2014 6th International Conference on Cyber Conflict
P.Brangetto, M.Maybaum, J.Stinissen (Eds.)
2014 © NATO CCD COE Publications, Tallinn

Inter-AS Routing Anomalies: Improved Detection and Classification*

Matthias Wübbeling
Fraunhofer FKIE
& University of Bonn
Bonn, Germany
wueb@cs.uni-bonn.de

Michael Meier
Fraunhofer FKIE
& University of Bonn
Bonn, Germany
mm@cs.uni-bonn.de

Till Elsner
Fraunhofer FKIE
& University of Bonn
Bonn, Germany
elsner@cs.uni-bonn.de

Abstract: Based on the interconnection of currently about 45.000 Autonomous Systems (ASs) the Internet and its routing system in particular is highly fragile. To exchange inter-AS routing information, the Border Gateway Protocol (BGP) is used since the very beginning, and will be used for the next years, even with IPv6. BGP has many weaknesses by design, of which the implicit trust of ASs to each other AS is the most threatening one. Although this has been topic on network security research for more than a decade, the problem still persists with no solution in sight. This paper contributes a solution to stay up to date concerning inter-AS routing anomalies based on a broad evidence collected from different publicly available sources. Such an overview is necessary to question and to rely on the Internet as a basis in general and must be a part of every cyber defense strategy. Existing methods of detecting inter-AS routing anomalies result in large sets of real time routing anomalies, based on the evaluation of routing announcements collected from different viewpoints. To decide, whether a detected anomaly is harmful or not, each of them has to be classified and correlated to others. We combine various detection methods and improve them with additional publicly available information. The improved outcome of the implemented routing anomaly detection system is used as input for our classification algorithms.

Keywords: *Internet, Routing, Anomaly Detection, BGP, Autonomous Systems*

* The work presented in this paper was partly funded by the German Federal Ministry of Education and Research under the projects MonIKA (BMBF 16BY1208A)

1. INTRODUCTION

The *Border Gateway Protocol* (BGP) [22] defines the exchange of IP routing information between interconnected Autonomous Systems (ASs) in computer networks. It is the only used routing protocol in the Internet and it is topic of security research since the late 90's. Therefore, itself and its inherent weaknesses are well known. Implicit trust between connected ASs results in the possibility for any AS to inject invalid and malicious routing information with very little effort. Wrong routing information, distributed from one or more ASs over the whole Internet could lead to large scale connectivity problems. The existence of contrary routing information at different locations is called a routing anomaly. Routing anomalies like Multiple Origin AS (MOAS) [25, 8] conflicts, where two or more ASs claim to own the same range of IP addresses, occur regularly. This situation is not only intended to cause harm, based on malicious intention. It can also happen as a result of misconfiguration inside an AS. Countermeasures against IP prefix hijacking, the advertisement of the same IP address space from a foreign AS, still do not exist. Legitimate owners of IP addresses are able to announce longer, more specific IP subnets than the causing/attacking AS, because they are preferred, when the route for a packet is chosen. Only few of these events are publicly known, usually those involving large internet companies such as YouTube or Google [23].

Although MOAS conflicts are easy to detect, they could be used intentionally by prefix owners to implement load balancing or to minimize the routing distance for connections to/from different locations. Thus, the distinction between legitimate and illegitimate conflicts is hard to make. Due to several reasons, e.g. performance issues on large routing systems or impracticability of approaches like S-BGP [14, 13], the threats still exist nowadays. The improvement of routing security brought by origin authentication [6] and asymmetric cryptography, e.g. RPKI [17] is currently small, because it is not yet implemented in broadly used hardware and business processes of ASs. Unless most parts of the Internet support origin authentication or RPKI, the routing system in general is as vulnerable as before. In contrast to prefix hijacking, routing anomalies, that are based on invalid topological information propagated in routing announcements, are significantly harder to detect and to classify.

Several approaches were made to detect and classify routing anomalies based on information gathered from inside the routing plane. They provide systems to identify prefix hijacking events [16, 7, 21]. None of those solutions really classify all found conflicts properly. Classification is necessary to determine whether an occurring conflict is legitimate or illegitimate to derive a level of criticality for a conflict. One common shortcoming of all these solutions is that the assumed ground truth, the data used to train and measure the detection and classification systems, is just based on inherent information exchanged via BGP itself [4, 18].

This is questionable because the exchanged routing information is not reliable, as discussed above. To determine facts of actually existing peering relations and legitimate IP address owners, it is necessary, to query other sources to increase the data used as ground truth evidence. Ground truth evidence in this context is the amount of reliable data to be used to find and to classify occurring routing anomalies.

Our contribution is (1) the collection of a broader data base of reliable information on peering relations between autonomous systems and therefore higher accuracy at finding and classifying routing anomalies, (2) an approach, based on existing systems named above, providing evidence to the ground truth used to find and classify routing anomalies, (3) a selection of reliable sources for this enrichment and (4) a crawling system, gathering information from different viewpoints inside the Internet routing layer, internet exchange points (where ASs can, an mostly do, peer with each other), and AS specific web services such as looking glass, a service to query information from running routers inside an AS.

This paper is structured as follows: first we describe the background and challenge of our research in section 2, then we present used detection and classification methods in section 3 and move on with the presentation of our approach to extend the assumed ground truth as argued above to improve the handling of routing anomalies in section 4. Our applied approach to classify routing anomalies is discussed in section 5 followed by the evaluation in section 6 and the last section 7 includes discussion and future work.

2. BACKGROUND

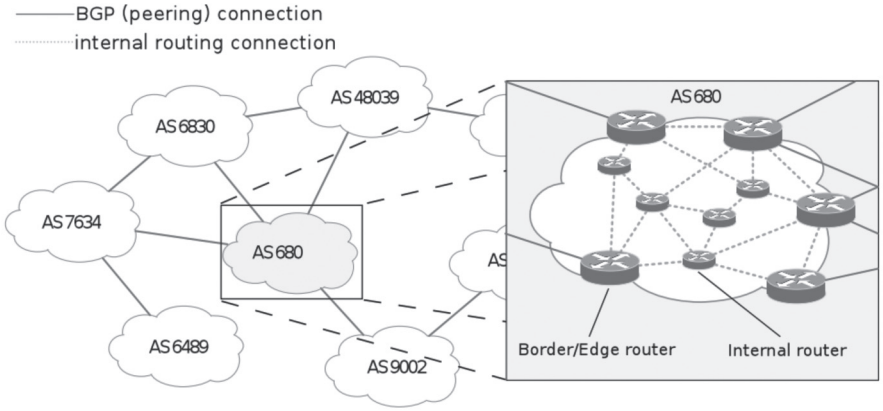
This section describes backgrounds of the Internet and its routing plane followed by an introduction and an explanation of Internet routing anomalies.

Internet routing

The current structure of the Internet is the result of massive growth in the last two decades, mostly driven by civil usage of the World Wide Web and other services such as IP-Telephony and IP-Television. The majority of Internet participants, either they use it for private or business purposes, has a limited view of the techniques behind the Internet. Although the Internet is seen as an abstract item, it is in fact just the interconnection of different independent networks, called Autonomous System (AS), as illustrated in Figure 1.

Each AS belongs to one administrative domain, most of them are large enterprises (e.g. ISPs, IT-Services), governments, organizations or universities. The interconnection of these networks is possible because of physical links between them. Routers connected to other ASs' routers are called border router or gateway. Thus, the Internet is not more than a network of networks. The connection between Autonomous Systems is called *neighborship* or *peering*. Each neighborship is related to at least one (commercial) agreement between the two parties. There are provider-customer, peering and transit relationships between ASs.

FIGURE 1: INTERNET ROUTING



To operate an AS as part of the Internet it is necessary to register a unique AS number. AS numbers are assigned by regional internet registries (RIRs) on behalf of the Internet Corporation for Assigned Names and Numbers (ICANN). Additionally, each AS needs at least one subnet of the global IP address space so it can be addressed by other ASs. These subnets are also regulated by ICANN and distributed by RIRs, e.g. RIPE NCC [1] for European customers. An AS announces owned IP addresses as *subnets* (also named *prefixes*; the prefix of an IP address defines the subnet to which an address belongs to) to each neighbor. The *announcement* (or *advertisement*) contains the served prefix in classless interdomain routing (CIDR) notation [11] together with the owners AS number as *origin* and additional information as described in the BGP [22]. Prefixes, reachable by a neighbor of a receiving AS, are then re-announced by that AS to all other neighbors, with the own AS number prepended to the origin AS. The concatenation of all AS numbers between an AS and the owner of a prefix builds the AS *path* for the prefix. When the receiving AS already knows another path to a prefix, only the best path will be chosen and sent to the neighbors.

Announcements of prefixes and AS paths are routing information used to deliver IP packets to their destination. To exchange routing information between AS border routers, BGP is used. BGP is the first and only routing protocol used in the Internet, so it is the de-facto standard, implemented in all participating border routers. Besides static routing protocols used inside ASs, BGP encounters the dynamics of a global and rapidly changing Internet. Border routers establish BGP connections to border routers of other ASs and exchange routing information with them. Since the Internet originally was an interconnection of trusted universities and research facilities, BGP assumes unlimited trust between neighbors regarding
routing information provided by them. Hence, BGP has no built-in verification mechanisms to check for the validity of routing announcements.

To provide access to the Internet, an AS needs routing information for every addressable Prefix. Because CIDR allows different prefix lengths, it is possible to aggregate them into shorter prefixes containing all subnets to decrease the number of necessary routing entries. To prevent routing failures, the most specific (longest) prefix determines the route to a destination address, if two or more prefixes overlap. Due to the implicit trust and a missing global authority, it is possible for ASs to provide invalid routing information. Thus, an AS can announce the reachability of an IP prefix, although it is not the legitimate owner nor does it have the advertised routing abilities.

BGP alongside Internet routing in general is subject of research activities for more than a decade [20, 15, 10]. Problems resulting from the weakness of implicit trust between neighbors, no matter whether they are in a provider-to-customer, a peering or transit relationship, cannot finally be solved. It is possible to filter announced routes from customers or peers but that is not sufficient to secure BGP routing as only few of the prefixes are originated by an AS's peers. Research projects and routing hardware vendors [14, 19] from time to time propose BGP optimizations or BGP successors to secure Internet routing [13] but none of them has been emerged to secure every day routing. Besides the goal to solve this issue, the research community accepts it as a fact and tries to find other ways to allow trustworthy inter AS routing. One of the main goals of network and internet security research is to provide reliable internet connectivity to end users, organizations and enterprises. To achieve this, the BGP-state of the Internet is continually monitored by different institutions and companies. BGPMon.net [2], as an example, offers services to inform victims of IP hijacking, in case of another AS illegitimately announcing any of their prefixes.

Most of the named research projects are built upon information collected by routing archives such as RIPE RIS [4] or routeviews [18]. Those archives peer with volunteer ASs and collect announced routes or received routing announcements from a route reflector, a border router that just reflects all received announcements to designated clients, inside different ASs around the globe. Relying on information derived from the routing layer itself is one of the handicaps all these projects have in common.

Routing anomalies

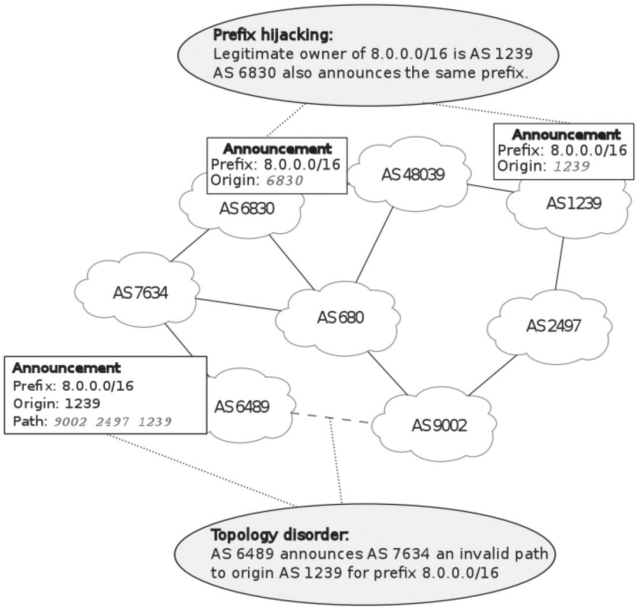
Anomalies within the routing plane of the Internet occur regularly and they last from only a few seconds to several months [7, 8, 16, 23]. This section will give a short summary of how anomalies can happen and how to react on them.

BGP is a message based protocol. Border, or *edge routers* of ASs send messages to their physically connected neighbors in other ASs to inform them about a) their own IP prefix and b) transitively reachable IP prefixes of other ASs. Beside other information, those messages at least contain the registered AS number of the peering AS, the AS number of the originating AS, meaning the AS that owns the announced prefix, and the AS numbers of all the ASs between the receiving and the originating AS, namely the AS path.

According to Lad et al. [16] and Qiu et al. [21], we consider a routing update an anomaly, when at least one of the following conditions is met:

- An invalid AS number is used.
- One or more invalid or reserved IP prefixes are used.
- The IP prefix is not owned by the originating AS.
- The same IP prefix is originated (announced) by two or more ASs.
- The given AS path has no physical equivalent.
- The provided AS path does not match common routing decisions.

FIGURE 2: ROUTING ANOMALIES: PREFIX HIJACKING AND TOPOLOGY DISORDER



The consequences (or incidents) of routing anomalies are commonly categorized into *blackholing, rerouting/path spoofing* and *hijacking* [21, 24]. For our research we only need to distinguish routing anomalies into two different types: prefix hijacking and *topology disorder*, as shown in Figure 2 and described below.

Prefix hijacking

Prefix hijacking occurs, when the given origin inside a BGP announcement, i.e. the owner of an IP prefix, is not the legitimated and registered AS itself. Prefix hijacking can affect a whole subnet or only parts of it with a larger prefix, which we then call *subprefix hijacking*. Subprefix hijacking differs from the common understanding of sub MOAS conflicts, as long as the subnet is legitimately assigned to another AS. The route selection process prefers paths with the longest prefix to determine the route to a specific IP address. In case of equal length, a MOAS conflict would match our understanding of prefix hijacking. Prefix hijacking could cause blackholing, when the wrongly announced prefix is not routed (or served) within the causing AS. It does not affect the whole Internet, it rather divides it concerning the announced prefix, one part of the Internet uses the benign and the other the bogus route, depending on AS specific routing decisions.

Topology disorder
Topology disorder happens, when an announced path is invalid, i.e. has no corresponding physical equivalent that could be traversed or violates reasonable routing decisions. Such a disorder could lead to longer and also shorter paths, hence influences the route selection process of other ASs. While prefix hijacking caused by accidental misconfiguration, a manipulated AS path can only happen intentionally save those caused by bugs in router firmware, but latter are rather unlikely.

Routing anomalies are not necessary harmful. Large service providers might enforce anomalies to realize geographical load balancing or multi homed ASs. An AS is multi homed, when it has two or more relations to ASs where it is customer in a provider-to-customer relationship, i.e. needs other ASs to address the rest of the Internet, to increase its own routing abilities and to have a backup path if one provider fails. That means, occurring anomalies caused by topology disorder have to be examined in a special way to classify them as legitimate or illegitimate ones. The results of this classification should be reliable enough to send proper alerts to legitimate owners of prefixes or administrators of causing and affected ASs to be informed about the anomaly and to solve it.

Conclusion
As a matter of fact, no real countermeasures to routing anomalies exist. It is thinkable, e.g. in case of a race for a specific IP prefix, to announce longer prefixes than the causing AS. This game stops at least at 24 Bits length because longer prefixes are not valid in the Internet routing. Thus, the conflict remains. Unless BGP could be totally replaced, AS operators and researchers have to deal with its weaknesses.

3. ANOMALY DETECTION

Since BGP routing weaknesses and anomalies are still topic of active research, various mechanisms and algorithms for anomaly detection have been proposed and developed by the research community. This section describes our applied approach to detect routing anomalies based on already existing solutions.

Our anomaly detection incorporates already existing approaches, which we combine to gain benefits from all solutions [26, 16, 21]. Based on these, we examine current routing announcements from the beginning of 2013 until the end of October 2013. We evaluate our results against a list of known anomaly routing events from Team Cymru and BGPMon.net [5, 2].

The named systems are mainly based on historical routing information derived from routing archives [4, 18]. To improve detection rates, detection runtime and in order to detect anomalies not yet in these lists, we filter the routing announcements prior giving them to the anomaly detection. Based on our broader knowledge we filter announcements that are reliably proper so that there is no need to run each classifier on it. Our contribution is that not only anomalies are classified on broader ground truth evidence, but additionally to confirm information found in the announcements prior the detection.

To improve the reliable data our solution is based on, we gather additional reliable routing (and especially peering) information from different (primary) sources of the Internet. How we achieved this is shown in the next section 4 of this paper.

While parsing retrieved BGP archives each contained announcement is evaluated before being inserted into the analysis database. As mentioned earlier, the database contains all announcements from the beginning of the regarded interval, i.e. January 2013 in this case, until the receive date of the examined announcement. If a database entry holds an announcement that is still vital and provides the same prefix but is originated by another AS, a MOAS conflict is detected. Such conflicts are calculated per prefix and reported with the affected prefix and all participating Autonomous Systems.

Afterwards, the AS path of each announcement is examined and checked for known and confirmed AS peering relations. Those peering relationships are derived from the database containing historical announcements. As this information is not sufficient, paths shall be examined based on the database created as a result of this papers research, see the following sections for further details on how the data is collected and evaluated. When no such peering relation can be confirmed for each contained AS link, an anomaly is raised with the affected announcement and the corresponding ASs. When an anomaly is detected, additional actions are triggered, such as querying the corresponding ASs or an internet exchange point both ASs are connected to.

Conclusion
Anomaly detection is primarily based on publicly available data and has to be improved by additional collected data as evidence of ground truth. Detected anomalies are stored inside a separate database for further use such like BGPMon.net or end user warning systems [27].

4. IMPROVING THE GROUND TRUTH EVIDENCE

This section describes our contribution and the steps we make to collect further information on routing relationships from other primary and reliable sources, in order to enlarge the assumed ground truth of our detection system.

To improve the basis of the classification of routing announcements, we need to obtain reliable information about peering and other business relationships of ASs, but unless such information is publicly available and it is known how to retrieve it, there is currently no way to take them into account. Confidential information aside, there is a lot of publicly available and usable information about AS relationships.
Existing approaches obtaining AS relationships [12, 9] use information gathered from within the routing system itself, based on collected BGP announcements and derived node degrees.

In the context of our project, the examined autonomous systems are restricted to those, located in countries of the European Union (EU). Having a number of 28 countries, we retrieved a number of about 11.500 ASs from the RIPE whois database located in respective countries.

This represents about 25% of the registered ASs worldwide. The number of registered and announced IPv4 prefixes is about 70.000 at the time of writing, what is around 14% of the globally assigned 510.000 prefixes we found in a recent table dump in October 2013 [4].

Our goal is to collect additional information on those EU-located ASs in order to improve the routing anomaly classification. Several sources exist, where such information could be found. We start with a naive approach and collect whois data from RIPE [1] first. A RIPE whois database entry contains information of a registered ASs, its AS number, the name, description, contacts and various other. The number of queries at RIPE is generally limited to 1.000 queries per 24 hours and IP address, when contact information is contained. Additional information on RIPE's whois database usage is given below.

Reaching this limit quickly leads us to look for other sources containing similar information. The website peeringdb.com [3] contains specific information about inter-AS peering and holds a list of known Internet Exchange Points (IXPs) in Europe. An Internet Exchange Point is a datacenter with special focus on network peering. To get an AS connected to many other ASs with little effort, AS operators rent special network ports in that datacenter. Depending on an ASs peering policy connections between different ASs can be established and used for BGP peering. Hence, it is feasible to establish peering connections to many other ASs located at the same datacenter with just one physically network connection. Due to the ease of establishing peering, ASs located in large IXPs commonly have many peers.

The peeringdb.com database does not claim to be complete but it gives a good starting point for further research. We extracted EU-located IXP datasets, 118 in number, including their website addresses and a list of ASs peering at them. An AS peering at a specific IXP is referred to as *member* of this IXP.

An AS entry from peeringdb also contains the address of looking glass servers provided by the AS, when it is publicly accessible. The utilization of information obtained from looking glasses is described in this section below. To get listed inside the peeringdb database it is necessary for an AS to register and provide sufficient information for the entry. Consequently, not all IXPs, ASs and peering relationships are listed there. As part of their peering policy, some ASs require the existence of a database entry of the peering partner at peeringdb.com to peer with them. Due to the fact that database entries are manually maintained by each AS itself, the database can't be regarded up to date.

Starting with the list of EU-located IXPs [3] we collect information from the IXPs' websites directly. Most of them provide a list of members and some additionally a detailed peering matrix. This is valuable and reliable information on actual peering relationships between listed ASs since IXPs get paid for peering services and therefore update information of their members regularly based on their business processes. If there is no peering matrix provided, the majority of IXPs at least list peering policies of their members showing whether it is open, selective or closed. The usage of peering policy information is also described below.

Looking glass

In order to get reliable information about peering relations, BGP specific information such as full routing tables and next hops for various routes are of interest. To ensure, that derived information is reliably and correct, we collect information from AS border routers directly by accessing them through their looking glass service. Looking glass servers provide access to live routing information of an AS itself. A looking glass service directly queries the routers involved in BGP operation to provide up to date information about actual relationships between BGP nodes. Based on settings and restrictions set by an AS's network operation center (NOC), different information can be requested from looking glass services.

Automated querying of looking glass servers is a great challenge. Where BGP routers provide direct access to the routing devices (e.g. via telnet), a more or less consistent interface is available to query the nodes participating in questionable routing by automatic means.
Looking glass servers, however, usually provide web interfaces to access the required information. Such a web interface provides access to at least partially the information available from border routers. The type of information differs between most of these web interfaces, as well as the web interfaces itself. Although usually optically similar, the technical differences of the provided web interfaces make automated querying and information parsing a complex task that often requires human intervention.

Our system queries as much EU located looking glass servers as possible and tries to reach a large coverage. Based on gathered looking glass information, the first hop of each route can be used to verify an indicated physical connection between ASs. Thus, the derived information will be used to mark peering relationships gathered from BGP announcements as confirmed. As another contribution, we provide additionally collected views on the Internet routing and can use them as another source for routes and AS paths for anomaly detection in addition to those, collected from routing archives. When an unavailable looking glass server is found in the list, we inform the provided AS's network operation center to inform them about the orphaned database entry and ask for an alternative address for looking glass access.

Neighbor information from a looking glass interface can help to decide whether routes are valid or not by providing information if those routes can actually exist. Invalid routes can be filtered, if the announced route has no physical counterpart, valid routes can be verified on each edge. Additional services provided from looking glasses such as ping or traceroute provide additional information about the connectivity and availability of BGP infrastructure outside of the Border Gateway protocol itself. As all of these information, BGP and non-BGP, are available from different viewpoints, verification or falsification of announced routes is easier and more reliable by making automated use of these information.

Peering policies

66 of 119 examined IXPs publish information about peering policies of their members. Information about peering policies are used if yet unknown AS paths are announced and examined by our anomaly detection tool. Topological details are required when it comes to a decision, whether a newly announced path is reasonable or not. It seems more likely that ASs peer with each other when they share the same IXP. The data from peeringdb [3] and the

information gathered from the EU IXPs directly indicate that many ASs are located at several IXPs. This increases the number of possible peering relationships in case of an open peering policy. If the peering policy is selective or closed, new peerings are less likely but more stable in general. When we find information on peering restrictions/conditions, e.g. up to date entries within the peeringdb.com database, they are additionally checked by our system. An open AS's peering policy indicates a smaller or non-profit AS, since larger Tier-1 or Tier-2 provider earn money to act as a smaller AS's upstream and have case-by-case or closed policies. We check this information with existing AS relationship classification [12] and obtain more reliable information on ASs' relationships to be used by our anomaly detection and classification.

Whois information

Whois information from regional registries such as RIPE [1] contain details about the owner of an AS. Data gathered from whois services is primarily used by our system to determine the country an AS is located in. For this purpose we firstly request all *descr* and *address* fields and parse them for country information. If no contrary information is found, the AS is counted to the corresponding country. There are few ASs with opposing country information in the address fields of the whois entry. This is likely, when a corporation that operates an AS has several business units, responsible for network operation, located in several countries. Those entries have to be checked and added to the database manually in the current implementation.

Secondly, whois information is used to classify occurring MOAS conflicts. For each AS participating in the conflict the given company or administrator is checked and an affiliation or relationship factor between them is calculated. If our heuristics indicate closer relationship, e.g. the same company name or equal responsible email addresses, a conflict is rather expected legitimate.

Thirdly some ASs provide peering hints in their whois entry. *Import* fields reflect which ASs and Prefixes are imported and accepted from which peer. *Export* fields name the corresponding outgoing rules. If paths are announced that violate those given rules, they are suspicious but not sufficiently illegitimate since a connection between the ASs might yet exist. The information contained in both fields are considered by the heuristic classifier.

In general it should be mentioned, that whois information from RIPE should be considered outdated and unreliable. Nonetheless, this information should not be ignored when classifying hijacking events as long as it is not contradicted by another more reliable source.

Conclusion

Based on the set of derived information we mark announcements (and especially contained routes within them) as confirmed, when it is evidence through our collected data, that such a peering really exists. Our database contains reliable information on peering from publicly available sources, facts on peering policies and historical data to be used by our heuristics.

5. CLASSIFICATION

This section provides a short overview on the improved classification and detection of routing anomalies.

The Classification of each routing announcement takes place before and after the anomaly detection itself. As stated earlier, anomalies can be legitimate or illegitimate. To differentiate between both classes, our classifier uses the additional information that has been gathered in the *improving the ground truth evidence* process.

Prefix hijacking
For prefix hijacking events, additional information from whois services provided by Internet registries is necessary to determine the legitimate owner of an affected prefix. When the legitimate owner is known, all other ASs, involved in this anomaly, will be checked for a (non-routing) relationship between each of them and the legitimate owner. It will be estimated from the information found in the internet registries whois database and on IXP websites as described earlier. If a relation is found and considered reasonable, i.e. AS operators are named similar or the contacts are equal, the anomaly will be classified as rather legitimate.

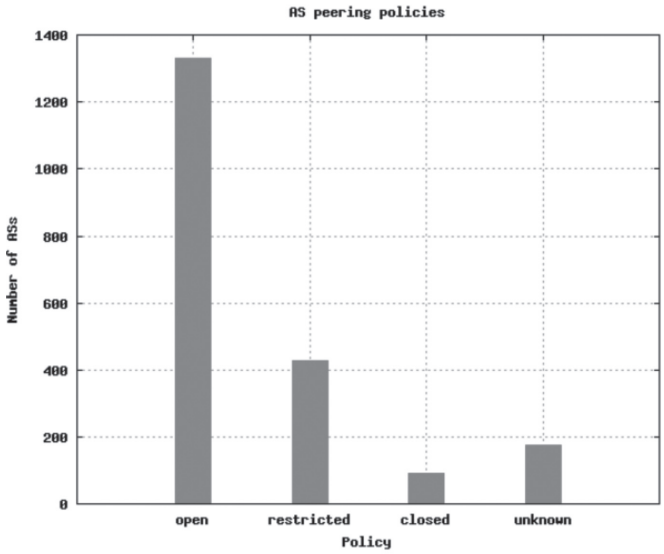
Topology disorder
If a topology disorder is detected in a BGP announcement, the corresponding path will be examined in a special way. First, all the ASs on the path will be checked for historical suspicious behavior and the relationship to predecessor and successor. Additional information about peering relationship between ASs can be used, to mark newly created links harmless. Wrong topological information like an attack against targeted ASs or prefixes can be used to influence routing decisions and lead to the usage of unpredicted paths for affected prefixes.

6. EVALUATION

We collect routing and peering information as described above for our studies. This section describes the collected data in detail and evaluates the impact of enriched ground truth evidence to existing methods and algorithms to observe Internet routing anomalies. One of the most valuable achievements is the decreased number of suspicious peering relationships through reliable evidence of actual connectivity between ASs.

The list of known IXPs located in the EU is used to gather peering relations of the participating members. The database at peeringdb.com lists 3065 ASs connected at these IXPs. We collected 66 lists from websites of 119 IXPs. Using our approach, we have found 5185 ASs as member of these. The difference between our AS list derived from IXPs directly and the list provided by peeringdb for large IXPs is presented in Table 1. Related to the IXP member provided by peeringdb, our system collected 74% more entries in total with assumable higher evidence.

FIGURE 3: AS PEERING POLICIES



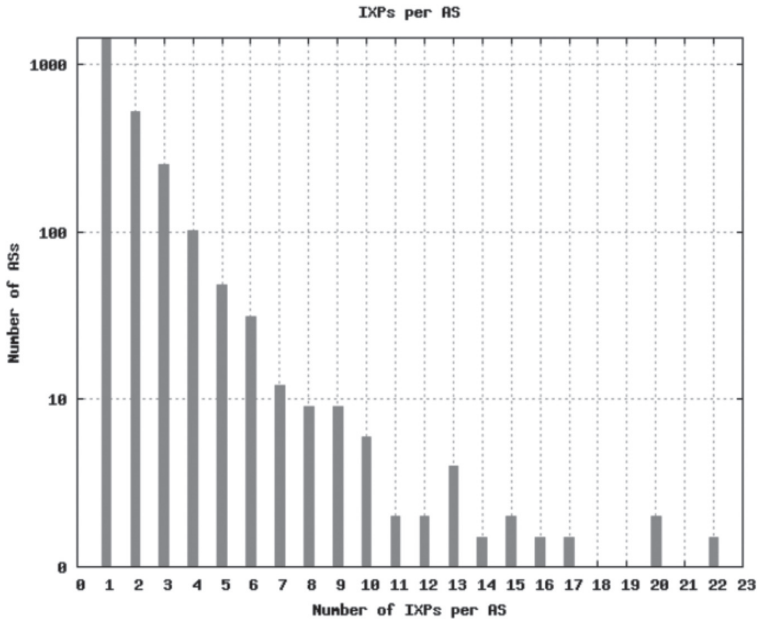
17 of those member lists contained the peering policy of 2024 different ASs used for the classification heuristics as described in the sections above. This additional information increases the number of conflicts being classified as legitimate ones.

TABLE 1

IXP	peeringdb.com	our database
AMS-IX	564	627
DE-CIX	453	515
France-IX	191	449
NL-IX	173	390
V-IX	87	120
Netnod	14	88
DIS-DK	41	42

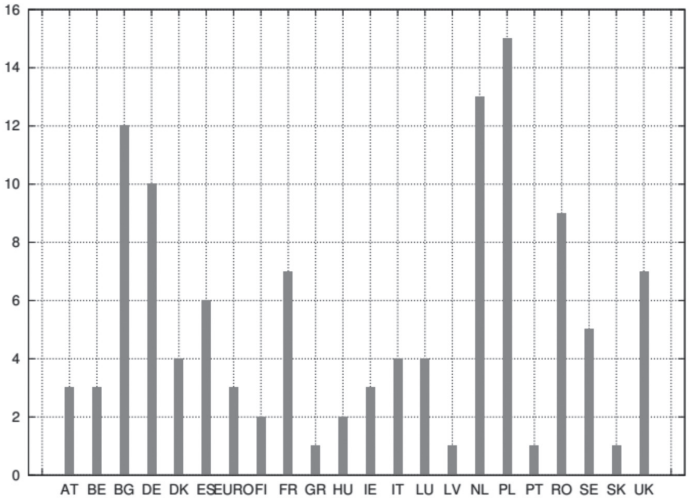
According to Figure 3, 1452 ASs have an open, 454 a restricted, 92 a closed and 175 a currently undeterminable peering policy. The number of IXPs, an AS is located at can be used as an indicator for its size or role inside the Internet routing system and increases the number of potential peering ASs. Therefore we determined the number of IXPs an AS is located at. As shown in Figure 4, most of the ASs are located at few IXPs.

FIGURE 4: NUMBER OF IXPs PER AS



The collection of looking glass URLs and the responses of them allows to gain evidence on direct AS peering relationships. When an unknown peering is found in the analyzed BGP announcements the looking glass servers of both related ASs are queried. If confirmed by the existing looking glasses the information is added into our database. During our research, we examined 116 looking glass servers run by EU-based network operations centers, from which we found 97 to be reachable for querying. Figure 5 shows the number of looking glasses by country. Although all operated within the European Union, most of these NOC's operate BGP nodes around the whole globe, having peering connections with major international operating ASs. Therefore, even a European effort to gather live BGP data from looking glasses provides a useful data base to make sense of anomalies detected throughout the whole Internet. The challenges of gathering those data could be overcome by a common approach supported by the network operations centers.

FIGURE 5: FOUND LOOKING GLASS SERVERS IN THE EU BY COUNTRY



The adoption of looking glass interfaces is still ongoing work.

Especially for the classification of prefix hijacking events the data gathered from RIPE's whois database is used. In case of a hijacking anomaly, the whois data regarding to the owners of affected ASs is considered. When equally or similarly named organizations own all those ASs, the conflict is rather classified as legitimate by the heuristics. Whois data of all 11687 ASs we located inside the EU has been pulled from RIPE to be used in our prefix hijacking classifier.

7. DISCUSSION AND FUTURE WORK

The state of Internet routing is still hard to determine continuously and thus, still vague. The number of involved autonomous systems increase and the number of IP prefixes will massively increase when IPv6 is implemented by all of them. That is why adjusting anomaly detection mechanisms is yet necessary. Our contribution is a larger data basis gathered from primary sources, that are trustworthier sources as those only based on information from within the examined routing system itself, i.e. routing archives, used for identifying and classifying routing anomalies. We created a system to increase evidence of routing information derived from these publicly available sources. This enrichment leads to more reliable detection and classification mechanisms and allows to decrease the number of decisions made on unreliable information. There is no final solution in sight to secure Internet routing at all. Thus network operators and security engineers have to work with continuously improved tools. The work on detection and classification of anomalies is not finally done and we will adjust our solution in the future to become more efficient and to collect more reliable information from primary sources such as IXPs and ASs themselves. To allow statements on the Internet routing state as a whole the restriction to EU-located ASs should be weakened and the number of monitored ASs shall be increased.

REFERENCES:

[1] Ripe network coordination centre (ncc). http://www.ripe.net (29. Nov 2013)
[2] BGP monitor, Nov 2013. http://www.bgpmon.net (29. Nov 2013)
[3] Peering database, http://www.peeringdb.com (29. Nov 2013)
[4] Ripe routing information service (RIS), http://www.ripe.net/data-tools/stats/ris (29. Nov 2013)
[5] Team cymru, http://www.team-cymru.org/Monitoring/BGP/ (9. Nov 2013)
[6] W. Aiello, J. Ioannidis, and P. McDaniel. Origin authentication in interdomain routing. In *Proceedings of the 10th ACM conference on Computer and communications security*, CCS '03, pages 165–178, New York, NY, USA, 2003. ACM.
[7] H. Ballani, P. Francis, and X. Zhang. A study of prefix hijacking and interception in the internet. *SIGCOMM Comput. Commun. Rev.*, 37(4):265–276, August 2007.
[8] K.W. Chin. On the characteristics of BGP multiple origin AS conflicts. In *Telecommunication Networks and Applications Conference*, 2007. ATNAC 2007. Australasian, pages 157 –162, December 2007.
[9] G. Di Battista, T. Erlebach, A. Hall, M. Patrignani, M. Pizzonia, and T. Schank. Computing the types of the relationships between autonomous systems. *IEEE/ACM Trans. Netw.*, 15(2):267–280, April 2007.
[10] R. Dube. A comparison of scaling techniques for BGP. SIGCOMM *Comput. Commun. Rev.*, 29(3):44–46, July 1999.
[11] V. Fuller and T. Li. Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan. RFC 4632 (Best Current Practice), August 2006.
[12] L. Gao. On inferring autonomous system relationships in the internet. *IEEE/ACM Trans. Netw.*, 9(6):733–745, December 2001.
[13] S. Goldberg, M. Schapira, P Hummon, and J. Rexford. How secure are secure interdomain routing protocols. In *Proceedings of the ACM SIGCOMM 2010 conference*, SIGCOMM '10, pages 87–98, New York, NY, USA, 2010. ACM.
[14] S. Kent, C. Lynn, and K. Seo. Secure Border Gateway Protocol (Secure-BGP). *IEEE Journal on Selected Areas in Communications*, 18(4):582–592, April 2000.
[15] C. Labovitz, G. R. Malan, and F. Jahanian. Internet routing instability. *IEEE/ACM Trans. Netw.*, 6(5):515–528, October 1998.
[16] 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*, USENIX-SS'06, Berkeley, CA, USA, 2006. USENIX Association.
[17] M. Lepinski and S. Kent. An Infrastructure to Support Secure Internet Routing. RFC 6480 (Proposed Standard), February 2012.
[18] David Meyer. Route views archive project.
[19] J. Ng. Extensions to bgp transport sobgp ertificates. Internet-Draft, May 2005.
[20] V. Paxson. End-to-end routing behavior in the Internet. In *Conference proceedings on Applications, technologies, architectures, and protocols for computer communications*, SIGCOMM '96, pages 25–38, New York, NY, USA, 1996. ACM.
[21] J. Qiu, L. Gao, S. Ranjan, and A. Nucci. Detecting bogus BGP route information: Going beyond prefix hijacking. In *Third International Conference on Security and Privacy in Communication Networks and the Workshops*, pages 381–390, 2007.
[22] Y. Rekhter, T. Li, and S. Hares. A Border Gateway Protocol 4 (BGP-4). RFC 4271 (Draft Standard), January 2006.
[23] T. Wan and P.C. van Oorschot. Analysis of BGP prefix origins during Google's May 2005 outage. In *Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International*, page 8 pp., april 2006.
[24] Z. Zhang, Y. Zhang, Y. C. Hu, and Z. M. Mao. Practical defenses against BGP prefix hijacking. In *Proceedings of the 2007 ACM CoNEXT conference*, CoNEXT '07, pages 3:1–3:12, New York, NY, USA, 2007. ACM.
[25] 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*, IMW '01, pages 31–35, New York, NY, USA, 2001. ACM.
[26] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, S. F. Wu, and L. Zhang. Detection of Invalid Routing Announcement in the Internet. In *Proceedings of the 2002 International Conference on Dependable Systems and Networks*, DSN '02, pages 59–68, Washington, DC, USA, 2002. IEEE Computer Society.
[27] M. Wübbeling. Visibility of Routing Anomalies for End Users. In Christoph Pohl, Sebastian Schinzel und Steffen Wendzel, editors. Proceedings of the Eight GI SIGSIDAR Graduate Workshop on Reactive Security (SPRING). Technical Report SR-2013-01, GI FG SIDAR, Munchen, Februar 2013