The BGP Visibility Scanner

Andra Lutu*†, Marcelo Bagnulo† and Olaf Maennel‡
*Institute IMDEA Networks, Spain
†University Carlos III of Madrid, Spain
‡Loughborough University, Loughborough, UK

***Abstract*—By tweaking the BGP configurations, the network operators are able to express their interdomain routing preferences, designed to accommodate a myriad goals. Given the complex interactions between policies in the Internet, the origin AS by itself cannot ensure that only by configuring a routing policy it can also achieve the anticipated results. Moreover, the definition of routing policies is a complicated process, involving a number of subtle tuning operations prone to errors. In this paper, we propose the *BGP Visibility Scanner* which allows network operators to validate the correct implementation of their routing policies, by corroborating the BGP routing information from approximatively 130 independent observation points in the Internet. We exemplify the use of the proposed methodology and also perform an initial validation for the BGP Visibility Scanner capabilities through various real operational use cases.**

I. Introduction

The Internet is the interconnection between multiple operationally independent networks, also known as Autonomous Systems (ASes). The Border Gateway Protocol (BGP) is responsible for the exchange of reachability information for IP prefixes and the selection of paths according to the routing policies specified by each network. By tweaking the BGP configurations, the network operators are able to express their interdomain routing preferences, designed to accommodate various operational, economic, and political factors. Thus, the originating ASes have the capability to influence the way the incoming and outgoing traffic flows in order to ultimately optimize the exploitation of their own network.

However, the origin AS by itself cannot ensure that only by configuring a routing policy it can also achieve the anticipated results [1]. The main reason behind this resides in the fact that the actual inter-domain routing is the result of the interaction of the routing policies of all the ASes involved, possibly bringing about a different outcome than the one expected by the different ASes. This situation is easily reflected in the case of the ASes using AS-Path prepending to express their routing policies, which may or may not lead to the expected result depending on the policies of the other ASes.

Moreover, the implementation of routing policies is a complicated process, involving subtle tuning operations serving all the origin's goals. Thus, it is an error-prone task and operators might end up with inaccurate configurations that could impact the efficacy of their interdomain strategies. For example, ill-defined outbound filters may lead to an AS unknowingly leaking internal routes to the Internet and impacting the effectiveness of its own active routing policies [2].

Consequently, in order to avoid the distortions of their routing policies due to accidental mis-configurations or adverse effects within the complex external netting of routing policies, ASes need to monitor the manner in which their preferences resonate in the global routing system. To this end, operators complement their internal perspective on routing with the information retrieved from external sources, e.g. publicly available looking-glasses. However useful, these tools have obvious limitations [3], e.g. allowing only for single per-route queries and not storing any historical information.

In this paper, we propose the ***BGP Visibility Scanner***[1] which allows network operators to validate the correct implementation of their routing policies, by corroborating the BGP routing information from approximatively 130 independent observation points in the interdomain. The tool allows networks to check how their own routes are being propagated in the Internet, verify the results of the implemented routing policies and identify possible cases where these policies backfired. By merging all the available information from the ASes enabled as active monitors active in the RIPE RIS [4] and RouteView [5] Projects, we create a *visibility scanner* for all the IPv4 prefixes active in the interdomain. The tool is subject to the limitations of the available public looking-glasses, which we further address accordingly. It is important to note that the properties of BGP do not allow us to get a complete picture on all policies, but nevertheless those public observations points provide a multi-angle perspective on the interdomain routing. Moreover, our tool has already proven its capability of triggering visibility alarms and helping networks deal with the problems caused by their own routing policies.

We focus our analysis on a particular expression of routing policy interaction, namely the interdomain route propagation process and the manner it is reflected in the interdomain global routing tables. We define the **Limited-Visibility Prefixes (LVPs)** as being stable long-lived Internet routes that are not present in all the global routing tables analyzed, but seen by at least two ASes. Contrariwise, we also define the **High-Visibility Prefixes (HVPs)** as the set of prefixes that are propagated within almost all the available full routing feeds. We note that *the limited visibility does not imply limited reachability*. There could be a *HV* less-specific prefix that provides reachability. In this sense, we also identify a set of so-called **Dark Prefixes (DPs)**, which represents a subset of

[1]The Visibility Scanner is publicly available at **visibility.it.uc3m.es**

the *LVPs* that are not covered by any *HV* less-specific prefix. These prefixes represent address space that, in the absence of a default route, may not be globally reachable.

There are several reasons behind the existence of *LV* prefixes, which can be classified as follows:

- *Intentional/Deliberate.* Some ASes create *LVPs* on purpose. There are several ways this can be done, including scoped advertisements (e.g. geographically scoped prefixes to offer connectivity only to networks located in a certain region) or advertisements only through (some) peering and not transit relationships.
- *Inflicted by third parties.* Some prefixes are announced by the origin ASes with the intention of being globally distributed, but some of the ASes receiving the prefix decide to filter them. A notable example of this is filtering by prefix length.
- *Unintentional/Accidental.* In many cases, *LVPs* are the result of errors in the configuration of filters of the origin or other ASes that have received the prefix announcement.

We perform a differential analysis to retrieve *LVPs* on a daily basis and we make the results of our study available on-line, thus creating the possibility for a close to real-time verification of the effectiveness of eventual modifications in the implemented routing preferences. We integrate in our analysis previously "cleaned" routing information, after the elimination of routes that do not represents an expression of routing policies, but of other network-specific operations, e.g. internal routes visible in only one monitor, converging routes, MOAS prefixes, bogon prefixes etc.

II. Related Work

BGP has been studied for more than 15 years. There is a magnitude of papers and knowledge in the community. Most of the work related to our efforts tackle the analysis of BGP raw data, which can be tricky and difficult [6]. First of all, the input data needs to be "cleaned" from artifacts [7], then the data needs to be carefully interpreted. There are three major research areas which rely on BGP raw data. The first one is interested in AS-topology inference [8], the second aims at detecting security related routing conditions, such as prefix hijacking (e.g., PHAS [9]). Finally, the third, tries to create tools that provide useful information for operators [10]. Multiple operational misconfiguration have been reported [2], but attempts go far beyond this. They include *RIPE Labs* [11], which has a whole section devoted to tools that assists operators or Renesys [12], which operates this type of services to operators for a fee.

We focus specifically in the monitoring of healthy deployment of policies focusing in *LVPs* which, to the best of our knowledge, is not covered by existing work [13]. Unlike such tools which integrate a vast amount of operational problems [13], we do not focus on inferring and/or monitoring the AS-level topology of the Internet, but on monitoring the healthy deployment of routing policies. In this sense our work is very closely related to the work on BGP wedgies by Griffin et al [14], [15]. However, none of those theoretical

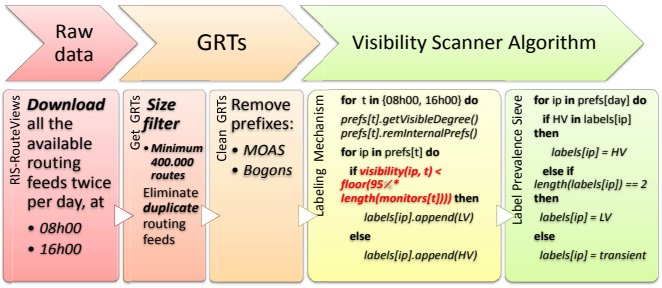

Fig. 1. Methodology used for determining the LVPs and the HVPs.

work is able to detect problematic routing conditions based on raw BGP observations. All of this work requires access to configuration files, which are typically not shared. Typically those are considered a company secret, and in fact BGP was designed to hide such policy information, making it hard to be inferred [3]. While we understand the limitations for the BGP protocol observations, we noticed that still a great deal that can be observed. In this sense, our work aims at reporting and aggregating the information to make it usable for operators.

III. Methodology

We collect the routing information from the two major publicly available repositories at RouteViews and RIPE RIS. The two repositories gather BGP data throughout the world, currently deploying 24 different collection points, which we further refer to as *collectors*. The collectors periodically receive BGP routing table *snapshots*, i.e. one time instance of a routing table, from over 400 active monitors. A *monitor* represents a network peering with the public RIS/RouteViews repositories and propagating its routing information.

We depict in Figure 1 all the required steps for molding the raw data into the *LV* and *HV* prefix sets, starting with the above-described data collection process until applying the visibility algorithm. We next expand on each of the different processing phases.

A. Refining the Raw Routing Data

We focus here on the second processing block of the flow chart in Figure 1, and we look at the steps we take to obtain the set of global routing tables (GRTs).

Conceptually, the so-called *Default Free Zone (DFZ)* represents the set of BGP-speaking routers that do not need a default route to forward packets towards any destination in the Internet. The routing table maintained in one of the DFZ routers is commonly known as the *global routing table*. Realistically speaking though, due to the current operational status of the Internet routing, such a GRT of the BGP routing is an idealized concept. However, Internet Service Providers (ISPs) do maintain their own version of the *global routing table*, which is propagated to customer networks upon request.

For the purpose of this paper, we loosely define the GRT as the entire routing table provided by a DFZ network to its customers requesting a full routing feed. This is not a formal definition, but it properly captures the main idea behind the type of data required for our study. We can identify the sets

of *HV* and *LV* prefixes only by comparing the GRTs from the active monitors. However, the monitors have different policies with respect to the public routing repositories, thus providing different types of routing feeds. We are able to identify three different types of feeds injected to collectors, namely:

Partial Routing Tables: this type of feed can be though as the result of establishing a peering-like business relationship between the monitor and the collector. By definition, these feeds are not GRTs, thus are not useful for our analysis.

Global Routing Tables: full routing feeds from the monitors. This is the main raw information that we want to keep.

Global Routing Table, including internal routes: in some cases, it may happen that the monitor announces, aside from the complete routing table, other additional internal information. This additional information is again of no interest for our study, since we do not focus on the internal operations of a network. Consequently, we need to identify and filter out these particular routes within the complete routing feed.

Filter routing tables based on minimum size restriction. In order to identify the feeds which constitute a GRT, the primary characteristic of the routing feeds on which we focus is the actual size of the routing table snapshot. Based on the BGP Analysis Report [16], we consider that *a complete routing feed from a monitor should have no less than **400,000** routing entries*. Consequently, we check over 500 routing feeds collected from the two repositories, and we discard all the BGP feeds that have less than the imposed lower-limit of prefixes.

In order to further verify the results of this heuristic, we verify the routing policy of the collector storing the routing information and the routing policy of the monitor offering its routing feed towards the collector. In particular, within the RIS project, for each collector it is specified the number of so-called *full routing feeds*, which is consistent with the number of tables with more than *400,000* entries.

Furthermore, we check in the `whois` database the public routing policies for the ASes peering with the two public repositories. We are able to retrieve information for 34 monitors feeding a routing table with more than *400,000* entries. We see that for 18 of them the public routing policy is *ANNOUNCE ANY*. This further confirms the assumption that the full routing feeds received by the collectors are actually consistent with propagating all the available routes maintained by the monitor. The rest 16 monitors are advertising the policy *ANNOUNCE AS_name*, which is not clearly defined.

In order to address the limitations of using the `whois` database, we check the publicly available topology maps [17] to infer the business relationship of the monitors towards the repositories. Consequently, we are able to check if the relationship with the repository is a *provider-to-customer* (p2c) relationship, meaning that the monitor may be exporting its complete routing table. We are however not able to verify this for AS6447 of the RouteViews Project. For AS12654 RIPE RIS project, we were able to validate 21 such relationships.

After applying the size filter to all the raw feeds, we are left only with the Complete Routing Tables and the Complete Routing Tables with internal information. We later deal with the additional refinements in order to eliminate the surplus data from these first results.

Eliminate duplicate routing feeds. After checking the content of duplicate feeds from the same AS and comparing them, we find that these multiple routing table snapshots are identical. Since analyzing the routing feeds including duplicates may trigger the generation of false positive *HVPs*, we only keep one unique instance of the routing table snapshots.

B. Sanitary Checks

After applying the previously described heuristics, we are able to identify the GRTs. We perform a couple of "sanitary" checks on the data contained in the GRTs, in order to further discard the information that is of no interest for our study. Hence, we apply the *bogon filter* and the *MOAS filter* on all the GRTs, as depicted in the third step in Figure 1.

Eliminate the bogon and martian routes from GRTs. Bogon prefixes are a class of routes that should never appear in the Internet. Bogons are defined as *Martians*, representing private and reserved address space or *Fullbogons*, which include the IP space that has been allocated to a Regional Internet Registry (RIR), but has not been assigned by that RIR to an actual Internet Service Provider (ISP) or other end-user.

We use the periodically updated filters from The Bogon Reference [18] in order to make sure that we eliminate any possible bogon route included in the GRTs. We usually identify just around 500 bogon prefixes within the routes injected in the Internet.

Not consider the MOAS prefixes. The Multiple-Originating AS (MOAS) [19] prefixes cannot be qualified within our study, since for these prefixes we are not able to identify which origin AS might be suffering/generating the reduced visibility of its prefixes. We plan to address this issue in the future work. Therefore, we identify and discard all the MOAS prefixes (i.e. 4.500 prefixes).

C. The Visibility Scanner Algorithm: the Labeling Mechanism

Having obtained the "clean" version of the GRTs, we proceed to applying the **Visibility Scanner Algorithm** for identifying prefixes with stable reduced visibility in the interdomain. At this point it is important to filter out the cases of reduced interdomain visibility caused by other factors unrelated to routing policies, e.g. BGP convergence or leaking internal routes to the collector. In order to avoid the problem of internal paths leaking towards the collectors, we remove all the routes learned from only one monitor which is also the route originating AS.

In order to address the confusion caused by converging prefixes emerging as false positive limited visibility prefixes in our results, we analyze two 8-hours apart samples of routing data and the per-prefix visibility. We focus on monitoring the propagation of routes, evaluate the *visibility degree* at every sampling moment and assign *visibility labels* based on our results. We define the *visibility degree* as the number of GRTs within the sample that contain (i.e. see) a certain prefix, and the *visibility label* as the visibility status of each prefix, i.e.

LV for Limited Visibility and *HV* for High Visibility. The visibility scanner algorithm is composed of the forth and fifth steps of the processing flow depicted in Figure 1, which we call *prevalence sieves*.

The Labeling Mechanism: assigning prefix visibility labels. Based on the visibility degree of the prefixes at each of the two sampling moments (i.e. 08h00 and 16h00), we assign a *visibility labels* at each sampling moment to all the prefixes discovered at each moment, as described in 1.

We use a 95% *minimum visibility rule* in order to assign the labels according with the observed visibility degrees. Consequently, *we define Limited Visibility prefixes as prefixes present in less than* 95% *of the active monitors at a sampling time.* Otherwise, the prefixes complying with the 95% minimum visibility rule are defined as High Visibility prefixes. Ideally, a *HV* prefix should be contained in absolutely all the routing tables contained in the sample. The choice of the 95% allows for a 5% error in the sampling also accommodating possible glitches that may appear in the data. Moreover, according to our threshold sensitivity analysis, we find that the set of *LVPs* is not highly sensitive to the values of the prevalence sieve threshold. We expand on this in section IV.

Visibility Label Prevalence Sieve. Eliminating Converging Prefixes. The visibility label *prevalence sieve* accounts for the dynamics of a prefix in time, as presented in the last block from Figure 1. After applying the *prevalence rule* integrated in the sieve, we decide the per-day label for the prefix. The high visibility of a prefix in at least one monitor sample hints the fact that the route could reach all the observed ASes. Should this change during the analyzed time, it might be a cause of, for example, topology changes or failures. Therefore, we consider that *the HV label always prevails*, i.e. if a prefix is tagged as *HV* in one of the samples, it is tagged as *HV* in the final set.

Otherwise, when no *HV* label is tagged, we analyze the cases of *LV* prefixes emerging in our results. If a prefix is tagged as *LVP* only once in the two sampling times, it might be a symptom of a prefix being withdrawn or, contrariwise, in the process of converging after just being injected. Having a single *LV* label means that the prefix is not present in the other sample, i.e. the prefix is no longer present in any of the routing tables, and there can be several explanations for this, including the prefix being withdrawn. In any case, these particular routes cannot be qualified within our study, thus we filter out any prefix with only one label in a day (and that label being *LV*). This helps us to eliminate routes that are not an expression of the routing policies, but a second-effect of other Internet operations. The only case where we can say a prefix has limited visibility and mark it accordingly, is when both labels assigned at each sampling time are *LV*.

D. Identifying Dark Prefixes.

Once we have identified the two main sets of prefixes, i.e. the *LVPs* and the *HVPs*, we move on to verifying the reachability of the LVPs in order to identify possible cases of reduced reachability. Consequently, for each of the prefix in the LVP category, we build the covering trie of less specific HV prefixes, from which we ultimately retrieve its root prefix (i.e. the smallest covering HV prefix). In the eventuality of not identifying any such globally visible less-specific prefix, we mark the LV prefix as *Dark* and continue our analysis.

IV. BGP Visibility Scanner: Prefix Visibility Analysis Example

We exemplify the usage of the proposed methodology by applying the algorithms proposed in the previous section every day during the month of October 2012. We present next the results of the per-day analysis and per-week analysis. We also underline several interest characteristics of the *LVPs*.

A. Applying the Visibility Scanner Algorithm

We begin by exemplifying the usage of the proposed methodology on a one-day complete routing data sample. We arbitrarily chose the date of 23rd of October 2012, from which we collect more than *500* routing feeds. After applying the *cleansing process* presented in section III-A, we identify *129 GRTs* forwarded to the public repositories by different ASes. We move on to performing the additional sanitary checks presented in section III-B. Consequently, we are able to overall identify and eliminate *499 bogon prefixes* and *4,796 MOAS prefixes* from all the GRTs. We observe that the overall total number of prefixes identified for the day is of *535,146 prefixes*. We evaluate the degree of visibility for every prefix present at each of the considered sampling times and we assign the visibility label according with the mechanism presented in III-C. Consequently, we identify and filter out approximatively *10,500 prefixes* that are thought to be leaked internal routes. In order to further remove the converging routes that may emerge in our study as limited visibility, we apply the prevalence sieve. Thus, in the case of the routing tables snapshots from October 23, 2012, we are able to identify and discard *7,800* converging prefixes. For the remaining prefixes, we apply the prevalence sieve and assigning per-day visibility tags. We are thus finally able to identify *98,253 prefixes* that are tagged *LVP* and *415,576* prefixes marked *HVP*. When checking how the two sets of prefixes overlap, we find that there are *2,400* LV prefixes without a covering high-visibility prefix, which we mark *DP*.

B. Characteristics of the Prefix Visibility Categories

We have previously defined the LVP set using a 95% prevalence rule. It is important to understand which is the sensitivity of the threshold to the actual data conditions. We represent in Figure 2 the distribution of prefixes on the possible degrees of visibility for the sample of data from 23 of October, 2012. We note that by varying the prevalence threshold value, the size of the two prefix sets does not suffer important changes (e.g. after changing the minimum threshold to 90%, only approximatively 800 prefixes are added to the *HVP* set). Also, due to the concentration of prefixes in the extremes values of the visibility degree, we conclude that with more routing feeds, the number of *LVPs* should increase.

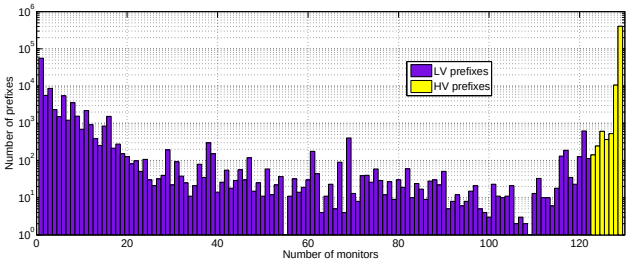

Fig. 2. **Distribution of prefixes on visibility degree.**

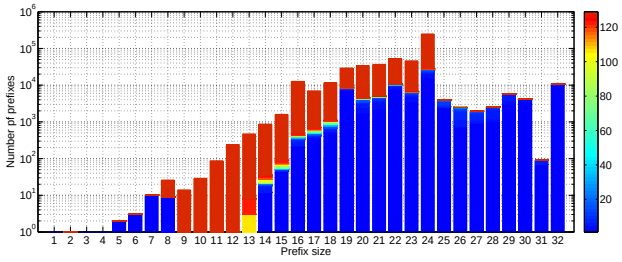

Fig. 3. **Prefix length visibility:** each bar shows the number of prefixes with a certain mask length. The color code represents the visibility distribution of the prefixes within each prefix-length category, according with the visibility degrees marked in the color legend in the right part of the plot.

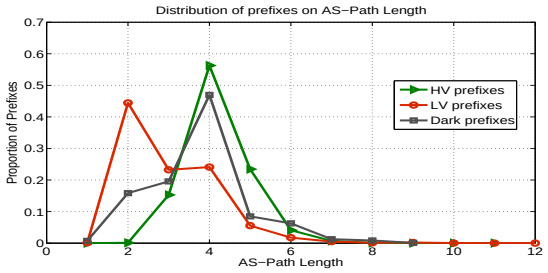

Fig. 4. PDF for the AS-Path length.

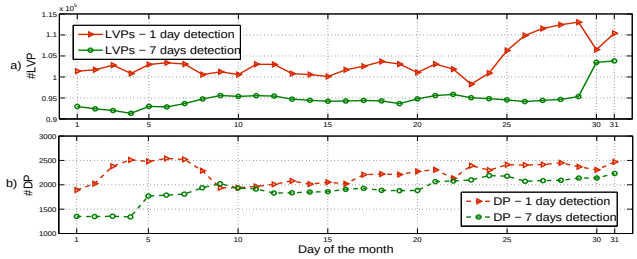

Fig. 5. Comparing the one-day and seven-days detection algorithms.

When comparing the three sets of data identifies, i.e. *LVP, HVP and DP*, we first observe that the limited visibility issue appears for prefixes of various lengths, from $/5$ to $/32$, as depicted in Figure 3. However, we do note the lack of prefixes more-specific than $/24$ in the *HVP* set, which is consistent with the best recommended BGP practices. Due to the fact that prefix length filters may be asymmetric or even missing in some cases, this type of interaction might derive in a generator of *LVPs*. We also observe the presence of *LV* prefixes less specific than $/8$, which, due to their small degree of visibility, may be accidentally leaked in the Internet.

Moreover, when we check the average `AS-Path` length of the *LVPs*, we observe a straightforward difference between the mean `AS-Path` length for *HVPs* of 4 and the mean `AS-Path` length for *LVPs* of 3. This is easily observed from the probability distribution function (PDF) in Figure 4. This information shows a more limited realm of expansion for the *LVPs* than for the general *HVPs*, restraining it closer to the prefix originating network. After applying the methodology every day during October 2012, we are then able to perform a visibility label stability analysis for the *LVPs* identified. In figure 5 we can observe the evolution during the whole month of the number of *LVPs* and *DPs* resulting from detecting the *LVPs* only in one day and from detecting the *LVPs* that were stable during the last 7 days. For the latter, we merely compare the labels tagged on the prefixes discovered during the latest seven days prior to the moment of analysis. Just like in the prevalence sieve in Section III-C, the *HV* label always prevails and we mark as *HV* any prefix with such a label. We discard any prefix with a number of labels lower than 5, i.e. which has been missing from the routing tables for more than 2 days. We assume a prefix is *stable-LV* only when it has at least 5 *LV* labels, no *DP* label and no *HV* label. Also, if a prefix happened to be labeled at some point during the 7 days as *DP*, it is sufficient to show the potential connectivity problems of that particular prefix and we further mark it *DP*. The number of *stable-LVPs* detected based on the latest seven days from the moment of analysis is not much smaller than the per-day number of *LVPs*, despite the implicit removal of possible false-positives and thus pointing to the fact that *LVPs* are a long lived effect.

V. Operational Use Cases

In this section, we perform an initial validation for the BGP Visibility Scanner capabilities through various real operational use cases meant to demonstrate the usage of the proposed tool. We have been in contact with several network operators which provided us with the means to verify and validate the efficiency of our proposed tool. We expand on a few operational examples that illustrate the variety of reasons behind the limited visibility of prefixes in the Internet.

We first provide real cases of ASes deliberately restricting the propagation of their prefixes and exemplify the manner in which there configuration reflect in the BGP Visibility Scanner. Using the BGP Visibility Scanner, we were able to verify and validate the routing policies of two of the Internet root-servers' operators. Consequently, for each root-server we have identified the presence of one more-specific *LV* prefix, which was meant for providing connectivity only to direct peers. The routing policy is correctly reflected in the limited visibility of the prefix. However, the *LV* prefix has global reachability due to the presence of HV less-specific prefixes which is used by the root-servers in order to avoid traffic fluctuations. Besides the two root-servers, the tool also validated the policy of a large content provider which deliberately limits the visibility of one of its prefixes in order to ensure that the incoming traffic is fed only through a geographically-specific local path.

The second type of use cases we present captures the results of unintentional routing policies mis-configurations. We

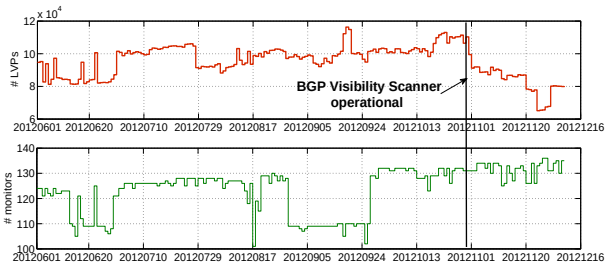

Fig. 6. Evolution of the number of LVPs between June and December 2012.

include here the case of a large and widely-spread ISP, which was able to identify and correct several issues regarding its defined routing policies. After checking the *BGP visibility scanner*, the ISP was able to isolate the case of a subset of prefixes with restricted visibility that were leaking through one of its direct peers. The ISP was able to recognize the mis-configuration of the outbound prefix-filter towards the neighbouring AS, which should have otherwise dropped the *LV* prefixes toward that peer AS. By only correcting this issue, the origin AS successfully eliminated *3,000 LV prefixes* of whose existence it was previously unaware. Also, the same operational AS was able to identify and correct other accidentally ill-configured routing policies on the provider edge devices which were causing the injection of static more-specific routes to the directly connected peers. The operator was able to eliminate an additional 200 leaked prefixes.

All the use cases come to highlight the different ways the BGP Visibility Scanner had been useful to real-world networks starting the day it became operational. Moreover, if we analyze the evolution in the number of *LVPs* over the past 6 months depicted in Figure 6, we observe that since the tool became operational at the beginning of November 2012, and operators became aware of its existence, the number of LV prefixes has been decreasing. This happens despite the slight increase in the number of sampled monitors, which should imply an increase in the number of *LVPs*, according with Figure 2.

VI. Conclusions and Future Work

Problematic routing conditions and complex interactions between policies in the Internet have been predicted manifold [1]. However, to detect them it is required that ISPs share their configurations, which appears to be unlikely in today's Internet. In this paper we investigated to what extent it is possible to discover the match between the *intended result* of applying routing policies and the *actual result* reflected in the global routing system. Just by using publicly available data, we present an initial methodology that scans raw BGP data, filters and analyzes it, so that we can extract potential problematic policy configuration. We have defined the terms of *limited visibility* and *dark prefixes*, which can be considered early warning signs for routing policies backfiring and not achieving their desired outcome. Despite many years of research on BGP data, such problems have not been sufficiently addressed [20]. We have presented our methodology to operators and received a lot of very promising feedback. For example, we found approximatively *90,000 stable LVPs* which, after talking to operators, decreased with approximatively *3,000 LVPs*. The latter prefixes were proven to be actual symptoms of ill-configured routing policies. The Visibility Scanner allows per origin-AS queries for the *LVPs* generated and provides additional information about them. As future work, we intend to improve the quality of our heuristics by continuing to validate our methodology with operators. Also, since the methodology can be applied on any set of similar data, we would like to integrate into the tool the private views from operators.

Acknowledgements

This work was supported by the European Communitys Seventh Framework Programme (FP7/2007-2013) grant no. 317647 (Leone). We would like to thank Shane Amante, Lars-Johan Liman and Joao Damas for valuable comments on the operational value of the proposed tool. We are also grateful to Cristel Pelsser, Pierre Francois, Alberto Garcia-Martinez and Randy Bush for the numerous discussions which helped improve this work.

References

[1] T. Griffin and G. Huston, "BGP Wedgies," 2005, RFC 4264.
[2] R. Mahajan, D. Wetherall, and T. Anderson, "Understanding bgp misconfiguration," *SIGCOMM Comput. Commun. Rev.*, vol. 32, no. 4, 2002.
[3] 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," *Selected Areas in Communications, IEEE Journal on]*, vol. 29, no. 9, 2011.
[4] "RIPE RIS Raw data." [Online]. Available: http://www.ripe.net/data-tools/stats/ris/ris-raw-data
[5] "University of Oregon Route Views Project." [Online]. Available: http://www.routeviews.org/
[6] G. Siganos and M. Faloutsos, "Analyzing bgp policies: methodology and tool," in *INFOCOM 2004*, vol. 3, 2004.
[7] B. Zhang, V. Kambhampati, M. Lad, D. Massey, and L. Zhang, "Identifying bgp routing table transfers," in *ACM SIGCOMM workshop on Mining network data*, 2005.
[8] R. Oliveira, M. Lad, B. Zhang, D. Pei, D. Massey, and L. Zhang, "Placing bgp monitors in the internet," *Technical No. UCLA, TR*, 2006.
[9] M. Lad, D. Massey, D. Pei, Y. Wu, B. Zhang, and L. Zhang, "PHAS: a prefix hijack alert system," in *Proceedings of the 15th conference on USENIX Security Symposium - Volume 15*, 2006.
[10] J. Wu, Z. M. Mao, J. Rexford, and J. Wang, "Finding a needle in a haystack: pinpointing significant bgp routing changes in an ip network," in *Symposium on Networked Systems Design & Implementation*, 2005.
[11] "RIPE Labs." [Online]. Available: https://labs.ripe.net/
[12] RENESYS, http://renesys.com/.
[13] Y.-J. Chi, R. Oliveira, and L. Zhang, "Cyclops: the as-level connectivity observatory," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 5, 2008.
[14] T. G. Griffin, F. B. Shepherd, and G. Wilfong, "The stable paths problem and interdomain routing," *IEEE/ACM Trans. Netw.*, vol. 10, no. 2, Apr. 2002.
[15] D. Perouli, T. Griffin, O. Maennel, S. Fahmy, C. Pelsser, A. Gurney, and I. Phillips, "Detecting Unsafe BGP Policies in a Flexible World," in *International Conference on Network Protocols (ICNP)*, 2012.
[16] "BGP Routing Table Analysis Report." [Online]. Available: http://bgp.potaroo.net/
[17] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the internet as-level topology," *ACM SIGCOMM CCR*, vol. 35, no. 1, 2005.
[18] "Team Cymru - The Bogon Reference." [Online]. Available: http://www.cymru.com/BGP/bogons.html
[19] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, S. F. Wu, and L. Zhang, "An analysis of BGP multiple origin AS (MOAS) conflicts," in *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement*, 2001.
[20] C. Labovitz, A. Ahuja, and M. Bailey, *Shining light on dark address space.* Arbor Networks, Incorporated, 2001.