initiative still has limitations. First, only a small fraction of ASes on the Internet participate in MANRS, and some participants, particularly some of the largest networks, are not fully conformant with its routing security requirements. Second, some MANRS participants rely solely on IRR-based BGP origin validation, which lacks cryptographic verification and is more vulnerable to inaccuracies and falsification compared to RPKI. As a result, MANRS alone is insufficient to ensure effective BGP origin validation, and RPKI deployment across the Internet remains limited. This gap has drawn the attention of the U.S. Federal Communications Commission (FCC), who has proposed regulations aimed at promoting RPKI adoption, especially among large and well-connected networks [51, 140, 183].

In this chapter, I show that it is possible to use public data sources to evaluate future deployment strategies for RPKI, providing policymakers and stakeholders with empirically grounded insights on improving routing security under the constraints of partial adoption. Because the decentralized nature of BGP makes large-scale, controlled ROV deployment experiments impractical, I use public measurement data in the following ways. First, I use Internet topology data inferred from public BGP measurements to simulate the propagation of RPKI-invalid announcements under current and hypothetical deployment scenarios. Second, I leverage

publicly available AS-ranking metrics to evaluate the routing influence of individual networks and identify strategic targets for RPKI deployment. The results of this framework can help inform policy proposals, such as those led by the U.S. Federal Communications Commission (FCC), and guide broader efforts to strengthen global routing security.

5.1 Overview

Despite MANRS's efforts to promote BGP origin validation mechanisms, RPKI adoption remains limited. As of May 2024, ROAs covered only 50% of publicly routed IPv4 address space [115]. ROV deployment is even more sparse as of 2024: Cloudflare has estimated that only 136 ASes on the Internet were actively filtering RPKI-invalid BGP routes [34]. As a result, some prominent BGP hijacks and misconfigurations could still affect the large portions of the Internet [35, 185]. To address the slow pace of RPKI deployment and the significant risks that remain, government agencies have increasingly turned their attention to promoting RPKI adoption. In particular, The White House and the U.S. National Institute of Standard and Technology (NIST) recommended routing security practices for organizations [140, 183] and the U.S. Federal Communications Commission (FCC) proposed reporting requirements for service providers [51]. Limited adoption of RPKI-ROV and growing policy interest raise two key questions: how can we estimate which networks still vulnerable under partial deployment, and how should we prioritize future ROA and ROV adoption to maximize protection?

Addressing these questions requires a systematic evaluation of hypothetical deployment scenarios across the Internet. Because of the decentralized nature of BGP, it is infeasible for researchers to control a large number of networks and conduct large-scale, controlled ROV deployment experiments. In contrast, simulation using public data offers a practical and repeatable alternative for evaluating deployment strategies. By modeling the Internet's topology as inferred from public BGP measurements and simulating the propagation of RPKI-invalid prefixes under different deployment scenarios, we can estimate the number of networks

vulnerable to hijacking under current ROV adoption and evaluate the potential benefits of future ROV deployment strategies.

While several existing studies have also used routing simulation to examine protection under partial ROV deployment, they have either relied on oversimplified assumptions such as random ROV deployment [55, 57] or focused on attack surfaces that remain even after full ROV adoption [61, 111]. In this work, we develop a lightweight routing simulation framework to comprehensively evaluate both the current effectiveness and future potential of ROV deployment scenarios. Our results can inform the networking community and other stakeholders to promote RPKI-ROV adoption to priority networks. We present two main analyses:

(1) We develop a simulation framework using public Internet topology data to analyze the propagation of RPKI-invalid BGP announcements under both the current ROV deployment landscape and potential future deployment strategies.

(2) By applying public AS ranking metrics, we identify target sets of ASes and simulate their potential impact: quantifying the reduction in both the number of ASes and the amount of address space exposed to RPKI-invalid announcements if they deploy ROV.

5.2 ROV Simulation Methodology

In this section, we introduce a metric to quantify the impact of partial ROV deployment and describe the assumptions underlying our BGP simulation. We then explain how we simulate the propagation of RPKI-invalid announcements under both current and future deployment scenarios and validate our simulation results.

5.2.1 Metric to quantify ROV deployment impact

To evaluate the influence of a given attacker AS on other networks in the Internet, we use metrics that align with the two granularities of our ranking methodology: AS-level and IP-level. At the AS level, we quantify the extent of propagation by counting the number of other ASes

that receive polluted routes. At the IP level, we capture the potential damage by counting the number of IPv4 addresses hosted by the ASes reached by the RPKI-invalid prefixes.

(1) The number of polluted ASes, a.k.a $|Polluted_m|$.

(2) The total number of polluted IPv4 addresses, defined as the sum of the effective sizes of all prefixes originated by the polluted ASes.

The effective size of a prefix p is computed as the number of IPv4 addresses in p, minus the total number of addresses covered by its direct sub-prefixes observed in the routing table:

$$\text{EffectiveSize}(p) = |\text{Addr}(p)| - \sum_{c \in \text{DirectChildren}(p)} |\text{Addr}(c)|$$

$$\text{PollutedAddresses} = \sum_{a \in Polluted_m} \sum_{p \in \text{Originated}(a)} \text{EffectiveSize}(p)$$

This adjustment ensures we do not double-count addresses that are covered by more-specific prefixes.

5.2.2 Simulation assumptions

We simulate the propagation of RPKI-invalid prefixes on the Internet. Our simulation makes the following assumptions:

(1) To evaluate the effectiveness of ROV deployment independent of ROA registration, we assume that all prefixes observed on the Internet are covered by a ROA. Without this assumption, the impact of ROV would be negatively affected by the incomplete coverage of ROAs in today's RPKI, where many prefixes remain in the NotFound state and therefore cannot be filtered by ROV. By assuming universal ROA coverage, we focus on the contribution of ROV enforcement itself regardless of the current limitations of RPKI adoption. This enables us to assess the maximum potential benefit of ROV deployment under ideal ROA coverage.

(2) We assume that an attacker seeks to maximize reach and therefore consider sub-prefix hijacks, which represent the worst-case scenario. Although exact-prefix hijacks also occur, their impact is more limited, and any ROV deployment that covers sub-prefix hijacks will by definition also mitigate exact-prefix hijacks (§1.2).

(3) We have surveyed large ROV-deploying ASes participating in MANRS and found that they originated only a few RPKI-invalid announcements (§4.6). These invalid announcements were not due to misconfigurations but were instead exceptional cases, such as customer-specific announcements or test prefixes, suggesting that these networks likely implement outbound filtering [102]. Based on this observation, we assume in our simulation that each ROV-deploying AS adopts outbound filtering and does not originate RPKI-invalid prefixes.

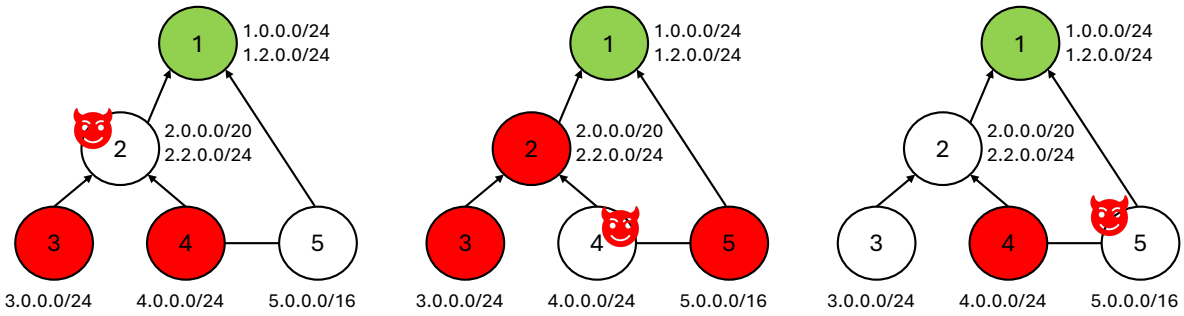
(4) We assume a threat model in which any AS may misconfigure its BGP announcements and in turn accidentally hijack arbitrary prefixes. RFC 6811 explicitly notes that RPKI should be understood as a system for preventing misconfigurations [122]. Thus, in our simulation, we assume that all ASes that are not deploying ROV originate RPKI-invalid prefixes.

5.2.3 Simulation construction

We define two actors when simulating the propagation of RPKI-invalid prefixes. We refer to the AS that originates the RPKI-invalid prefix as the *polluter* AS, denoted as m. We refer to ASes that receive the RPKI-invalid prefix as *polluted* ASes, denoted as $Polluted_m$. For a given set of n ASes that deploy ROV, we conduct a simulation of the following steps:

(1) We construct the graph using the observed Internet topology derived from the *AS relationship dataset*, where each vertex $v \in V$ is an AS and each edge between any two vertices $(u, v) \in E$ represents 2 interconnected ASes.

(2) We pick a set of n ASes as the set of ROV-deploying ASes *ROV* in this simulation iteration (see next section on simulation parameters).



(a) AS2 can pollute 2 ASes (AS3, AS4) and their 512 addresses.

(b) AS4 can pollute 3 ASes (AS2, AS3, AS5) and 70,144 addresses.

(c) AS5 can pollute only 1 AS (AS4) and its 256 addresses.

Figure 5.1. Example AS-level topology and the prefixes originated by each AS. The devil icon marks the polluting AS (the attacker). Red nodes represent ASes that are polluted (i.e., accept the RPKI-invalid prefix), green indicates an AS that deploys ROV, and white indicates ASes that are not polluted. Assuming AS1 enforces ROV, we simulate hijacks by iteratively selecting each remaining AS as a polluter originating RPKI-invalid prefixes. For each case, we count the number of polluted ASes and the size of their originated IP space.

(3) For each non-ROV-deploying AS $m \in V \setminus ROV$ in the graph, we assume it originates an RPKI-invalid prefix (a.k.a. m is a polluter).

(4) Following the valley-free routing principle [56], we simulate the route propagation starting from m and identify each AS i that receives the RPKI-invalid prefix. Given that m is conducting a sub-prefix hijack, any AS i that receives the hijacked prefix is guaranteed to select the route originated from m due to the longest-prefix-matching rule (§1.1. We add all such i to $Polluted_m$.

Figure 5.1 illustrates a simulation outcome in which AS1 is the ROV-deploying AS. In one iteration (Figure 5.1a), AS2 is the polluter, and AS3 and AS4 become polluted. AS5, however, does not receive the hijacked prefix due to the valley-free routing principle: AS4 receives the RPKI-invalid route from its provider (AS2) and exports it only to its customers, but AS5 is a peer. Similarly, Figure 5.1b and 5.1c show that AS4 can pollute AS2, AS3, and AS5; AS5 can only pollute AS4.

5.2.4 Approximating currently ROV-deploying ASes

To estimate the impact of current ROV-deploying ASes, we first need to identify which ASes actually deploy ROV. However, reliably determining ROV deployment remains an open challenge (§1.7.6). To approximate the set of ROV-deploying ASes, we adopt a methodology similar to Li et al.[101]. We first use the RoVista dataset[178] to identify ASes that are likely protected by ROV. RoVista considers an AS to have fully protected by ROV if either it enforces ROV itself or all of its upstream providers do. Since smaller networks such as stub ASes are less likely to have the resources to deploy ROV (§1.5.1), we treat ASes with a RoVista protection score of 100 that are not stub ASes as currently deploying ROV.

5.2.5 AS rankings for selecting future ROV-deploying ASes

To evaluate the potential impact of targeted ROV deployment, we need a way to identify networks whose adoption could effectively limit the propagation of RPKI-invalid prefixes. AS ranking metrics provide a systematic means to prioritize candidate ASes by estimating their influence in the global routing system. These rankings provide some indication of the impact a network has on routing by quantifying either its customer reach (*customer cone*) or its centrality in BGP paths (*AS hegemony*), which we introduce in detail below. In our simulations, we use both global and country-level rankings: global rankings highlight networks with the greatest reach across the entire Internet, while country-level rankings capture influential networks within specific country boundaries, which can be important for country-specific deployment strategies.

Customer Cone ranking. In 2013, Luckie *et al.* [109] proposed the concept of customer cone (CC) to count the number of ASes that a given AS can reach through only customer relationships. Similarly, we can obtain a IP-level customer cone by counting the number IP addresses that an AS can reach through its customers. In 2023, Huffaker et al. [75] adapted this global customer cone definition to create country-level customer cone rankings. We download the April 1, 2025 snapshot of the AS-leven and IP-level global customer cone rankings through the

CAIDA AS Rank [22] website and refer them as AS-level and IP-level CCG (Global Customer Cone). We also compute the AS-level country-level customer cone for each AS by grouping its customer-cone ASes by their registered countries, which we denote as AS-level CCC (Country-level Customer Cone).

AS Hegemony ranking. In 2017, Fontugne *et al.* [52] proposed AS Hegemony (AH), a metric that quantifies how central an AS is in terms of routing dependency. It calculates how often an AS appears on BGP paths used to reach other ASes (both globally and within a country), including via peering and provider relationships. We download the April 1 2025 snapshots of global and country-level IP-level AH rankings on the IIJ Internet Health Report archive [84]. IIJ provides the AS-level AH rankings (also called hegemony cone) through Internet Yellow Pages [53]. We query the IYP Console [96] using the Cypher query below to obtain the AS-level global AH rankings for April 1 2025.

```
MATCH (up:AS)<-[:DEPENDS_ON]-(down:AS)
RETURN up.asn, COUNT(DISTINCT down.asn) AS customer_cone
ORDER BY customer_cone DESC
```

We run the simulation for various values of n, where $n \in 10, 20, \ldots, 100, 200, \ldots, 1000$. For each n, we select the top n ASes according to four AS-level ranking metrics (CCG, CCC, AHG, and AHC) and evaluate the propagation of RPKI-invalid prefixes using the number of polluted ASes as the metric. Using AS-level rankings treats all ASes equally; however, in practice not all ASes are equal as some ASes serve more customers or host more services than others. Since our public data cannot directly capture these differences, we approximate them using the amount of address space each AS originates. To incorporate this approximation, we also apply two IP-level metrics (CCG and AHG) to select ROV-deploying ASes, and in this case evaluate the impact of RPKI-invalid propagation by the number of polluted IPv4 addresses.

5.2.6 Validating simulation with in-the-wild RPKI-Invalid prefixes

We validate our simulation against RPKI-invalid prefixes observed in the wild to assess how closely it reflects real-world propagation. This step is necessary because both our input datasets and methodology have limitations. For instance, the AS-relationship dataset may miss certain links due to incomplete BGP visibility, and the RoVista dataset may fail to capture some ASes that deploy ROV in practice. By comparing simulated propagation to observed propagation, we can assess the impact of these limitations on our results and develop a clearer understanding of how our simulation diverges from real-world behavior. Our validation process consists of the following steps:

(1) **Retrieve RPKI statuses:** We use the April 2, 2025 snapshot of the IHR-ROV dataset to obtain the RPKI validation status of each announced prefix.

(2) **Filter for *clean* invalids:** We obtain RPKI-invalid prefixes without multi-origin AS (MOAS) conflicts. This ensures that each prefix has a single origin AS, so that if no AS filters the announcement via ROV, the prefix would propagate broadly across the Internet. This avoids interference from routing policies such as AS path length preference.

(3) **Extract RIB data:** Using the April 2, 2025 snapshot of the RIB from the `route-views2` collector, we extract all observed AS paths to those *clean* invalid prefixes.

(4) **Simulate invalid propagation:** We run our simulation to model the propagation of each selected RPKI-invalid prefix, assuming no filtering except by ROV-deploying ASes.

(5) **Compare observed and simulated paths:** We compare the AS paths observed in the wild with the simulated propagation paths to evaluate how closely the simulation reflects real-world behavior.

5.3 Simulation Results

In this section, we validate our simulation framework and present an evaluation of both the current ROV deployment landscape and potential future deployment strategies. We first note that our framework is constrained by the limitations of the public datasets on which it relies. We then find that the estimated ROV deployment landscape as of April 2025 had only limited effectiveness in reducing the propagation of RPKI-invalid prefixes. Next, we compare different AS ranking metrics as strategies for prioritizing future ROV adoption to estimate which approaches can more effectively limit propagation. Finally, we examine the FCC's proposed significant service providers and show that, within the U.S., their deployment of ROV has only limited impact, as the overall reduction in RPKI-invalid propagation remains small.

5.3.1 Simulation validation

Following our validation procedure, we analyze the consistency between our simulation and real-world observations of RPKI-invalid prefix propagation:

- We first retrieved 582 RPKI-invalid prefixes from the IHR-ROV dataset.
- We filter out 52 prefixes that had conflicting origin ASes (i.e., multi-origin announcements), leaving 530 RPKI-invalid prefixes with a single origin AS (i.e., no multi-origin AS conflicts).
- We query the Route Views RIB (route-views2 collector, April 2, 2025) for these 530 prefixes and find that 452 of them appeared in the RIB. These were received from 22 collector peers, who learned the prefixes via 1,088 unique AS paths originating from 247 ASes.
- We compare these 1,088 AS paths with the simulated paths from each origin AS to the corresponding collector peers. Of these, 494 paths matched exactly between the real-world

data and our simulation. The remaining 594 paths could not be matched for the following reasons:

- **Uncertainty of ROV-deploying ASes:** We observe 363 unmatched paths traversing 25 ASes that we inaccurately infer as likely ROV-deploying. This limitation stems from RoVista's methodology, which infers ROV protection based only on a subset of RPKI-invalid prefixes [102]. As a result, those 25 ASes may not have been fully protected by ROV, and because they had customers, our approximation incorrectly classified them as deploying ROV. For instance, RoVista inferred that AS701 was fully protected by ROV, yet we observe it to have received and propagated RPKI-invalids. Our evaluation of real-world RPKI-invalid propagation therefore inherits the limitations of the RoVista dataset, underscoring the broader challenge of relying on public datasets for evaluating the effectiveness of current ROV-deployment landscape (§1.7.6).
- **Missing topology data:** 231 unmatched paths included links not present in our constructed AS-level topology. These discrepancies likely stem from the AS relationship dataset, which only retains links stably observed for at least five days, whereas the AS paths we extract from the RouteViews RIB come from a single snapshot. As a result, short-lived or less frequently observed connections may be missing. For example, upon manual inspection, we identify AS paths traversing IXPs that had no recorded neighbors in the relationship dataset. Our evaluation of RPKI-invalid propagation therefore inherits the incompleteness of the AS relationship dataset, highlighting the broader challenge of relying on public datasets with limited visibility into transient or unobserved links (§1.7.2).

Having evaluated the accuracy and limitations of our public-data-based simulation framework, we next apply it to evaluate the effectiveness of the current ROV deployment landscape.

5.3.2 Current ROV deployment does not sufficiently reduce RPKI-invalids propagation

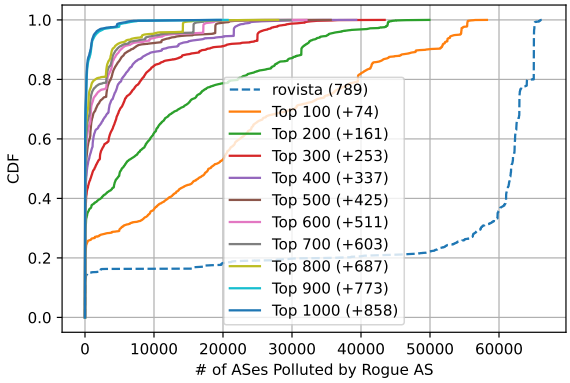
To estimate the set of ASes that currently deploys ROV on the Internet, we first look at the ASes that are protected by ROV according to RoVista. There were 30,340 ASes in the RoVista snapshot we retrieved, of which 3,623 ASes had a ROV score of 100%, meaning that either they have deployed ROV themselves or were protected by their upstream ASes that deployed ROV. We then remove ASes with no customers or peers and obtain 789 ASes that likely actively deployed ROV at the time of analysis. Of those 789 ASes, 230 were registered in the U.S. and included U.S.-based large transit providers such as Lumen, Cogent, and NTT America. We find that 16 of the 19 tier-1 providers (ASes with no providers [23]) in the RoVista dataset and thus have likely deployed ROV. The three remaining non-ROV-deploying ASes were AS286 (GTT), AS1239 (Sprint/T-Mobile), and AS6762 (Telecom Italia). We find that AS286 belonged to KPN, which was acquired by GTT. Interestingly, GTT's main AS number AS3257 did actively deploy ROV according to our inference.

We then simulate the propagation of RPKI-invalid announcements when those ASes deployed ROV and found that the current set of 789 ROV-deploying ASes still allows the widespread propagation of RPKI-invalid announcements. The blue dashed line in Figure 5.2a shows that the median, and 90th-percentile polluter AS can pollute 62,255 and 65,055 ASes, showing that most ASes can pollute many other ASes by originating an RPKI-invalid prefix. Of the top 20 polluter ASes that polluted the most ASes, we find that 13 belonged to a large CDN that obtained transit from 14 of the top 20 transit providers (according to CAIDA's AS Rank). We then look at the remaining top polluter ASes: AS17060 was a stub AS with 10 providers but 5 did not deploy ROV, and could reach all top 10 transit providers within 2 hops. This case highlights that even stub ASes, especially ones with many upstream providers, can act as highly effective polluters if their upstream providers do not deploy ROV, as their RPKI-invalid announcements can propagate through multiple major transit networks and spread widely across the Internet.

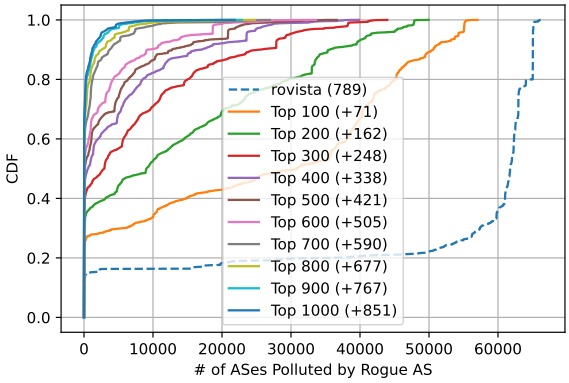
These results show that current ROV deployment is insufficient to limit the propagation of RPKI-invalid prefixes across the Internet. Although many large U.S.-based transit providers and most tier-1 ASes have likely deployed ROV, the vast majority of polluter ASes can still reach more than 60k ASes on the Internet. This result emphasizes that, even with 789 ASes likely deploying ROV, we are still in the early stages of adoption. Strategically guided future ROV deployment is needed to strengthen global routing security and reduce the risks posed by origin hijacks. However, it is important to note that these results are affected by the limitations of the public data used in our analysis. Because we may have overestimated the number of ASes deploying ROV and our constructed topology omits AS links not visible to our limited set of BGP collectors, our simulation likely underestimates the true extent of RPKI-invalid propagation. As a result, the real-world ROV deployment landscape may be even less effective than it appears here.

In addition, using private data to analyze the effectiveness of current ROV deployment landscape can paint a very different picture. For example, Kentik's analysis of NetFlow traffic shows that less than 1% of observed traffic was directed toward RPKI-invalid prefixes, suggesting that current deployment is already effective in practice [114]. However, their approach relies on private traffic data from a limited set of customer networks that allow them to publish aggregated statistics, which may not be representative of the Internet as a whole. The contrast between our conclusion from public data and Kentik's conclusion from private data highlight both the strengths and limitations of public-data–driven methods: while public data enables independent, repeatable measurement of the Internet, it cannot capture alternative perspectives, such as traffic flows, that require private datasets.

While such traffic-level insights are valuable, our work emphasizes what can be learned from public data. In the next section, we apply our simulation framework to compare different public-data–driven metrics as future ROV deployment strategies and their effectiveness in limiting the spread of RPKI-invalid prefixes.



(a) Estimated propagation of RPKI-invalid prefixes under a deployment scenario where the top 100 ASes ranked by Global Customer Cone deploy ROV. In this case, the median originating AS can reach 17,946 ASes.



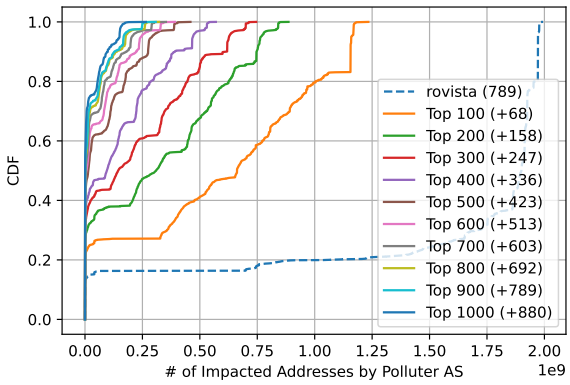
(b) Estimated propagation of RPKI-invalid prefixes under a deployment scenario where the top 100 ASes ranked by Global AS Hegemony deploy ROV. In this case, the median originating AS can reach 30,860 ASes.

Figure 5.2. Comparison of deployment strategies using AS-level metrics: selecting the top N ROV-deploying ASes using the AS-level Global Customer Cone metric limits the propagation of RPKI-invalid prefixes more effectively than selecting them using the AS-level Global AS Hegemony metric.

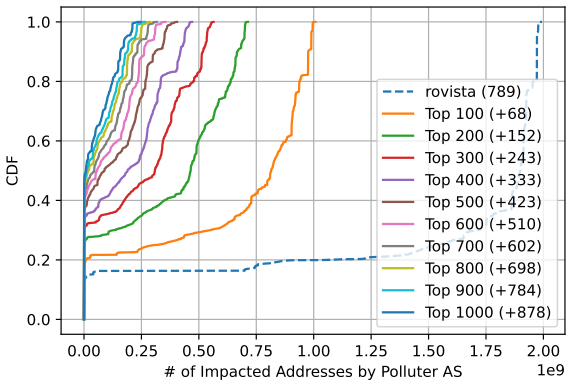
5.3.3 Difference in effectiveness in using different AS Ranking metrics

In this section, we evaluate the effectiveness of the Customer Cone and AS Hegemony metrics as strategies for prioritizing future ROV deployment. Figure 5.2 compares the distribution of the number of ASes each polluter can reach under different ROV deployment strategies, specifically when prioritizing the top N ASes ranked by AS-level CCG and AS-level AHG. Figure 5.2a (orange line) shows that if the top 100 ASes ranked by AS-level CCG deploy ROV, the median polluter can reach 17,946 ASes. In contrast, Figure 5.2b shows that if the top 100 ASes ranked by AS-level AHG deploy ROV, the median polluter can reach 30,860 ASes. Comparing Figures 5.2a and 5.2b, we observe that polluters consistently reach more ASes under the AHG-based strategy than under the CCG-based strategy. This suggests that prioritizing ASes for ROV deployment using the Customer Cone metric is more effective at reducing polluters' AS reach than using the AS Hegemony metric.

Figure 5.3 evaluates the number of *polluted addresses* each polluter can reach under



(a) Estimated propagation of RPKI-invalid prefixes when the top 100 ASes ranked by Global Customer Cone metric deploy ROV. The median polluter can reach 661 million IPv4 addresses, while the 75th percentile polluter can reach 946 million.



(b) Estimated propagation of RPKI-invalid prefixes when the top 100 ASes ranked by the Global AS Hegemony metric deploy ROV. The median polluter can reach 815 million IPv4 addresses, while the 75th percentile polluter can reach 920 million.

Figure 5.3. Comparison of deployment strategies using IP-level metrics: selecting the top *N* ROV-deploying ASes with the Global Customer Cone metric is more effective at limiting RPKI-invalid propagation from the median polluter, while the Global AS Hegemony metric is more effective for the 75th and higher percentiles.

IP-level ranking strategies: IP-level CCG and IP-level AHG. Figure 5.3a (orange line) shows that if the top 100 ASes ranked by IP-level CCG deploy ROV, the median polluter can still reach 661 million IPv4 addresses. In contrast, Figure 5.3b shows that if the top 100 ASes ranked by IP-level AHG deploy ROV, the median polluter can reach 815 million. This indicates that IP-level CCG is more effective at reducing the median reach of RPKI-invalid prefixes. However, at the 75th percentile, polluters can reach 946 million addresses under CCG and 920 million under AHG, suggesting that IP-level AHG is more effective at limiting the reach of high-impact polluters.

To better understand these differences, we manually inspect the top polluter ASes under both strategies. Under the CCG metric, 14 of the top 20 polluters were major content delivery networks (CDNs), each with hundreds of peers, allowing them to propagate RPKI-invalid prefixes to a large number of ASes. In contrast, none of the top 20 polluters under the AHG metric were CDNs. This difference is because the CCG metric only accounts for customer relationships, which prevents large CDNs from being ranked highly. By contrast, AHG ranks heavily peered CDNs near the top (some within the top 20), and since these CDNs deploy ROV, they neither

originate nor propagate RPKI-invalid prefixes.

We then conduct a broader comparison of polluter ASes under the IP-level CCG and IP-level AHG strategies when the top 100 ASes deploys ROV. Among the 77,723 potential polluter ASes in the topology, 39.9% (31,015 ASes) could polluted more IP addresses under CCG, while 44.0% (34,210 ASes) could pollute more under AHG. However, the most damaging polluter AS under CCG could pollute 1.78 billion addresses, while the top polluter AS under AHG could reach 989 million. This suggests that future ROV deployment strategy based on CCG is more effective at broadly reducing the reach of polluters, but less effective at mitigating the worst polluters.

To further analyze differences between the ranking strategies, we compare the top 100 ASes ranked by AS-level AHG and IP-level AHG and found notable divergences. The IP-level AHG rankings included more cloud providers such as Akamai and Amazon, and excluded smaller ISPs with large customer degrees but limited prefix origination. In one extreme case, AS749 (U.S. Department of Defense) ranked fourth in IP-level AHG despite being a stub AS with no customers or peers. This is because AS749 originates a very large amount of legacy DoD IPv4 address space (more than 200 million IPv4 addresses). Since AS hegemony measures how often an AS appears in BGP paths, weighted by the size of the address space reached along each path, networks that originate a large volume of address space are ranked highly even if they have few customers.

These differences also apply to evaluation metrics. Counting polluted ASes versus polluted addresses can produce divergent outcomes, since not all ASes originate the same amount of address space. In the IP-level metrics, the upper tails of the CDFs (Figure 5.3b) are farther away from the baseline than in AS-level metrics (Figure 5.2b). This reflects the fact that a small number of ASes originate disproportionately large portions of the global IPv4 address space, making it easier to prevent the spread of RPKI-invalid traffic to those address blocks.

These findings highlight that the choice of both evaluation and ranking metrics should be guided by the specific goals and priorities of stakeholders. When selecting an evaluation metric,

Table 5.1. Mapping of significant service providers listed by the FCC to their primary AS numbers.

FCC-listed significant service providers	Primary AS numbers
AT&T Inc.	7018
Altice USA (Optimum and Cablevision)	19108, 53488, 6128
Charter Communications	7843, 10796, 11427, 20001, 11351, 11426
Comcast Corporation	7015, 7922
Cox Communications, Inc.	22773
Lumen Technologies, Inc.	3356
T-Mobile USA, Inc.	21928, 22140, 13644, 11461, 25877
Telephone & Data Systems (incl. US Cellular)	4181, 21864
Verizon Communications, Inc.	701, 702, 703

one must consider whether the objective is to protect a larger number of ASes regardless of their size or to prioritize protecting ASes that originate more address space, which may serve a greater number of users. Similarly, when choosing a ranking metric for ROV deployment, there is a tradeoff between reducing the overall pollution capability of many ASes versus targeting high-impact polluters that can reach large portions of the Internet. Metrics like AS-level CCG are more effective at broad suppression of propagation, while metrics like AHG may better reduce the influence of individual heavy polluters.

5.3.4 U.S. FCC-Listed service providers offer limited marginal benefit

So far, our analysis has focused on global prioritization strategies for ROV deployment. We now narrow down to a country-level perspective and examine the case of the United States, where policy discussions have targeted specific networks as particularly important for routing security.

We analyze the marginal benefit provided by the nine significant service providers listed in the FCC NPRM Paragraph 38 [51]. We identify the primary AS numbers that originated address space corresponding to the nine service providers. Due to mergers and acquisitions, some providers operate under multiple organization names which may have multiple AS numbers each. For instance, Altice USA owns both Cablevision and Optimum, which are its two main ISPs.

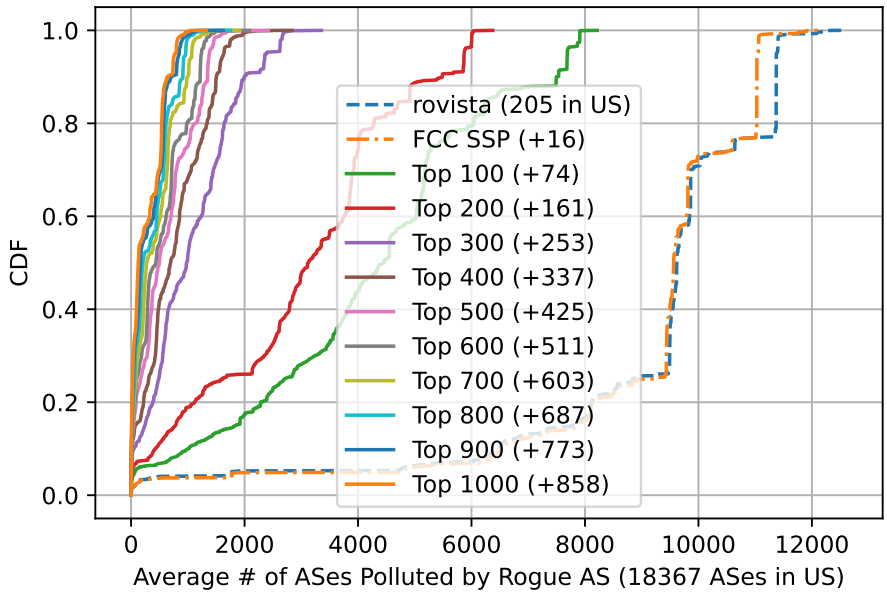


Figure 5.4. Impact of adding FCC-listed significant service providers and other top ranked U.S. ASes as ROV-deploying ASes. The median polluter could reach 9,617 U.S.-based ASes with only RoVista ASes deploying ROV, compared to 9,561 when the additional 16 FCC SSPs were included.

We manually look up the subsidiaries of each provider through the U.S. Securities and Exchange Commission (SEC) and identified the underlying ISPs. We then manually look up the ISP organization names to obtain their AS numbers, and obtained all the sibling AS numbers from the *as2org dataset*. Through this process, we identify 24 ASes operated by the nine significant service providers (Table 5.1), 16 of which likely have not yet deployed ROV according our approximation using the *RoVista dataset* (§5.2.4).

Figure 5.4 shows the capability of polluter ASes to reach U.S-registered ASes. The April 1, 2025 snapshot of the *as2org dataset* includes 18,367 ASes registered in the U.S. Comparing the blue dashed line and orange dash-dotted line, we see that the distribution of pollution capability of ASes barely reduced with the addition of the 16 ASes belonging to the FCC-listed significant service providers (SSPs). In our simulation, the median polluter AS could pollute 9,617 U.S-based ASes when only RoVista ASes deployed ROV and the number reduced to 9,561 ASes when the additional 16 FCC SSP ASes deployed ROV. We also see a noticeable shift on the upper

right corner of the CDFs in Figure 5.4. We find 13,286 polluter ASes experienced a reduction of 342 polluted ASes. We investigate those 342 ASes and found that they were single-homed stub customers of the 16 ASes.

This result highlights two observations. First, the FCC-listed SSP ASes have limited topological influence. Many of those ASes serve primarily stub customers. Such a business model constrains their ability to broadly reduce the propagation of RPKI-invalid prefixes, because filtering at these ASes only protects their direct stub customers and no further downstream networks. On the flip side, these SSP ASes are very effective in protecting their stub customers from receiving RPKI-invalid prefixes, as stubs typically receive all routes solely from their immediate upstream (in this case, the SSP AS that deploys ROV). For example, we examine the direct customers of Optimum (AS19108) and found that 22 of its 65 customer ASes were critical infrastructure organizations, including school districts, local government agencies, hospitals, and aviation companies. From a national security or public service perspective, government agencies may prioritize protecting these types of networks over optimizing global ROV coverage. However, if the goal is to broadly protect U.S-registered ASes against RPKI-invalid prefix propagation, then ROV deployment by a small set of ASes has limited effect due to their low topological reach.

5.4 Limitations

Our simulation technique inherits the limitations by the incompleteness of topology data and uncertainty in ROV deployment status. While our simulations and analysis provide insights into the potential benefits of strategically guided ROV deployment, several practical limitations in real-world ROV behavior may affect the accuracy of our modeling:

Complex ROV deployment policy. Some transit providers deliberately choose not to enforce ROV on routes received from their customers. This decision is often driven by concerns that filtering RPKI-invalid announcements could disrupt business relationships or reduce revenue,

as some customers may misconfigure their routes. To provide operational flexibility, providers may allow invalid announcements from customers to propagate, even if they technically deploy ROV. As a result, these networks may unintentionally facilitate the spread of accidental hijacks.

Partial ROV deployment within an AS. ROV is implemented at the router level rather than at the AS level. There are various reasons why a network may only enable ROV on some routers but not others. Older ISPs and transit providers may have old equipment that does not support ROV. Sometimes routers may malfunction or be misconfigured so the network may inadvertently fail to enforce ROV. Some larger networks may be in the process of deploying ROV and thus only enable ROV on some of those routers. Consequently, within a single AS, some routers may enforce ROV while others may not. Our simulation operates on an AS-level, and thus does not consider such intra-AS inconsistency in ROV deployment.

Default Route Usage. Networks may use default routes when forwarding packets if the destination prefix is not in the routing table. In this case, if a network's upstream ASes deploy ROV and filter out the RPKI-invalid route, the downstream network will not receive the route, and will forward traffic destined to the RPKI-invalid prefix via the default route.

5.5 Summary

This chapter presents a lightweight simulation framework to evaluate the effectiveness of the current partial ROV deployment landscape by quantifying how many ASes can be reached by potential RPKI-invalid prefixes using Internet topology data inferred from public BGP data. We find that, under the current state where 789 ASes likely deploy ROV, the median polluter AS can still reach 62,255 ASes, indicating that the existing deployment is insufficient to prevent widespread pollution. We further evaluate the impact of different prioritization strategies for incremental ROV deployment. We compare customer cone and AS hegemony rankings and found that customer cone is more effective at broadly reducing the reach of polluters, while AS hegemony is more effective at targeting heavy polluters. Our marginal benefit analysis shows

that the largest gains occur in the early stages of deployment—an important insight given the current early status of RPKI adoption.

These findings demonstrate that public data sources can be effectively leveraged to evaluate both the current effectiveness and future potential of ROV deployment, providing stakeholders with actionable insight into the impact of different adoption strategies. The simulation framework developed in this study can be used to guide empirically informed policy decisions and technical interventions, helping to maximize the benefits of RPKI in future deployment and strengthen the resilience of the global routing system.

Chapter 5 in part, is a reprint of the material as it was submitted to *International Conference on emerging Networking EXperiments and Technologies (CoNEXT) 2025*. Ben Du, Shivani Hariprasad, Romain Fontugne, Cecilia Testart, Alex C. Snoeren, and kc claffy. The dissertation author was the primary investigator and author of this paper.

Chapter 6

Conclusion

This dissertation shows that it is possible to evaluate the effectiveness of BGP origin validation mechanisms using public measurement data. As independent researchers, we do not have access to the internal operational practices of networks, but it is necessary to evaluate the state of deployment and hold networks accountable for adopting origin validation mechanisms that improve routing security. By using public data, I developed repeatable methodologies that can be adopted not only by researchers, but also by network operators, industry members, and policymakers to continuously monitor the effectiveness of BGP origin validation mechanisms and make empirically supported recommendations to improve routing security. However, one open question remains: how accurate are the results we obtained through public data? To answer this question, I review the limitations of the public datasets I have used and then point out how future work can improve upon these limitations.

A core limitation is that some public datasets I have used are incomplete. However, such incompleteness does not always negatively impact the evaluation of the effectiveness of BGP origin validation mechanisms. In several cases, datasets with limited coverage were sufficient to reveal weaknesses in the IRR and RPKI. For example, even though RPKI does not cover all of IRR, I was able to compare a subset of IRR records and identify stale and attacker-falsified IRR entries, showing that improper hygiene in the IRR can undermine the effectiveness of IRR-based origin validation. Similarly, while WHOIS records capture only part of the leasing ecosystem and

blocklists record only a subset of abusive networks, they were sufficient to reveal that attackers have exploited the IP leasing market to obtain legitimate RPKI authorizations. These cases show that incomplete data can provide concrete evidence of vulnerabilities in the BGP origin validation mechanisms.

In other cases, however, incompleteness does affect confidence in the results. For example, based on public BGP data, I observed few false announcements from MANRS networks, suggesting that most were conformant to their commitments. But public BGP data may not capture all prefixes originated by a network, and unobserved non-conformant prefixes could lead to an overestimation of MANRS conformance. Because current BGP collectors peer with only about 1% of active ASes and the routes they observe are largely redundant [6], the resulting restricted and overlapping views of collectors make it difficult to know whether our observations are representative of the Internet as a whole. Future work could address this uncertainty by incorporating private BGP data such as collaboration with large content providers or ISPs, since prior research with access to such data has vaguely mentioned the large number of announcements not visible in public datasets [31]. Access to private BGP data would allow researchers to better estimate the number of unobserved prefixes and provide a more comprehensive measurement of the conformance of MANRS networks.

Another example of incomplete data limiting our confidence is that public AS relationship datasets are incomplete and may miss certain links, which could lead to underestimating the reach of RPKI-invalid prefixes under a partial deployment model. Since it is infeasible to obtain a complete record of all interconnections to construct a topology that fully reflects the real Internet, future work could assess how missing links affect estimates of invalid-route propagation. One approach is to supplement existing datasets with additional public evidence, such as peering relationships recorded at Internet Exchanges (IXPs). Another is to apply machine-learning–based AS link prediction [163] to infer likely missing links and then re-run simulations of RPKI-invalid propagation on these enriched topologies. Together, these strategies would provide a clearer picture of how topology incompleteness impacts the evaluation of future partial RPKI deployment

scenarios.

Although incomplete public data does not compromise the accuracy when identifying specific vulnerabilities in BGP origin validation mechanisms, it does introduce uncertainty when evaluating their deployment on the global Internet. Nevertheless, the analyses in this dissertation provide actionable insights for improving routing security. Several of these recommendations have since gained attention within the operator community. For IRR, I recommend that database operators remove records conflicting with RPKI and that networks registering in the IRR delete outdated entries and ensure consistency across databases. In recent years, RADB has begun removing conflicting records [147], and IP holders have directed their customers to delete outdated IRR entries [70].

In the case of IP leasing, I recommend that brokers monitor blocklist activity and revoke authorizations for abusive customers, and that address holders track the use of their resources and verify the identities of their leasing customers. Our research has gained the attention of at least one large broker, which has committed to making the leasing market more resilient to abuse [69].

In evaluating MANRS networks' conformance to BGP origin validation practices, I recommend that both the initiative and its participating networks conduct regular conformance checks against their commitments, promptly correct misconfigurations, and further promote RPKI deployment to accelerate the transition away from IRR-based validation. MANRS has also broadened its engagement by launching fellowship programs that support researchers across technical and policy domains [29, 30]. Finally, I recommend that policymakers and industry groups empirically evaluate future deployment strategies and prioritize networks with high routing influence, thereby maximizing the security gains of incremental RPKI adoption.

When evaluating the effectiveness of BGP origin validation mechanisms with public data, it is helpful to think about two ends of a spectrum. On one end, public data may be too limited to provide meaningful answers, requiring access to private operational data to obtain accurate results. On the other end, public data is sufficient to confidently support our conclusions,

making it an essential tool for independent researchers and stakeholders to assess BGP origin validation deployment. The results in this dissertation fall in between these two ends. Public data is incomplete and sometimes constrains the confidence we can place in results, but it is also often sufficient to reveal concrete vulnerabilities, provide actionable guidance, and even influence operator practices. In this sense, public data is not a substitute for private operational information, but it remains a useful and important resource for evaluating deployment, monitoring progress, and holding networks accountable in improving routing security.

Bibliography

[1] AFRINIC. 2020. Consolidated Policy Manual. https://afrinic.net/policy/manual

[2] AFRINIC. 2021. AFRINIC's Internet Routing Registry. https://afrinic.net/internet-routing-registry

[3] Gautam Akiwate, Raffaele Sommese, Mattijs Jonker, Zakir Durumeric, KC Claffy, Geoffrey M. Voelker, and Stefan Savage. 2022. Retroactive Identification of Targeted DNS Infrastructure Hijacking. In *Proceedings of the 22nd ACM Internet Measurement Conference* (Nice, France) *(IMC '22)*. Association for Computing Machinery, New York, NY, USA, 14–32. https://doi.org/10.1145/3517745.3561425

[4] C. Alaettinoglu, C. Villamizar, E. Gerich, D. Kessens, D. Meyer, T. Bates, D. Karrenberg, and M. Terpstra. 1999. *Routing Policy Specification Language (RPSL)*. RFC 2622. RFC Editor.

[5] Thomas Alfroy, Thomas Holterbach, Thomas Krenc, KC Claffy, and Cristel Pelsser. 2023. Internet Science Moonshot: Expanding BGP Data Horizons. In *Proceedings of the 22nd ACM Workshop on Hot Topics in Networks* (Cambridge, MA, USA) *(HotNets '23)*. Association for Computing Machinery, New York, NY, USA, 102–108. https://doi.org/10.1145/3626111.3628202

[6] Thomas Alfroy, Thomas Holterbach, Thomas Krenc, K. C. Claffy, and Cristel Pelsser. 2024. The Next Generation of BGP Data Collection Platforms. In *Proceedings of the ACM SIGCOMM 2024 Conference* (Sydney, NSW, Australia) *(ACM SIGCOMM '24)*. Association for Computing Machinery, New York, NY, USA, 794–812. https://doi.org/10.1145/3651890.3672251

[7] Thomas Alfroy, Thomas Holterbach, and Cristel Pelsser. 2022. MVP: Measuring Internet Routing From the Most Valuable Points. In *Proceedings of the 22nd ACM Internet Measurement Conference* (Nice, France) *(IMC '22)*. Association for Computing Machinery, New York, NY, USA, 770–771. https://doi.org/10.1145/3517745.3563031

[8] AMS-IX. 2023. AMS-IX Route Servers. https://www.ams-ix.net/ams/documentation/ams-ix-route-servers

[9] APNIC. 2024. Registered IPv4 Brokers. https://www.apnic.net/manage-ip/manage-resources/transfer-resources/transfer-of-unused-ip-and-as-numbers/transfer-facilitators/

[10] APNIC. 2024. Understanding Address Management Hierarchy. https://www.apnic.net/manage-ip/manage-resources/address-management-objectives-2/address-management-objectives/

[11] Arelion. 2021. Routing Security BGP Prefix Filter Updates. https://www.arelion.com/our-network/bgp-routing/routing-security

[12] ARIN. 2024. Find a Qualified Facilitator. https://www.arin.net/resources/registry/transfers/facilitators/qualifiedfacilitators/

[13] ARIN. 2024. Introduction to ARIN's Database. https://www.arin.net/resources/guide/account/database/#delegation-records

[14] Hitesh Ballani, Paul Francis, and Xinyang Zhang. 2007. A Study of Prefix Hijacking and Interception in the Internet. In *Proceedings of the 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications* (Kyoto, Japan) *(SIGCOMM '07)*. Association for Computing Machinery, New York, NY, USA, 265–276.

[15] Barclay Ballard. 2020. Amazon Has Hoarded Billions of Dollars' Worth of IPv4 - but Why? https://www.techradar.com/news/amazon-has-hoarded-billions-of-dollars-worth-of-ipv4-but-why

[16] Jeff Barr. 2023. New – AWS Public IPv4 Address Charge + Public IP Insights. https://aws.amazon.com/blogs/aws/new-aws-public-ipv4-address-charge-public-ip-insights/

[17] T. Bates, E. Gerich, L. Joncheray, J-M. Jouanigot, D. Karrenberg, M. Terpstra, and J. Yu. 1995. *Representation of IP Routing Policies in a Routing Registry (ripe-81++)*. RFC 1786. RFC Editor.

[18] Robert Beverly. 2023. IP Neo-colonialism: Geo-auditing RIR Address Registrations. *arXiv preprint arXiv:2308.12436* (2023).

[19] Jay Borkenhagen. 2019. AT&T/AS7018 Now Drops Invalid Prefixes From Peers. https://mailman.nanog.org/pipermail/nanog/2019-February/099501.html

[20] Tom Brightbill. 2020. Bringing Your Own IPs to Cloudflare (BYOIP). https://blog.cloudflare.com/bringing-your-own-ips-to-cloudflare-byoip

[21] R. Bush and R. Austein. 2017. *The Resource Public Key Infrastructure (RPKI) to Router Protocol, Version 1*. RFC 8210. RFC Editor.

[22] CAIDA. 2021. AS Rank. https://asrank.caida.org/

[23] CAIDA. 2021. AS Relationships. https://www.caida.org/catalog/datasets/as-relationships/

[24] CAIDA. 2021. Inferred as to Organization Mapping Dataset. https://www.caida.org/catalog/datasets/as-organizations/

[25] CAIDA. 2021. Routeviews Prefix to as Mappings Dataset (pfx2as) for IPv4 and IPv6. https://www.caida.org/catalog/datasets/routeviews-prefix2as/

[26] Massimo Candela. 2022. A One-Year Review of RPKI Operations, RIPE 84. https://ripe84.ripe.net/archives/video/741/

[27] Esteban Carisimo, Carlos Selmo, J. Ignacio Alvarez-Hamelin, and Amogh Dhamdhere. 2019. Studying the Evolution of Content Providers in IPv4 and IPv6 Internet Cores. *Computer Communications* 145 (2019), 54–65. https://doi.org/10.1016/j.comcom.2019.05.022

[28] Telia Carrier. 2021. Routing Security. https://www.teliacarrier.com/our-network/bgp-routing/routing-security.html

[29] Harish Chowdhary. 2023. MANRS Compliance Increasing, More Work Needed in India. https://manrs.org/2023/03/manrs-compliance-increasing-more-work-needed-in-india/

[30] Harish Chowdhary and Naveen Kumar Chaudhary. 2023. MANRS Statistical analysis and adoption in india as a collaborative security tool. *Journal of Cyber Security Technology* 7, 4 (2023), 181–198. https://doi.org/10.1080/23742917.2023.2175527

[31] Taejoong Chung, Emile Aben, Tim Bruijnzeels, Balakrishnan Chandrasekaran, David Choffnes, Dave Levin, Bruce M. Maggs, Alan Mislove, Roland van Rijswijk-Deij, John Rula, and Nick Sullivan. 2019. RPKI Is Coming of Age: a Longitudinal Study of RPKI Deployment and Invalid Route Origins. In *Proceedings of the Internet Measurement Conference* (Amsterdam, Netherlands) *(IMC '19)*. Association for Computing Machinery, New York, NY, USA, 406–419. https://doi.org/10.1145/3355369.3355596

[32] Luca Cittadini, Wolfgang Mühlbauer, Steve Uhlig, Randy Bush, Pierre Francois, and Olaf Maennel. 2010. Evolution of Internet Address Space Deaggregation: Myths and Reality. *IEEE Journal on Selected Areas in Communications* 28, 8 (2010), 1238–1249.

[33] David Clark, Cecilia Testart, Matthew Luckie, and kc claffy. 2024. A Path Forward: Improving Internet Routing Security by Enabling Zones of Trust. *Journal of Cybersecurity* 10, 1 (12 2024), tyae023. https://doi.org/10.1093/cybsec/tyae023 arXiv:https://academic.oup.com/cybersecurity/article-pdf/10/1/tyae023/61182465/tyae023.pdf

[34] Cloudflare. 2022. Is BGP Safe Yet? https://isbgpsafeyet.com/

[35] Cloudflare. 2025. How Verizon and a BGP Optimizer Knocked Large Parts of the Internet Offline Today. https://blog.cloudflare.com/how-verizon-and-a-bgp-optimizer-knocked-large-parts-of-the-internet-offline-today/

[36] Federal Communications Commission. 2022. FCC Launches Inquiry Into Internet Routing Vulnerabilities. https://www.fcc.gov/document/fcc-launches-inquiry-internet-routing-vulnerabilities

[37] Danny Cooper, Ethan Heilman, Kyle Brogle, Leonid Reyzin, and Sharon Goldberg. 2013. On the Risk of Misbehaving RPKI Authorities. In *Proceedings of the Twelfth ACM Workshop on Hot Topics in Networks* (College Park, Maryland) *(HotNets-XII)*. Association for Computing Machinery, New York, NY, USA, Article 16, 7 pages. https://doi.org/10.1145/2535771.2535787

[38] Omar Darwich, Hugo Rimlinger, Milo Dreyfus, Matthieu Gouel, and Kevin Vermeulen. 2023. Replication: Towards a Publicly Available Internet Scale IP Geolocation Dataset. In *Proceedings of the 2023 ACM on Internet Measurement Conference* (Montreal QC, Canada) *(IMC '23)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3618257.3624801

[39] DE-CIX. 2023. Frankfurt Route Server Guide. https://www.de-cix.net/en/locations/frankfurt/route-server-guide

[40] Bernhard Degen, Ben Du, Ricky K. P. Mok, Raffaele Sommese, Mattijs Jonker, Roland van Rijswijk-Deij, and Kc Claffy. 2025. From Scarcity to Opportunity: Examining Abuse of the IPv4 Leasing Market. In *2025 9th Network Traffic Measurement and Analysis Conference (TMA)*. 1–11. https://doi.org/10.23919/TMA66427.2025.11097001

[41] Amogh Dhamdhere and Constantine Dovrolis. 2011. Twelve Years in the Evolution of the Internet Ecosystem. *IEEE/ACM Transactions on Networking* 19, 5 (Oct 2011), 1420–1433.

[42] Ben Du, Gautam Akiwate, Thomas Krenc, Cecilia Testart, Alexander Marder, Bradley Huffaker, Alex C Snoeren, and KC Claffy. 2022. IRR Hygiene in the RPKI Era. In *International Conference on Passive and Active Network Measurement*. Springer, 321–337.

[43] Ben Du, Massimo Candela, Bradley Huffaker, Alex C. Snoeren, and kc claffy. 2020. RIPE IPmap Active Geolocation: Mechanism and Performance Evaluation. *SIGCOMM Comput. Commun. Rev.* 50, 2 (may 2020), 3–10. https://doi.org/10.1145/3402413.3402415

[44] Ben Du, Romain Fontugne, Cecilia Testart, Alex C. Snoeren, and kc claffy. 2024. Sublet

Your Subnet: Inferring IP Leasing in the Wild. In *Proceedings of the 2024 ACM on Internet Measurement Conference* (Madrid, Spain) *(IMC '24)*. Association for Computing Machinery, New York, NY, USA, 328–336. https://doi.org/10.1145/3646547.3689010

[45] Ben Du, Katherine Izhikevich, Sumanth Rao, Guatam Akiwate, Cecilia Testart, Alex C. Snoeren, and kc claffy. 2023. IRRegularities in the Internet Routing Registry. In *Proceedings of the 2023 ACM on Internet Measurement Conference* (Montreal QC, Canada) *(IMC '23)*. Association for Computing Machinery, New York, NY, USA, 104–110. https://doi.org/10.1145/3618257.3624843

[46] Ben Du, Alex C Snoeren, et al. 2021. Poster: IRR Hygiene in the RPKI Era. In *ACM SIGCOMM Internet Measurement Conference 2021*.

[47] Ben Du, Cecilia Testart, Romain Fontugne, Gautam Akiwate, Alex C. Snoeren, and kc claffy. 2022. Mind Your MANRS: Measuring the MANRS Ecosystem. In *Proceedings of the 22nd ACM Internet Measurement Conference* (Nice, France) *(IMC '22)*. Association for Computing Machinery, New York, NY, USA, 716–729. https://doi.org/10.1145/3517745.3561419

[48] Ben Du, Cecilia Testart, Romain Fontugne, Alex C. Snoeren, and Kc Claffy. 2023. Poster: Taking the Low Road: How RPKI Invalids Propagate. In *Proceedings of the ACM SIGCOMM 2023 Conference* (New York, NY, USA) *(ACM SIGCOMM '23)*. Association for Computing Machinery, New York, NY, USA, 1144–1146. https://doi.org/10.1145/3603269.3610843

[49] Alain Durand. 2020. *Resource Public Key Infrastructure (RPKI) Technical Analysis*. Technical Report. ICANN. https://www.icann.org/en/system/files/files/octo-014-02sep20-en.pdf

[50] J. Durand, I. Pepelnjak, and G. Doering. 2015. *BGP Operations and Security*. BCP 194. RFC Editor.

[51] Federal Communications Commission. 2024. *Promoting Secure and Reliable Internet Routing: Notice of Proposed Rulemaking*. Technical Report. https://docs.fcc.gov/public/attachments/DOC-402609A1.pdf

[52] Romain Fontugne, Anant Shah, and Emile Aben. 2018. The (thin) Bridges of as Connectivity: Measuring Dependency Using as Hegemony. In *Passive and Active Measurement: 19th International Conference, PAM 2018, Berlin, Germany, March 26–27, 2018, Proceedings 19*. Springer, 216–227.

[53] Romain Fontugne, Malte Tashiro, Raffaele Sommese, Mattijs Jonker, Zachary S. Bischof, and Emile Aben. 2024. The Wisdom of the Measurement Crowd: Building the Internet Yellow Pages a Knowledge Graph for the Internet. In *Proceedings of the 2024 ACM on*

Internet Measurement Conference (Madrid, Spain) *(IMC '24)*. Association for Computing Machinery, New York, NY, USA, 183–198. https://doi.org/10.1145/3646547.3688444

[54] Sir Foxy. 2023. Hosting Providers Are Outraged at IPXO Over 'Two Incident' Abuse Report Policy. https://lowendbox.com/blog/hosting-providers-are-outraged-at-ipxo-over-two-incident-abuse-report-policy/

[55] Justin Furuness, Cameron Morris, Reynaldo Morillo, Amir Herzberg, and Bing Wang. 2023. BGPy: the BGP Python Security Simulator. In *Proceedings of the 16th Cyber Security Experimentation and Test Workshop* (Marina del Rey, CA, USA) *(CSET '23)*. Association for Computing Machinery, New York, NY, USA, 41–56. https://doi.org/10.1145/3607505.3607509

[56] Lixin Gao and J. Rexford. 2001. Stable Internet Routing Without Global Coordination. *IEEE/ACM Transactions on Networking* 9, 6 (2001), 681–692. https://doi.org/10.1109/90.974523

[57] Yossi Gilad, Avichai Cohen, Amir Herzberg, Michael Schapira, and Haya Shulman. 2016. Are We There yet? on RPKI's Deployment and Security. Cryptology ePrint Archive, Paper 2016/1010. https://eprint.iacr.org/2016/1010

[58] Yossi Gilad, Sharon Goldberg, Kotikalapudi Sriram, Job Snijders, and Ben Maddison. 2021. *The Use of MaxLength in the RPKI*. Internet-Draft draft-ietf-sidrops-rpkimaxlen-09. IETF Secretariat. https://www.ietf.org/archive/id/draft-ietf-sidrops-rpkimaxlen-09.txt

[59] Yossi Gilad, Omar Sagga, and Sharon Goldberg. 2017. MaxLength Considered Harmful to the RPKI. In *Proceedings of the 13th International Conference on Emerging Networking EXperiments and Technologies* (Incheon, Republic of Korea) *(CoNEXT '17)*. Association for Computing Machinery, New York, NY, USA, 101–107. https://doi.org/10.1145/3143361.3143363

[60] Vasileios Giotsas, Ioana Livadariu, and Petros Gigis. 2020. A First Look at the Misuse and Abuse of the IPv4 Transfer Market. In *Passive and Active Measurement*, Anna Sperotto, Alberto Dainotti, and Burkhard Stiller (Eds.). Springer International Publishing, Cham, 88–103.

[61] Sharon Goldberg, Michael Schapira, Peter Hummon, and Jennifer Rexford. 2010. How Secure Are Secure Interdomain Routing Protocols. In *Proceedings of the ACM SIGCOMM 2010 Conference* (New Delhi, India) *(SIGCOMM '10)*. Association for Computing Machinery, New York, NY, USA, 87–98. https://doi.org/10.1145/1851182.1851195

[62] Geoffrey Goodell, William Aiello, Timothy Griffin, John Ioannidis, Patrick D McDaniel, and Aviel D Rubin. 2003. Working Around BGP: an Incremental Approach to Improving Security and Accuracy in Interdomain Routing.. In *NDSS*, Vol. 23. Citeseer, 156.

[63] Dan Goodin. 2018. How 3ve's BGP Hijackers Eluded the Internet—and Made $29M. https://arstechnica.com/information-technology/2018/12/how-3ves-bgp-hijackers-eluded-the-internet-and-made-29m/

[64] Dan Goodin. 2022. How 3 Hours of Inaction From Amazon Cost Cryptocurrency Holders $235,000. https://arstechnica.com/information-technology/2022/09/how-3-hours-of-inaction-from-amazon-cost-cryptocurrency-holders-235000/

[65] Google. 2021. Peering with Google. https://peering.google.com/#/options/peering

[66] Google. 2024. External IP Address Pricing. https://cloud.google.com/vpc/network-pricing#vpc-pricing

[67] Google. 2024. Planning for Bring Your Own IP Addresses. https://cloud.google.com/vpc/docs/byoip-planning

[68] GoSSDHosting. 2023. Terrible Experience with IPXO. https://www.webhostingtalk.com/showthread.php?t=1882869

[69] Vincentas Grinius. 2024. Proud Moment for IPXO. https://www.linkedin.com/posts/grinius_sublet-your-subnet-inferring-ip-leasing-activity-7249351702196195330-P_DT/

[70] Brander Group. 2025. Removing Internet Routing Registry (IRR) Objects. https://brandergroup.net/removing-internet-routing-registry-objects-irr/

[71] Ethan Heilman, Danny Cooper, Leonid Reyzin, and Sharon Goldberg. 2014. From the Consent of the Routed: Improving the Transparency of the RPKI *(SIGCOMM '14)*. Association for Computing Machinery, New York, NY, USA, 51–62. https://doi.org/10.1145/2619239.2626293

[72] Remi Hendriks, Tim Betzer, Ben Du, Raffaele Sommese, Mattijs Jonker, and Roland van Rijswijk-Deij. 2025. Locating and Enumerating Anycast: a Comparison of Two Approaches. In *Proceedings of the 2025 Applied Networking Research Workshop* (Madrid, Spain) *(ANRW '25)*. Association for Computing Machinery, New York, NY, USA, 99–105. https://doi.org/10.1145/3744200.3744783

[73] Thomas Holterbach, Thomas Alfroy, Amreesh Phokeer, Alberto Dainotti, and Cristel Pelsser. 2024. A System to Detect Forged-Origin BGP Hijacks. In *21st USENIX Symposium on Networked Systems Design and Implementation (NSDI 24)*. USENIX Association, Santa Clara, CA, 1751–1770. https://www.usenix.org/conference/nsdi24/presentation/holterbach

[74] Takafusa Hori. 2021. *IIJ's Efforts with RPKI*. Technical Report. https://www.iij.ad.jp/en/dev/iir/pdf/iir_vol50_focus1_EN.pdf

[75] Bradley Huffaker, Romain Fontugne, Alexander Marder, and kc claffy. 2023. On the Importance of Being an as: an Approach to Country-Level as Rankings. In *Proceedings of the 2023 ACM on Internet Measurement Conference* (Montreal QC, Canada) *(IMC '23)*. Association for Computing Machinery, New York, NY, USA, 52–65. https://doi.org/10.1145/3618257.3624798

[76] G. Huston and G. Michaelson. 2012. *Validation of Route Origination Using the Resource Certificate Public Key Infrastructure (PKI) and Route Origin Authorizations (ROAs)*. RFC 6483. RFC Editor.

[77] Geoff Huston, Mattia Rossi, and Grenville Armitage. 2011. Securing BGP — a Literature Survey. *IEEE Communications Surveys Tutorials* 13, 2 (2011), 199–222. https://doi.org/10.1109/SURV.2011.041010.00041

[78] IPv4.Global. 2024. IPv4 Address Auctions. https://auctions.ipv4.global/

[79] IPXO. 2022. Abuse Incident Handling Process at the IPXO Marketplace. https://www.ipxo.com/kb/abuse-prevention-and-reputation/abuse-incident-handling-marketplace/

[80] IPXO. 2022. RPKI Management at IPXO (IP Holder). https://www.ipxo.com/kb/technical-guides/rpki-management-at-ipxo-ip-holder/

[81] IPXO. 2024. Bring Your Own IP – a Cost-Efficient Way to Expand Your Infrastructure. https://www.ipxo.com/bring-your-own-ip/

[82] IPXO. 2024. IPXO Market Place. https://www.ipxo.com/ip-marketplace/

[83] Katherine Izhikevich, Ben Du, Sumanth Rao, Alisha Ukani, and Liz Izhikevich. 2024. Trust, But Verify, Operator-Reported Geolocation. *arXiv preprint arXiv:2409.19109* (2024).

[84] Internet Initiative Japan. 2022. Internet Health Report. https://ihr.iijlab.net/ihr/en-us/rov

[85] Job Snijders. 2018. Routing Security Roadmap. https://nlnog.net/static/nlnogday2018/9_routing_security_roadmap_nlnog_2018_snijders.pdf

[86] Josh Karlin, Stephanie Forrest, and Jennifer Rexford. 2006. Pretty Good BGP: Improving BGP by Cautiously Adopting Routes. In *Proceedings of the Proceedings of the 2006 IEEE International Conference on Network Protocols (ICNP '06)*. IEEE Computer Society, USA, 290–299. https://doi.org/10.1109/ICNP.2006.320179

[87] Stephen Kent, Charles Lynn, and Karen. Seo. 2000. Secure Border Gateway Protocol (S-BGP). *IEEE Journal on Selected Areas in Communications* 18, 4 (2000), 582–592.

[88] Akmal Khan, Hyun-chul Kim, Taekyoung Kwon, and Yanghee Choi. 2013. A Comparative Study on IP Prefixes and Their Origin Ases in BGP and the IRR. *SIGCOMM Comput. Commun. Rev.* 43, 3 (July 2013), 16–24. https://doi.org/10.1145/2500098.2500101

[89] Etienne Khan, Anna Sperotto, Jeroen van der Ham, and Roland van Rijswijk-Deij. 2023. Stranger VPNs: Investigating the Geo-Unblocking Capabilities of Commercial VPN Providers. In *International Conference on Passive and Active Network Measurement*. Springer, 46–68.

[90] Mohammad Taha Khan, Joe DeBlasio, Geoffrey M. Voelker, Alex C. Snoeren, Chris Kanich, and Narseo Vallina-Rodriguez. 2018. An Empirical Analysis of the Commercial VPN Ecosystem. In *Proceedings of the Internet Measurement Conference 2018* (Boston, MA, USA) *(IMC '18)*. Association for Computing Machinery, New York, NY, USA, 443–456. https://doi.org/10.1145/3278532.3278570

[91] E-yong Kim, Li Xiao, Klara Nahrstedt, and Kunsoo Park. 2008. Secure Interdomain Routing Registry. *IEEE Transactions on Information Forensics and Security* 3, 2 (2008), 304–316. https://doi.org/10.1109/TIFS.2008.922050

[92] Thomas Krenc and Anja Feldmann. 2016. BGP Prefix Delegations: a Deep Dive. In *Proceedings of the 2016 Internet Measurement Conference* (Santa Monica, California, USA) *(IMC '16)*. Association for Computing Machinery, New York, NY, USA, 469–475. https://doi.org/10.1145/2987443.2987458

[93] John Kristoff, Randy Bush, Chris Kanich, George Michaelson, Amreesh Phokeer, Thomas C. Schmidt, and Matthias Wählisch. 2020. On Measuring RPKI Relying Parties. In *Proceedings of the ACM Internet Measurement Conference* (Virtual Event, USA) *(IMC '20)*. Association for Computing Machinery, New York, NY, USA, 484–491. https://doi.org/10.1145/3419394.3423622

[94] Ching-Heng Ku. 2020. 98% of Taiwan's IP Address Holders Have Signed RPKI ROAs. https://blog.apnic.net/2020/10/16/98-of-taiwans-ip-address-holders-have-signed-rpki-roas/

[95] Brenden Kuerbis and Milton Mueller. 2017. Internet Routing Registries, Data Governance, and Security. *Journal of Cyber Policy* 2, 1 (2017), 64–81.

[96] IIJ Labs. 2025. IYP Console. https://iyp.iijlab.net/

[97] LACNIC. 2024. IPv4 Address Allocation and Assignment Policies. https://www.lacnic.net/682/2/lacnic/

[98] Matt Lehwess. 2018. Introducing Bring Your Own IP (BYOIP) for Amazon VPC. https://aws.amazon.com/blogs/networking-and-content-delivery/introducing-bring-your-own

-ip-byoip-for-amazon-vpc/

[99] M. Lepinski and S. Kent. 2012. *An Infrastructure to Support Secure Internet Routing*. RFC 6480. RFC Editor. http://www.rfc-editor.org/rfc/rfc6480.txt

[100] Matt Lepinski and Kotikalapudi Sriram. 2017. *BGPsec Protocol Specification*. RFC 8205.

[101] Weitong Li, Yuze Li, and Taejoong Chung. 2025. ImpROV: Measurement and Practical Mitigation of Collateral Damage of RPKI Route Origin Validation. In *Proceedings of the USENIX Security Symposium (Security'25)*. Seattle, USA.

[102] Weitong Li, Zhexiao Lin, Md. Ishtiaq Ashiq, Emile Aben, Romain Fontugne, Amreesh Phokeer, and Taejoong Chung. 2023. RoVista: Measuring and Analyzing the Route Origin Validation (ROV) in RPKI. In *Proceedings of the 2023 ACM on Internet Measurement Conference* (Montreal QC, Canada) *(IMC '23)*. Association for Computing Machinery, New York, NY, USA, 73–88. https://doi.org/10.1145/3618257.3624806

[103] NANOG Mailing List. 2011. "Leasing" of Space via Non-connectivity Providers (was: Re: and so It Ends...). https://mailman.nanog.org/pipermail/nanog/2011-February/032158.html

[104] Ioana Livadariu, Thomas Dreibholz, Anas Saeed Al-Selwi, Haakon Bryhni, Olav Lysne, Steinar Bjørnstad, and Ahmed Elmokashfi. 2020. On the Accuracy of Country-Level IP Geolocation. In *Proceedings of the 2020 Applied Networking Research Workshop* (Virtual Event, Spain) *(ANRW '20)*. Association for Computing Machinery, New York, NY, USA, 67–73. https://doi.org/10.1145/3404868.3406664

[105] Ioana Livadariu, Ahmed Elmokashfi, and Amogh Dhamdhere. 2017. On IPv4 Transfer Markets: Analyzing Reported Transfers and Inferring Transfers in the Wild. *Computer Communications* 111 (2017), 105–119. https://doi.org/10.1016/j.comcom.2017.07.012

[106] Ioana Livadariu, Ahmed Elmokashfi, Amogh Dhamdhere, and kc claffy. 2013. A First Look at IPv4 Transfer Markets. In *Proceedings of the Ninth ACM Conference on Emerging Networking Experiments and Technologies* (Santa Barbara, California, USA) *(CoNEXT '13)*. Association for Computing Machinery, New York, NY, USA, 7–12. https://doi.org/10.1145/2535372.2535416

[107] Ioana Livadariu, Kevin Vermeulen, Maxime Mouchet, and Vasilis Giotsas. 2024. Geofeeds: Revolutionizing IP Geolocation or Illusionary Promises? *Proc. ACM Netw.* 2, CoNEXT3, Article 15 (Aug 2024), 21 pages. https://doi.org/10.1145/3676869

[108] Matthew Luckie, Robert Beverly, Ryan Koga, Ken Keys, Joshua A. Kroll, and k claffy. 2019. Network Hygiene, Incentives, and Regulation: Deployment of Source Address Vali-

dation in the Internet. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security* (London, United Kingdom) *(CCS '19)*. Association for Computing Machinery, New York, NY, USA, 465–480.

[109] Matthew Luckie, Bradley Huffaker, Amogh Dhamdhere, Vasileios Giotsas, and kc claffy. 2013. AS Relationships, Customer Cones, and Validation. In *Proceedings of the 2013 Conference on Internet Measurement Conference* (Barcelona, Spain) *(IMC '13)*. Association for Computing Machinery, New York, NY, USA, 243–256. https://doi.org/10.1145/2504730.2504735

[110] Lumen Technologies 2021. *Lumen Technologies Guide to the LEVEL3 Internet Routing Registry*. Lumen Technologies.

[111] Robert Lychev, Sharon Goldberg, and Michael Schapira. 2013. BGP Security in Partial Deployment: Is the Juice Worth the Squeeze?. In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM* (Hong Kong, China) *(SIGCOMM '13)*. Association for Computing Machinery, New York, NY, USA, 171–182. https://doi.org/10.1145/2486001.2486010

[112] Doug Madory. 2018. BGP Hijack of Amazon DNS to Steal Crypto Currency. https://dyn.com/blog/bgp-hijack-of-amazon-dns-to-steal-crypto-currency/

[113] Doug Madory. 2020. Visulizing Routing Incidents in 3D. https://ripe80.ripe.net/presentations/14-3dleak_viz_madory_ripe.pdf

[114] Doug Madory and Job Snijders. 2022. Measuring RPKI ROV adoption with NetFlow. https://www.kentik.com/blog/measuring-rpki-rov-adoption-with-netflow/

[115] Doug Madory and Job Snijders. 2024. RPKI ROV Deployment Reaches Major Milestone. https://www.kentik.com/blog/rpki-rov-deployment-reaches-major-milestone/

[116] MANRS. 2021. Mutually Agreed Norms for Routing Security. https://www.manrs.org/

[117] MANRS. 2022. MANRS Observatory. https://observatory.manrs.org/#/overview

[118] Apostolaki Maria, Zohar Aviv, and Vanbever Laurent. 2017. Hijacking Bitcoin: Routing Attacks on Cryptocurrencies. In *Security and Privacy (SP), 2017 IEEE Symposium On*. IEEE.

[119] Merit Network. 2021. The Internet Routing Registry - RADb. https://www.radb.net/

[120] Inc Merit Network. 2018. Internet Routing Registry. http://www.irr.net

[121] George Michaelson. 2017. IRR and RPKI: a Problem Statement. https://conference.apn

ic.net/44/assets/files/APCS549/Global-IRR-and-RPKI-a-problem-statement.pdf

[122] Prodosh Mohapatra, John Scudder, David Ward, Randy Bush, and Rob Austein. 2013. *BGP Prefix Origin Validation*. RFC 6811. http://www.rfc-editor.org/rfc/rfc6811.txt

[123] Milton Mueller, Vagisha Srivastava, and Brenden Kuerbis. 2021. IA Fight Over Crumbs: the AFRINIC Crisis. https://www.internetgovernance.org/2021/08/19/a-fight-over-crumbs-the-afrinic-crisis/

[124] S. Murphy. 2006. *BGP Security Vulnerabilities Analysis*. RFC 4272. RFC Editor.

[125] RIPE NCC. 2008. YouTube Hijacking: a RIPE NCC RIS Case Study. https://www.ripe.net/publications/news/industry-developments/youtube-hijacking-a-ripe-ncc-ris-case-study

[126] RIPE NCC. 2023. Brokers. Internet Archive. https://web.archive.org/web/20231216062402/https://www.ripe.net/manage-ips-and-asns/resource-transfers-and-mergers/brokers

[127] RIPE NCC. 2023. Routing Information System (RIS). https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris

[128] RIPE NCC. 2024. Decommissioning Recognised IPv4 Transfer Brokers List and IPv4 Transfer Listing Service. https://www.ripe.net/publications/news/decommissioning-recognised-ipv4-transfer-brokers-list-and-ipv4-transfer-listing-service/

[129] RIPE NCC. 2024. IPv4 Address Allocation and Assignment Policies. https://www.ripe.net/publications/docs/ripe-822/#70-types-of-address-space

[130] RIPE NCC. 2024. Legacy Space in the RIPE Registry. https://www.ripe.net/manage-ips-and-asns/legacy-resources/legacy-space-in-the-ripe-registry/

[131] RIPE NCC. 2024. RIPE NCC Services to Legacy Internet Resource Holders. https://www.ripe.net/manage-ips-and-asns/legacy-resources/ripe-ncc-services-to-legacy-internet-resource-holders/

[132] Hacker News. 2022. IPv4 Address Auctions. https://news.ycombinator.com/item?id=32416043

[133] NLNOG. 2021. IRR Explorer. https://irrexplorer.nlnog.net/

[134] NLNOG. 2025. Rejecting RPKI Invalid BGP Routes. https://bgpfilterguide.nlnog.net/guides/reject_invalids/

[135] Ola Nordström and Constantinos Dovrolis. 2004. Beware of BGP Attacks. *SIGCOMM*

Comput. Commun. Rev. 34, 2 (April 2004), 1–8. https://doi.org/10.1145/997150.997152

[136] NTT. 2022. Routing Registry - RPKI Based BGP Origin Validation. https://www.gin.ntt.net/support-center/policies-procedures/routing-registry/#RPKI

[137] NTT. 2023. RPKI Suppression of Conflicting IRR Information. https://www.gin.ntt.net/support-center/policies-procedures/routing-registry/

[138] NTT. 2025. IRRd version 4 Questions and Answers. https://www.gin.ntt.net/support-center/policies-procedures/routing-registry/

[139] University of Oregon. 2023. Route Views Project. http://www.routeviews.org/routeviews

[140] National Institute of Standards and Technology (NIST). 2025. *Special Publication 800-189 Revision 1 (Initial Public Draft): Guide for the Mapping of Identified Internet Assets to National Security Systems*. Technical Report. National Institute of Standards and Technology. https://csrc.nist.gov/pubs/sp/800/189/r1/ipd

[141] Ricardo Oliveira, Dan Pei, Walter Willinger, Beichuan Zhang, and Lixia Zhang. 2010. The (in)completeness of the Observed Internet AS-level Structure. *IEEE/ACM Trans. Netw.* 18, 1 (Feb. 2010), 109–122. https://doi.org/10.1109/TNET.2009.2020798

[142] Leo Oliver, Gautam Akiwate, Matthew Luckie, Ben Du, and kc claffy. 2022. Stop, DROP, and ROA: Effectiveness of Defenses Through the Lens of DROP. In *Proceedings of the 22nd ACM Internet Measurement Conference* (Nice, France) *(IMC '22)*. Association for Computing Machinery, New York, NY, USA, 730–737. https://doi.org/10.1145/3517745.3561454

[143] P.C. van Oorschot, Tao Wan, and Evangelos Kranakis. 2007. On Interdomain Routing Security and Pretty Secure BGP (PsBGP). *ACM Trans. Inf. Syst. Secur.* 10, 3 (July 2007), 11–es.

[144] Eric Osterweil, Shane Amante, Dan Massey, and Danny McPherson. 2011. The Great IPv4 Land Grab: Resource Certification for the IPv4 Grey Market. In *Proceedings of the 10th ACM Workshop on Hot Topics in Networks* (Cambridge, Massachusetts) *(HotNets-X)*. Association for Computing Machinery, New York, NY, USA, Article 12, 6 pages. https://doi.org/10.1145/2070562.2070574

[145] Lars Prehn and Anja Feldmann. 2021. How Biased Is Our Validation (data) for as Relationships?. In *Proceedings of the 21st ACM Internet Measurement Conference* (Virtual Event) *(IMC '21)*. Association for Computing Machinery, New York, NY, USA, 612–620. https://doi.org/10.1145/3487552.3487825

[146] Lars Prehn, Franziska Lichtblau, and Anja Feldmann. 2020. When Wells Run Dry: the

2020 IPv4 Address Market. In *Proceedings of the 16th International Conference on Emerging Networking EXperiments and Technologies* (Barcelona, Spain) *(CoNEXT '20)*. Association for Computing Machinery, New York, NY, USA, 46–54. https://doi.org/10.1145/3386367.3431301

[147] RADB. 2025. IRRD v4 Migration FAQ. https://www.radb.net/support/informational/irrdv4-migration-faq.html

[148] Anirudh Ramachandran and Nick Feamster. 2006. Understanding the Network-level Behavior of Spammers. In *Proceedings of the 2006 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications* (Pisa, Italy) *(SIGCOMM '06)*. Association for Computing Machinery, New York, NY, USA, 291–302. https://doi.org/10.1145/1159913.1159947

[149] Y. Rekhter, T. Li, and S. Hares. 2006. *A Border Gateway Protocol 4 (BGP-4)*. RFC 4271. RFC Editor. http://www.rfc-editor.org/rfc/rfc4271.txt http://www.rfc-editor.org/rfc/rfc4271.txt.

[150] Andreas Reuter, Randy Bush, Italo Cunha, Ethan Katz-Bassett, Thomas C. Schmidt, and Matthias Wählisch. 2018. Towards a Rigorous Methodology for Measuring Adoption of RPKI Route Validation and Filtering. *SIGCOMM Comput. Commun. Rev.* 48, 1 (April 2018), 19–27. https://doi.org/10.1145/3211852.3211856

[151] Philipp Richter, Mark Allman, Randy Bush, and Vern Paxson. 2015. A Primer on IPv4 Scarcity. *SIGCOMM Comput. Commun. Rev.* 45, 2 (apr 2015), 21–31. https://doi.org/10.1145/2766330.2766335

[152] RIPE NCC. 2019. Managing Route Objects in the IRR. https://www.ripe.net/manage-ips-and-asns/db/support/managing-route-objects-in-the-irr

[153] RIPE NCC. 2019. RIPE NCC IRR Database Non-Authoritative Route Object Clean-up. https://www.ripe.net/publications/docs/ripe-731/

[154] RIPE NCC. 2019. Route Object Creation. https://www.ripe.net/support/training/material/bgp-operations-and-security-training-course/route-object-creation-flowchart.pdf

[155] RIPE NCC. 2019. The RIPE NCC Has Run Out of IPv4 Addresses. https://www.ripe.net/publications/news/about-ripe-ncc-and-ripe/the-ripe-ncc-has-run-out-of-ipv4-addresses

[156] RIPE NCC. 2021. Ending Support for the RIPE NCC RPKI Validator. https://www.ripe.net/publications/news/announcements/ending-support-for-the-ripe-ncc-rpki-validator

[157] RIPE NCC. 2021. RPKI Validator. https://rpki-validator.ripe.net/

[158] Andrei Robachevsky. 2025. Improving Routing Security Through Concerted Action. https://www.first.org/resources/papers/conf2025/1130-1205-Routing-Security-for-Enterprises-Andrei-Robachevsky.pdf

[159] Nils Rodday, Ítalo Cunha, Randy Bush, Ethan Katz-Bassett, Gabi Dreo Rodosek, Thomas C Schmidt, and Matthias Wählisch. 2021. Revisiting RPKI Route Origin Validation on the Data Plane. In *Proc. of Network Traffic Measurement and Analysis Conference (TMA). IFIP. Accepted for Publication*.

[160] Nurul Islam Roman. 2014. Multihoming Design Options with Portable IP Address- Pros and Cons. https://blog.apnic.net/2014/09/22/multihoming-design-options-with-portable-ip-address-pros-and-cons/

[161] Sojun Ryu. 2022. Post Mortem of KlaySwap Incident Through BGP Hijacking. https://medium.com/s2wblog/post-mortem-of-klayswap-incident-through-bgp-hijacking-en-3ed7e33de600

[162] Loqman Salamatian, Todd Arnold, Ítalo Cunha, Jiangchen Zhu, Yunfan Zhang, Ethan Katz-Bassett, and Matt Calder. 2023. Who Squats IPv4 Addresses? *SIGCOMM Comput. Commun. Rev.* 53, 1 (apr 2023), 48–72. https://doi.org/10.1145/3594255.3594260

[163] Loqman Salamatian, Kevin Vermeulen, Italo Cunha, Vasilis Giotsas, and Ethan Katz-Bassett. 2024. metAScritic: Reframing AS-Level Topology Discovery as a Recommendation System. In *Proceedings of the 2024 ACM on Internet Measurement Conference* (Madrid, Spain) *(IMC '24)*. Association for Computing Machinery, New York, NY, USA, 337–364. https://doi.org/10.1145/3646547.3688429

[164] Khwaja Zubair Sediqi, Lars Prehn, and Oliver Gasser. 2022. Hyper-specific Prefixes: Gotta Enjoy the Little Things in Interdomain Routing. *SIGCOMM Comput. Commun. Rev.* 52, 2 (jun 2022), 20–34. https://doi.org/10.1145/3544912.3544916

[165] Pavlos Sermpezis, Vasileios Kotronis, Petros Gigis, Xenofontas Dimitropoulos, Danilo Cicalese, Alistair King, and Alberto Dainotti. 2018. ARTEMIS: Neutralizing BGP Hijacking Within a Minute. *IEEE/ACM Transactions on Networking* 26, 6 (2018), 2471–2486.

[166] Pavlos Sermpezis, Lars Prehn, Sofia Kostoglou, Marcel Flores, Athena Vakali, and Emile Aben. 2023. Bias in Internet Measurement Platforms. In *2023 7th Network Traffic Measurement and Analysis Conference (TMA)*. 1–10. https://doi.org/10.23919/TMA58422.2023.10198985

[167] Amazon Web Services. 2023. Settlement Free Peering Policy. https://aws.amazon.com/peering/policy/

[168] Kris Shrishak and Haya Shulman. 2020. Limiting the Power of RPKI Authorities. In *Proceedings of the Applied Networking Research Workshop* (Virtual Event, Spain) *(ANRW '20)*. Association for Computing Machinery, New York, NY, USA, 12–18. https://doi.org/10.1145/3404868.3406674

[169] Aftab Siddiqui. 2017. Google Buys a /12 IPv4 Address Block. https://www.internetsociety.org/blog/2017/05/google-buys-a-12-ipv4-address-block/

[170] Aftab Siddiqui. 2020. Big Route Leak Shows Need for Routing Security. https://www.manrs.org/2020/07/big-route-leak-shows-need-for-routing-security

[171] Georgos Siganos and Michalis Faloutsos. 2004. Analyzing BGP Policies: Methodology and Tool. In *IEEE INFOCOM 2004*, Vol. 3. 1640–1651 vol.3. https://doi.org/10.1109/INFCOM.2004.1354576

[172] Georgos Siganos and Michalis Faloutsos. 2007. Neighborhood Watch for Internet Routing: Can We Improve the Robustness of Internet Routing Today?. In *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*. 1271–1279. https://doi.org/10.1109/INFCOM.2007.151

[173] Internet Society. 2018. *Routing Security for Policymakers: an Internet Society White Paper*. White Paper. MANRS. https://www.manrs.org/wp-content/uploads/2018/10/Routing-Security-for-Policymakers-EN.pdf

[174] Internet Society. 2020. Making the Most of Our MANRS Partnerships – NIC.br and Brazil Lead the MANRS Pack. https://www.manrs.org/2020/06/making-the-most-of-our-manrs-partnerships-nic-br-and-brazil-lead-the-manrs-pack

[175] Spamhaus. 2024. Don't Route or Peer Lists (DROP). https://www.spamhaus.org/blocklists/do-not-route-or-peer/

[176] Kotikapaludi Sriram, Oliver Borchert, Okhee Kim, Patrick Gleichmann, and Doug Montgomery. 2009. A Comparative Analysis of BGP Anomaly Detection and Robustness Algorithms. In *2009 Cybersecurity Applications & Technology Conference for Homeland Security*. 25–38. https://doi.org/10.1109/CATCH.2009.20

[177] John Sweeting. 2023. ARIN and IPv4 Address Leasing. https://www.arin.net/blog/2023/02/23/ipv4-leasing/

[178] Virginia Tech. 2025. RoVista. https://rovista.netsecurelab.org/

[179] Cecilia Testart. 2018. Reviewing a Historical Internet Vulnerability: Why Isn't BGP More Secure and What Can We Do About It? TPRC.

[180] Cecilia Testart, Philipp Richter, Alistair King, Alberto Dainotti, and David Clark. 2019. Profiling BGP Serial Hijackers: Capturing Persistent Misbehavior in the Global Routing Table. In *Proceedings of the Internet Measurement Conference* (Amsterdam, Netherlands) *(IMC '19)*. Association for Computing Machinery, New York, NY, USA, 420–434.

[181] Cecilia Testart, Philipp Richter, Alistair King, Alberto Dainotti, and David Clark. 2020. To Filter or Not to Filter: Measuring the Benefits of Registering in the RPKI Today. In *International Conference on Passive and Active Network Measurement*. Springer, 71–87.

[182] Cecilia Testart, Josephine Wolff, Deepak Gouda, and Romain Fontugne. 2024. Identifying Current Barriers in RPKI Adoption. In *Proceedings of the 52nd Research Conference on Communication, Information and Internet Policy (TPRC)*. https://doi.org/10.2139/ssrn.4948317

[183] The White House. 2024. *Roadmap to Enhancing Internet Routing Security*. Technical Report. Executive Office of the President of the United States. https://www.whitehouse.gov/wp-content/uploads/2024/09/Roadmap-to-Enhancing-Internet-Routing-Security.pdf

[184] Andree Toonk. 2014. Using BGP Data to Find Spammers. https://www.bgpmon.net/using-bgp-data-to-find-spammers/

[185] Bill Toulas. 2024. Cloudflare Blames Recent Outage on BGP Hijacking Incident. https://www.bleepingcomputer.com/news/security/cloudflare-blames-recent-outage-on-bgp-hijacking-incident/

[186] Pierre-Antoine Vervier, Olivier Thonnard, and Marc Dacier. 2015. Mind Your Blocks: on the Stealthiness of Malicious BGP Hijacks.. In *The Network and Distributed System Security Symposium (NDSS)*.

[187] Matthias Wählisch, Robert Schmidt, Thomas C. Schmidt, Olaf Maennel, Steve Uhlig, and Gareth Tyson. 2015. RiPKI: the Tragic Story of RPKI Deployment in the Web Ecosystem. In *Proc. of Fourteenth ACM Workshop on Hot Topics in Networks (HotNets)*. ACM, New York.

[188] Joe Westover. 2021. Are Your IP Addresses Portable? https://www.arin.net/blog/2021/01/04/are-your-ip-addresses-portable/

[189] Russ White. 2003. Securing BGP through secure origin BGP (soBGP). *Business Communications Review* 33, 5 (2003), 47–53.

[190] Jon Worley. 2020. IRR and ARIN-NONAUTH: How Do I Use It? https://www.arin.net/blog/2020/06/25/irr-and-arin-nonauth-how-do-i-use-it/

[191] Christopher S Yoo and David A Wishnick. 2019. Lowering Legal Barriers to RPKI

Adoption. *U of Penn Law School, Public Law Research Paper* 19-02 (2019).

[192] Luciano Zembruzki, Raffaele Sommese, Lisandro Zambenedetti Granville, Arthur Selle Jacobs, Mattijs Jonker, and Giovane C. M. Moura. 2022. Hosting Industry Centralization and Consolidation. In *NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium*. 1–9. https://doi.org/10.1109/NOMS54207.2022.9789881