IRRegularities in the Internet Routing Registry

Ben Du
UC San Diego
bendu@ucsd.edu

Katherine Izhikevich
UC San Diego
kizhikev@ucsd.edu

Sumanth Rao
UC San Diego
svrao@ucsd.edu

Gautam Akiwate
Stanford University
gakiwate@cs.stanford.edu

Cecilia Testart
Georgia Tech
ctestart@gatech.edu

Alex C. Snoeren
UC San Diego
snoeren@cs.ucsd.edu

kc claffy
CAIDA/UC San Diego
kc@caida.org

ABSTRACT

The Internet Routing Registry (IRR) is a set of distributed databases used by networks to register routing policy information and to validate messages received in the Border Gateway Protocol (BGP). First deployed in the 1990s, the IRR remains the most widely used database for routing security purposes, despite the existence of more recent and more secure alternatives. Yet, the IRR lacks a strict validation standard and the limited coordination across different database providers can lead to inaccuracies. Moreover, it has been reported that attackers have begun to register false records in the IRR to bypass operators' defenses when launching attacks on the Internet routing system, such as BGP hijacks. In this paper, we provide a longitudinal analysis of the IRR over the span of 1.5 years. We develop a workflow to identify *irregular* IRR records that contain conflicting information compared to different routing data sources. We identify 34,199 irregular route objects out of 1,542,724 route objects from November 2021 to May 2023 in the largest IRR database and find 6,373 to be potentially suspicious.

CCS CONCEPTS

• **Networks** → **Network security**; **Network measurement**.

KEYWORDS

BGP, Internet Routing Registry, Routing Security.

ACM Reference Format:
Ben Du, Katherine Izhikevich, Sumanth Rao, Gautam Akiwate, Cecilia Testart, Alex C. Snoeren, and kc claffy. 2023. IRRegularities in the Internet Routing Registry. In *Proceedings of the 2023 ACM Internet Measurement Conference (IMC '23), October 24–26, 2023, Montreal, QC, Canada.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3618257.3624843

1 INTRODUCTION

The Internet Routing Registry (IRR) is a conglomerate of distributed databases that facilitate sharing of routing policy information. Network operators can use the Routing Policy Specification Language (RPSL) to register routing policies of their networks in one or more of the 18 currently operational IRR databases as of May 2023 [2, 28]. Different types of organizations (i.e., commercial, non-profit, and Internet registries) operate IRR databases, and IRR database operators have different procedures to authenticate database users and validate information registered in the IRR. Due to the diversity of IRR databases and a lack of coordination among them, information can be inconsistent, and the quality of IRR databases varies. In addition, networks may register routing information in multiple IRR databases but only maintain a subset of them, further contributing to the inconsistency of across IRRs [12, 21, 41].

Operators use information in the IRR for several purposes, such as constructing prefix lists and building inbound BGP filters. Transit providers may use information registered by customers in IRR databases to decide which IP prefixes to provide transit for [4, 22]. Some cloud providers and Internet Exchange Points (IXPs) require peering networks to register their IP prefixes and AS numbers in an IRR [3, 11, 18, 37]. Some networks allow their customers and peers to register in any IRR database, while others support only a subset. Although there is a more recent and secure alternative of IRR, the Resource Public Key Infrastructure (RPKI), many networks have not yet deployed RPKI-based filtering [10, 14, 15, 44]. 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 [13, 23].

The high popularity of IRR-based filtering makes inconsistency across IRR databases a problem for routing security. 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. Such vulnerabilities have allowed malicious actors to falsify records in the IRR and subsequently falsely originate the prefix in BGP. The upstream provider may be unaware of those falsified IRR records and may proceed to propagate the hijacked prefix to the rest of the Internet, thus completing the life cycle of a BGP hijack. Recent news articles and blog entries [16, 17, 43] have reported increasingly more complicated methods to falsify IRR records that have caused significant financial loss to victims.

We analyze the IRR ecosystem to help the Internet community understand the current strengths and weaknesses of the IRR, propose a pipeline to identify irregular IRR records, and compile a list of IRR records that we validated as irregular and potentially suspicious. Our main contributions are as follows:

(1) We quantify the inconsistencies among IRR databases, between IRR databases and BGP, between IRR databases and RPKI, and show the evolution of IRR database quality.
(2) We develop an automated workflow to filter IRR records with origin ASes that are inconsistent with these sources (i.e. the *irregular* objects).

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

ACM ISBN 979-8-4007-0382-9/23/10.
https://doi.org/10.1145/3618257.3624843

(3) We validate the irregular IRR records against RPKI and reported BGP hijackers, manually inspect some suspicious cases, and compile a list of irregular IRR records.

2 BACKGROUND

In this section, we introduce the structure of the IRR and relevant IRR records. We also discuss a selection of abuse cases in the IRR.

2.1 Internet Routing Registry

The Internet Routing Registry was first introduced in 1995 to facilitate the sharing of routing policy information among networks [6]. Two of the largest and oldest IRR databases are RADB and the RIPE IRR, operated by Merit Network and the RIPE NCC [24, 34], respectively. Over the past few decades, newer IRR databases have emerged operated by commercial companies (*e.g.,* Lumen, NTT), Regional Internet Registries (RIRs), and Local Internet Registries (LIRs). Regardless of the organizational type of the IRR operators, they all serve the same purpose of sharing routing information so that others can use the databases to construct BGP route filters.

Each IRR database is managed independently under different policies and registration processes. The five RIRs (RIPE, ARIN, APNIC, AFRINIC, and LACNIC) manage *authoritative IRR databases*. Routing information registered in those IRR databases undergoes a validation process against the address ownership information to ensure correctness [35]. IRR databases operated by other institutions are *non-authoritative IRR databases* and are not validated [19].

The relevant IRR records in this paper are `route`, `inetnum`, `mntner`, and `as-set` objects. To register routing information in the IRR, an organization first needs to register its authentication information and network operator email in a maintainer (`mntner`) object. The organization can then create and modify IRR records such as the `route` object. The `route` object contains the IP prefix and ASN that the organization intends to use to originate the prefix in BGP. The authoritative IRRs (*e.g.,* RIPE, ARIN) contain the `inetnum` object (or its equivalent `NetHandle`), which contains address ownership information, but this object is generally not present in other IRRs. The `as-set` object can be used to denote the customers, peers, or providers of an AS [5].

2.2 Falsified IRR Records

In response to the increasing operator use of the IRR to validate BGP announcements, malicious actors have begun to inject false IRR records in an effort to increase the likelihood of launching a successful BGP hijack attack.

False records in RADB. In one instance an abuse report received by UCSD reported that AS207427 (`GoHosted.eu`) hijacked 3 UCSD prefixes in BGP for ≈45 days through the end of 2020 and into the beginning of 2021. The postmortem report reveals that the attacker registered `route` objects containing those prefixes and AS207427 as the origin AS in RADB. The upstream provider of AS207427 propagated the announcement to the rest of the Internet because they were able to validate the malicious announcement against RADB records. RADB later deleted the false `route` object after being contacted by the true address space owner.

False records in ALTDB. In August 2022, Celer Network (AS209243), a blockchain technology company, lost $235,000 USD worth of cryptocurrency as a victim of BGP hijacking. The attacker hijacked the Amazon address space that was used to host Celer Network's website and rerouted Celer's customers to their phishing page [17]. The attacker pretended to be an upstream provider of AS16509 (Amazon) by registering a `route` object in ALTDB with the hijacked prefix 44.235.216.0/24 with AS16509 (Amazon) and an `as-set` object containing AS209243 and AS16509 as members.

3 RELATED WORK

Previous works have studied the accuracy of the IRR and its consistency with BGP. Khan et al. [20] 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. Siganos and Faloutsos [38] compared the business relationships of networks extracted from the IRR to that from BGP data and found 83% of the routing policies were consistent. Less than a year later, Siganos and Faloutsos [39] expanded the study to compare the consistency of BGP prefixes 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 validation method matched maintainers of IRR `route` objects to maintainers of RIR WHOIS database records (`inetnum` objects), 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. [40] enhanced Siganos and Faloutsos' validation algorithm to analyze the consistency between `route` and `inetnum` (or `net-handle`) objects in all authoritative IRRs and RADB. They found APNIC to be the most consistent and RADB the least consistent. However, RADB was not designed to store address ownership information and hence has few `inetnum` objects. We need another approach to evaluate the consistency of RADB.

The increasing deployment of RPKI allows a more comprehensive and rigorous consistency analysis of RADB. In 2021, Du et al. [12] found significant inconsistency between RADB and RPKI. They suggested that network operators should not trust all IRR databases equally given the uneven hygiene across IRRs. In 2022, Oliver et al. [31] found evidence of attackers abusing the IRR to circumvent IRR-based filters.

These studies show the difficulty of comprehensively validating IRR information as there is limited ground truth for Internet routing information. In this paper, we provide a first look of the inconsistencies across all IRR databases and propose a workflow to identify irregular IRR records without external sources of ground truth. We then adapt the methodology by Du et al. [12]—i.e. using RPKI as a source of ground truth—to validate our result, which we use to further refine our list of inferred irregular IRR records.

4 DATASET

We use the following datasets to study the behavior of IRR objects and identify irregularities.

IRR archive. We downloaded daily snapshots of IRR databases [28] between November 2021 and May 2023 and aggregated the `route` objects from each IRR database into a separate longitudinal database. We refer to this dataset as the *IRR dataset.* In November 2021, we were able to access 21 IRR databases from 17

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

Table 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.

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 [28]. `Canarie` stopped responding to FTP requests before May 2023 but was still listed as active on `irr.net`. Table 1 describes the size of each IRR database in November 2021 and May 2023.

BGP dataset. We used the CAIDA BGPView tool [32] to read BGP updates collected from Routeviews [30] and RIPE RIS [27] between November 2021 and May 2023. We created BGP snapshots in 5-minute increments to capture transient BGP announcements. We constructed a database of these snapshots, which we call the *BGP dataset*.

RPKI archive. RIPE NCC publishes daily lists of validated Route Origin Authorization (ROA) payloads from the five RPKI trust anchors (APNIC, ARIN, RIPE NCC, AFRINIC, LACNIC) [36]. We sampled daily snapshots of this dataset to create our *RPKI dataset*.

Serial BGP hijackers. Testart et al. [42] provide a list of BGP serial hijackers based on their routing behavior. We refer to this list at the *serial hijacker dataset*.

Supporting datasets. We used the CAIDA AS Rank dataset [7], CAIDA AS Relationship dataset [8], and the CAIDA AS-to-Organizations dataset (as2org) [9] to analyze the ASes registered in the IRR databases.

5 METHODOLOGY

We present our methodology to study the baseline characteristics of route objects in IRR databases and describe the steps we use to identify irregular IRR records.

5.1 IRR Characteristics

We describe the following three metrics we use to characterize the data quality in the IRR database.

5.1.1 Inter-IRR Consistency. 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.

5.1.2 RPKI Consistency. We employ the methodology used by Du et al. [12] to update the RPKI consistency of route objects in the 17 IRR databases that are still active as of May 2023.

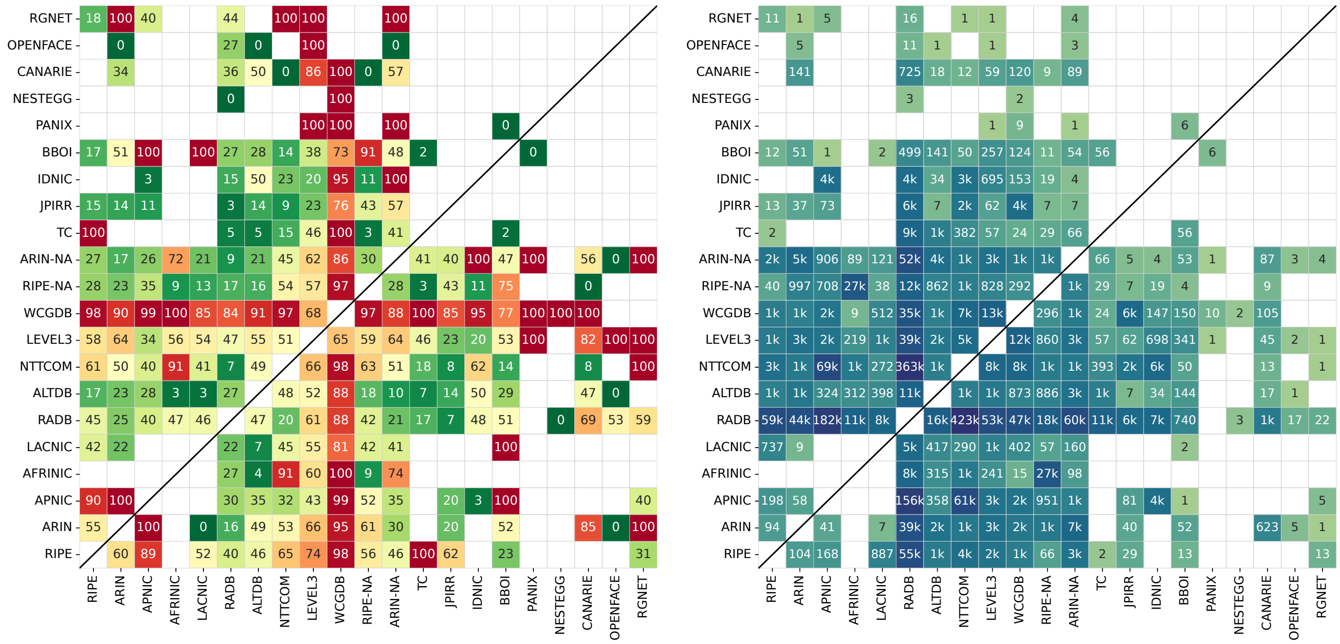
5.1.3 BGP Overlap. For each IRR database, we count the number of route objects with the exact same prefix and origin AS in BGP between November 2021 and May 2023 and calculate its percentage over all route objects.

5.2 Identifying Irregular Route Objects

We outline the steps to identify route objects that may be registered for malicious purposes.

5.2.1 Mismatching origin AS with authoritative IRRs. We consider the information in authoritative IRR databases to be more trustworthy than other IRRs (§ 2.1), 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 § 5.1.1 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 § 5.1.1 step (1) to the following: P_i^B is a covering prefix of P^A.

5.2.2 Matching IRR objects to BGP. An adversary forging IRR records to circumvent filtering [31] would announce in BGP the prefix from their falsified route object to achieve their goal (e.g. route hijacking, spam, phishing, etc.). Therefore, for the route objects we classified as **inconsistent** above, we check whether their prefixes appeared in BGP during the same time period. We classify the **inconsistent** route objects into three categories:



(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. 1b.**

(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 1: Inconsistency between IRR databases (§5.1). 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.

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

5.2.3 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.

6 IRR BASELINE CHARACTERISTICS

To understand the baseline characteristics of the IRR databases, we analyze the inter-IRR consistency, examine trends in IRR consistency with RPKI since 2021 [12], and calculate the overlap between BGP and IRR databases.

6.1 Consistency across IRR Databases

Figure 1a shows the percentage of route objects with the same prefix but different origin ASes between pairs of IRRs. We found 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 [12]. We also noticed 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 were the mismatching records between pairs of authoritative IRR databases, since each RIR only allows registration of route objects containing address blocks managed by that RIR, which do not overlap with each other. We speculate 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.

6.2 IRR Consistency with RPKI

We found 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 shows the percentage of route objects that were RPKI consistent (green) and RPKI inconsistent (red) in November 2021 and May 2023. Since we were able to compare more route objects to RPKI in 2023, we discovered 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 found that 4 IRR databases (LACNIC, BBOI, TC, NTTCOM) were 100% consistent with RPKI, likely due to a policy

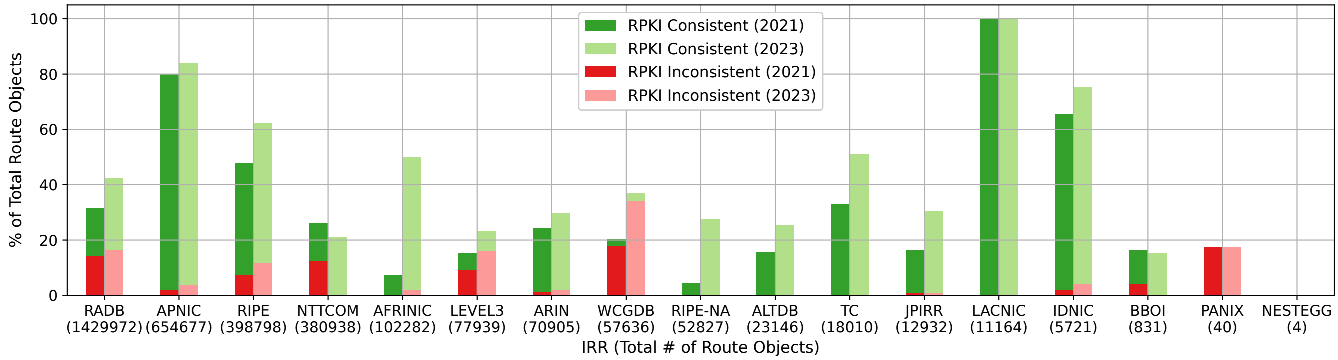


Figure 2: In May 2023, 13 of 17 IRR databases had more RPKI-consistent route objects than RPKI-inconsistent ones.

to reject route objects that are RPKI inconsistent [29]. Conversely, we found no RPKI-consistent records in PANIX and NESTEGG, and we recommend operators not to use them in route filtering.

6.3 Overlap with BGP Announcements

We calculated 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. [20] 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% in 2013 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 studied the inconsistencies between authoritative IRRs and BGP (§6.3) and found 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, 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 [23]). Overall it is a challenge to identify outdated records in authoritative IRRs as there is limited ground truth.

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)

Table 2: IRR overlap with BGP. The middle column shows the number of route objects present between November 2021 and May 2023. The right shows the percentage of route objects that had the same prefix and origin AS in BGP over the same 1.5-year period.

7 IRREGULAR ROUTE OBJECTS

We identify the irregular route objects in RADB and ALTDB, where falsified records have been reported (§2.2).

7.1 RADB Analysis

Filtering irregular objects We applied our workflow to filter irregular route objects in RADB. We started with 1,218,946 unique prefixes from route objects in RADB between November 2021 and May 2023. We found 196,664 unique prefixes that have a mismatching origin AS with any of the five authoritative IRRs (RIPE, ARIN, APNIC, AFRINIC, LACNIC). Of these prefixes, we found the mismatching ASes for 46,262 prefixes had a sibling or customer-provider relationship. We removed those prefixes, leaving 150,402 prefixes that were inconsistent with IRR (Table 3 line 2).

Comparing those inconsistent prefixes to prefix origins announced in BGP (§5.2.2), we found 3,382 prefixes *fully overlap*, 23,353 *partially overlap*, and 32,289 prefixes have *no overlap* (Table 3 line 3-5). We focused 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 [42]. We were able to match 23,353 *partial overlapped* prefixes to 34,199 prefix origins in BGP announcements (Table 3 line 5). We found 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 classified those 34,199 prefix origins as *irregular* and further analyzed them.

Validation We first validated the *irregular* route objects using automated steps (§5.2.3). We performed Route Origin Validation (ROV) [26] on the *irregular* route object using the *RPKI dataset.*

		20.4% (249,725/1,218,946)	39.8% (99,323/249,725) consistent
	1,218,946	Appear in Auth IRR	60.2% (150,402/249,725) **inconsistent**
RADB	Total	39.2% (59,024/150,402)	54.7% (32,289/59,024) no overlap
	Prefixes	Appear in BGP	5.7% (3,382/59,024) full overlap
		and inconsistent	39.6% (23,353/59,024) **partial overlap** → 34,199 *irregular* route objects

Table 3: 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 (§5.2.1). An inconsistent RADB prefix partially overlaps with BGP if there is a common origin AS between RADB and BGP (§5.2.2). 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.*

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

We then compared 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 found that the mismatching RPKI records were created recently, possibly due to the network adopting RPKI and properly creating all of the records.

Of the 13,676 *irregular* route objects that were RPKI-inconsistent/unknown, we removed the ones whose AS appear 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 compared our list of 34,199 (Table 3) route objects with the list of serial hijackers from Testart et al. [42] and found 5,581 route objects registered by 168 serial hijacker ASes. We found one of those ASes (AS35916) to be a small US-based ISP with 10 customers according to CAIDA's AS Rank [7]. 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 [1]. 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 found 30.4% (10,408 / 34,199) irregular route objects registered by ipxo.com, an IP leasing company. They present a challenge to automating inference of irregular objects. They registered 738 ASes under different maintainers in RADB, none of which had a sibling or customer-provider/peering relationship [9]. Those ASes displayed sporadic BGP activity, with announcement duration spanning from 10 minutes to more than 500 days. Prehn et al. [33] 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.

7.2 ALTDB Analysis

Similar to RADB, we applied our workflow to ALTDB and identified 1,206 unique prefixes that were inconsistent with the authoritative IRR databases. Of these, we found 918 *fully overlapped* with BGP, 5 *partially overlapped*, and 12 had no overlap. We mapped the 5 *partially overlapped* 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 inspected those 11 prefixes and found 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 identified a benign case where the route object was registered by Akamai who could have originated the prefix on their business customer's behalf.

8 DISCUSSION AND FUTURE WORK

The IRR is a decade-old routing policy sharing platform that many networks still use for security purposes. The lack of a standardized validation mechanism and coordination across IRR providers allows outdated and misconfigured records to persist, giving adversaries the opportunity to forge IRR records for malicious purposes. Increasing incentives to abuse the Internet routing system (e.g. stealing cryptocurrency) has motivated increasingly complicated BGP hijacking techniques [25], and we should expect attackers trying to exploit the IRR in future attacks. We provided a first look at inconsistencies across IRR databases and proposed an approach to infer irregular activities in the IRR without external sources of ground truth. We found IRR databases prone to staleness and errors, confirming the importance of operators transitioning to RPKI-based filtering. In addition, we found inconsistencies between IRR databases, suggesting opportunities for improved coordination across IRR providers to improve routing security. Finally, we described the challenges of inferring the suspiciousness of such irregular objects and compiled a list of 6,373 suspicious route objects. We hope this work inspires new directions in automating the detection of abuse of IRRs, such as a multilateral comparison across IRR databases, ideally in time to prevent or thwart an attacker's ultimate objective.

ACKNOWLEDGMENTS

The authors thank Liz Izhikevich, Alisha Ukani, and our shepherd, Ralph Holz, for providing insightful comments on various versions of this work. This work is based on research sponsored by U.S. NSF grants OAC-2131987 and CNS-2120399. The views and conclusions are those of the authors and do not necessarily represent endorsements, either expressed or implied, of NSF.

REFERENCES

[1] 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

[2] Cengiz Alaettinoglu, Curtis Villamizar, Elise Gerich, David Kessens, David Meyer, Tony Bates, Daniel Karrenberg, and Marten Terpstra. 1999. *Routing Policy Specification Language (RPSL)*. RFC 2622.

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

[4] Arelion. 2021. Routing Security BGP Prefix filter updates. https://www.arelion.com/our-network/bgp-routing/routing-security

[5] ARIN. 2023. ARIN IRR-online User Guide. https://www.arin.net/resources/manage/irr/userguide/

[6] Tony Bates, Elise Gerich, Laurent Joncheray, J-M. Jouanigot, Daniel Karrenberg, Marten Terpstra, and Jessica Yu. 1995. *Representation of IP Routing Policies in a Routing Registry (ripe-81++)*. RFC 1786. RFC Editor.

[7] CAIDA. 2022. AS Rank. https://asrank.caida.org

[8] CAIDA. 2022. AS Relationships. https://www.caida.org/catalog/datasets/as-relationships

[9] CAIDA. 2022. Inferred AS to Organization Mapping Dataset. https://www.caida.org/catalog/datasets/as-organizations

[10] 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.

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

[12] 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.

[13] 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

[14] 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

[15] Yossi Gilad, Avichai Cohen, Amir Herzberg, Michael Schapira, and Haya Shulman. 2017. Are we there yet? on rpki's deployment and security. In *Network and Distributed System Security Symposium*.

[16] 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/

[17] 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/

[18] Google. 2022. Peering with Google. https://peering.google.com/#/options/peering

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

[20] 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

[21] Brenden Kuerbis and Milton Mueller. 2017. Internet routing registries, data governance, and security. *Journal of Cyber Policy* 2, 1 (2017), 64–81.

[22] Lumen Technologies 2021. *Lumen Technologies guide to the LEVEL3 Internet Routing Registry*. Lumen Technologies.

[23] MANRS. 2023. Mutually Agreed Norms for Routing Security. https://www.manrs.org

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

[25] Alexandros Milolidakis, Tobias Bühler, Kunyu Wang, Marco Chiesa, Laurent Vanbever, and Stefano Vissicchio. 2023. On the Effectiveness of BGP Hijackers That Evade Public Route Collectors. *IEEE Access* 11 (2023), 31092–31124. https://doi.org/10.1109/ACCESS.2023.3261128

[26] 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

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

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

[29] NTT. 2023. RPKI suppression of conflicting IRR information. https://www.gin.ntt.net/support-center/policies-procedures/routing-registry/

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

[31] 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

[32] Chiara Orsini, Alistair King, Danilo Giordano, Vasileios Giotsas, and Alberto Dainotti. 2016. BGPStream: A Software Framework for Live and Historical BGP Data Analysis. In *Proceedings of the 2016 Internet Measurement Conference* (Santa Monica, California, USA) *(IMC '16)*. Association for Computing Machinery, New York, NY, USA, 429–444. https://doi.org/10.1145/2987443.2987482

[33] 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

[34] 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

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

[36] RIPE NCC. 2022. RPKI Dataset. https://ftp.ripe.net/ripe/rpki

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

[38] 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

[39] 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

[40] 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

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

[42] 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.

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

[44] 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.

A ETHICAL CONSIDERATIONS

All IRR datasets used in this paper are publicly available on https://www.irr.net/docs/list.html. All CAIDA datasets are available on https://catalog.caida.org/. Our analysis is available on https://github.com/CAIDA/IRR-IRRegularities-Analysis. This work does not raise any ethical issues.