Replication: A Two Decade Review of Policy Atoms - Tracing the Evolution of AS Path Sharing Prefixes

Weili Wu
weili.wu@gatech.edu
Georgia Institute of Technology
Atlanta, USA

Zachary S. Bischof
z@chary.io
Georgia Institute of Technology
Atlanta, USA

Cecilia Testart
ctestart@gatech.edu
Georgia Institute of Technology
Atlanta, USA

Alberto Dainotti
dainotti@gatech.edu
Georgia Institute of Technology
Atlanta, USA

ABSTRACT

Afek et al. characterized the formation and stability of *policy atoms*, groups of prefixes that share the same Autonomous System (AS) paths as observed by BGP collectors, a concept initially defined by Broido and Claffy in 2001. Policy atoms provide a valuable perspective on the inter-domain routing policies in the Internet. With the rapid growth and increasing complexity of the Internet since these studies, we believe it is important to reassess the implications and applicability of policy atoms.

In this paper, we revisit the policy atom concept after two decades and replicate the study performed by Afek et al. to assess the current state of AS path sharing and shed light on the evolution of policy atoms. We demonstrate that the Internet still operates on the level of policy atoms rather than individual ASes, as prefixes within the same atom tend to experience changes in AS path simultaneously. We apply the concept of policy atoms in IPv6 and find that this observation also holds true for IPv6 prefixes. We also relate trends in the characteristics of policy atoms with the development of inter-domain routing policies. We highlight new insights generated by the perspective of policy atoms and their potential for further applications. Our code is publicly available to support reproducibility and to encourage future research on this topic.

CCS CONCEPTS

• **Networks** → **Public Internet**; **Network dynamics**.

KEYWORDS

BGP, Inter-domain Routing, Routing Policy

ACM Reference Format:
Weili Wu, Zachary S. Bischof, Cecilia Testart, and Alberto Dainotti. 2025. Replication: A Two Decade Review of Policy Atoms - Tracing the Evolution of AS Path Sharing Prefixes. In *Proceedings of the 2025 ACM Internet Measurement Conference (IMC '25), October 28–31, 2025, Madison, WI, USA.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3730567.3764432

IMC '25, October 28–31, 2025, Madison, WI, USA

ACM ISBN 979-8-4007-1860-1/2025/10
https://doi.org/10.1145/3730567.3764432

1 INTRODUCTION

The Internet consists of tens of thousands of Autonomous Systems (ASes), each broadcasting its unique set of prefixes to its neighbors. ASes communicate through the Border Gateway Protocol (BGP) to exchange routing messages, which include route information for an AS to reach a prefix. Such routes are sequences of AS hops, known as AS paths. Often, multiple prefixes are advertised together, leading to shared AS paths.

In 2001, Broido and Claffy introduced the concept of "policy atoms" [**?** **?**], a group of prefixes that, given a set of routers with global routing tables (vantage points), share the same AS path within each of these routers. In other words, a policy atom manifests when a (sub)set of prefixes originated by the same AS experience the same hop-by-hop BGP path selection, likely due to identical policies applied by its origin AS and all the ASes along the paths to each vantage point. In 2002, Afek et al. further characterized and validated policy atoms [**?**] and studied how this concept could theoretically be leveraged to reduce the number of updates sent in the routing system.

Since the introduction of policy atoms in the early 2000s, Internet infrastructure and routing practices have undergone dramatic changes, such as the flattening of the Internet hierarchy, the rise of private and cloud network interconnections, the adoption of Resource Public Key Infrastructure (RPKI), and the development of new traffic engineering techniques. Each of these changes has the potential to impact the status and significance of policy atoms.

Recent studies suggest the concept of policy atoms remains useful. For example, Darwich et al. recently used BGP policy atoms as the core building block of an algorithm that identifies traffic engineering events [**?**]. This usage of policy atoms underscores the need for revisiting policy atoms with an up-to-date characterization to understand the applicability of the policy and for guiding better and appropriate usage of policy atoms in today's Internet.

In this paper, we assess the current status of policy atoms to bridge this gap in understanding policy atoms and exploring their potential applicability by replicating the study of Afek et al. Using historical BGP data from RIPE RIS [**?**] and RouteViews [**?**], spanning from 2004 to 2024, we apply the methodology used by Afek et al. [**?**] to study the characteristics and composition of policy atoms, and identify correlations between the changes and trends in Internet infrastructure. We make our code publicly available to facilitate further studies of policy atoms [**?**].

The contributions of our replication study include:

- An analysis of publicly available BGP data that characterizes policy atoms and how they relate to routing decisions currently made by ASes in the Internet.
- A study of the evolution of policy atoms across more than two decades that identifies and correlates changes in atom behavior with changes in inter-domain routing.
- An analysis of the applicability of policy atoms in IPv6 and how they compare to policy atoms in IPv4 routing.
- We make our code and revised methodology publicly available to encourage further replication and inform measurement target selection.

The rest of the paper is organized as follows. Section 2 defines policy atoms and describes data sanitization methods that suit the current Internet. Section 3 details the methodology we inferred to conduct the analysis and shows that it reproduces the results in the original paper. Section 4 presents the analysis on the longitudinal data, combining our data sanitization and validated methodology, and Section 5 extends the analysis to IPv6.

2 REVISITING POLICY ATOMS

In this section, we first present an overview of policy atoms and their relationship with routing policies on the Internet. We then describe the details of our methodology, including the workflow and data sanitization techniques. Where relevant, we compare our choices to those in prior papers.

2.1 Concept of policy atom

In the prior work by Broido and Claffy [**?** **?**], the concept of policy atoms was introduced to capture routing policies on the Internet. Elaborated by Afek et al. [**?**], a *policy atom* is defined as a group of prefixes $\{P\}$ such that for any pair of prefixes $P_i, P_j \in \{P\}$ and for any router that holds a full BGP table on the Internet, the AS path for P_i, equals the AS path for P_j. In other words, policy atoms are groups of prefixes that share the same AS paths within each global vantage point.

2.2 Policy atoms and routing policy

BGP [**?**], the main protocol used for inter-domain routing in the Internet, is a policy-based protocol where ASes can independently consider factors impacting their border routers' decision making process, and determine what to propagate to their neighbors. This process is largely influenced by the operator's import and export routing policies, which are often driven by business relationships between ASes and their performance, security, and traffic engineering goals. Due to these factors, ASes might not have a uniform routing policy across all the prefixes they originate [**?** **?**].

Policy atoms emerge as the result of the interplay among routing policies of the AS originating a given set of prefixes and the ASes transiting those prefixes towards the rest of the Internet. AS policies are implemented through router-level import and export filters that directly influence the formation of policy atoms. For example, a multi-homed AS that announces a different set of prefixes to each of its upstream providers will result in the prefixes originated by that AS to be in different policy atoms.

The relevance of policy atoms is connected to their ability to capture routing policies. The number of policy atoms is related to the number of distinct routing policies, and hence can be used as a proxy for routing policy complexity. For example, a higher number of atoms for prefixes originated by the same AS indicates a larger set of policies and more fine-grained traffic control applied to the prefixes.

2.3 Methodology from the original paper

To compute policy atoms, Afek et al. [**?**] collected BGP routing table snapshots from multiple vantage points and processed the data to extract the AS paths for each prefix. Then, they grouped prefixes based on their AS path across all vantage points, forming the policy atoms. When some prefixes are missing in the tables of some vantage points, they consider those prefixes to have an "empty" path. As a consequence, prefixes that do not appear in the routing table of a given vantage point will not be part of policy atoms with other prefixes with which they share the AS path on all other routers.

The original paper methodology provides the foundation for understanding the structure and dynamics of Internet routing policies, but is not well-equipped for the current state of BGP data collection. Since the publication of the policy atom papers [**?** **?** **?**], the global routing table has 10 times more prefixes and the number of BGP collector peers has increased from less than a dozen to over a thousand different ASes. However, a significant share of prefixes are only visible by one or two BGP collector peers and many peers only share a partial routing table with BGP collectors. These two artifacts of the BGP collector infrastructure or of very localized route advertisements that are not intended for global routing, impact the inference of policy atoms and their stability. Thus, building on top of the original methodology, we add data sanitization steps to compute policy atoms for prefixes intended for global routing, and we use BGP collector peers that provide their full routing tables as vantage points.

2.4 Methodology to infer global policy atoms

This section outlines our updated approach for computing policy atoms using BGP data of the last 20+ years.

2.4.1 Data snapshots. To study more than 20 years of policy atoms, we compute the atoms once per quarter using 4 snapshots to enable the study of short-, mid-, and long-term stability of the computed policy atoms. More specifically, using BGPStream [**?**], we download the RIBs of all RIPE RIS [**?**] and RouteViews [**?**] BGP collectors on the 15th at 8 am, 15th at 4 pm, 16th at 8 am, and 22nd at 8 am of January, April, July, and October, from 2004 to 2024. In addition, we also analyze the updates for 4 hours after the first snapshot of every quarter.

2.4.2 Vantage points selection. As RIPE RIS [**?**] and RouteViews [**?**] increased the number of BGP peers providing routing data to their BGP collector infrastructures over the past decades, they started having peers that sent only a reduced set of prefixes and not the full routing table. To compute policy atoms, we are interested in grouping together all prefixes that share the same paths from a global perspective. Thus, we use BGP data from peers that share

their full routing tables, which we call *full-feed peers*. Unfortunately, BGP collector infrastructures do not explicitly track which of their peers share a full routing table. Therefore, we infer if a peer is sending its full routing table to BGP collectors based on the number of prefixes it shares data for, compared to the maximum count of prefixes any peer shared data for. We consider a peer as a full-feed peer if it shares BGP data for more than 90% of the maximum count of unique prefixes a peer shares with BGP collectors in a snapshot.

By excluding the data from *partial-feed peers*, we may be missing route announcements that are unique to those peers. As a result, we may be excluding from our analysis prefixes that are only seen by those peers. In addition, we may be bundling together in policy atoms prefixes that some partial feeders have different paths to them. However, if these prefixes and paths do not show up in the routing data from the hundreds of full-feed peers, they need to be very localized. As our main focus is the study of routing policies impacting the routing of prefixes intended to have global reachability, using the data of full-feed peers provides enough coverage.

2.4.3 Prefix filtering. To compute policy atoms, prior papers have proposed different approaches in terms of which prefixes to consider or filter:

- **Broido and Claffy** [**?** **?**]: In this initial paper, the methodology considers only prefixes in *all* routing tables shared with BGP collectors.
- **Afek et al.** [**?**]: In this consecutive paper, the methodology considers all prefixes, from *any* routing table shared to BGP collectors.

The two divergent approaches taken by the previous papers are not well-suited for our longitudinal analysis. First, many prefixes with high visibility are still not present in all full-feed peers' routing data. For instance, a network may aggregate prefixes or have only received an aggregated prefix for traffic engineering purposes. Second, considering prefixes in any full-feed peer routing table includes many prefixes that are only in a few peers' routing tables due to very localized routing, misconfigurations, or BGP collector platform artifacts (*e.g.,* a stuck route).

To capture as many distinct routes as possible while limiting misconfiguration and other artifacts in the data, we filter out from BGP data snapshots prefixes that *(i)* are not seen in at least two route collectors and *(ii)* are not in routing tables from at least four ASes. The first condition removes artifacts in the data potentially caused by a misconfigured route collector.[1] The second condition filters out very localized prefixes, where less than a handful of ASNs see such announcements. In Appendix A8.5 we share the details of the sensitivity analysis of the thresholds.

We note that we only consider IPv4 (IPv6) prefixes with prefix lengths less than or equal to /24 (/48). We also note that we do not filter out Multi-Origin AS (MOAS) prefixes. Afek et al. [**?**] do not remove MOAS during policy atom computation. Nonetheless, they do not consider atoms with MOAS conflict during one of their analysis. More generally, the authors state that MOAS prefixes are less than 5% of total prefixes. We identified MOAS prefixes and verified that between 2004 and 2024, the percentage of MOAS prefixes present in BGP data snapshots is consistently below 5%. Therefore, we do not remove MOAS prefixes during policy atom computation. We highlight that MOAS prefixes do not impact the computation of policy atoms. Indeed, MOAS and non-MOAS prefixes would not be in the same atom because to be in the same atom, prefixes need to have the same AS path to vantage points and thus the same origin AS.

2.4.4 Additional data cleaning steps. To prevent other data artifacts from impacting policy atom computation, we identify AS sets in AS paths and BGP peers with ADD-PATH feature to remove that data when needed.

The BGP AS-SET attribute is used when aggregating paths. The aggregating router places all the ASes removed from the path into a set, resulting in an AS-PATH that may look like "1 2 [3 4 5]". This aggregation removes all the information about the AS path after the aggregation point. Afek et al. [**?**] preserve AS-SET in their work because they have no information regarding the aggregated path. We expand the AS-SET only if it contains only one element, and remove other cases in our study. In our analysis, the percentage of paths containing AS-SETs is less than 1%.

The ADD-PATH feature allows routers to advertise multiple paths for the same prefix. However, the lack of uniform support for this feature on all routers and BGP collectors can lead to compatibility issues, where routers that support ADD-PATH may not properly communicate with collectors or other routers that do not. Similarly, collectors that do not support ADD-PATH may not correctly interpret or handle updates from routers that utilize this feature. This issue particularly affects certain peers connected to RouteViews collectors. To identify problematic peers or collectors, we check for specific patterns indicative of ADD-PATH parsing errors. We removed peers from 4 ASNs to maintain the correctness of the data, as they connected to RouteViews collectors that are not compatible with the ADD-PATH feature. We detail the patterns and ASNs we removed in Appendix A8.3. Additionally, we identified and removed one peer that appeared to be misconfigured, as it has a private ASN (AS65000) in the AS path of numerous prefixes. We documented the details of this case in Appendix A8.3.

Lastly, we removed peers that share excessive duplicate prefixes, defined as cases where more than 10% of the prefixes shared with BGP collectors were identical.

Using the data described in § 2.4.1, selecting AS peer vantage points as reported in § 2.4.2, filtering prefixes according to § 2.4.3 and following additional steps described in § 2.4.4, we then apply the initial paper's policy atom definition (see Section 2.1) to compute the policy atoms.

2.5 Analysis and Metrics

To demonstrate the continued relevance and applicability of policy atoms as analytical tools in the modern Internet, there are four analyses performed by Afek et al. [**?**]: *(i)* General statistics for ASes and atoms, *(ii)* Correlation of atom structure to Internet update records, *(iii)* Formation of policy atoms, and *(iv)* Stability of policy atoms. Below, we summarize the four replicated analyses and their significance to our research goals:

- **General statistics for ASes and atoms:** We quantify basic properties of policy atoms, such as atom count, size distribution,

[1] As an example, various RouteViews collectors have at different times generated spurious records due to their inability to process announcements with the `ADD_PATH` BGP extension.

and the ratio of atoms to ASes and prefixes. Understanding these fundamental metrics helps to assess how the complexity and granularity of routing policy have evolved over the last two decades.
- **Correlation of atom structure to Internet update records:** We examine if prefixes within the same atom appear together in the same update record. This analysis helps us to examine whether Internet routing still operates at the level of policy atoms.
- **Formation of policy atoms:** We measure the point in an AS path where policy atoms are formed. This analysis helps us understand whether policy atoms are created due to the routing policies of the origin AS.
- **Stability of policy atoms:** We compute the percentage of policy atoms that remained unchanged and the percentage of prefixes that remained grouped over time. This analysis validates that policy atoms are applicable with a stable structure.

3 REPRODUCING POLICY ATOM ANALYSIS

Though the original paper published by Afek et al. [**?**] provides high-level descriptions of the methodology they used for determining policy atoms, the paper does not mention making their code publicly available. Thus, we attempt to replicate their methodology to the best of our abilities. This includes making assumptions or inferences regarding details not directly specified in the paper, such as which route collector and/or VPs they chose to include in the study, which (if any) prefixes were filtered and removed from the analysis, and the approximate time frame of the data they used in their analysis. In this section, we describe our methodology to replicate this study and explain our assumptions in more detail.

We note that by applying our inferred methodology to BGP data from a comparable time period, we were able to obtain results that were largely in line with the findings published in the original study. However, there was one notable way in which our initial results diverged: the characterization of the formation distance of policy atoms. To address this, we describe multiple approaches for calculating formation distance, including both our interpretation of the methodology used in [**?**] and the methodology that produced results similar to those shown in [**?**] and then discuss the method that we think is most appropriate in the context of policy atoms.

3.1 Reconstructing the dataset

To ensure that our replicated methodology for creating policy atoms was consistent with the approach described in the original paper [**?**], we wanted to apply our inferred methodology to the same dataset and compare results. However, because the original paper [**?**] does not provide precise details about the input data, we attempted to reverse-engineer key aspects of their setup, such as which route collector peers were used, whether any prefixes were filtered, and the time period of data used in the analysis.

3.1.1 Time window. Because the Internet Measurement Workshop submission deadline was in May 2002, the dataset used in the original paper must be collected before that month. As the authors noted they used data at the time of writing, we thus limited potential snapshot candidates to the first five months of 2002. We then selected snapshots that resulted in a similar number of unique prefixes and atoms as described in Section 3 of [**?**]. Based on these numbers, we selected 2002-01-15 at 8 am UTC as it resulted in similar values.

3.1.2 Collectors and peers. The original paper refers to using 13 peers in Section 2 [**?**]. Based on our investigation, in 2002, RIPE RIS collector RRC00 was the only collector that operated in a global scope, while the remaining RIS collectors (RRC01-RRC08) were connected to specific IXPs and had only local peering scopes. During this time period, RRC00 had exactly 13 full-feed BGP peers, aligning with the details provided in the original paper.

3.1.3 Prefix filtering. Afek et al. [**?**] did not mention whether or not any prefix length filtering was applied. We therefore tested two common options: *(i)* accepting all prefixes and *(ii)* excluding prefixes more specific than /24. The unrestricted option yields roughly 115K distinct prefixes, which is in line with the number of prefixes mentioned in the original paper. In contrast, filtering removed approximately 5K prefixes, resulting in approximately 110K prefixes. Therefore, we elected to include all prefixes in our reproduced dataset.

3.1.4 Final input dataset. Based on the findings and assumptions described above, we elected to use a BGP snapshot collected on 2002-01-15 at 8 am UTC, from the 13 full-feed peers connected to RIS collector RRC00, with no prefix-length filtering as our input dataset. We then applied our inferred methodology for constructing policy atoms to this dataset and compared our summary statistics to those presented in the original paper, which we discuss in the rest of this section.

3.2 General Statistics for ASes and Atoms

Using the aforementioned assumptions, we were able to confirm that the general statistics of policy atoms shown in Appendix A8.4.1 match what was reported in the original paper. We found that the number of ASes, prefixes and atoms is similar to the original paper, with 12.5K ASes, 115K prefixes, and 26K atoms. This confirms the correctness of inferred assumptions.

3.3 Correlation of Atom Structure to Internet Update Records

We analyzed BGP update records captured by RRC00 during the four hours following the snapshot at 2002-01-15 8 am UTC.

3.3.1 Methodology. We compute the likelihood of prefixes within an AS or an atom being seen in full within a single BGP update. Formally speaking, let $\mathcal{A}_t = \{a_1, \ldots, a_n\}$ be the set of atoms computed from one BGP snapshot, and let a be an atom with prefix set $Prefix(a)$ of size k. Note that an AS with k prefixes is processed in the same way. For every update record r, let $Prefix(r)$ be the set of prefixes inside the update record. Then for every atom a and every update record r, there are three possible cases:

(1) None of the prefixes within atom a is in update record r: $Prefix(a) \cap Prefix(r) = \emptyset$
(2) All prefixes within atom a are in update record r: $Prefix(a) \subseteq Prefix(r)$
(3) A partial subset of the prefixes within atom a are in update record r: $Prefix(a) \cap Prefix(r) \neq \emptyset$ and $Prefix(a) \nsubseteq Prefix(r)$.

Let $N_{\text{all}}(a)$ be the number of BGP updates in which all of the k prefixes of a appear (case 2), and $N_{\text{partial}}(a)$ be the number of BGP updates in which at least one but not all of the k prefixes of a appear (case 3).

Then for atom or AS with k prefixes, the likelihood of prefixes within an AS or an atom being seen in full within a single BGP update is computed as

$$\Pr_{\text{full}}(k) = \frac{\sum_{a:\,|P(a)|=k} N_{\text{all}}(a)}{\sum_{a:\,|P(a)|=k} \left(N_{\text{all}}(a) + N_{\text{partial}}(a)\right)}.$$

i.e., the percentage of times that an atom or AS with k prefixes, which has at least one of its prefixes in an update message, appears with all k of its prefixes included in that same update.

We confirmed our methodology can yield similar results as the original paper, shown in Appendix A8.4.2

3.4 Formation of Policy Atoms

3.4.1 Concept and definition. Following the definition proposed by Afek et al. [**?**], we first define the splitting point between two atoms as the first AS that differs along the AS path, starting from the origin AS. Stated differently, it is the length of the minimal AS path starting from the origin AS not shared by the two atoms. Let $\mathcal{P}$ be the set of BGP peers in the snapshot. For an origin AS o, let $\mathcal{A}_o = \{a_1, \ldots, a_m\}$ be the atoms whose prefixes originate at o. For each peer $p \in \mathcal{P}$ and atom $a_i \in \mathcal{A}_o$, let

$$\mathbf{Path}_p(a_i) = \langle p, \ldots, v_3, v_2, o \rangle,$$

denote the AS path for atom a_i seen by peer p. If peer p did not see atom a_i, we treat $\mathbf{Path}_p(a_i)$ as the empty string.

For two atoms $a_i, a_j \in \mathcal{A}_o$ the splitting point for peer p is

$$\text{split}_p(a_i, a_j) = \begin{cases} 1 & \text{if } \mathbf{Path}_p(a_i) = \varnothing \text{ or } \mathbf{Path}_p(a_j) = \varnothing, \\ \min\{\, i > 1 \mid v_i \neq w_i \,\} & \text{otherwise.} \end{cases}$$

where $\mathbf{Path}_p(a_j) = \langle p, \ldots, w_3, w_2, o \rangle$.

The overall splitting point between atom a_i and atom a_j is the earliest divergence seen from any peer:

$$\text{split}(a_i, a_j) = \min_{p \in \mathcal{P}} \text{split}_p(a_i, a_j).$$

Thus, a missing path at any peer forces $\text{split}(a_i, a_j) = 1$, matching the description in the original paper.

Then, for an arbitrary origin AS, the formation distance of an atom is the shortest distance from the origin AS that makes the atom distinguishable from any other atoms from the same origin AS. Equivalently, it is the maximum splitting point length between this atom and every other atom from the same origin AS:

$$d(a_i) = \max_{a_k \in \mathcal{A}_o \setminus \{a_i\}} \text{split}(a_i, a_k).$$

From these, we derive two origin AS level metrics. The first split point for origin AS o is the first AS going from the origin AS that is not shared by all atoms, defined as $d_{\min}(o) = \min_{a \in \mathcal{A}_o} d(a)$. In other words, it is the shortest distance for at least one atom to be distinguishable. The last split point is the shortest length of the AS path such that the partial AS path of this length is different in each atom, defined as $d_{\max}(o) = \max_{a \in \mathcal{A}_o} d(a)$. In other words, it is the shortest distance for all atoms to become distinguishable.

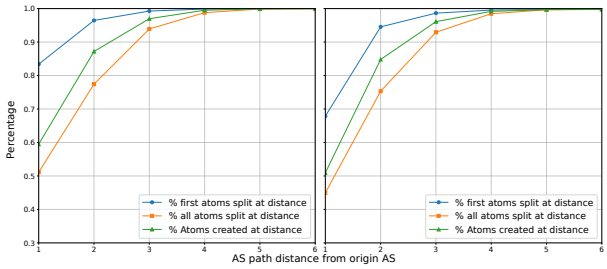

Figure 1: Formation distance of policy atoms computed with method *(iii)* (left) and method *(ii)* (right). We found that method *(iii)* results in more atoms forming at distance 1 compared to the original paper, while method *(ii)* results in a curve similar to the original paper.

3.4.2 Methodology. Afek et al. do not explicitly mention how they remove consecutive duplicates of ASes introduced by AS-path prepending while keeping knowledge of duplication. AS path prepending represents policy imposed by the AS to the atom, yet counting every duplicate AS can inflate formation distances. We considered three plausible methods to handle it. *(i)* Remove prepended ASes before grouping prefixes into atoms, *(ii)* Construct atoms with raw AS path and remove prepended ASes before computing formation distance, *(iii)* Keep full AS paths when computing atoms and ignore duplicate copies of the same AS when counting AS hops when computing formation distance.

We believe the first two options are not appropriate: *(i)* discards the policy information entirely, while *(ii)* will create two atoms that are not distinguishable. For example, if we have two atoms with AS path $\langle AS1, AS2, AS3 \rangle$ and $\langle AS1, AS2, AS2, AS3 \rangle$, after removing duplicates, two atoms with AS path $\langle AS1, AS2, AS3 \rangle$ and $\langle AS1, AS2, AS3 \rangle$, will become indistinguishable.

We therefore adopted method *(iii)* , handling AS duplication during the formation distance analysis to not inflate the formation distance. We compute atoms based on the raw AS path. We then count in terms of unique ASes in the stripped AS path to determine the split point.

3.4.3 Comparison of results. Figure 1 shows the cumulative distribution of formation distances (% atoms formed at distance) produced by method *(iii)* (left) and method *(ii)* (right). Relative to Fig. 3 in the original paper [**?**], our adopted method, shown in the left plot in Figure 1, is about 10 percentage points higher at distance 1. Figure 1 shows that atoms with distance 1 constitute 61% of total atoms. We break down the results and find atoms with distance 1 are constituted of *(i)* there is only one atom from this origin AS (38%), *(ii)* the atom is observed by a unique set of peers compared to other atoms from the same origin AS (13%), *(iii)* as a result of a difference in AS path prepending (10%). There are 10% of atoms that are formed at distance 1 due to AS path prepending, matching the difference between the two figures. We also found that if we adopt method *(ii)* , removing AS duplication in the prepending after grouping based on AS path with prepending, we will get a similar curve as the original paper, shown in the right plot in Figure 1. Based on the reasoning above, changing the AS path after grouping based on the AS path will cause atoms to be indistinguishable in the worst case. We therefore retain our procedure above through the rest of our analysis.

3.5 Stability of Policy Atoms

Let $\mathcal{A}_t$ be the set of atoms computed based on the BGP snapshot taken at time t, and let $Prefix(a)$ denote the set of prefixes within atom $a \in \mathcal{A}_t$. Following Afek et al. [**?**], we quantify the stability of policy atoms using two metrics:

(a) **Complete atom match (CAM)**

$$\mathrm{CAM}(t_1, t_2) = \frac{\left|\{\, a \in \mathcal{A}_{t_1} \mid \exists\, b \in \mathcal{A}_{t_2} : Prefix(a) \equiv Prefix(b) \,\}\right|}{|\mathcal{A}_{t_1}|}.$$

This represents the fraction of atoms at t_1 whose entire prefix set appears unchanged at t_2. An atom is stable only if the exact composition of prefixes is present in the first and second atoms being compared.

(b) **Maximized prefix match (MPM)**

Let $\phi : \mathcal{A}_{t_1} \to \mathcal{A}_{t_2}$ be a one-to-one mapping that maximizes the total prefix overlap using a greedy approach.
The ratio is

$$\mathrm{MPM}(t_1, t_2) = \frac{\sum_{a \in \mathcal{A}_{t_1}} |Prefix(a) \cap Prefix(\phi(a))|}{\sum_{a \in \mathcal{A}_{t_1}} |Prefix(a)|}.$$

This represents the percentage of prefixes that remained in the same atom, even if the atom itself has split or merged with another atom.

To assess the stability of policy atoms over time, we employ the two metrics *(i)* complete atom match and *(ii)* maximized prefix match used in the prior paper [**?**]. We found that our methodology can yield similar results as the original paper, shown in Appendix A8.4.3.

4 CURRENT STATUS OF POLICY ATOMS AND LONGITUDINAL STUDY

We now evaluate the current status of policy atoms by comparing their characteristics from 2004 to 2024. Our goal is to demonstrate the continued relevance and applicability of policy atoms as analytical tools in the modern Internet. Our analysis is based on the methodology validated in Section 3. We adapt filtering criteria for VPs and prefixes to account for modern Internet dynamics, as detailed in Section 2.4.2 and 2.4.3. Specifically, we employ updated selection criteria to mitigate biases introduced by RPKI deployment and to capture a broader spectrum of routing events.

For the rest of this section, we compare the atoms computed using the first and last snapshots that we captured, which are Jan 15 2004 8 am UTC and Oct 15 2024 8 am UTC, respectively.

4.1 General Statistics

To contextualize our subsequent analyses, we begin by comparing fundamental statistics of policy atoms between 2004 and 2024 (Table 1). The data reveals two concurrent observations: *(i)* an increase in the share of smaller atoms and *(ii)* the emergence of larger atoms.

Table 1 compares the general statistics of policy atoms between 2004 and 2024. We observe a 7.8-fold increase in the total number of prefixes, from 131,526 in 2004 to 1,028,444 in 2024. This increase in the number of prefixes is primarily driven by the trend of prefix fragmentation. At the same time, we observe a 14.1-fold increase in the total number of policy atoms, from 34,261 in 2004 to 483,117 in 2024. The difference in the growth rates of total atoms and prefixes can be attributed to an effect of the adoption of more granular routing policies by network operators which will be discussed further on the formation distance of policy atoms in Section 4.3.

Table 1: Compare general stats of atoms in 2004 and 2024.

Year	Jan 2004	Oct 2024
Number of prefixes	131,526	1,028,444
Number of ASes	16,490	76,672
Number of ASes with one atom	9,818 (59.5%)	31,009 (40.4%)
Number of atoms	34,261	483,117
Number of atoms with one prefix	19,772 (57.7%)	355,197 (73.5%)
Mean atom size	3.84	2.13
99th percentile of atom size	40	17
Largest atom size	1,020	3,072

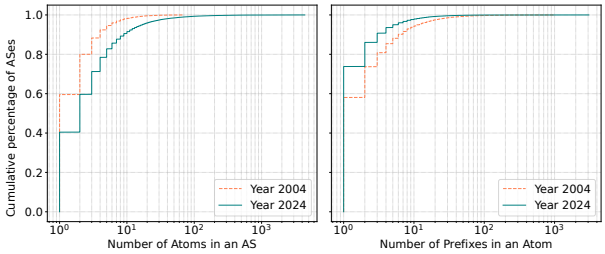

Figure 2: Distribution of the number of policy atoms per AS (left) and the number of prefixes per policy atom (right) in 2004 and 2024. 2024 tends to have more atoms per AS and fewer prefixes per atom, indicating the split of policy atoms.

The number of atoms containing only one prefix has risen from 19,772 (57.7%) in 2004 to 355,197 (73.5%) in 2024, suggesting that there are many more small atoms, caused by more complex routing policies. Consequently, the mean atom size has fallen from 3.84 to 2.13 prefixes per atom, and the 99th percentile of atom size has dropped from 40 to 17 prefixes per atom. Figure 2 illustrates the changes in two characteristics of policy atoms between 2004 and 2024. The CDF for the number of prefixes in an atom (right) is left-skewed, comparing 2024 to 2004, indicating more atoms with fewer prefixes. The CDF for the number of atoms in an AS (left) is right-skewed, comparing 2024 to 2004, indicating ASes tend to have more atoms. This supports the observation from the table that the percentage of atoms with one prefix indicates the splitting of atoms. At the same time, the maximum observed atom size nearly tripled from 1,020 prefixes (2004) to 3,072 prefixes (2024). While atoms have fragmented, with small-sized atoms becoming more common, the emergence of a few exceptionally large atoms highlights the growing complexity of routing policy in the current Internet.

4.2 Correlation with BGP Update Records

To investigate whether policy atoms are indeed driven by routing policies, we examine if prefixes within the same atom appear together in the same update record. The idea is that if prefixes

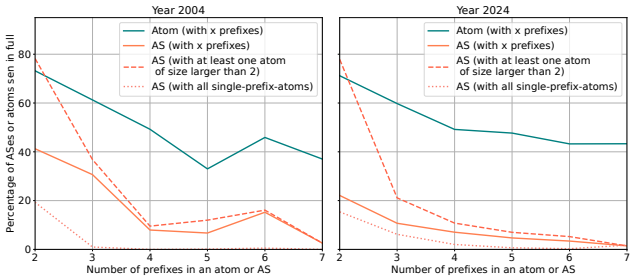

Figure 3: Compare the likelihood of AS and Atom seen in full in one BGP Update in Year 2004 (left) and Year 2024 (right). Atoms are much more likely to be seen in full in one BGP update compared to AS with the same amount of prefixes.

belonging to an atom are frequently updated together, it suggests that they are subject to the same routing policies. As a comparison, we compute the likelihood of prefixes within an AS being updated together.

Figure 3 compares the likelihood of ASes and atoms with a certain number of prefixes being seen in full within a single BGP update in 2004 (left) and 2024 (right). Within these two figures, the solid lines are drawn based on the methodology described in Section 3.3.

We demonstrate results for atoms with fewer than or equal to 7 prefixes for simplicity. This threshold already covers 95% of the total atoms in 2024, as shown in the atom size distribution in Figure 2 (right).

Of all updates that contain at least one prefix of a k-prefix atom, the percentage of those updates that include all k prefixes from that atom is more than 40% in 2024, for atom sizes of 2, 3, 4, 5, and 6, indicating that prefixes within the same atom are frequently updated together. As a comparison, of all updates that contain at least one prefix of a k-prefix AS, the percentage of updates containing all prefixes in an AS is significantly lower by around 30% in 2024. The consistently higher percentage for atoms compared to ASes (solid orange line above the solid teal line) suggests that prefixes move at the atom level rather than the AS level or the prefix level.

To further emphasize the point that the Internet operates at the level of atoms, we expand the analysis and categorize ASes into the following two categories to remove the co-effect of atoms inside the AS curve in the first category.

(1) AS with only single-prefix-atoms
(2) The other ASes (AS with at least one atom of size more than 1)

For ASes with only single-prefix atoms (coral dotted line), we observe nearly zero percent of update records that have at least one prefix of the AS, containing all prefixes in these ASes. The observation that these ASes are highly unlikely to be seen in full reinforces the conclusion that prefixes move at the atom level rather than the AS level.

In summary, the BGP update records analysis provides strong evidence that policy atoms are indeed governed by routing policies, as prefixes within the same atom are frequently updated together. This observation is emphasized after categorizing ASes, reinforcing the conclusion that the Internet operates at the policy atom level.

4.3 Formation distance

Afek et al. [**?**] leveraged formation distance to investigate whether policy atoms are created due to the routing policies of the origin AS. This section investigates how formation distance has changed from 2004 to 2024, explores the underlying reasons for these changes, and relates these trends to changes in Internet routing policies.

We compare the formation distance of policy atoms between 2004 and 2024 and present the results in Table 2. We observe two key trends from this comparison: *(i)* atoms formed at distance 1 (i.e., at the origin AS) dramatically decreased from 2004 to 2024, *(ii)* atoms are now formed further away from the origin AS. We will explore the causes of these changes.

	2004	2024
Atom formed at dist 1	45%	20%
Atom formed at dist 2	30%	30%
Atom formed at dist 3	17%	33%
Atom formed at dist 4	6%	12%

Table 2: Formation distance distribution in 2004 and 2024

Atoms formed at distance 1 dramatically decreased from 2004 to 2024. Atoms formed at the origin (distance 1) fell from 45% to 20%. There are three reasons for an atom to be formed at the origin AS (has distance 1): *(i)* there is only one atom from this origin AS, *(ii)* the atom is observed by a unique set of peers compared to other atoms from the same origin AS, *(iii)* as a result of a difference in AS path prepending. In 2004, there were 9,818 ASes with 1 atom, either they only advertise one prefix, or they have multiple prefixes, but all prefixes share the same AS path, this contributes 29% of the atoms formed at distance 1. In 2024, there are 31,009 ASes with 1 atom, a similar share among the total number of ASes that advertise prefixes, the percentage of atoms that are the only atom of an AS decreased to 6%. This causes the percentage of atoms formed at distance 1 to decrease dramatically from 2004 to 2024. To this end, we plot formation distance with AS with a single atom excluded in Figure 4, we observe that the dotted line (excluding single atom ASes) for distance 1 is relatively stable over the years.

Atoms are now formed farther from the origin AS. Looking at the trend of formation distance of atoms over the years shown in Figure 4, more atoms are formed at ASes farther away from the origin AS at a quick speed from 2004 through 2008, the rate slowed down from 2008 to 2016, while the percentage of atoms formed at each distance remains stable from 2016 to 2024. Note that we plot the percentage of atoms formed from distance 1 to distance 5, following the original paper, because 99% of atoms are formed at a distance of at most 5 from the origin AS. One reason for atoms to form far away from the origin AS is multiple sibling ASes that belong to the same organization. For example, it takes some atoms from AS1501 (DoD), 6 ASes (all belong to the DoD network) to reach the first split point (Tier 1). Over the period, BGP communities have been widely adopted and potentially allow the intermediate AS to impose routing policy on other networks. Streibelt et al. reported a 200% increase in the number of ASes that use communities as seen in BGP advertisements and a 250% increase in the number of unique communities between 2010 and 2018 [**?**]. For example,

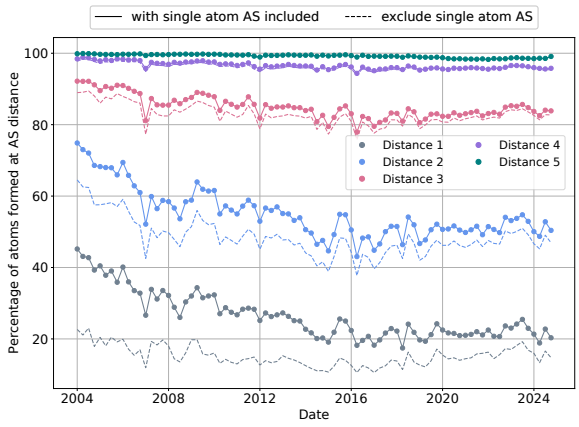

Figure 4: The trend of the percentage of atoms created at different numbers of ASes away from the origin AS over the last 20 years. The average formation distance of atoms over the last 20 years has increased.

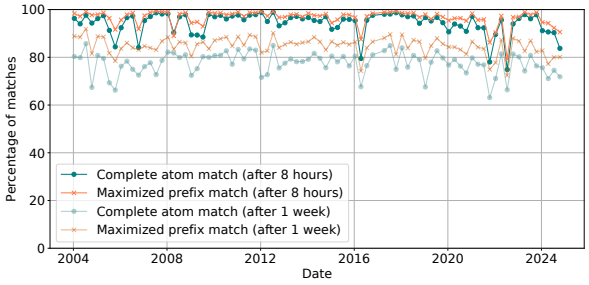

Figure 5: The short-term stability (after 8 hours) and long-term stability (after a week) of atoms in two metrics remain stable and consistently high over the last 20 years.

GTT (AS3257) published in their documents that 3257:2990 means do not announce in North America, and 3257:2592 means prepend 2x only in Asia. [**?**] Orange (AS5511), on the other hand, exposes communities that tune announcements to US peers and specific large neighbors [**?**]. However, as community values can either be used to influence routing decisions made by other networks or to annotate their networks' routing policies [**?**], it is difficult to determine the percentage of atom breaks associated with the use of action communities. We left the analysis of BGP community value and atoms for future work.

Section 4.1 showed that the mean atom size has decreased from 3.84 to 2.13 prefixes per atom, suggesting a more complex set of routing policies applied to prefixes, causing atoms to split. What is highlighted by the formation distance of policy atoms is that those atoms are not split due to the routing policy of the origin AS, but rather by the intermediate ASes in the path. This suggests that not only the origin AS is responsible for the formation of policy atoms, but also the intermediate ASes in the path impose policy decisions that lead to the formation of policy atoms.

Elaborating on the definition of formation distance of atoms described in Section 3.4, the formation distance of an atom depends on the first AS whose policy differs for this atom compared to the most similar atom. Thus, if a transit applies selective export, either as requested by the prior AS through community values or due to its own policy, it might create a split of atoms at the AS after the transit. The formation distance for the split atoms will be at least the number of ASes right after this transit. Consider a single-homed origin AS o that announced two prefixes to a transit AS T, and T has exported one prefix to $AS1$ and the other prefix to $AS2$. Then one vantage point VP might observe different AS paths for these two prefixes, $[VP, AS1, T, o]$ and $[VP, AS2, T, o]$, resulting in two atoms, even though the single-homed origin AS announced them to the same transit AS.

Kastanakis et al. [**?**] studied the prevalence of selective advertisement in the Internet, and showed that selective export as one of the causes of selectively announced prefixes is widely used by large transit providers compared to 20 years ago. For large transits (e.g., AS3257, AS3292, AS3549, AS5511, AS7018), they examined for prefixes that appeared to be selectively advertised, and found that at most one third of their customers are single-homed, which intuitively cannot apply selective announcement, and thus they conclude that an intermediate AS in the path applies selective export policies.

Consistently, in our policy atom dataset, compared to 20 years ago, we observe that the average formation distances have increased with the percentage of atoms formed at distance 1 dropping from 45% to 20%, and the percentage of atoms formed at distance 3 or more increases from 17% to 33%. The majority of atoms are formed at a distance of 3. To summarize, these results indicate that atoms are formed not only due to the routing policy of the origin AS, but also due to the routing policies of the intermediate transits, evidencing the increased complexity of inter-domain routing policy.

4.4 Stability of Policy Atoms

For policy atoms to be helpful in a distributed environment, it is essential to maintain a consistent view of the atom structure across different vantage points. Rapid changes in atom composition would decrease policy atoms' applicability. To assess the stability of policy atoms over time, we compute the percentage of atoms that remained unchanged and the percentage of prefixes that remained grouped after specific time intervals, as described in Section 3.5.

Table 3 compares the stability levels of policy atoms between 2004 and 2024. We observe that both the short-term stability (after 8 hours) and long-term stability (after a week) of atoms dropped dramatically. In January 2004, 96.3% of atoms remained after 8 hours, whereas by October 2024, only 83.7% remained, a drop of 12 percentage points for complete atom match. A similar, but smaller decrease is observed for maximized prefix match (from 98.3% to 90.6%). Despite this overall decline, the relative drop from 8 hours to 24 hours remains nearly identical (around 3 percent in both years), and the additional drop from 24 hours to a week is less than 10 percent in both years. This suggests that atom breaks happen quickly, prefixes that remain in the same atom after 8 hours are more likely to remain in the same atom after 24 hours and after a week.

Figure 5 reveals that the atom stability has stayed consistently high, in both short term (after 8 hours, teal) and long term (after a week, coral), using the two metrics discussed in Section 4.4.

Table 3: Comparison of stability among atoms in 2004 and 2024.

	Jan 2004		Oct 2024	
	Complete Atoms (%)	Maximized Prefixes (%)	Complete Atoms (%)	Maximized Prefixes (%)
After 8 hours	96.3	98.3	83.7	90.6
After 24 hours	91.4	95.0	79.3	87.2
After 1 week	80.3	88.8	71.9	80.1

The short-term stability (after 8 hours) of atoms remains around 98%, and the long-term stability (after a week) is around 80%, This indicates that although prefixes may move between atoms, the overall composition of atoms does not change significantly. This observation suggests that it is still feasible to maintain the atom membership information. While the longitudinal study suggests the decrease in stability in 2024 is not a persistent trend, such occasional drops in stability still require further investigation to understand the underlying causes.

With changes that we have seen in the formation distance, which indicates a growing complexity in Internet routing policies, the overall stability of atoms has not significantly altered, especially considering the increase in the absolute number of atoms and prefixes over the past two decades. This dramatically illustrates the applicability of atoms in the current Internet. We should also note that the observed decrease in stability over long periods suggests that regular updates to the atom dataset may be necessary to maintain its accuracy.

4.4.1 Implications on Vantage Point Selection. A key practical question is which VP should be employed when policy atoms are used to detect routing-policy changes. To answer this, we further investigate the dynamics of policy atoms and study how widely an atom split is visible across VPs.

We process one snapshot per day at 8 am UTC from 1 January 2018 to 31 October 2020, and look into the changes in atoms over time.

(1) **Detect splits.** An atom present (identified by prefix composition) in snapshots t and t+1 is flagged as split if, in snapshot t+2, any of its prefixes are present in a different atom. We ignore atom merge because there is no VP changes view on prefixes grouping for this atom.
(2) **Count observers.** For each split, we examine all VPs in snapshot $t+2$ and count how many of them report the post-split atoms (or more, if the split yields > 2 atoms). Note that we do not compare the AS paths across snapshots, as it is totally likely that the entire AS path set has changed, but the prefixes are still grouped in the same atom, following the definition of policy atoms. We count the number of VPs that previously observed all prefixes with the same AS path, but now observe them in different atoms.
(3) **Aggregate.** Beginning 3 January 2018, for every day, we record (i) a list of split events and (ii) the number of VPs observing the event, for each event.

Figure 6 shows that most splits are seen by very few VPs. 80% of the events are visible to at most three BGP VPs, and 60% to only one VP. This suggests that the scope of the majority of the atom split events is localized. Figure 7 aligns all days in the time window and plots, for each day, the distribution of observer counts among that day's splits. Due to space limits, we only show a random 100-day period out of the 1000 days in the time window. The whole time window shows a similar pattern, and we put it in Appendix A8.6. We found that most of the atom split events happen on a day that is observed by a single VP, seems to be driven by one single VP rather than distributed evenly across all VPs. Through manual investigation of the top VP who observed the most splits, we found that it is likely to be the VP's own policy change (change in provider) that causes the split.

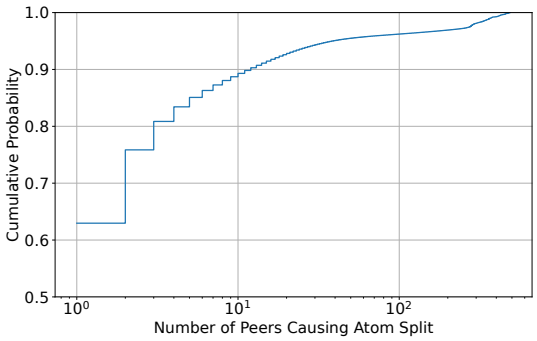

Figure 6: Distribution of the number of observers for all atom split events, suggesting that the scope of most of the splits is localized.

We believe this highlights one of the contributions of our replication study, which is to select VPs wisely when applying policy atoms in practice, depending on different use cases. For example, for understanding global routing policies, one may want to select VPs that are less likely to change their peering policies, Otherwise, the localized change of a single VP may be interpreted as a network-wide event. However, for use cases such as reducing measurement overhead by probing per atom rather than per prefix, one may want to include all VPs to capture as many policies as possible.

4.5 Impact of Internet Changes on Policy Atoms

It is reasonable to expect that the continuing growth of the Internet, for example, the increase in the number of ASes, prefixes and BGP vantage points (VPs), could affect the number and prevalence of atoms on the Internet. We briefly discuss how ASes, prefixes, and VPs jointly contribute to the change of policy atoms in this section. On one hand, the growth of content and cloud origins introduces export choices that are region and peer scoped. As the Internet flattens, there are more peering links established at IXPs and private interconnections. This increases multi-homing and creates more opportunities for intermediate AS to perform routing control (e.g.,

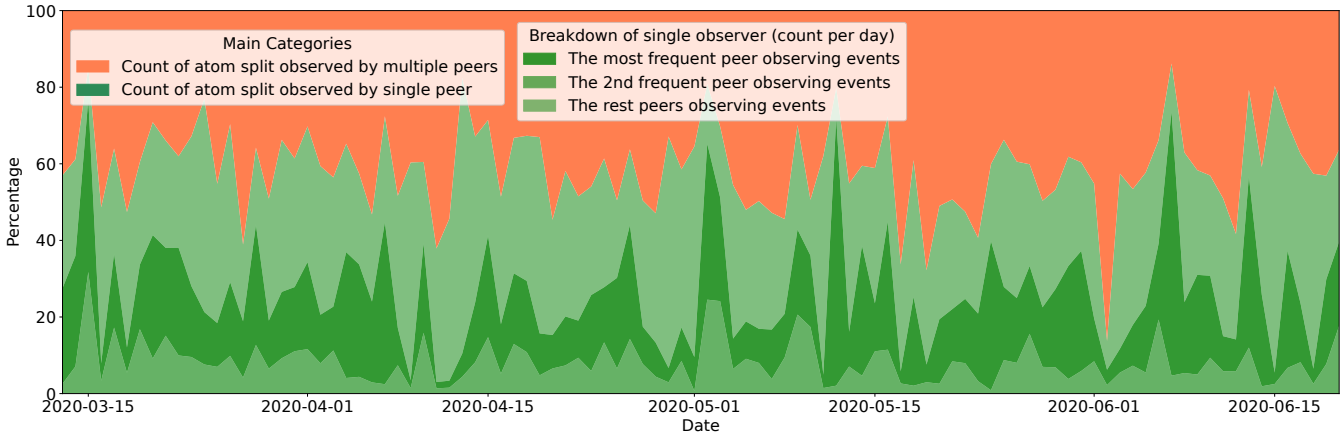

Figure 7: The distribution of observer counts for each day's atom split events, shows the splits are also driven by one single VP rather than distributed evenly across all VPs.

selective export and regional based policies). Consequently, this change contributes to the positive growth of the number of policy atoms and has a negative effect on the atom size. On the other hand, the exhaustion of the Internet address space encourages more specific prefixes and fine-grained traffic engineering. This imposes a positive impact on the number of policy atoms. Another impact is the expansion of BGP collectors, as shown in Figure 13, we have around 600 full-feed peers in 2024 compared to less than 50 full-feed peers in 2004. Each full-feed peer contributes their own view of the Internet, which helps us to capture more diverse routing policies. We learned from the BGP community values that ASes' export policy could be different for different geographical locations and peering ASes. The increased coverage of vantage points helps to capture more policies.

4.6 Takeaways

The analysis of policy atoms reveals several key observations:

- The general statistics demonstrate an effect of the adoption of more granular routing policies, with a simultaneous increase in the size of large policy atoms, reflecting the growing complexity of routing policies.
- Our analysis of BGP updates provides strong evidence that policy atoms are governed by routing policies, as prefixes within the same atom are frequently updated together. This observation is emphasized after categorizing ASes, reinforcing the idea that the Internet operates at the level of policy atoms rather than ASes.
- Formation distance analysis shows that a significant fraction of policy atoms is still formed due to the routing policy of the origin AS. The shift in atom formation distances, with more policy atoms forming further away from the origin AS, suggests a more complex interplay between the routing policies of different ASes.
- Policy atoms maintain high short-term stability and reasonable long-term stability over the years, supporting their validity in understanding BGP routing policies and other applicability. We also found that when atom splits occur, they are usually localized, and sometimes depend on vantage points' specific policies rather than policies closer to the origin AS.

Table 4: Compare general statistics of atoms between IPv4 and IPv6.

Year	v4 (2024)	v6 (2024)	v6 (2011)
Number of prefixes	1,028,444	227,363	4,178
Number of ASes	76,672	34,164	2,938
# single-atom ASes	31,009 (40.4%)	22,297 (65.3%)	2,558 (87.1%)
Number of atoms	483,117	94,494	3,486
# single-prefix atoms	355,197 (73.5%)	73,327 (77.6%)	3,223 (92.5%)
Mean atom size	2.13	2.41	1.20
99th percentile of atom size	17	20	5
Largest atom size	3,072	2,317	32

To conclude, with the increased complexity of routing policies over the past two decades, policy atoms remain a valuable concept for analyzing BGP routing policies. We believe that the concept of policy atoms will continue to be relevant in the future and will be more widely applicable in the context of the modern Internet, while also needing to be updated frequently and computed carefully.

5 POLICY ATOMS IN IPV6

In the nearly two and a half decades since Broido and Claffy's [**?**] and Afek et al.'s [**?**] analysis of policy atoms, adoption and support for IPv6 have increased substantially. According to Google, the fraction of its users that access its services via IPv6 recently surpassed 45% (compared to roughly 1% in late 2012) [**?**]. Additionally, the IPv6 BGP table reached 227,363 advertised prefixes and 34,164 originating ASes in October 2024, according to our calculation. Motivated by the increased importance of IPv6 on today's Internet, we apply the concept of policy atoms to IPv6 routing and compare it to IPv4.

5.1 Atom characteristics

Table 4 shows a high-level summary of the characteristics of the policy atoms identified in IPv4 and IPv6 routing. Overall, we see a significant decline in the share of single-atom ASes in IPv6, falling

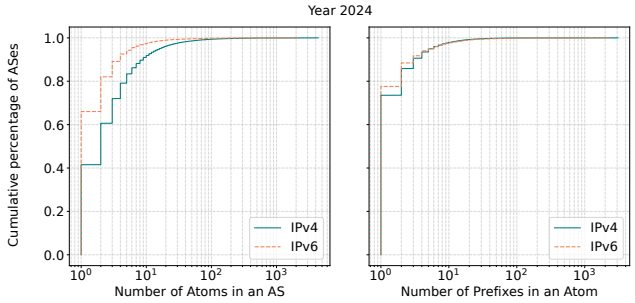

Figure 8: Distribution of the number of policy atoms per AS (left) and the number of prefixes per policy atom (right) in IPv4 and IPv6. IPv6 tends to have fewer atoms per AS and a similar number of prefixes per atom distribution compared to IPv4.

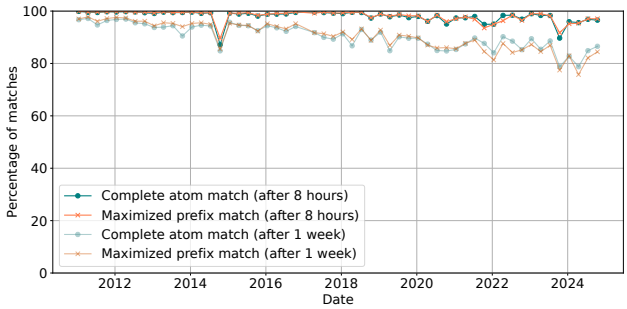

Figure 9: The short-term stability (after 8 hours) and long-term stability (after a week) of atoms in two metrics remains stable and consistently high for IPv6.

from 87% (2,558 of 2,938 ASes) in 2011 to 65% (22,297 of 34,164 ASes) in 2024, reflecting finer-grained policy decisions as the adoption of IPv6 increased. We also observe a consistent increase in the mean atom size (i.e., number of prefixes per atom) for IPv6 from 1.20 in 2011 to 2.41 in 2024, surpassing the average atom size of IPv4 (2.13). Similarly, we see that the largest IPv6 atoms and the 99th percentile of atom size have increased dramatically, with the largest atom reaching a size that is of a similar order of magnitude to IPv4 (3,072 prefixes for IPv4 and 2,317 prefixes for IPv6).

Digging into the data, we observed that a large fraction of single-prefix ASes were associated with China's research network CERNET and, more specifically, with its Future Internet Technology Infrastructure (FITI) project [**?**] testbed. Starting in 2021, the FITI project created 4,096 new ASNs (11% of all ASNs observed) and 4,096 /32 prefixes that were subnets of a single /20 block (240a:a000::/20). As they are legitimate prefixes, we do not remove them from our analysis.

Figure 8 illustrates the distribution of policy atoms between IPv4 and IPv6, where we observe that IPv6 has a similar distribution of prefixes per atom and fewer atoms per AS (largely comprised of ASes that are part of the FITI project) as IPv4. We conclude from the basic statistics that IPv6 has matured over the past decade, reflecting increasingly fine-grained routing policy during the adoption. However, IPv6 routing appears to have coarser-grained traffic engineering than IPv4, indicated by increasing mean atom size. The rest of this section investigates these trends in greater detail.

5.2 Policy atom stability

Figure 9 presents the stability of policy atoms over time for IPv6 after 8 hours (teal) and after a week (coral) in two metrics discussed in Section 3.5. In general, we observe a slight decline in atom stability as IPv6 becomes more widely deployed throughout the years, but the overall stability remains higher and more consistent than in IPv4, as shown in Figure 9.

5.3 Atom structure and BGP updates

Figure 10 presents the trend in the likelihood of ASes (coral) and atoms (teal) being seen in full within a single BGP update for IPv6. Following the same methodology as in Sections 3.3 and 4.2, we plot the solid, dotted, and dashed lines. The results confirm a consistently

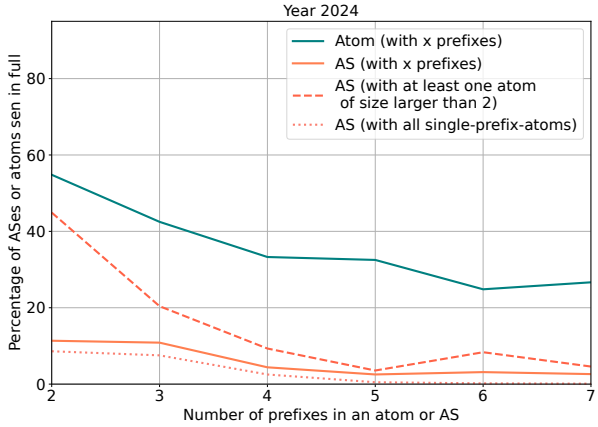

Figure 10: Likelihood of AS and Atom seen in full in one BGP Update for IPv6. The likelihood of atoms seen in full is consistently higher than AS seen in full.

strong correlation between atom structure and BGP updates for IPv6, similar to the findings for IPv4, as shown in Figure 10.

5.4 Formation distance of atoms

Figure 11 shows the trend in the percentage of atoms created at various formation distances from the origin AS for IPv6. We follow the same definition as in Section 3.4 and additional analysis as in Section 4.3 to plot the solid and dashed lines.

Indicated by the solid line, there is a significant and consistent decrease in the percentage of atoms created at distance one between 2011 and 2021, attributed to the decrease in percentage of single prefix ASes as in IPv4. The percentage of atoms created at distance 2 has also decreased, suggesting atoms are being formed farther away from the origin AS due to the policy imposed by intermediate ASes, shown by the dashed line after removing the effect of single atom ASes. Overall, the formation distance for IPv6 is becoming stable since 2020.

Comparing Figure 4 and Figure 11, we find that the average atom formation distance for IPv6 is smaller than that for IPv4, highlighted by more atoms formed at distances 1 and 2 for IPv6. This observation aligned with our observation in the basic statistics,

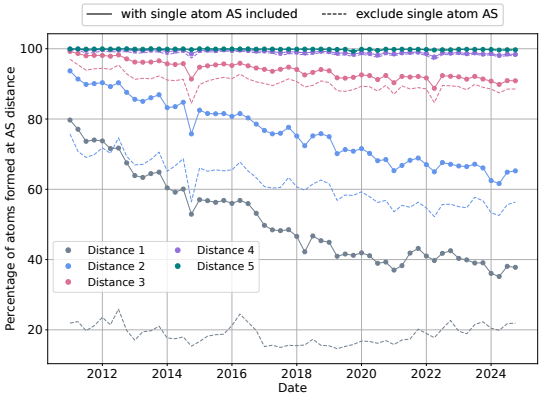

Figure 11: The trend of the percentage of atoms created at different numbers of ASes away from the origin AS for IPv6. More atoms are formed away from the origin AS.

IPv6 routing appears to have coarser-grained traffic engineering than IPv4, indicated by the smaller average formation distance of atoms.

5.5 Takeaways

First, our stability and update-correlation metrics for IPv6 atoms remain as strong as in IPv4, confirming that routing in IPv6 operates at the atom level and the concept of policy atom is fully applicable in the IPv6 context. Second, ASes seem to have coarser-grained routing policy under IPv6, indicated by larger mean atom sizes and smaller average formation distances compared to IPv4. Finally, both short-term and long-term stability of IPv6 atoms exceed that of IPv4, indicating a more stable atom structure and hence could potentially support applications of policy atoms in IPv6 such as selecting routing targets to reduce probing overhead [**? ?**].

6 RELATED WORK

The concept of policy atoms, introduced by Broido and Claffy in 2001 [**? ?**] and further validated and characterized by Afek et al. [**?**] in 2002, has been a valuable tool for understanding routing policies in the Internet. However, to the best of our knowledge, there has been no recent work that reproduces or replicates the full methodology of this paper to examine the current status of this concept.

There are studies seeking explanations for the potential causes of policy atoms [**?**] and we recognize a recent replication work [**?**] that reviews the selective announcement policy of ASes, which can directly influence the creation of policy atoms. Our work provides a comprehensive analysis of policy atoms in light of these findings.

Building on policy atoms, prior research in 2006 and 2008 (Netdiff and iPlane), explored applications of policy atoms: selecting routing targets to reduce probing overhead [**? ?**]. They placed measurement nodes inside edge networks to measure performance, to potentially not overload the edge, and they leveraged policy atoms to significantly reduce the number of destinations to probe. The authors concluded that probing by policy atoms instead of by prefixes considerably reduces the use of resources while maintaining good levels of accuracy. Although probing within an atom may reveal new routes, it generally does not alter the routing information significantly, therefore they updated the policy atom list every two weeks [**?**]. Very recently, there was a study in 2025 that used policy atoms in its core methodology to identify traffic engineering events [**?**].

Despite the applications explored in prior research, there remain gaps in understanding how to safely apply policy atoms in practice. We found there are no recent works investigating changes in policy atom membership to support these applications. We acknowledge these questions remain unsolved, however, our work contributes to a more comprehensive understanding of policy atoms and takes a first step to addressing these gaps.

There have been many works built on top of this concept to examine the routing policies and facilitate other applications. The lack of a recent understanding of the concept, coupled with the potential applicability of policy atoms in the context of new challenges, highlights the need for a replication study to re-examine and explore the concept of policy atom. Our work aims to bridge the gap by replicating the study by Afek et al. and providing an up-to-date analysis of policy atoms and exploring their applications in the current Internet.

7 DISCUSSION AND FUTURE WORK

Building on our longitudinal analysis of BGP policy atoms, we identify several interesting applications of the policy atom concept.

7.1 Detecting unreliable vantage points

Our study reveals that many atom splits are highly localized to specific VPs, suggesting that the majority of atom splits are not caused by a widely perceived change in routing policies and are more closely localized to the VP. By applying the concept of policy atoms, we can identify VPs that are more likely to "break" atom stability. Tracking this behavior over time would let researchers more accurately determine which routing changes are due to changes in routing policies in networks closer to the VP rather than the origin AS.

7.2 Policy atoms as a lens on BGP dynamics

Our analysis confirms that prefixes within the same atom are often updated at the same time. Because prefixes inside an atom have a high likelihood of changing AS path together in UPDATE bursts, policy atoms are a useful tool for understanding BGP dynamics. Unstable routes that affect an entire atom reflect a policy change or a network event, whereas churn associated to one prefix inside an atom is far more likely to be noise, leakage or transient misconfiguration. As a result, we believe it may be possible to identify and filter spurious single-prefix flaps and prioritize events that affect historically stable atoms.

7.3 Mapping IPv4 and IPv6 policy atoms within the same AS

In our analysis, we found that the distribution of prefixes per atom in IPv6 is similar to that of IPv4. Building on these findings, we believe that it is possible to leverage the concept of policy atoms–and the structure of these atoms (e.g., their structure, formation

distance, etc.)–to characterize IPv4 and IPv6 prefixes and identify "sibling prefixes" (i.e., prefixes that serve similar purposes in IPv4 and IPv6) [**?**].

8 CONCLUSIONS

This study revisits the concept of policy atoms after two decades, focusing on the longitudinal analysis of BGP data from 2004 to 2024 to assess the evolution and the current applicability of this concept within inter-domain routing. By reconstructing and applying the methodology described in Afek et al. [**?**], we not only validated prior results but also clarified ambiguities in the original approach, providing a more precise and documented methodology for future research that leverages policy atoms.

Our findings confirm that policy atoms remain a useful concept, despite a significant increase in network scale and complexity in routing policies. We further extend our evaluation to IPv6, confirming that policy atoms are a useful tool for analyzing routing behavior in IPv6.

Furthermore, we show that most atom split events are highly localized to specific peers, highlighting that studies that incorporate policy atoms into their methodology must carefully select which collector peers to include as to avoid mistaking local artifacts for routing changes.

Looking further, future research could further explore the dynamic characteristics of policy atoms with technologies such as machine learning for atom membership prediction, and investigate the application of policy atoms in the context of Internet routing security. We make our code and results publicly available to facilitate future research and replication studies to verify and expand upon our findings [**?**].

REFERENCES

[] [n. d.]. Google's IPv6 Statistics. https://www.google.com/intl/en/ipv6/statistics.html.

[] Yehuda Afek, Omer Ben-Shalom, and Anat Bremler-Barr. 2002. On the structure and application of BGP policy atoms. In *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet Measurement* (Marseille, France) *(IMW '02)*. Association for Computing Machinery, New York, NY, USA, 209–214. https://doi.org/10.1145/637201.637234

[] Robert Beverly and Arthur Berger. 2015. Server siblings: Identifying shared IPv4/IPv6 infrastructure via active fingerprinting. In *International Conference on Passive and Active Network Measurement*. Springer, 149–161.

[] Andre Broido. 2001. Analysis of RouteViews BGP data: policy atoms. In *Proceedings of network-related data management (NRDM) workshop Santa Barbara*. Cooperative Association for Internet Data Analysis-CAIDA, San Diego .

[] A Broido and k claffy. 2001. *Complexity of global routing policies*. Technical Report. Cooperative Association for Internet Data Analysis (CAIDA). https://doi.org/paper/2001_cgr

[] Omar Darwich, Cristel Pelsser, and Kevin Vermeulen. 2025. Detecting Traffic Engineering from Public BGP Data. In *Passive and Active Measurement: 26th International Conference, PAM 2025, Virtual Event, March 10-12, 2025, Proceedings*. Springer-Verlag, Berlin, Heidelberg, 307–334. https://doi.org/10.1007/978-3-031-85960-1_13

[] GTT Communications. 2025. BGP Communities (AS3257). https://www.as3257.net/communities.txt.

[] Savvas Kastanakis, Vasileios Giotsas, Ioana Livadariu, and Neeraj Suri. 2023. Replication: 20 Years of Inferring Interdomain Routing Policies. In *Proceedings of the 2023 ACM on Internet Measurement Conference* (Montreal QC, Canada) *(IMC '23)*. Association for Computing Machinery, New York, NY, USA, 16–29. https://doi.org/10.1145/3618257.3624799

[] Thomas Krenc, Matthew Luckie, Alexander Marder, and kc claffy. 2023. Coarse-grained Inference of BGP Community Intent. In *Proceedings of the 2023 ACM on Internet Measurement Conference* (Montreal QC, Canada) *(IMC '23)*. Association for Computing Machinery, New York, NY, USA, 66–72. https://doi.org/10.1145/3618257.3624838

[] Harsha Madhyastha, Tomas Isdal, Michael Piatek, Colin Dixon, Thomas Anderson, Arvind Krishnamurthy, and Arun Venkataramani. 2006. iPlane: An Information Plane for Distributed Services. In *7th USENIX Symposium on Operating Systems Design and Implementation (OSDI 06)*. USENIX Association, Seattle, WA. https://www.usenix.org/conference/osdi-06/iplane-information-plane-distributed-services

[] Ratul Mahajan, Ming Zhang, Lindsey Poole, and Vivek Pai. 2008. Uncovering performance differences among backbone ISPs with Netdiff. In *Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation* (San Francisco, California) *(NSDI'08)*. USENIX Association, USA, 205–218.

[] Wolfgang Mühlbauer, Anja Feldmann, Olaf Maennel, Matthew Roughan, and Steve Uhlig. 2006. Building an AS-topology model that captures route diversity. *SIGCOMM Comput. Commun. Rev.* 36, 4 (2006), 195–206. https://doi.org/10.1145/1151659.1159937

[] Wolfgang Mühlbauer, Steve Uhlig, Bingjie Fu, Mickael Meulle, and Olaf Maennel. 2007. In search for an appropriate granularity to model routing policies. 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, 145–156. https://doi.org/10.1145/1282380.1282398

[] Orange S.A. 2021. BGP Best Practices for IP Transit Customer. https://wholesale.orange.com/portail/resources/other/Orange_BGP_Best_Practices_for_IPT_Customers.pdf.

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

[] Yakov Rekhter, Susan Hares, and Tony Li. 2006. A Border Gateway Protocol 4 (BGP-4). RFC 4271. https://doi.org/10.17487/RFC4271

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

[] Florian Streibelt, Franziska Lichtblau, Robert Beverly, Anja Feldmann, Cristel Pelsser, Georgios Smaragdakis, and Randy Bush. 2018. BGP Communities: Even more Worms in the Routing Can. In *Proceedings of the Internet Measurement Conference 2018* (Boston, MA, USA) *(IMC '18)*. Association for Computing Machinery, New York, NY, USA, 279–92. https://doi.org/10.1145/3278532.3278557

[] Tsinghua University. [n. d.]. Launch Ceremony of Future Internet Technology Infrastructure held at Tsinghua University. https://www.tsinghua.edu.cn/en/info/1420/10191.htm. Accessed: 12 May 2025.

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

[] Feng Wang and Lixin Gao. 2003. On inferring and characterizing internet routing policies. In *Proceedings of the 3rd ACM SIGCOMM Conference on Internet Measurement* (Miami Beach, FL, USA) *(IMC '03)*. Association for Computing Machinery, New York, NY, USA, 15–26. https://doi.org/10.1145/948205.948208

[] Weili Wu, Zachary S. Bischof, Cecilia Testart, and Alberto Dainotti. 2025. *IMC25-Policy-Atom-Replication*. https://doi.org/10.5281/zenodo.17230538

APPENDIX

A8.1 Ethics

This study does not raise any ethical issues. The datasets used are publicly available.

A8.2 Threshold for determining full-feed peers and the number of peers determined by the threshold

In Figure 12, we illustrate the change in the threshold over the past 20 years, increasing from 100K to 1M.

In Figure 13, we present the change in the number of full-feed peers given the above threshold.

A8.3 Identifying abnormal BGP peers

In this section, we describe our method to identify abnormal BGP peers for removal. We exclude peers from the following 5 ASes: AS136557, AS57695, AS42541, AS47065, and AS25885. Table 5 lists the affected period with these peer ASNs. We choose to remove these peers entirely from our analysis instead of only removing the affected period, as we believe this provides a cleaner dataset for our longitudinal analysis.

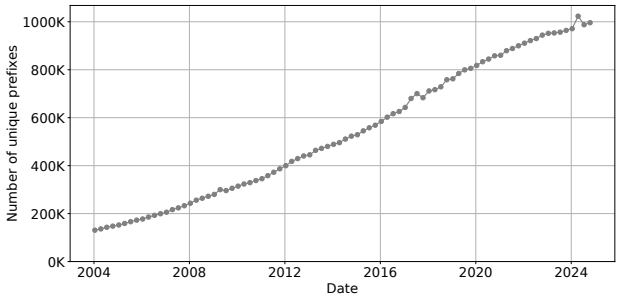

Figure 12: Threshold for determining full-feed peers

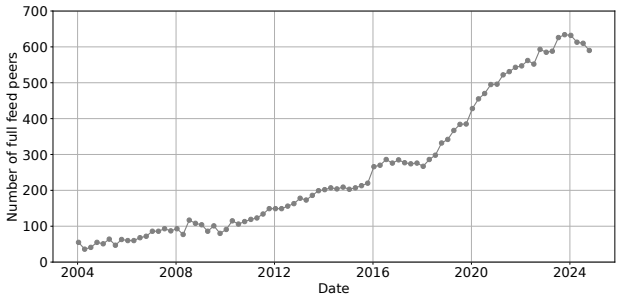

Figure 13: The number of full-feed peers determined by the threshold

A8.3.1 Peers with ADD-PATH issues. ADD-PATH parsing issues are known on certain Routeviews collectors. We identified manually four such peers, AS136557, AS57695, AS42541, and AS47065. During the affected period listed in Table 5, BGPStream output for these peers shows warnings such as *(i)* "unknown BGP4MP record subtype 9", *(ii)* "Duplicate Path Attribute", and *(iii)* "Invalid MP(UN)REACH NLRI". To verify, we inspect from the specific collector and time window using `bgpreader`:

```
bgpreader -j Peer_ASN -w start_time,end_time
```

Corresponding raw MRT file from the Route Views archive can also be checked to verify the problem. [2]

A8.3.2 Peer with misconfiguration issues. We also removed another peer, AS25885, because it exhibits abnormal behavior. We notice that, including AS25885 inflates the number of policy atoms by approximately 30% (from ~350,000 to ~450,000), which far exceeds the normal atom count fluctuation. More importantly, for the affected prefixes, the majority of the AS paths reported by peers from AS25885 have AS65000 immediately before AS25885 (i.e., 25885 65000 …), with more than 150,000 atoms with such paths. Because AS65000 is reserved for private usage, this is a clear signal that this peer is misconfigured. We thus removed this Peer AS entirely from our analysis.

[2]Example file indicating issues for AS136557: https://archive.routeviews.org/routeviews.perth/bgpdata/2022.09/UPDATES/updates.20220909.0615.bz2.

Table 5: Abnormal BGP peers removed from our analysis

Peer ASN	Affected Period
136557	Sep 2022
57695	Jan 2022
42541	May 2020 to Feb 2021
47065	Feb 2021
25885	Nov 2020 to Mar 2023

A8.4 Reproducing the results

A8.4.1 General statistics.

In Figure 14, we reproduce the AS and atom distribution in 2002.

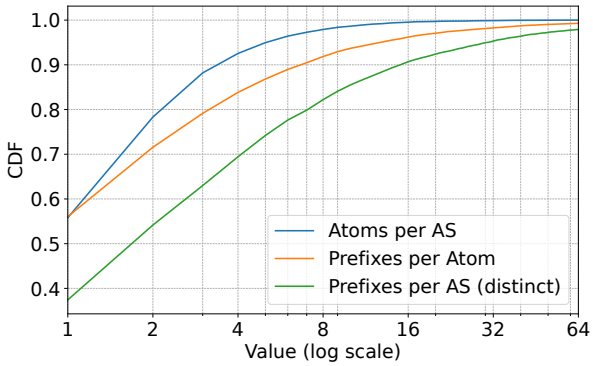

Figure 14: Distribution of AS and Atoms matches the original paper.

A8.4.2 Correlation of Atom Structure to Internet Update Records. In Figure 15, we reproduce the correlation of atom structure to Internet update records, which aligns with the original paper.

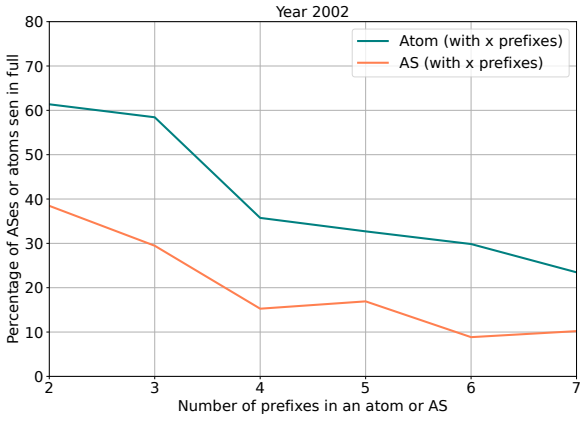

Figure 15: Update analysis result matches with the original paper.

A8.4.3 Stability of Policy Atoms.

In Table 6, we reproduce the stability of the policy atoms, which aligns with the original paper.

Table 6: Reproduced stability of policy atoms over time

Time span	Original paper		Reproduced result	
	CAM	MPM	CAM	MPM
8 Hours	95.3%	97.7%	94.2%	97.5%
1 Day	91.6%	97%	91.8%	96.2%
1 Week	77.5%	86%	77.6%	87%

A8.5 Prefix filtering sensitivity analysis

In this section, we evaluate the thresholds we use to filter out prefixes that we do not consider when computing policy atoms and in our further analysis, using the data from the BGP collectors snapshot of October 15, 2025 at 8am UTC.

As explained in § 2.4.3, in our data cleaning steps we remove from BGP data snapshots prefixes that are not seen *(i)* by at least two BGP route collectors and *(ii)* by at least four different peer ASes. In other words, we adopt the prefix visibility thresholds of $\geq$ 2 collectors and $\geq$ 4 peer ASes. To select the thresholds and assess their impact, we compute the total count of prefixes under different combinations of thresholds. Table 7 presents the count of prefixes for different [collectors (rows), peer ASes (columns)] thresholds pairs. We include results for thresholds up to one more than our selected values. The highlighted cell corresponds to the selected threshold: $\geq$ 2 collectors and $\geq$ 4 peers.

Table 7 reveals that the number of prefixes seen by 4 or more peer ASes is relatively stable, with less than 0.5% reduction by increasing the threshold. Similarly, increasing the minimum number of BGP collector threshold has minimal consequences in the reduction of prefixes. We note though that as we know that during our measurement window some collectors had configuration problems impacting the routes in the BGP data from their peers (see Appendix A8.3), we filter out prefixes that are not seen by at least two collectors even if in most days there is almost no difference when considering prefixes seen by 4 or more peer ASes.

We repeat the sensitivity analysis for the last 7 years, computing the number of prefixes for October 14 and 15 of each year, and we observe similar results. In every year, the number of prefixes on the 14th and 15th for each threshold pair differs by less than 0.1%, confirming our results are consistent.

Therefore, to compute policy atoms, we use routing data of prefixes that are seen by at least 4 peer ASes and are shared with at least 2 distinct BGP collectors. The choice of $\geq$ 4 peer ASes removes misconfigurations and very local prefixes that are not relevant to our study. The choice of $\geq$ 2 collectors reduces data artifacts from issues with a single collector.

Table 7: Count of valid prefixes under different thresholds for collectors (rows) and peer ASes (columns). We report values up to one more than our threshold values ($\geq$ 2 collectors, $\geq$ 4 peers). The highlighted cell corresponds to the adopted threshold: $\geq$ 2 collectors and $\geq$ 4 peers.

Peer ASes → Collectors ↓	1	2	3	4	5
1	1083140	1038893	1031541	1028448	1027292
2	1042682	1038153	1031476	**1028444**	1027287
3	1034520	1032946	1030279	1028027	1027259

A8.6 1000 days of atom split breakdown

We present the atom split breakdown for the full time window in Figure 16.

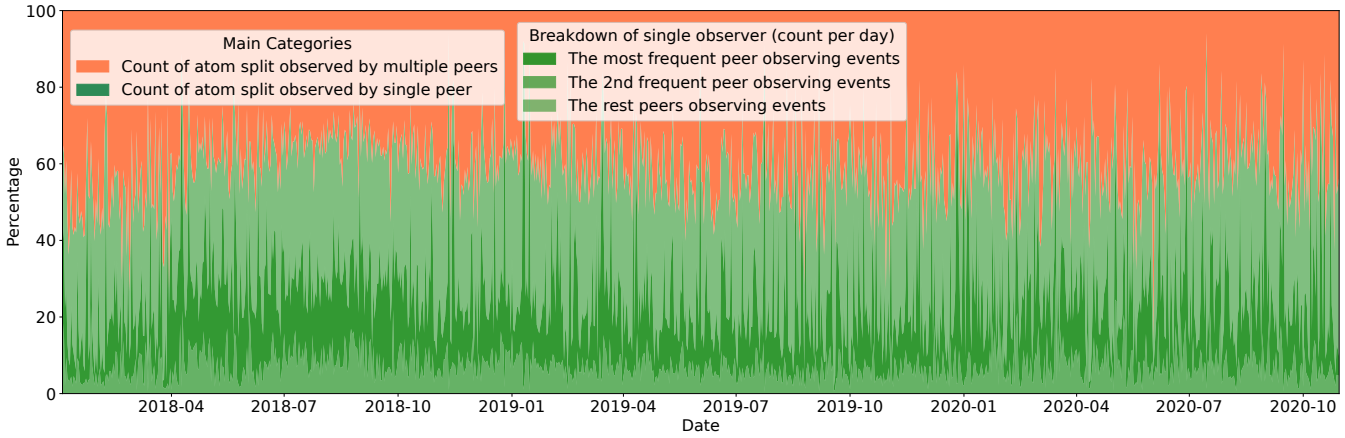

Figure 16: Atom split breakdown for the full time window.