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:

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

5 METHODOLOGY

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

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 $R_i$ in $IRR_A$ consisting of prefix $p_i$ and origin $AS_i$, we classify the route object in the following steps:

1. Find in $IRR_B$ a list of route objects $R_i^B$ with prefixes $p_i^B$ such that every prefix $p_i^B$ is the same as $p_i$.
2. If there does not exist such a list $R_i^B$, then we classify $R_i$ as no overlap.
3. (Continued).

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.

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 inconsistent if any AS $AS_i$ corresponding to $R_i$ is inconsistent with respect to $IRR_B$.

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 period. We classify the inconsistent route objects into three categories:

![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 RGENET databases were retired before May 2023.](image-url)
Figure 1: Inconsistency between IRR databases (§5.1). For example, when comparing RIPE IRR to ARIN IRR, 104 route objects from the steps above have a matching RPKI record in our RPKI dataset, 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
We identify the irregular route objects in RADB and ALTDB, where 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 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 1,597 (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.

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, 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.

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.

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.
We found that of the 34,199 irregular route objects, 20,523 are consistent, 4,082 have a 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.

**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.