Coarse-grained Inference of BGP Community Intent

Thomas Krenc
CAIDA, UC San Diego
La Jolla, CA, USA
tkrenc@caida.org

Matthew Luckie
CAIDA, UC San Diego
La Jolla, CA, USA
mjl@caida.org

Alexander Marder
CAIDA, UC San Diego
La Jolla, CA, USA
amarder@caida.org

kc claffy
CAIDA, UC San Diego
La Jolla, CA, USA
kc@caida.org

ABSTRACT

BGP communities allow operators to influence routing decisions made by other networks (*action* communities) and to annotate their network's routing information with metadata such as where each route was learned or the relationship the network has with their neighbor (*information* communities). BGP communities also help researchers understand complex Internet routing behaviors. However, there is no standard convention for how operators assign community values, and significant efforts to scalably infer community meanings have ignored this high-level classification. We discovered that doing so comes at significant cost in accuracy, of both inference and validation. To advance this narrow but powerful direction in Internet infrastructure research, we design and validate an algorithm to execute this first fundamental step: inferring whether a BGP community is *action* or *information*. We applied our method to 78,480 community values observed in public BGP data for May 2023. Validating our inferences (24,376 action and 54,104 informational communities) against available ground truth (6,259 communities) we find that our method classified 96.5% correctly. We found that the precision of a state-of-the-art location community inference method increased from 68.2% to 94.8% with our classifications. We publicly share our code, dictionaries, inferences, and datasets to enable the community to benefit from them.

CCS CONCEPTS

• **Networks** → *Network protocols*; **Network management**.

KEYWORDS

Border Gateway Protocol (BGP), BGP communities.

ACM Reference Format:
Thomas Krenc, Matthew Luckie, Alexander Marder, and kc claffy. 2023. Coarse-grained Inference of BGP Community Intent. In *Proceedings of the 2023 ACM Internet Measurement Conference (IMC '23), October 24–26, 2023, Montreal, QC, Canada.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3618257.3624838

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

ACM ISBN 979-8-4007-0382-9/23/10…$15.00
https://doi.org/10.1145/3618257.3624838

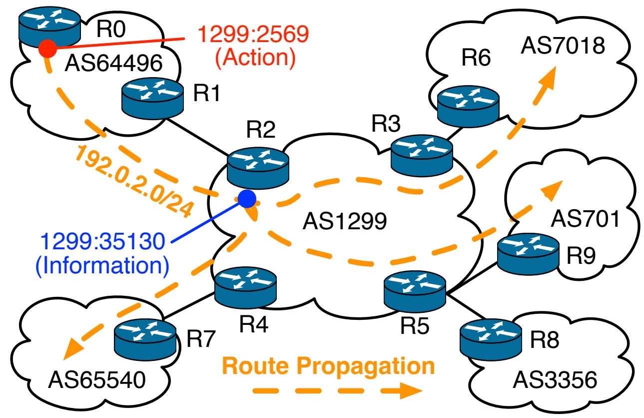

Figure 1: Arelion (AS1299) communities and their effect on routing. AS64496 sets action community 1299:2569 on the route it originates at R0, instructing AS1299 to not export the route to AS3356 in Europe. AS1299 sets information community 1299:35130 on routes received by R2, which records that AS1299 learned the route in Boston, MA, USA.

1 INTRODUCTION

The Border Gateway Protocol (BGP) [40] is the routing protocol used by network operators to organize inter-domain routing among the >75K Autonomous Systems (ASes) on the Internet. BGP has evolved to meet operator needs in an increasingly complex innovation ecosystem. One driving component of BGP is the *communities* route attribute [34], a signaling mechanism used in BGP routers to implement scalable policies at a prefix granularity.

Operators use BGP communities to influence routing decisions made by other networks, or to record metadata that operators can use when applying BGP policy [12]. Figure 1 illustrates some community values observed in the wild. By convention, the ASN represented by the first number in the community string is the ASN that assigns the meaning of the second number. In this case, Arelion (AS1299) provides *action* communities that their customers may use to control how Arelion propagates their routes. For example, Arelion's convention is to not export a route marked with 1299:2569 to Level3 in Europe [1]. Arelion also uses *information* communities to annotate their routes with metadata for internal purposes. For example, 1299:35130 means that Arelion learned the route in Boston, MA, US [1].

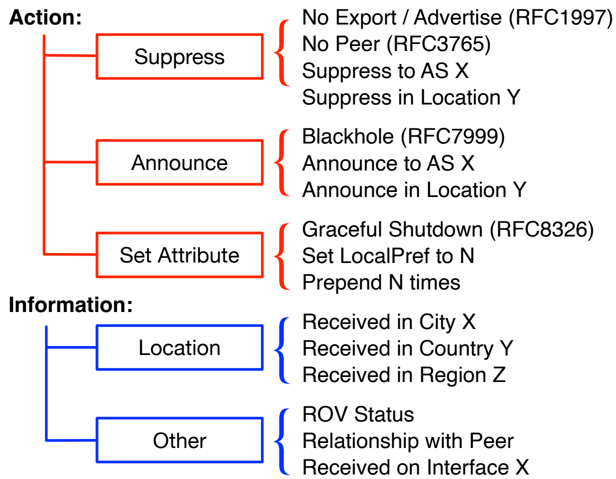

Figure 2: Taxonomy of categories for regular and large BGP communities. Our long-term goal is to automatically infer the precise meanings of community values. Classifying a community as action or information is a critical first step toward that goal.

A comprehensive dictionary of BGP community meanings would be useful to operators [44], and also help researchers understand complex Internet routing behaviors (§3). Apart from a few standardized community values [16, 26, 30, 34], each network operator can decide the meaning of each community value they use, and there is enormous heterogeneity in how each operator designs and implements their approach. Some operators publicly document their dictionary that maps each community value to a route outcome, which allows operators and researchers to interpret their routes. As of May 2023, NLNOG maintains the largest repository of *regular* (32-bit, see §2) community mappings, consisting of 46 ASes. In the wild, we observed 5,491 ASNs with 78,480 unique undocumented communities. Where there is no community mapping available, knowing whether it is action or information can help to understand (1) which network (ISP or customer) likely set a community, (2) which routes are subject to modification through an action community, and (3) whether a route is anomalous (e.g., sudden absence of information communities). Further, this classification is an important first step towards fine-grained inference of community meanings.

In this work, we design, implement, and validate a method to infer whether a community is action or information. We used our inferences to improve a recently-published algorithm that inferred if a BGP-observed community was a location community, increasing the precision of that method from 68.2% to 94.8%. We publicly share our code, dictionaries, inferences, and datasets [33].

2 BACKGROUND

The BGP communities attribute is an optional extension that augments BGP announcements with metadata [12, 34]. A *regular* 32-bit BGP community has the form α:β, where the first 16 bits (α) contain the AS number that defines the meaning of the remaining 16 bits (β). Because regular communities do not provide space for a 32-bit ASN [6], in 2009 the IETF standardized *extended* 48-bit BGP communities, where the first 32 bits contain a 32-bit ASN [47]. To allow

	Europe	N. America	Asia Pacific
Level3	1299:256x	1299:556x	1299:756x
Orange	1299:254x	1299:554x	1299:754x
Verizon	1299:257x	1299:557x	1299:757x
GTT	1299:269x	1299:569x	1299:769x

x: 1, 2, 3 = prepend 1, 2, 3 times; 9 = do not export.

Figure 3: Example Arelion action communities that affect route export in Europe (2), North America (5), and Asia Pacific (7) for Level3 (56), Orange (54), Verizon (57), and GTT (69).

for more powerful community attributes, in 2017 the IETF standardized *large* 96-bit BGP community attributes in RFC8092 [24], which have the form α:β:γ, where the first 32 bits (α) contains the AS number that defines the remaining pair (β:γ) of 32-bit values. RFC8092 states that operators can use BGP communities either to convey *information* or induce an *action*. Figure 2 provides a taxonomy of common BGP community types within these broad categories.

Information communities: An operator can configure their routers to attach information communities to routes to enable sophisticated routing decisions within their AS. These communities help operators overcome scaling issues in BGP configuration. If an operator annotates a route with an information community that describes whether the route was received from a peer, provider, or customer, then the operator can use the community value to simplify BGP configuration of route export policy. For example, an AS can configure its BGP sessions to only export routes tagged with a "learned from customer" community to peers and providers, preventing route leaks. Similarly, if an operator annotates a route with an information community that describes the geographic location where an AS received an announcement, then the operator can use the community value to select geographically optimal paths. Finally, if an operator annotates a route with an information community that records the ROV status of a route, then the operator can use the community value to prefer routes from a valid origin, and only accept an ROV-invalid route when no other route is available.

Action communities: A neighbor AS attaches action communities to a route to influence routing in an upstream AS. Examples include instructing an upstream AS to (1) prepend itself to an AS path [22], (2) selectively advertise a route at specific Internet Exchange Points (IXPs) [23, 41], (3) blackhole a prefix to mitigate denial of service attacks [7, 9, 21, 30]), and (4) change the upstream's local preference value for the route. Sophisticated community-enabled services offered by providers can be important for prospective customers seeking to form new transit relationships.

Operator practices: While each operator is free to define the meaning of individual community values for their network, operators typically bring structure to their own community values by grouping values that have a similar outcome to aid comprehension. For example, while the value 1299:2569 causes Arelion to not export the route to AS3356 in Europe (Figure 1), community values 1299:2561, 1299:2562, and 1299:2563 cause Arelion to prepend its ASN 1 to 3 times, respectively, before exporting the route to Level3 in Europe [1]. That is, community values of the form 1299:256x involve Level3 (56) in Europe (2) in some way. Figure 3 illustrates similar Arelion action communities for other regions and ASes (underlined digits).

3 RELATED WORK

Operational use of BGP communities has expanded considerably over the past decade. Using publicly available BGP route archives, Streibelt *et al.* reported that approximately twice as many ASes defined communities in 2018 (≈5K) than did so in 2010 (≈2.5K). They found that the total number of unique community values increased from ≈19K to ≈63K, and that at least 75% of routes in April 2018 contained at least one community [46].

BGP communities in Internet infrastructure research: BGP communities have been used to validate algorithmic inference of AS relationships [10, 17, 20, 25, 27, 35], identify peering locations [2] and density [23], analyze how specific ASes cause AS-level path instability [15], infer the prevalence of DDoS attacks or server outages [18, 21, 28], detect path changes when network interconnection infrastructure fails [18], infer the number of sessions between two ASes [31], investigate how routers internal to an AS change [19], analyze the visibility DNS root servers [13], and investigate routing policies at IXPs [23, 37]. Each of these efforts required manually compiling a BGP community dictionary that covered a subset of observed community values. Our automated classification is a step towards a central repository of BGP community meanings.

Security implications of BGP communities: BGP provides no mechanism to validate communities, so any AS along the AS path may add, modify, or delete communities. In 2018, Streibelt *et al.* described attacks that use BGP communities to blackhole legitimate traffic [46]. In 2019, Birge-Lee *et al.* [3] showed the potential for sophisticated attacks that could use communities for fine-grained hijacking, circumventing conventional hijack detection techniques.

Classification of BGP communities: In 2002, Quoitin and Bonaventure provided a taxonomy of the early uses of communities that they publicly observed in public BGP data, producing an IETF Internet Draft [39]. Their taxonomy covered *route tagging* and *route redistribution* communities, which correspond to *information* and *action* communities. In 2007, Donnet and Bonaventure manually built a dictionary of BGP communities using information published on network operator websites and in Internet Routing Registry (IRR) records [12]. Their dictionary was comprised of 73.5% *inbound* and 26.3% *outbound* communities, which also correspond to *information* and *action* communities. When they examined BGP data, they found that their dictionary covered 22% of the observed communities. Recognizing that network and IXP operators often cannot interpret communities observed in announcements, in 2019 EuroIX proposed action and information community standards, along with sub-categories like filtering and origin validation [14], which has not yet gained operational traction. Our approach is a first step towards automated fine-grained BGP community classification. A fine-grained classification would allow for automated interpretation of currently deployed BGP communities, so that operators will not have to incur the operational cost of modifying and deploying new BGP community configurations.

In 2022, Da Silva Jr. *et al.* developed a technique to infer if a community signals a *location* [8] – a specific sub-category of information communities shown in Figure 2. Their method left inferring *what* location was signaled in the community as future work. Their method examined each community in isolation from other communities used by the AS, and their paper reported that their method had a "high number of false positives for action communities" – action communities that their method falsely inferred to be location communities. We report in §6 that first classifying communities as action or information, and then excluding action communities from their inferences increases the precision of their method from 68.2% to 94.8%.

4 DATA

BGP routing data: We extracted unique AS path and BGP Community tuples observed in RIBs and updates collected by RouteViews [43] and RIPE RIS [42]. For a single week (May 1-7, 2023), we collected ≈174M tuples. In total, there were 100,506 unique communities, consisting of 88,982 regular and 11,524 large communities. In this work, we focus on regular communities owing to their prevalence. We also used CAIDA's AS relationship [4] and organization [5] inferences for this week to provide context to our method.

BGP community dictionaries: During 2022 and 2023, we manually collected dictionaries for 59 ASes using (1) dictionaries published on their websites and in their Internet Routing Registry (IRR) records, (2) dictionaries collected by NLNOG to inform their looking glass [45], and (3) dictionaries provided by the One Step website [38]. We labeled each community as *action* or *information* according to the taxonomy in Figure 2. Because operators contiguously number communities that have a similar purpose (§2), we identified patterns in community values and summarized these patterns using regular expressions. For example, the regular expression 1299:[257]\d\d[1239] matches action communities illustrated in Figure 3 that affect Arelion's export policy to different ASes (the two underlined digits matched with \d\d) in different regions (2, 5, 7). For these 59 ASes, we created 199 information and 133 action regexes that covered the contiguous information and action communities that we collected, respectively.

Our assembled dictionary uses data sources of unknown and likely varying age, which could negatively impact our inferences if the dictionary entries are stale. To understand the degree to which network operators change their community assignments, we compared our dictionary against one published in 2007 [11, 12] which we did not use to assemble our dictionary. Among 19 ASes with communities in both the 2007 dictionary and ours, the coarse-grained categories had not changed. This stability is consistent with the fact that changing community assignments require potentially costly changes to operational network configurations that use communities.

5 METHOD

5.1 Intuition

We first describe what we should expect to observe in public BGP data using a first principles approach, and demonstrate that those expectations manifest using our ground truth.

Information and action communities appear in distinct clusters: Figure 4 summarizes BGP data observations for 30 of the 59 ASes in our ground truth dictionary. These 30 ASes had both

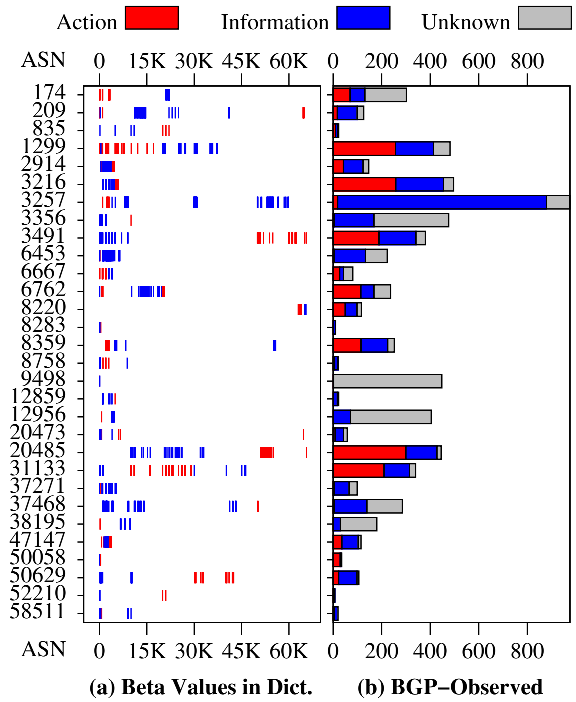

Figure 4: AS dictionaries paired with communities seen in BGP data. Operators contiguously number communities that have a similar purpose (action or information). Many ASes have communities that they have have not publicly documented (unknown).

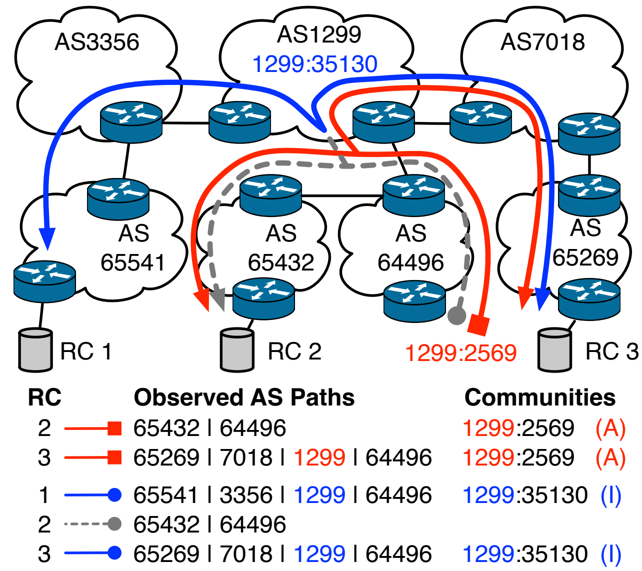

Figure 5: Action communities are set by customers to influence routing of the identified AS, so the signaled AS may not appear in the route's AS path. Information communities are set by the AS itself, so the signaled AS should appear in the route's AS path.

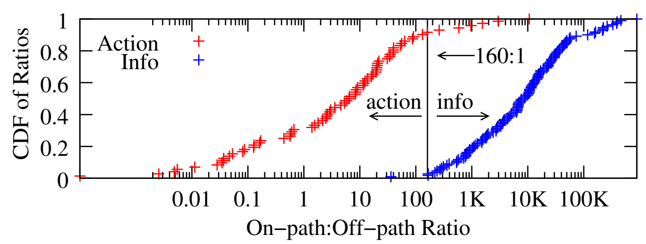

Figure 6: CDF of on-path:off-path ratios of baseline clusters. The optimal ratio of 160:1 yields 98% accuracy.

information and action communities; the remainder had only information or action communities. The left plot shows the clustering of communities in our dictionary that we also observed in BGP data, with ASes devoting ranges of values that have similar purpose. For example, Arelion defines two action communities (1299:50, 1299:150) that cause Arelion to set 50 or 150 as the local preference on the route [1], followed by information communities (1299:430, 1299:431) that signal ROV status, action communities (1299:666, 661, 999) for blackholing, followed by action communities (1299:2000 – 1299:7999) that control route export (Figure 3), action communities (1299:10050 – 1299:17150) that set the local preference of a route in different regions, followed by ranges of information communities (1299:20000 – 1299:39999) that record where the route was learned. The following analyses (and our method) examine BGP routing properties of distinct BGP community clusters.

Action communities are more likely to appear *off-path* than information communities: Figure 5 illustrates what we should expect to observe in public BGP data for routes with action or information communities. AS64496 originates two routes, tagging one (colored red with square end) with the action community 1299:2569, instructing AS1299 to not propagate the route to AS3556. Because AS64496 propagates the route to both AS1299 and AS65432, route collector RC2 records a route without AS1299 in the AS path but with the action community set. That is, the signaled AS (1299) does not necessarily appear in the AS path (can be *off-path*) when a neighbor signals an action community. The other route from AS64496, initially grey with round end, is tagged by AS1299 with 1299:35130 to signal that it received the route in Boston; we color the route blue after AS1299 tags the route. All collectors record routes with AS1299 in the path when they observe community 1299:35130 because AS1299 set that information community (the signaling AS is *on-path*). We calculated the on-path:off-path ratio of a community by counting the number of unique AS paths the community appeared on-path and off-path, respectively.

Figure 5 illustrates the ideal case; a peer of AS64496 happened to receive a route with an action community intended for AS1299, and it relayed the route to a route collector. There is no guarantee that other ASes signaling action communities to the same provider AS would have the same behavior, or have their announcements recorded by route collectors. Further, a small number (≈400) ASes filter all communities from routes before they further announce them [32]. By clustering community values, we aggregate the path

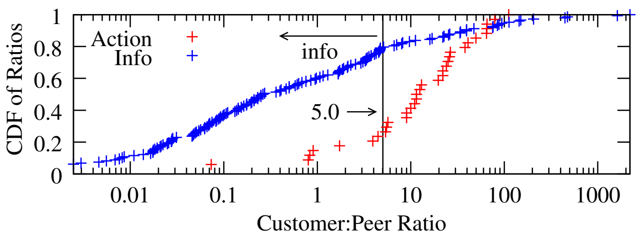

Figure 7: CDF of customer:peer ratios of baseline clusters. The optimal ratio of 5:1 yields 80% accuracy.

features of each individual community inside a cluster. We calculated the on-path:off-path ratio of a cluster by first calculating the on-path:off-path ratios of individual communities in the cluster, and then assigned the average of these to the cluster.

Using the regular expressions that we created from ground truth data, we obtained 332 community clusters covering 6,259 communities in BGP data. 937 communities appeared in on-path clusters, and 66 communities appeared in off-path clusters, which we classified as informational and action, respectively. The majority of communities (5,256) appeared in 183 mixed clusters, i.e., clusters that contained both on-path and off-path counts, and figure 6 plots a CDF of on-path:off-path ratios of those 111 information and 72 action clusters. Nearly all information community clusters had a on-path:off-path ratio of at least 160:1, while few action community clusters had a ratio that was more than 160:1, and we can use this simple ratio to classify clusters as action or information.

Action communities are mostly signaled from customer to provider: Providers offer action communities to enable customers to influence provider policy on customer routes. Therefore, we should expect that if we observe a route with an action community attached, and the AS α is observed in the AS path, then the subsequent AS in the path is a customer of AS α. Using the example in Figure 5, we expect that an AS relationship inference algorithm would infer the AS link between 1299 and 64496 to be a p2c link, because 64496 signaled the 1299:2569 action community. Figure 7 demonstrates that intuition to hold in public BGP data: CAIDA AS relationship inferences show that action community clusters involved more inferred customers than inferred peers. However, this property is not a useful distinguishing feature, as ASes typically set information communities regardless of the relationship they have with their neighbor, and public BGP data records many customers downstream of providers that set information communities. At best, if we categorized clusters with a customer:peer ratio less than 5:1 as information, we would achieve a maximum accuracy of 80%.

5.2 Inference Algorithm

Figure 8 provides an overview of our method. For each AS α, we cluster the observed β values into numeric ranges using an unsupervised clustering algorithm. We then use the on-path:off-path ratio for observed clusters to infer if the communities in the cluster are information or action. Finally, we apply those inferences to observed communities.

Estimating community clusters: Our goal of estimating clusters is to approximate the performance of using regular expressions. If the estimated clusters are too general, we might incorrectly merge action and information clusters. If they are too specific, then a sparsely populated cluster may also lead to misclassification of the β values within the cluster. To that end, our method identifies sequences of community values where the gap between any pair of adjacent β values is not more than a defined gap value. Figure 9 shows the accuracy of our inferences given different gap parameters. A gap value of 140 approximates the clustering behavior in the wild, yielding an overall accuracy of 96.5% over all labeled communities in our ground truth. If we had not clustered communities – i.e. we considered each community in isolation, then our method would have yielded 73.7% accuracy.

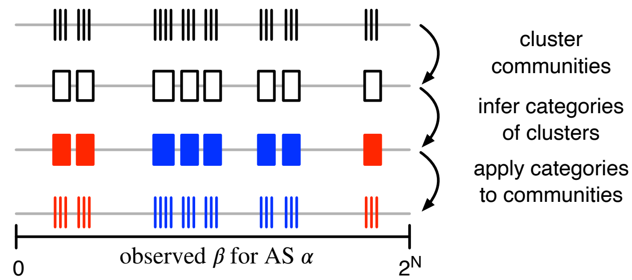

Figure 8: Workflow of method to infer information and action communities. (i) cluster communities, (ii) infer labels, (iii) apply to communities.

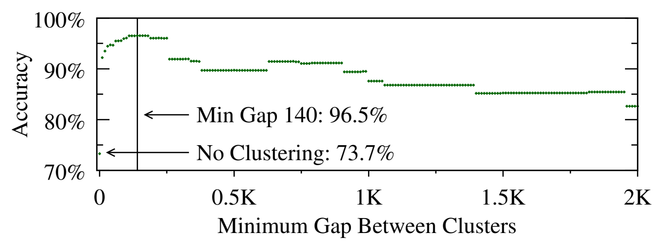

Figure 9: Inference accuracy given different minimum gap parameters. Gap parameters 100–250 yield accuracy values above 96%; we used 140.

Inferring action or information labels for clusters: We examine the on-path:off-path ratio of each cluster, using the intuition from §5.1, where info communities occur on-path more frequently than action communities do. We did not classify communities where the first 16 bits were from the private ASN range, or when the ASN (or a sibling thereof) did not appear in any AS path in our dataset. This excluded IXP communities from classification, where operators configured the IXP route servers to not embed the route server's ASN in the AS path. In this configuration, all communities associated with the IXP would appear off-path regardless of their purpose, therefore our method cannot correctly classify them.

When communities in a cluster were never observed off-path, or when the on-path:off-path ratio was above 160:1, we inferred the community cluster contained information communities. When communities in a cluster were always observed off-path, or when the on-path:off-path ratio was below 160:1, we inferred the community cluster contained action communities.

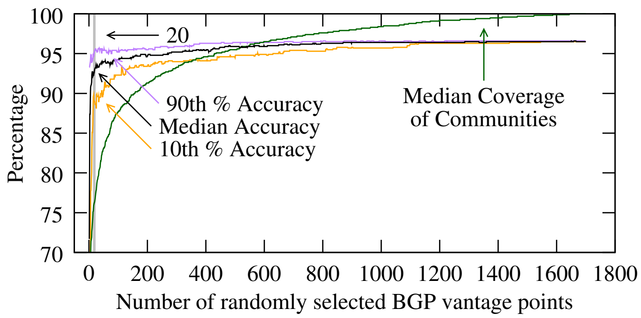

Figure 10: Impact on accuracy through accumulation of vantage points that send communities. We used a fixed on-path:off-path ratio (160:1) and minimum gap (140) between clusters, and ran 50 experiments. Using 20 vantage points, the median accuracy stabilized above 93%.

6 EVALUATION

Using the 7 days of BGP data (May 1-7, 2023), we applied our inference method to 78,480 of 88,982 regular BGP communities. Our method inferred 54,104 information and 24,376 action communities defined by 5,491 ISPs observed in BGP. For the 6,259 communities covered by the regexes in our dictionary, we observed an overall accuracy of 96.5%.

Impact of using a subset of input BGP data: The thresholds that we used to infer community categories, i.e., on-path:off-path ratio and minimum gap, represent the best case for our input data – an aggregate of all vantage points available. We therefore tested the accuracy of our method using a subset of available vantage points. Figure 10 shows the change in accuracy as we increase the number of BGP vantage points available to our method, where we add randomly-selected BGP vantage points from those available. Using 20 vantage points, the median accuracy stabilized above 93%, covering 76.5% of the communities observed with all vantage points.

Benefits of additional days of input BGP data: We evaluated the accuracy of our results as a function of the input data size, by starting with one day of BGP data, incrementally adding more days to the input data, and finishing with seven days of BGP data. The accuracy stabilized with two or more days between 96.4% and 96.6%.

Accuracy of inferences over time: We evaluated the accuracy of our results over time, by using a single day of BGP data from the first day of every month from June 2022 to May 2023, inclusive. The accuracy was stable, ranging from 92.6% to 95.4%. Over this time, the number of inferred communities increased by 5%, primarily due to newly observed information communities.

Improving BGP location community inference: As discussed in §3, Da Silva Jr. *et al.* reported that their method has a "high number of false positives for action communities" – action communities that their method falsely inferred to be location communities [8]. Using the location community inferences and ground truth dataset that they publicly released [29], we examined the benefit of filtering out communities that our method inferred to be action communities from their December 2020 inferences. We emphasize that the type column of Table 1 uses the labels from their ground truth dictionary, applied to the BGP location community inferences that

Class	Type	Before	After
Info	Geolocation	476 (68.2%)	472 (94.8%)
Action	Traffic Engineering	206	12
Info	Route Type	13	11
Info	Internal Routes	3	3
Total:		698	498

Table 1: Comparison of classification results of Da Silva *et al.*, before reducing false positives using our inferences, and after. Precision for correctly predicting a location community increased from 68.2% to 94.8%.

they publicly released [29]. Using their released data, we computed a precision ($TP/(TP+FP)$) of 68.2% for their location inferences. Most (206 of 222) of the error was from traffic engineering action communities that their method inferred were location communities. Our method filtered out all but 12, increasing the overall precision of their inferences to 94.8%.

7 FUTURE DIRECTIONS

We have taken a first step toward the challenging task of automatically inferring fine-grained BGP community meanings, which are behind a significant volume of work (§3). This first step will enable the measurement community to pursue automated inference of these dictionaries, advancing our ability to reason about BGP routing observed in the wild.

We imagine several future directions for this work.

Improving geolocation research. Future work could correlate information communities observed in routes with geolocation properties of routes, using geolocated routers [36] or originated prefixes [48], enhancing the community's ability to understand the geographic nature of paths taken by traffic.

Improving AS relationship inference. The community could pursue techniques that correlate AS relationship inferences with observed information communities to better understand complex AS relationships. The community could also use these improved AS relationship and geolocation datasets to interpret action communities that suppress or promote announcements to specific ASes or locations that would otherwise receive a route.

Improving situational awareness for operators. Current BGP community dictionaries cover only a small fraction of observed BGP communities, an operational challenge [44] that has led to public calls for data from operators to populate these dictionaries [45]. We believe that the research community can work in conjunction with the operations community to accelerate inference and validation of BGP community dictionaries, of mutual benefit to both communities.

ACKNOWLEDGMENTS

The authors express their gratitude to the anonymous reviewers and our shepherd Kyle Schomp for their thoughtful input. This material is based on research sponsored by the National Science Foundation (NSF) grants CNS-2120399 and OAC-2131987. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of NSF.

REFERENCES

[1] Arelion. 2023. BGP Communities. https://www.arelion.com/our-network/bgp-routing/bgp-communities.

[2] Brice Augustin, Balachander Krishnamurthy, and Walter Willinger. 2009. IXPs: Mapped?. In *ACM IMC*.

[3] Henry Birge-Lee, Liang Wang, Jennifer Rexford, and Prateek Mittal. 2019. SICO: Surgical interception attacks by manipulating bgp communities. In *ACM CCS*.

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

[5] CAIDA. 2022. CAIDA. Mapping Autonomous Systems to Organizations. https://www.caida.org/research/topology/as2org/.

[6] Enke Chen and Quaizar Vohra. 2007. BGP Support for Four-octet AS Number Space. RFC 4893. https://www.rfc-editor.org/info/rfc4893

[7] CISCO. 2005. Remotely Triggered Black Hole Filtering - Destination Based and Source Based. Cisco White Paper, http://www.cisco.com/c/dam/en_us/about/security/intelligence/blackhole.pdf.

[8] Brivaldo A. Da Silva Jr, Paulo Mol, Osvaldo Fonseca, Ítalo Cunha, Ronaldo A. Ferreira, and Ethan Katz-Bassett. 2022. Automatic Inference of BGP Location Communities. In *ACM SIGMETRICS*.

[9] Christoph Dietzel, Anja Feldmann, and Thomas King. 2016. Blackholing at IXPs: On the effectiveness of DDoS mitigation in the wild. In *PAM*.

[10] Xenofontas Dimitropoulos, Dmitri Krioukov, Marina Fomenkov, Bradley Huffaker, Young Hyun, kc claffy, and George Riley. 2007. AS relationships: Inference and validation. In *ACM SIGCOMM CCR*.

[11] Benoit Donnet. 2007. BGP Communities Classification. https://web.archive.org/web/20080423050836/http://inl.info.ucl.ac.be:80/communities.

[12] Benoit Donnet and Olivier Bonaventure. 2008. On BGP Communities. In *ACM SIGCOMM CCR*.

[13] Emile Aben. 2022. BGP Community propagation. https://observablehq.com/@emileaben/bgp-community-propagation.

[14] EuroIX. 2019. Large BGP Communities. https://www.euro-ix.net/en/forixps/large-bgp-communities/.

[15] Anja Feldmann, Olaf Maennel, Z. Morley Mao, Arthur Berger, and Bruce Maggs. 2004. Locating Internet Routing Instabilities. *ACM SIGCOMM CCR*.

[16] Pierre Francois, Bruno Decraene, Cristel Pelsser, Keyur Patel, and Clarence Filsfils. 2018. Graceful BGP Session Shutdown. RFC 8326. https://www.rfc-editor.org/info/rfc8326

[17] Lixin Gao. 2001. On Inferring Autonomous System Relationships in the Internet. *IEEE/ACM ToN*.

[18] Vasileios Giotsas, Christoph Dietzel, Georgios Smaragdakis, Anja Feldmann, Arthur Berger, and Emile Aben. 2017. Detecting Peering Infrastructure Outages in the Wild. In *ACM SIGCOMM*.

[19] Vasileios Giotsas, Thomas Koch, Elverton Fazzion, Ítalo Cunha, Matt Calder, Harsha V. Madhyastha, and Ethan Katz-Bassett. 2020. Reduce, Reuse, Recycle: Repurposing Existing Measurements to Identify Stale Traceroutes. In *ACM IMC*.

[20] Vasileios Giotsas, Matthew Luckie, Bradley Huffaker, and kc claffy. 2014. Inferring complex AS relationships. In *ACM IMC*.

[21] Vasileios Giotsas, Georgios Smaragdakis, Christoph Dietzel, Philipp Richter, Anja Feldmann, and Arthur Berger. 2017. Inferring BGP Blackholing Activity in the Internet. In *ACM IMC*.

[22] Vasileios Giotsas and Shi Zhou. 2013. Improving the Discovery of IXP Peering Links through Passive BGP Measurements. In *IEEE INFOCOM WKSHPS*.

[23] Vasileios Giotsas, Shi Zhou, Matthew Luckie, and kc claffy. 2013. Inferring multilateral peering. In *ACM CoNEXT*.

[24] Jakob Heitz, Job Snijders, Keyur Patel, Ignas Bagdonas, and Nick Hilliard. 2017. BGP Large Communities Attribute. RFC 8092. https://rfc-editor.org/rfc/rfc8092

[25] Bradley Huffaker and Vaseilios Giotsas. 2016. AS Relationships – with geographic annotations. http://www.caida.org/data/as-relationships-geo/.

[26] Geoff Huston. 2004. NOPEER Community for Border Gateway Protocol (BGP) Route Scope Control. RFC 3765. https://www.rfc-editor.org/info/rfc3765

[27] Yuchen Jin, Colin Scott, Amogh Dhamdhere, Vasileios Giotsas, Arvind Krishnamurthy, and Scott Shenker. 2019. Stable and Practical AS Relationship Inference with ProbLink. In *USENIX NSDI*.

[28] Mattijs Jonker, Aiko Pras, Alberto Dainotti, and Anna Sperotto. 2018. A First Joint Look at DoS Attacks and BGP Blackholing in the Wild. In *ACM IMC*.

[29] Brivaldo A. Da Silva Junior, Paulo Mol, Osvaldo Fonseca, Ítalo Cunha, Ronaldo A. Ferreira, and Ethan Katz-Bassett. 2021. BGP Communities – Supplemental Material. https://github.com/TopoMapping/bgp-communities.

[30] Thomas King, Christoph Dietzel, Job Snijders, Gert Doering, and Greg Hankins. 2016. BLACKHOLE Community. RFC 7999. https://rfc-editor.org/rfc/rfc7999.txt

[31] Thomas Krenc, Robert Beverly, and Georgios Smaragdakis. 2020. Keep your Communities Clean: Exploring the Routing Message Impact of BGP Communities. In *ACM CoNEXT*.

[32] Thomas Krenc, Robert Beverly, and Georgios Smaragdakis. 2021. AS-Level BGP Community Usage Classification. In *ACM IMC* (Virtual Event).

[33] Thomas Krenc, Matthew Luckie, Alexander Marder, and kc claffy. 2023. Data supplement for "Coarse-grained Inference of BGP Community Intent". https://publicdata.caida.org/datasets/supplement/2023-imc-bgpcomms/.

[34] Tony Li, Ravi Chandra, and Paul S. Traina. 1996. BGP Communities Attribute. RFC 1997. https://rfc-editor.org/rfc/rfc1997.txt

[35] Matthew Luckie, Bradley Huffaker, Amogh Dhamdhere, Vasileios Giotsas, and kc claffy. 2013. AS relationships, customer cones, and validation. In *ACM IMC*.

[36] Matthew Luckie, Bradley Huffaker, Alexander Marder, Zachary Bischof, Marianne Fletcher, and kc claffy. 2021. Learning to Extract Geographic Information from Internet Router Hostnames. In *ACM CoNEXT*.

[37] Fabricio Mazzola, Pedro Marcos, and Marinho Barcellos. 2022. Light, Camera, Actions: characterizing the usage of IXPs' action BGP communities. In *ACM CoNEXT*.

[38] OneStep. 2015. BGP Communities Guide. https://onestep.net/communities.

[39] Bruno Quoitin and Olivier Bonaventure. 2002. *A survey of the utilization of the BGP community attribute*. Internet-Draft draft-quoitin-bgp-comm-survey-00. Internet Engineering Task Force. https://datatracker.ietf.org/doc/html/draft-quoitin-bgp-comm-survey-00

[40] Yakov Rekhter, Susan Hares, and Tony Li. 2006. A Border Gateway Protocol 4 (BGP-4). RFC 4271. https://rfc-editor.org/rfc/rfc4271.txt

[41] Philipp Richter, Georgios Smaragdakis, Anja Feldmann, Nikolaos Chatzis, Jan Boettger, and Walter Willinger. 2014. Peering at peerings: On the role of IXP route servers. In *ACM IMC*.

[42] RIPE. 2021. RIS - RIPE Network Coordination Centre. http://ris.ripe.net/.

[43] RouteViews. 2021. University of Oregon RouteViews project. http://www.routeviews.org/.

[44] Job Snijders. 2022. Request for BGP Community-to-text mappings for BGP Looking Glass. https://mailman.nanog.org/pipermail/nanog/2022-September/220625.html.

[45] Stichting NLNOG. 2023. NLNOG Looking Glass. https://github.com/NLNOG/lg.ring.nlnog.net.

[46] 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 *ACM IMC*.

[47] Dan Tappan, Yakov Rekhter, and Srihari R. Sangli. 2009. 4-Octet AS Specific BGP Extended Community. RFC 5668. https://www.rfc-editor.org/info/rfc5668

[48] Philipp Winter, Ramakrishna Padmanabhan, Alistair King, and Alberto Dainotti. 2019. Geo-locating BGP prefixes. In *TMA*.

A APPENDIX

A.1 Ethics

We use publicly available BGP data and publicly available documentation to build our dictionaries that inform our method. We believe that our coarse-grained inferences and the content of our paper raise no ethics concerns.