On the Quality of BGP Route Collectors for iBGP Policy Inference

Luca Cittadini* Stefano Vissicchio† Benoit Donnet§
*Roma Tre University †Université catholique de Louvain § Université de Liège
*ratm@dia.uniroma3.it †stefano.vissicchio@uclouvain.be §benoit.donnet@ulg.ac.be

***Abstract*—A significant portion of what is known about Internet routing stems out from public BGP datasets. For this reason, numerous research efforts were devoted to (i) assessing the (in)completeness of the datasets, (ii) identifying biases in the dataset, and (iii) augmenting data quality by optimally placing new collectors. However, those studies focused on techniques to extract information about the AS-level Internet topology.**

In this paper, we show that considering different metrics influences the conclusions about biases and collector placement. Namely, we compare AS-level topology discovery with iBGP policy inference. We find that the same datasets exhibit significantly diverse biases for these two metrics. For example, the sensitivity to the number and position of collectors is noticeably different. Moreover, for both metrics, the marginal utility of adding a new collector is strongly localized with respect to the proximity of the collector. Our results suggest that the "optimal" position for new collectors can only be defined with respect to a specific metric, hence posing a fundamental trade-off for maximizing the utility of extensions to the BGP data collection infrastructure.

I. Introduction

The Internet is a global network connecting different *domains* or *Autonomous Systems* (ASes) together. Each AS is typically administered by a different *Internet Service Provider* (ISP). In the following, we will use the terms AS and ISP interchangeably. To exchange inter-domain routing information on Internet destinations with other ASes, each ISP needs to configure the *Border Gateway Protocol* [1] (BGP). Actually, BGP consists of two protocols: eBGP and iBGP. *External BGP* (eBGP) is used between routers in different ASes. Conversely, *internal BGP* (iBGP) is used by routers in a single AS to distribute the routes learned via eBGP.

One key feature of both eBGP and iBGP is the support for highly customizable *routing policies*. ISPs configure eBGP policies to tune inter-domain routing choices at a fine level of granularity [2]. Moreover, to meet traffic engineering goals within their own network, they can and do [3] use *iBGP policies*, e.g., policies on internal BGP routers. While eBGP routing policies are deeply studied (e.g., [2], [4], [5]), few things are known about iBGP policies. Nevertheless, estimating the popularity of iBGP policies has important implications [6] on the applicability of previously proposed theoretical insights (e.g., [7]), protocol enhancements (e.g., [8]), and tools (e.g., [9]), which assume that policies are applied solely in eBGP.

The presence of iBGP policies can be detected based on public BGP data [3]. BGP datasets such as RIPE RIS [10] and Oregon Routeviews [11] consist of sets of BGP messages stored by *collector peers* (CPs).

In this paper, we analyze the biases of public BGP datasets with respect to iBGP policy inference. In particular, we compare iBGP policy inference biases with those affecting Internet topology discovery techniques [12], [13]. We evaluate three main factors that can bias inter-domain link detection, iBGP policy inference, or both metrics: *prefix visibility from the CPs*, *number of available CPs*, and *position of the CPs*. The importance of those factors derives from their direct impact on the number and kind of BGP messages received by CPs, hence on their information on the BGP routing system. To analyze the impact of the considered factors on both inter-domain link detection and iBGP policy inference, we systematically evaluate the respective sensitivity of the two metrics. Due to the unavailability of ground truth, we quantify the sensitivity by carefully slicing an initial dataset into multiple smaller datasets, in such a way that the bias is exposed.

Our results show that the factors affecting the quality of iBGP policy inference significantly differ from those identified by previous work for topology discovery. Interestingly, the ability to detect iBGP policies is more sensitive than topology discovery to the per-CP prefix visibility. Moreover, the quality of iBGP policy inference is also more sensitive to the number of CPs, especially because BGP datasets tend to exhibit less redundancy for iBGP policy inference than for topology discovery. This suggests that, unless new CPs are added deliberately to improve the quality of link detection, extensions of the monitoring infrastructure are likely to be more beneficial for iBGP policy inference. More in general, our findings suggest that the quality of BGP datasets vary significantly depending on the considered metric.

Finally, complementing recent research efforts [13], [14], we study the marginal utility of new CPs for inter-domain link and iBGP policy inference. We observe that, for both cases, the marginal utility is highly localized in the proximity of the CP, even though the localization effect is more evident for link detection than for iBGP policy inference. Our results also show that new CPs can hardly be placed in such a way to maximize the increased accuracy for both metrics at the same time.

The rest of this paper is organized as follows. Sec. II presents the required background for this paper. Sec. III provides details on our dataset and on our bias comparison methodology. Sec. IV evaluates the impact of prefix visibility at the CPs. Sec. V discusses the impact of the number of CPs. Sec. VI analyses the impact of the CP position. Sec. VII relates this paper to existing literature on BGP datasets. Finally, conclusions are drawn in Sec. VIII.

Stefano Vissicchio is a postdoctoral researcher of the Belgian fund for scientific research (F.R.S.-FNRS)

ISBN 978-3-901882-58-6 © 2014 IFIP

Step	Criterion
1	Prefer routes with higher `local-preference`
2	Prefer routes with lower `as-path` length
3	Prefer routes with lower `origin`
4	Among the routes received from the same AS neighbor, prefer those having lower `MED`
5	Prefer routes learned via eBGP to those learned via iBGP
6	Prefer routes with lower IGP metric to the egress point
7	Prefer the route having the lowest `router-id`

TABLE I. Steps in the BGP decision process.

II. Background

This section presents the required background for this paper. We first focus on the basics of BGP (Sec. II-A). Next, we introduce inference techniques for AS-level Internet topology (Sec. II-B) and for iBGP policies (Sec. II-C).

A. BGP and BGP Collector Peers (CPs)

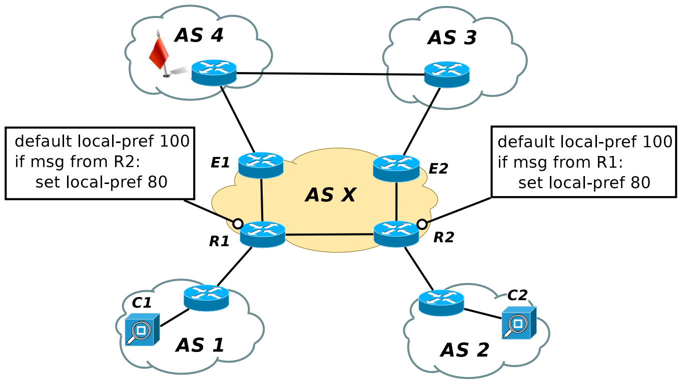
In BGP [1] (both eBGP and iBGP), reachability information of IP destinations is represented by *routes*. Each route associates a destination prefix to a set of *attributes*. When a BGP router learns a route to a given prefix, it possibly modifies some of the attributes in the route, picks its *best* route for that prefix, and sends its best route to all its *BGP neighbors*. The best route is selected by executing the *BGP decision process* summarized in Table I. Essentially, a route is selected as best based on the associated attribute values. By configuring routers to modify BGP attributes, operators can accommodate specific high-level goals, e.g., achieving fine-grained control of outgoing traffic [15].

In this paper, we apply inference techniques to public BGP datasets. Those datasets consist of BGP routes collected by geographically-distributed vantage points deployed within RIPE RIS [10] and Oregon RouteViews [11] projects. BGP vantage points are technically called *collector peers* (CPs).

As an illustration, consider the BGP network shown in Fig. 1. In this figure, each AS is depicted as a cloud, router icons show the placement of BGP routers, and square magnifier icons represent CPs. Black segments represent BGP sessions, and the flag icon indicates a prefix p originated by $AS\ 4$. In the example, ASX uses iBGP route reflection [16] as indicated by the absence of a full-mesh of iBGP sessions inside ASX. For simplicity, we assume that the `local-preference` attribute is set to a default value of 100 by each BGP router in the figure, i.e., we temporarily ignore the callouts in the example. As soon as $E1$ receives a BGP route to p by its eBGP neighbor in $AS\ 4$, it runs the BGP decision process and selects it as the best route to p. In fact, even if $E1$ receives a route to p traversing $E2$ and $AS\ 3$, the direct route through its eBGP neighbor in $AS\ 4$ has a shorter `as-path` length, hence it will be preferred. $E1$ will then propagate its best route to its iBGP neighbor $R1$, which will select it as its best. Eventually, the route with `as-path` $(AS\ 1, AS\ X, AS\ 4)$ will reach $C1$, which will store it in the standard MRT format [17].

B. AS-level Topology Discovery

While hiding the internal structure of ISPs, eBGP routes contain explicit information about traversed inter-domain paths at the AS-level granularity. Indeed, one of the attributes carried in each eBGP route is the `as-path`, i.e., the sequence of ASes traversed by the given route. Thanks to the `as-path` attribute, we can partially reconstruct the *Internet AS-level topology* [18], where inter-domain links are represented by edges in a graph. Internet AS-level topology discovery can then be based on BGP routes only. Given a set of BGP routes, an inter-domain link between two ASes $AS\ 1$ and $AS\ 2$ can be detected each time $AS\ 1$ and $AS\ 2$ appear as consecutive ASes in the `as-path` of a BGP route. This approach has a one-side error, i.e., it never reports a non-existing link as existing but it may not be able to detect all the inter-domain links in the Internet. In the following, we refer to inter-domain links inferred through this technique as *detected links*.



default local-pref 100
if msg from R2:
set local-pref 80

default local-pref 100
if msg from R1:
set local-pref 80

Fig. 1. Example BGP network with two vantage points (i.e., $C1$ and $C2$) and an AS deploying iBGP policies (i.e., $AS\ X$).

Consider the example in Fig. 1. Assume that $C1$ stored, at a given time, a BGP route to prefix p with AS path $(AS\ 1, AS\ X, AS\ 4)$, e.g., as a result of the BGP route propagation described in Sec. II-A. From this route, we can detect inter-domain links $(AS\ 1, AS\ X)$ and $(AS\ X, AS\ 4)$. Unfortunately, it provides no evidence of the presence of $AS\ 3$ nor of any of its inter-domain links with $AS\ X$ and $AS\ 4$. We might be able to reconstruct the full AS-level topology in the example if $C1$ collected routes to other destinations (e.g., prefixes originated by $AS\ 3$), or if other public CPs (e.g., $C2$) provided different BGP routes traversing $AS\ 3$.

C. iBGP Policies

Most of the existing literature (e.g., [2], [7], [9]) assumes that iBGP acts as a simple dispatcher of inter-domain routes, with routing policies applied exclusively to eBGP sessions. However, a significant amount of ISPs do apply iBGP policies [3], in the sense that they intentionally modify attributes on iBGP messages, e.g., to increase internal routing flexibility and improve traffic engineering. Consider again the example in Fig. 1. As shown by the callouts in the figure and highlighted by the filled-in cloud, $AS\ X$ is applying iBGP policies. In particular, those policies ensure that BGP routers $R1$ and $R2$ respectively prefer routes through $E1$ and $E2$, e.g., for intra-domain traffic load balancing.

To estimate the number of ISPs that apply iBGP policies, Cittadini et al. [3] proposed a technique computing a conservative lower bound. The inference is based on looking in public BGP datasets for evidences of BGP routers in the same AS that steadily select distinct routes that are not equally good up to the first three steps of the BGP decision process (see

Table I). Indeed, assuming connected iBGP topologies inside each AS, only routes that are equally good up through the first three BGP decision steps can be steadily selected by iBGP routers in the absence of iBGP policies (see, e.g., [19]). Hence, whenever two or more routes with different `as-path` lengths are simultaneously active at a given $AS\ X$, it can be inferred that $AS\ X$ is applying iBGP policies. To compute the set of routes that are simultaneously active at each AS, Cittadini et al. leverage the technique by Mühlbauer et al. [20] and compare BGP routes gathered by different vantage points. Observe that, by definition, this technique has a one-side error, that is, it never misreports ASes deploying iBGP policies. In the following, we refer to ASes which we infer applying iBGP policies with this technique as *inferred ASes*.

As an illustration of the iBGP policy inference technique, consider again Fig. 1. Two CPs $C1$ and $C2$ are respectively located in $AS\ 1$ and $AS\ 2$. Consider the BGP routes collected by those two CPs after BGP route propagation over the network. Given the iBGP policies applied by $AS\ X$, $C1$ and $C2$ will store a BGP route with `as-path` $(AS\ 1, AS\ X, AS\ 4)$ and $(AS\ 2, AS\ X, AS\ 3, AS\ 4)$, respectively. From those two routes, it is possible to infer that routers in $AS\ X$ simultaneously select two different routes to prefix p, a condition known as *route diversity* [20]. Moreover, the two routes have different `as-path` length. Since this cannot happen in a connected iBGP topology without iBGP policies, we can conclude that $AS\ X$ is applying iBGP policies.

III. Methodology

In this section, we describe the public BGP dataset that we study (Sec. III-A) and the methodology that we adopt to expose biases for the considered inference techniques (Sec. III-B).

A. Dataset

Our dataset consists of BGP routing table dumps from RIS [10] collectors on September 16th, 2012, i.e., a random day reasonably far from any Western festivity. Since we additionally use table dumps on the same day (i.e., September 16th) for every year between 2009 and 2011 as validation datasets (i.e., to validate our results), we exclude all the collectors that were not up and running during any of the selected dates, namely, `rrc02`, `rrc07`, `rrc08`, `rrc09`, `rrc14`, and `rrc16`. Furthermore, we perform sample tests to cross-check our results with datasets from few other random days in September and October.

Observe that our dataset is much smaller than the one used in [3]: not only we select a subset of route collectors in order to have consistent snapshots over time, but we restrict ourselves to BGP table dumps, disregarding routing updates. It is known [3] that a significant amount of real-world ISPs do apply iBGP policies. Of course, the estimates obtained in our experiments are much lower than the one described in [3]. In this paper, however, we specifically focus on understanding what are the intrinsic limitations of the data provided by public BGP collectors with respect to our ability to pinpoint ISPs that apply iBGP policies. To this end, considering only routing tables does not introduce biases, as the inference technique is based on snapshots of routing information at a given point in time. In fact, considering BGP updates can be seen as just a way to extract multiple snapshots near in time.

B. Assessing Biases Without Ground Truth

It is known that public BGP data are biased. For instance, they tend to capture certain types of links between ASes much better than others [13], [21]. For both AS-level topology discovery and iBGP policies inference, estimating the bias is nontrivial due to the lack of ground truth. Indeed, ISPs are very reluctant to disclose details about their BGP configurations. For this reason, we take the full dataset as a baseline, and study each inference technique on selected subsets of the dataset, which we call *sub-datasets*. Namely, we extract sub-datasets to expose a specific bias with respect to the full dataset.

We analyze multiple potential biasing factors (i.e., number of CPs, position of CPs, prefix visibility for each CP). For each factor, we pick specific sub-datasets, we perform link discovery and iBGP policy inference on them, and we compare the respective results with each other, and with those obtained using the full dataset. Such a differential approach provides us with a repeatable and statistically significant comparison of the relative influence of the different biasing factors. Moreover, it allows us to compare the two inference techniques with each other. However, our methodology cannot be used to quantify the absolute biases that are possibly present in the full dataset. Given that the ground truth is not publicly available, the same drawback also applies to any other measurement methodology.

To make sure that our analyses are not biased by the specific selection of the full dataset, we cross-validate our results by applying the same methodology to our *validation datasets*, consisting of RIS table dumps taken on September 16th of years 2009, 2010, and 2011. For the sake of brevity, since we got extremely similar results across all validation datasets, we only report the results on the full 2012 dataset in the following.

IV. Prefix Visibility

To encourage wide adoption, BGP data collection infrastructures [10], [11] do not enforce standardization across CPs. For example, no restrictions are imposed on the quantity and quality of BGP messages that CPs are expected to provide. As a consequence, not all CPs provide information about the full set of globally routable prefixes in the Internet, i.e., the full Routing Information Base (*RIB*). Ideally, the full RIB contains all the prefixes in the "default free zone" of the Internet, which is the portion of the Internet that cannot use a default route to ensure global connectivity. However, such a definition has been shown to be not applicable in practice [21]. Throughout the paper, we define the full RIB to be the maximum number of unique prefixes seen by a single CP.

Consider Fig. 2. The line with square points indicates the cumulative distribution function (CDF) of the percentage of the full RIB seen by CPs in our dataset. The distribution is bimodal, with roughly 80% of the CPs having visibility of less than 30% of the RIB, and roughly 20% of the CPs having visibility of more than 85%. We denote CPs having visibility less than or greater than 50% of the RIB as "partial CPs" or "full CPs", respectively. The distribution is dominated by CPs deployed in Internet Exchange Points (IXPs)[1], which are more likely to have visibility of only a restricted subset of prefixes.

[1] information on the location of CPs is extracted from the RIS Web site [10]

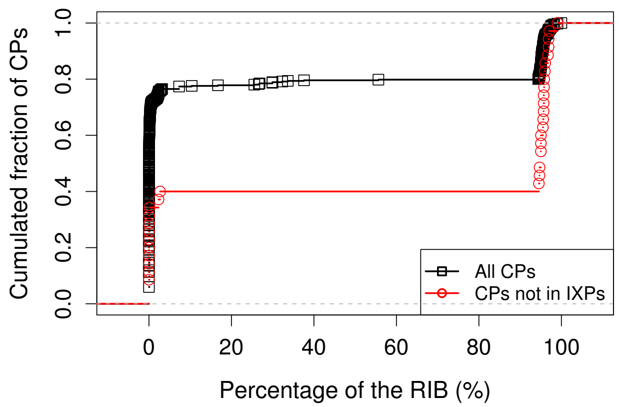


Fig. 2. CDF of the percentage of the global RIB seen by each CP.

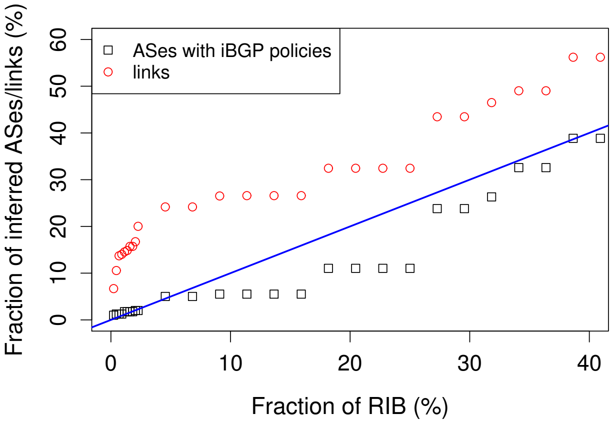


Fig. 3. Comparison of the impact of prefix visibility on link detection (circles) and iBGP policy inference (squares).

The line with circle points in Fig. 2 shows the distribution of the subset of CPs that are not located in an IXP. A plausible explanation is that operators often configure CPs with the same routing policies that they apply to other members of the same IXP. The smaller differences among CPs in the same class can be explained by route aggregation and specific routing policies applied by ISPs [13].

It is clear that having CPs with higher prefix visibility is beneficial for topology discovery. In fact, there is a very high correlation between seeing more prefixes and being aware of a bigger portion of the Internet AS-level graph. Our ability to infer iBGP policies also increases as the number of prefixes seen by CPs increase, because there is a higher chance that two CPs can observe route diversity. While both metrics benefit from increased visibility, there is a fundamental difference: link detection benefits from CPs seeing *different* prefixes, whereas iBGP policy inference benefits from CPs seeing the *same* prefix. We now study whether this difference leads to different sensitivity of the two metrics.

To quantify the sensitivity of both metrics to the number of prefixes seen by CPs, we generate sub-datasets (as defined in Sec. III) by filtering out CPs that see more than k prefixes, for $k = 1,000, 2,000, \ldots, 10,000, 20,000, \ldots 200,000$. Each sub-dataset contains CPs tracking a fraction, ranging from roughly 0.2% to about 42%, of the full RIB. For each sub-dataset, we compute the ratio of the number of inferred ASes in the sub-dataset and the number of inferred ASes in the full dataset. Moreover, we compute the same ratio considering detected links instead of inferred ASes.

Fig. 3 shows the results of those experiments. Each circle (resp. square) data point (x, y) indicates that CPs with a visibility of at most $x\%$ of the full RIB are able to detect $y\%$ of the links (resp., ASes) that can be discovered by considering the full dataset. To ease comparison, we also plot a straight line representing linear increase (i.e., $y = x$). Fig. 3 highlights different absolute values but similar trends for link detection and iBGP policy inference. Both metrics exhibit a growing trend which is roughly linear but quantized. The quantization effect can be easily explained by the skewed distribution of prefix visibility (see Fig. 2). The difference between the two curves is more significant when we constrain the visibility to be less than 4% of the RIB. A plausible explanation is that, for those sub-datasets, there are so few CPs and so few prefixes that it is highly unlikely that two CPs collect information on the same prefix, which is a precondition to infer iBGP policies.

V. Number and Position of CPs

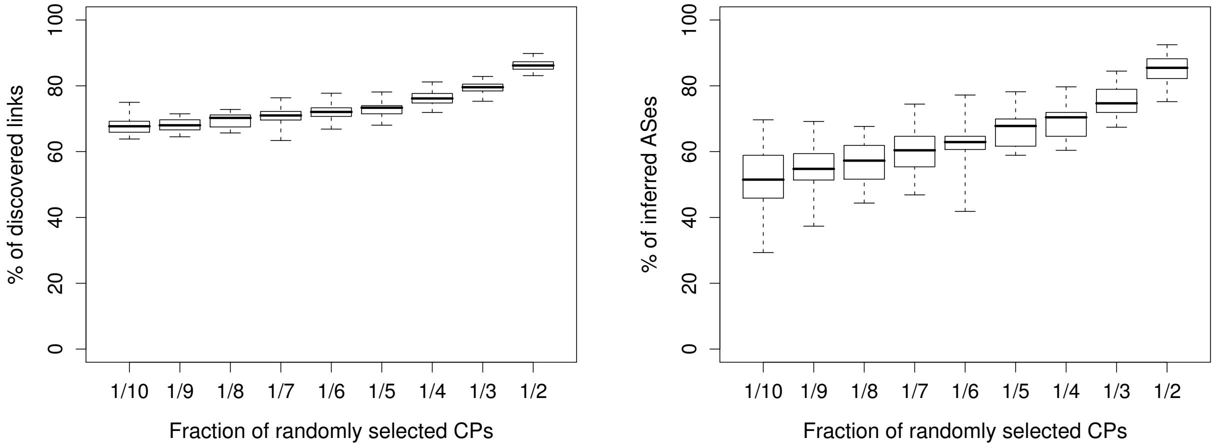
In this section, we analyze the sensitivity of our metrics with respect to the number and position of CPs in our datasets. In particular, Sec. V-A contains analyses on the absolute number of CPs, and Sec. V-B studies the impact of the position of CPs in the hierarchy of ASes in the Internet.

A. Number of CPs and Information Redundancy

Intuitively, a higher number of CPs corresponds to more BGP routes in the dataset, hence increased accuracy for both link detection and iBGP policy inference. We now study whether the impact of the number of CPs is proportional for both metrics. To this end, we build sub-datasets by selecting CPs from the full dataset uniformly at random. More precisely, for $n = 1, \ldots, 10$, we randomly build a sub-dataset containing $1/n$ of the CPs in the dataset. For statistical relevance, we repeat the process 30 times, resulting in 30 sub-datasets for each value of n. Then, we compute the number of detected links and of inferred ASes on each sub-dataset.

The boxplot in Fig. 4 summarizes the results of these experiments. The lower (resp. upper) end of the whisker represents the minimum (resp. maximum) value measured across the 30 sub-datasets sharing the same value of n. The lower (resp. upper) end of the box represents the first (resp. third) quartile and the thick horizontal line within the box represents the median.

The plot highlights that, as the number of CPs in each sub-dataset decreases, the number of detected links exhibits a smoother decreasing trend with respect to the inferred ASes. For example, we are able to detect roughly 70% of the links using just $1/10^{\text{th}}$ of the available CPs (see Fig. 4a). With the same sub-dataset, we can only infer about 50% of the ASes (see Fig. 4b). As a matter of comparison, we need at least three times more CPs ($1/3^{\text{rd}}$ as opposed to $1/10^{\text{th}}$) to infer a comparable proportion (roughly 70%) of the ASes with iBGP policy. In addition, the variability of both metrics increases as the cardinality of sub-datasets decreases. However, for sub-datasets with the same cardinality, iBGP policy inference exhibits more variability than link detection. This can be



(a) Links.

(b) iBGP Policies.

Fig. 4. Distribution of the number of the detected links and inferred ASes for different random selections of CPs. The full dataset (i.e., the 100% mark on the y-axis) contains 122160 unique links and 399 ASes for which we were able to detect iBGP policies.

explained as a consequence of the disproportional sensitivity of the two metrics to prefix visibility (see Sec. IV). Indeed, CPs with high visibility are more likely to be excluded from smaller sub-datasets.

Fig. 4 suggests that our dataset contains more redundant information for link detection than for iBGP policy inference. However, by comparing those sub-datasets, we can only extract qualitative information. Indeed, the considered sub-datasets do not take into account that specific detected links (or, respectively, inferred ASes) have potentially different redundancy. For example, there might be links (or ASes) that are intrinsically easier to infer, hence more likely to be detected in each sub-dataset.

In order to remove this uncertainty and quantify the difference in redundancy between the two metrics, we introduce the concept of *critical CPs*. A set of CPs is critical for a link (resp. AS) if removing them from the dataset makes us unable to detect that link (resp. AS). For inter-domain links, computing the set of critical CPs corresponds to find all CPs that can see each link. For ASes with iBGP policies, however, computing the set of critical CPs requires a deeper understanding of the technique [20] that we use to capture route diversity. For an AS A and a prefix p, let $\mathcal{R}_{A,p}$ be the set of distinct routes to p from A. Note that each route in $\mathcal{R}_{A,p}$ is seen by at least one CP. If $\mathcal{R}_{A,p}$ contains a single route ($|\mathcal{R}_{A,p}| = 1$) or no routes ($|\mathcal{R}_{A,p}| = 0$), then we are unable to infer iBGP policies for AS A using routes to prefix p. The set of critical CPs for A is then defined as the minimum number of CPs removing which we have $|\mathcal{R}_{A,p}| < 2$ for all prefixes p.

For each detected link and inferred AS, we compute the corresponding set of critical CPs, and we calculate the ratio of the number of critical CPs to the total number of available CPs. For both detected links and inferred ASes, a higher value of the ratio corresponds to a greater redundancy in the dataset, by definition of critical CP.

Fig. 5 shows the CDF of the critical CP ratio for link discovery and iBGP policy inference. Each circle (resp. square) data point (x, y) indicates that $y \times 100\%$ of detected links (resp. inferred ASes) have a critical CP set of at most $x \times 100\%$ of the available CPs. This analysis confirms that the information provided by the CPs in the dataset for link discovery is more redundant than for iBGP policy inference. Indeed, the size of the critical CP set is generally bigger for link detection than it is for iBGP policy inference. For example, the most redundant inferred ASes have a set of critical CPs of circa 10% of the available CPs, while roughly half of the detected links show more redundancy. It is also interesting to observe that, for both metrics, roughly 20% of the links or ASes have very few critical CPs (less than 1% of the CPs). This suggests that the position of a CP have a strong correlation with the amount of information that we can extract from it. We deepen the study of this aspect in Section VI.

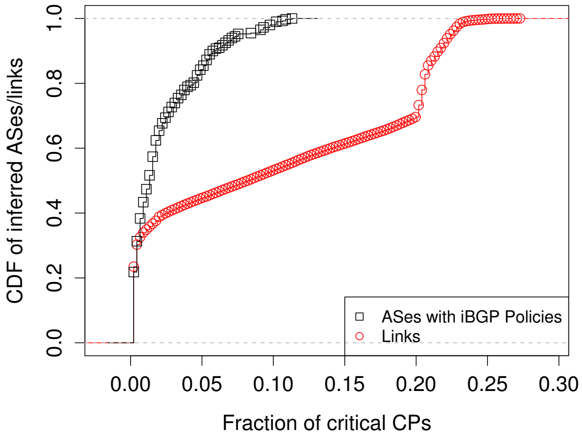
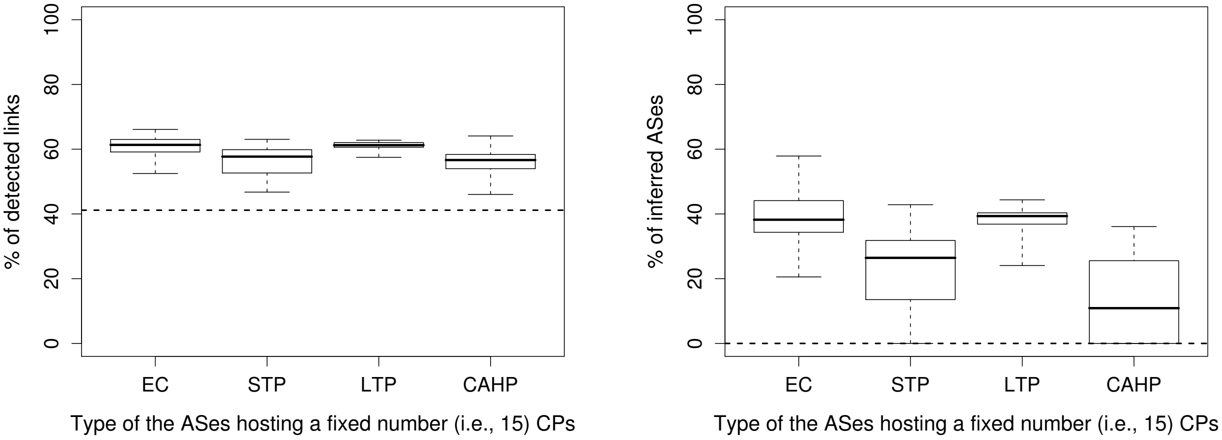


Fig. 5. Comparison of the fraction of critical CPs for topology discovery (circles) and iBGP policy inference (squares).

B. Position in the Internet Hierarchy

It is known that, due to how ISPs apply BGP policies, the position of CPs in the Internet hierarchy affects their ability to measure the AS-level graph. For example, Oliveira et al. [12] classify each missing link as either "invisible" or "hidden" to distinguish whether the dataset contains at least one CP that should theoretically be able to detect that link. *Customer-provider* links (i.e., links between two ASes where one is buying transit service from the other) cannot be invisible,



(a) Links.

(b) iBGP Policies.

Fig. 6. Breakdown of CP inference power per class of ASes. The dashed lines represent the number of links or ASes with iBGP policies detected in the 95% of the experiments.

as opposed to *peer-peer* links (i.e., links between two ASes based on mutual exchange of selected traffic free of charge). In particular, a peer-peer link between ASes A and B can only be detected by CPs located in ASes that are (possibly indirect) customers of either A or B. This bias in the dataset is known, and a number of techniques have been proposed to estimate the real number of peer-peer links in the Internet. The most recent estimates (see, e.g., [22]) are several times higher than the number that can be inferred from public BGP datasets. Based on this observation, existing literature concurs in placing new CPs at the bottom of the Internet hierarchy in order to maximize the number of discovered peer-peer links.

We perform experiments to understand the correlation between the position of CPs in the Internet hierarchy and our ability to infer links and ASes. Based on the classification of ASes proposed by Dhamdhere and Dovrolis [23], we classify each AS in one of four categories: *Enterprise Customers* (EC – mostly users that are not Internet access, transit, or content providers), *Small Transit Providers* (STP – regional ISPs), *Large Transit Providers* (LTP – international ISPs), and *Content, Access and Hosting Providers* (CAHP – ISPs offering Internet access and/or server hosting and/or content without providing any kind of transit or access). For each category, we build sub-datasets by randomly extracting CPs located in an AS belonging to that category. Observe that the distribution of CPs is not uniform across the categories, with 14% of CPs being in ECs, 26% in STPs, 6% in LTPs, and 54% in CAHPs. To avoid being biased by the differences in the distribution, we extract the same number of CPs for each category. We select the number of CPs by taking the half of the number of CPs in the category having the least, resulting in sub-datasets of 15 CPs each. For statistical significance, we repeat the random extraction of sub-datasets 30 times.

Fig. 6 shows the results of our experiments on those sub-datasets. The boxes in the two plots show median, first and third quartile, maximum and minimum values, with the same graphical convention as in Fig. 4. Additionally, to highlight the differences in the overlap of information between different sub-datasets, we plot an horizontal line, which we refer to as the *baseline*, and represent the number of links or ASes that we were able to infer in at least 95% of all sub-datasets.

First of all, we note the difference in the percentages between inferred ASes and detected links. For every AS category, the percentage of inferred ASes is significantly smaller than the corresponding one of detected links, indicating that our ability to infer iBGP policies is more dependent on the presence of CPs in multiple AS categories. This is especially evident by comparing the baselines for the two metrics. Note that sub-datasets based on CPs in STPs and CAHPs show more variability for both metrics. This could be an artifact of how we extracted sub-datasets. In fact, since STPs and CAHPs host the majority of CPs, extracting 15 CPs at random gives a higher likelihood of including partial CPs in the sub-datasets. Median values are less affected than variability by such an extraction artifact. By looking at median values, we conclude that for both metrics CPs in ECs and LTPs contribute more to our inference ability than CPs in the other categories.

VI. Marginal Utility of New CPs

In this section, we study the *marginal utility* of CPs for both link detection and iBGP policy inference. Such a study can be at the basis of CP placement strategies that try to optimize the position of new CPs.

Computing the marginal utility for any given CP is actually straightforward. Indeed, we can simply build a sub-dataset by removing that CP, and compare the inference results on the sub-dataset with the ones on the full dataset. Unfortunately, when computed with this direct approach, the marginal utility is close to zero for all CPs, mainly because of redundant information in the dataset (see Sec. V-A). Interestingly, Barford et al. [24] observed a similar effect for the marginal utility of adding new traceroute sources. Thus, direct indicators of marginal utility are almost meaningless for further analyses, e.g., optimal CP placement.

To overcome this difficulty, we resort to using two indirect indicators. In Sec. VI-A, we study how the marginal utility of a CP is influenced by the AS-path distance between the CP and the links (or ASes with iBGP policies) that it can infer. In Sec. VI-B, we consider the diversity between the BGP routes that it collects and those collected by the other CPs in the dataset, and we study the relationship between such a diversity and the marginal utility of that CP.

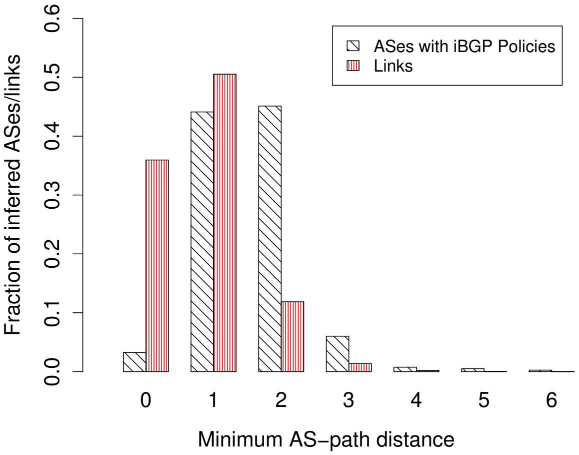


Fig. 7. Distribution of the distance of ASes to the closest CP.

A. AS-Path Distance

Intuitively, the marginal utility of CPs for topology discovery depends on the distance from the link to be inferred. In particular, a CP C which is close to a link l has high probability to collect a BGP route traversing that link. Conversely, as the distance between C and l increases, not only the probability that C can detect l decreases, but also the probability that other CPs can detect l increases. For these reasons, a new CP is unlike to provide useful information on a distant link. For iBGP policy inference, instead, there is no intuitive reason why the ability of a CP to detect an AS should be higher or lower depending on its distance to that AS. In order to study the relationship between marginal utility and proximity for both link detection and iBGP policy inference, we define the concept of *AS-path distance*. The AS-path distance $dist(C, A)$ between a CP C and an AS A is the minimum number of hops from the AS hosting C to A in the `as-path` attribute of any BGP route collected by C. In particular, if C collects a BGP route such that the first hop in the `as-path` is X, then $dist(C, X) = 0$. If X appears as the second hop, $dist(C, X) = 1$, and so on. Similarly, we define the AS-path distance between a CP C and a link $l = (A_1, A_2)$ as $dist(C, l) = \min(dist(C, A_1), dist(C, A_2))$.

Note that AS-path distance is not simply the length of the shortest path between two nodes in the Internet topology, because it takes into account only `as-paths` that actually appear in BGP routes collected by a given CP. This has three important consequences. First, by definition, AS-path distance honors BGP policies, i.e., it never reports a distance based on a path that would not be allowed by BGP policies. Second, AS-path distance takes into account route diversity: if two different CPs are hosted by the same AS but they collect different routes, their AS-path distances will differ accordingly. Third, AS-path distance is not biased by potentially missing links in the AS-level graph, because it depends only on collected BGP routes.

We leverage AS-path distance to build sub-datasets. For each link l and integer $k = 1, \ldots, 5$, we extract a sub-dataset containing only CPs at a distance less than or equal to k from l. We build sub-datasets for iBGP policy inference with a similar procedure that considers distances from inferred ASes instead of links. Using those sub-datasets, we finally compute the minimum AS-path distances for each link l (resp. AS A) as the minimum value of k such that l (resp. A) can be inferred.

Fig. 7 shows the distribution of the minimum AS-path distances for all links (bars with vertical filling) and for ASes with iBGP policies (bars with oblique filling). For example, the vertical bar with coordinate $x = 0$ indicates that for roughly 36% of the links, a link can be detected from a CP at distance 0, i.e., a CP in one of the two ASes forming the link.

Generally speaking, the plot shows that only a tiny fraction (less than 6%) of links and ASes with iBGP policies have been inferred thanks to information provided by distant CPs, i.e., at a distance greater than 2. Given the average length of AS paths in the Internet (i.e., about 4 and growing very slowly over time [23]), it is quite surprising that a CP can give useful information on ASes (or links) that are 4 or 5 hops away. On the contrary, for the vast majority of both discovered links and inferred ASes, CPs providing non-redundant information are at a distance smaller than 3 from links and ASes, respectively. This suggests that, with high probability, the marginal utility of new CPs is *localized*. As the Internet becomes denser [23] and as more CPs are added to the measurement infrastructure, we expect this localization effect to be amplified, leading to diminishing marginal utility of single CPs.

B. Routing State Distance

The marginal utility of adding a new CP intuitively depends on the diversity of BGP routes that it collects with respect to the routes already in the dataset. For example, if each and every BGP route collected by the new CP were already present in the dataset, the CP would have no marginal utility at all.

Recently, Gürsun et al. [25] have proposed a metric, called *routing state distance* (RSD), to quantify the diversity of BGP routes. In particular, RSD considers all the collected BGP routes to each prefix pair to compute and compare the routing states of the two prefixes. Gürsun et al. define the routing state $RS(p)$ of a prefix p as the directed graph obtained by merging all AS-paths in BGP routes for p. Then, they define the routing state distance between two prefixes p_1 and p_2 as $|RS(p_1) \oplus RS(p_2)|$, where $\oplus$ denotes the XOR between the two graphs (i.e., their union minus their intersection).

Since we need to compare BGP routes collected by different CPs rather than towards different prefixes, we adapt the definition of routing state. Namely, we define the routing state $RS(C)$ of a CP C as the directed graph obtained by merging all AS-paths in the routes collected by C. Unfortunately, the vanilla definition of RSD between two CPs C_1 and C_2, i.e., $|RS(C_1) \oplus RS(C_2)|$ has important intrinsic limitations. First, we cannot compare RSD across sub-datasets because the size of the routing state depends on the size of the dataset. Even worse, CPs with low prefix visibility would have smaller routing states compared to CPs with high prefix visibility. In other words, with the definition above, RSD would be an absolute, non-normalized value having high sensitivity to both the size of the dataset and the difference in the number of prefixes seen by each CP. We adapt the definition of the metric to mitigate the impact of this unwanted sensitivity by *(i)* restricting our analysis to only "full" CP, i.e., CPs that see at least half of the global RIB; and *(ii)* normalizing the metric according to the size of the union between the two

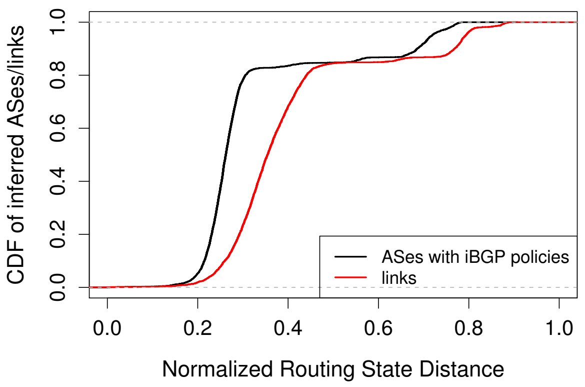


Fig. 8. CDF of the Routing State Distance when computed on all the ASes and only on the portion of the Internet topology relative to iBGP policy inference.

considered routing states. The normalized metric RSD^* is ultimately defined as:

$$RSD^*(C_1, C_2) = \frac{|RS(C_1) \oplus RS(C_2)|}{|RS(C_1) \cup RS(C_2)|}$$

Observe that this definition of RSD^* is equivalent to the *Jaccard distance* between $RS(C_1)$ and $RS(C_2)$. As opposed to RSD, RSD^* does not tend to report lower distances as the size of $RS(C_1)$ or $RS(C_2)$ decreases. Moreover, it ranges from 0 (the two CPs see exactly the same AS-paths) to 1 (the two CPs see completely disjoint AS-paths), irrespective of the size of the routing states.

In order to compare detected links with inferred ASes, we compute routing states on *(i)* the full dataset, and *(ii)* a sub-dataset without all BGP routes that do not contain any AS performing iBGP policy. Observe that, in the latter case, the excluded routes are irrelevant to the iBGP policy inference technique, that is, applying the technique to either the full dataset or the sub-dataset yields identical results.

Fig. 8 shows the CDF of the normalized RSD between all possible pairs of CPs for both detected links (black curve) and inferred ASes (red curve). The normalized RSD is generally higher for links. This means that distant CPs are able to provide more information on links than on inferred ASes, and vice versa for close CPs. This matches the intuition that collecting different BGP routes enables us to detect larger portions of the Internet. Hence, maximizing RSD^* with respect to existing CPs is a good placement strategy for new CPs if we want to improve link detection. On the contrary, the same strategy does not look effective to improve our ability to infer iBGP policy. Indeed, while link detection can benefit from CPs observing different and potentially unrelated portions of the network, inferring iBGP policy *requires* a minimum overlap among collected BGP routes. That is, in order to detect an AS applying iBGP policies, we need CPs to collect different AS paths that traverse that AS for the same prefix.

C. Placing New CPs

Sec. V suggests that, if we added new CPs randomly following the same distribution of the existing infrastructure, the marginal utility of new CPs would be higher for iBGP policy inference than it would be for link detection (see Fig. 4 and 5). However, Sec. VI-A indicates that, for both metrics, the marginal utility will decrease as more CPs are added and as the Internet becomes denser. One might ask whether placement strategies exist that maximize the marginal utility of new CPs for both metrics.

Numerous contributions tackled the problem of extending BGP collection infrastructures in such a way to maximize the marginal utility of new CPs with respect to link detection (see, e.g., [13], [26]). Those contributions highlight the need for deploying CPs in stub ASes and IXPs to improve the detection of peer-peer links. However, CPs in IXPs have low prefix visibility, and Sec. IV (see Fig. 3) suggests that CPs with low prefix visibility are more likely to improve link detection rather than iBGP policy inference. Hence, CP placement techniques proposed for link detection are likely not to maximize the marginal utility of new CPs for iBGP policy inference.

This section generalizes those observations, suggesting that a placement strategy can hardly maximize the marginal utility of new CPs for inference of both link and ASes with iBGP policies. Sec. VI-A suggests that the impact of new CPs is limited to the proximity of the deployment location (see Fig. 7), even though this happens more frequently for link detection than for iBGP policy inference. Moreover, the analysis in Sec. VI-B shows that, contrary to topology discovery, iBGP policy inference poses a fundamental trade-off about the amount of overlap in the BGP routes collected by new CPs. In fact, both a perfect overlap and no overlap lead to a null marginal utility. Ideally, a placement strategy for iBGP policy inference would then need some heuristics to find a good balance between those two limit cases. Devising such a strategy remains an interesting open problem.

VII. Related Work

Previous works used public BGP dataset in a wide variety of manners, e.g., for analyses on AS-level topology discovery [18], commercial relationships between ASes [27], BGP update churn [28], prefix allocation and aggregation [29], prefix reachability [21], and many more. The significance of contributions that tried to match BGP control-plane data with data-plane measurements was questioned by Bush et al., who showed in [21] that BGP data are inadequate for this goal. Indeed, not only the `as-path` attribute correlates poorly with traceroute-like measurements, but even prefix reachability is independent of the presence of BGP route for that prefix. Our study focuses on biases of collected control-plane data.

Among the studies that use BGP data, techniques to discover the AS-level Internet topology are definitely the most common. It is widely accepted that the BGP route collection infrastructure is reasonably accurate to study links between a customer and its provider, while links between eBGP peers are not adequately represented in the dataset [12], [30]. Such a bias is so relevant that most contributions on the optimal placement of new CPs [13], [26] try to counter it by adding new CPs in ASes at the bottom of the Internet hierarchy. While we confirm both the existence of the bias and the effectiveness of this placement strategies for topology inference, our results show that other metrics are not equally affected by the same bias and, therefore, placement strategies designed for one metric are unlikely to achieve good marginal utility for others.

Mühlbauer et al. [20] noted that, by comparing BGP routes to the same destination prefix and simultaneously active for a given AS (i.e., route diversity), it is possible to infer information about the internal structure of that AS. By looking deeper at the BGP attributes of those simultaneously active routes, we can detect iBGP policies [3]. In this paper, we investigate whether iBGP policy inference is affected by the same biases that hold for AS-level topology discovery.

VIII. Conclusions

In this paper, we study the biases of public BGP datasets with respect to the inference of iBGP policies inside ISPs. We define a methodology to expose biases even in the absence of ground truth, and we apply our methodology to compare the biases for iBGP policies with the ones affecting AS-level topology discovery. Our study is focused on the sensitivity of iBGP policy inference and AS-level topology discovery along three main dimensions: (i) visibility of prefixes, (ii) number of available collector peers, and (iii) location of the collector peers. Moreover, we investigate the marginal utility of newly added collector peers and we discuss the implications in terms of optimal placement of new vantage points.

Our results suggest that the BGP dataset is differently biased for iBGP policy inference than it is for AS-level topology discovery. Consistent with this, we also found that algorithms to place new collector peers for AS-level topology discovery, e.g., [12], [13], hardly maximize the benefit for other metrics (i.e., iBGP policy inference in the studied case). Hence, extensions of the BGP monitoring infrastructure need to be driven by a fixed set of objectives, and different infrastructures can deserve different purposes. Finding a strategy that optimizes the marginal utility of new collector peers for iBGP policy inference (or, *a fortiori*, for multiple metrics at the same time) remains an interesting open problem.

Acknowledgments

Mr. Donnet's work is partially funded by the European Commission funded mPlane ICT-318627 project.

References

[1] Y. Rekhter, T. Li, and S. Hares, "A border gateway protocol 4 (BGP-4)," Internet Engineering Task Force, RFC 4271, January 2006.

[2] M. Caesar and J. Rexford, "BGP routing policies in ISP networks," *IEEE Network*, vol. 19, no. 6, pp. 5–11, November–December 2005.

[3] L. Cittadini, G. Di Battista, and S. Vissicchio, "Doing don'ts: Modifying BGP attributes within an autonomous system," in *Proc. IEEE/IFIP Network Operations and Management Symposium (NOMS)*, April 2010.

[4] B. Quoitin, S. Tandel, S. Uhlig, and O. Bonaventure, "Interdomain traffic engineering with redistribution communities," *Computer Communications*, vol. 27, no. 4, pp. 335–363, March 2004.

[5] T. Griffin, F. B. Shepherd, and G. Wilfong, "The stable paths problem and interdomain routing," *IEEE/ACM Transactions on Networking*, vol. 10, no. 2, pp. 232–243, April 2002.

[6] S. Vissicchio, L. Cittadini, and G. Di Battista, "On iBGP routing policies," *IEEE/ACM Transactions on Networking*, 2014, to appear.

[7] T. Griffin and G. Wilfong, "On the correctness of iBGP configuration," *ACM SIGCOMM Computer Communication Review*, vol. 32, no. 4, pp. 17–29, October 2002.

[8] D. Walton, A. Retana, E. Chen, and J. Scudder, "Advertisement of multiple paths in BGP," Internet Engineering Task Force, Internet Draft draft-ietf-idr-add-paths-09, October 2013.

[9] A. Flavel, J. McMahon, a. Shaikh, M. Roughan, and N. Bean, "BGP route prediction within ISPs," *Computer Communications*, vol. 33, no. 10, pp. 1180–1190, June 2010r.

[10] RIPE, "Routing information service (RIS)," see: http://www.ripe.net/ris.

[11] University of Oregon, "Oregon routeviews project," see http://www.routeviews.org.

[12] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "The (in)completeness of the observed Internet AS-level structure," *IEEE/ACM Transactions on Networking*, vol. 18, no. 1, pp. 109–122, February 2010.

[13] E. Gregori, A. Improta, L. Lenzini, L. Rossi, and L. Sani, "On the incompleteness of the AS-level graph: a novel methodology for BGP route collector placement," in *Proc. ACM/USENIX Internet Measurement Conference (IMC)*, November 2012.

[14] Z. Ying, Z. Zheng, Z. Morley, Y. C. Mao, B. Hu, and M. Maggs, "On the impact of route monitor selection," in *Proc. ACM/USENIX Internet Measurement Conference (IMC)*, November 2007.

[15] B. Donnet and O. Bonaventure, "On BGP communities," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 2, pp. 55–59, April 2008.

[16] T. Bates, E. Chen, and R. Chandra, "BGP route reflection: An alternative to full mesh internal BGP (iBGP)," Internet Engineering Task Force, RFC 4456, April 2006.

[17] L. Blunk, M. Karir, and C. Labovitz, "Multi-threaded routing toolkit (MRT) routing information export format," Internet Engineering Task Force, RFC 6396, October 2011.

[18] B. Donnet and T. Friedman, "Internet topology discovery: a survey," *IEEE Communications Surveys and Tutorials*, vol. 9, no. 4, pp. 2–15, December 2007.

[19] N. Feamster and J. Rexford, "Network-wide prediction of BGP routes," *IEEE/ACM Transactions on Networking*, vol. 15, no. 2, pp. 253–266, April 2007.

[20] W. Mühlbauer, A. Feldmann, O. Maennel, M. Roughan, and S. Uhlig, "Building an AS-topology model that captures route diversity," in *Proc. ACM SIGCOMM*, August 2006.

[21] R. Bush, O. Maennel, M. Roughan, and S. Uhlig, "Internet Optometry: Assessing the Broken Glasses in Internet Reachability," in *Proc. ACM/USENIX Internet Measurement Conference (IMC)*, November 2009.

[22] V. Giotsas, S. Zhou, M. Luckie, and k. claffy, "Inferring multilateral peering agreements," in *Proc. ACM CoNEXT*, December 2013.

[23] A. Dhamdhere and C. Dovrolis, "Ten years in the evolution of the Internet ecosystem," in *Proc. ACM/USENIX Internet Measurement Conference (IMC)*, November 2008.

[24] P. Barford, A. Bestavros, J. W. Byers, and M. Crovella, "On the marginal utility of network topology measurements," in *Proc. Internet Measurement Workshop (IMW)*, November 2001.

[25] G. Gürsun, N. Ruchansky, E. Terzi, and M. Crovella, "Routing state distance: A path-based metric for network analysis," in *Proc. ACM/USENIX Internet Measurement Conference (IMC)*, November 2012.

[26] K. Chen, D. Choffnes, R. Potharaju, Y. Chen, F. Bustamante, D. Pei, and Y. Zhao, "Where the sidewalk ends: Extending the Internet AS graph using traceroutes from P2P users," in *Proc. ACM SIGCOM CoNEXT*, December 2009.

[27] L. Gao, "On inferring autonomous system relationships in the Internet," *IEEE/ACM Transactions on Networking*, vol. 9, no. 6, pp. 733–745, December 2001.

[28] A. Elmokashfi, A. Kvalbein, and C. Dovrolis, "BGP churn evolution: A perspective from the core," *IEEE/ACM Transactions on Networking*, vol. 20, no. 2, pp. 571–584, April 2012.

[29] L. Cittadini, W. Muhlbauer, S. Uhlig, R. Bush, P. Franois, and O. Maennel, "Evolution of Internet address space deaggregation: Myths and reality," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 8, pp. 1238–1249, October 2010.

[30] M. Roughan, W. Willinger, O. Maennel, D. Perouli, and R. Bush, "10 lessons from 10 years of measuring and modeling the internet's autonomous systems," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 11, pp. 1810–1821, October 2011.