Saarland University

Faculty of Science and Technology I

Department of Computer Science

(Bachelorthesis)

Correlating BGP and Traceroute Data

submitted by
Simon Koch

submitted
June 3, 2015

Reviewers
1. Prof. Dr. Michael Backes
2. Dr. Christian Rossow

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Statement in Lieu of an Oath

I hereby confirm that I have written this thesis on my own and that I have not used any other media or materials than the ones referred to in this thesis.

Einverständniserklärung

Ich bin damit einverstanden, dass meine (bestandene) Arbeit in beiden Versionen in die Bibliothek der Informatik aufgenommen und damit veröffentlicht wird.

Declaration of Consent

I agree to make both versions of my thesis (with a passing grade) accessible to the public by having them added to the library of the Computer Science Department.

Saarbrücken, ____________________ ____________________________
(Datum/Date) (Unterschrift/Signature)

Acknoledgements

I want to thank my advisers Esfandiar Mohammadi and Sebastian Meiser for their valuable and motivating support as well as for the time they spent scrutinizing my attempts on writing this thesis. They provided me with insights into scientific writing and scientific work I would have never gained on my own. I hope they will continue supervising bachelor and master students as I for myself learned a lot from them.

Abstract

The Internet, as a network of networks, routes traffic by relaying the data from one network to another, thus enabling the owner of those networks to observe the contained information or relay it to intelligence agencies. In this thesis, we evaluate the program *Traceroute* as a means to identify individual steps in a connection. We present a tool chain to correlate *Traceroute* detected paths with paths derived from the *Border Gateway Protocol* and show a strong correlation between them. Additionally, we present multiple reasons why mismatches between Traceroute paths and the Border Gateway Protocol paths have to be expected and categorize our observations accordingly. We conclude by arguing that *Traceroute* is a valid path detection tool.

Contents

1 Introduction

We are under surveillance. Even though there have been reports about intelligence agencies monitoring Internet communication [13, 28], the revelations of 2013 by Edward Snowden showed a whole new scale of global surveillance [19]. The revelations showed that intelligence agencies, first and foremost the NSA, are trying to gain complete access to any communication in the Internet. There are new articles on a regular basis that reveal new information about the surveillance efforts and how our everyday Internet traffic is observed.

The Internet is a collaboration of thousands of different, smaller networks called *autonomous systems* (ASes, sg. AS). ASes by themselves are composed of interconnected machines that are assigned IPs from a set of IPs (called prefixes) exclusively owned by the AS owner. Every AS is connected to their neighboring ASes, called peers. For communication in the Internet, a chain between such peers is used, where each AS relays received traffic one step further to the target AS that contains the target machine.

Since any traffic going to or leaving an AS is relayed through its machines, an owner of an involved AS has, as a consequence, a certain degree of access to the transmitted data. Even if the traffic is encrypted, meta data, such as the connection duration, data volume, and involved parties can still be observed. Meta data can reveal the type and content of a connection, hence it is highly sensitive, as for example the news sites that a person visits can reveal his or her political and social opinions. Thus, we are interested in understanding to whom our meta data is leaked. The present thesis focusses on the aspect of understanding to whom meta data is leaked[1].

To gain information about involved parties in a connection we need to find out the connection path on the AS level. The involved ASes can then easily be translated to real life entities, as AS ownership is public knowledge. There are three possible approaches we can attempt: we can *ask the ISP*, use the *Border Gateway Protocol (BGP)*, or use a program called *Traceroute*.

Asking the ISP. The first approach is to ask the AS owner where the traffic is routed to, and then query all subsequent AS owners on a per connection basis. This approach is obviously unfeasible, as carriers do not provide tools for such requests to be answered in a timely manner. Furthermore, this information is usually considered a business secret.

BGP. BGP connects ASes with each other and represents today's main approach of exchanging reachability and routing information in-between ASes. BGP allows ASes to notify their peers about which connections they provide by exchanging messages. Multiple BGP messages for the same routing information can be received by a peer. However, the peer only selects one message that it incorporates in its own set of routing information. This

[1] Research showed, that anonymization networks such as Tor also need to be aware of path participants as a network attacker can deanonimize Tor users in certain cases [14].

set is then advertised to the peers of the AS. When further advertising a received routing information the AS that previously received the message, appends its own number onto a already contained list of previous steps, thus a BGP message also contains the propagation path. Therefore, it is possible to derive the path of the traffic based on BGP communication data.

The expected usage of BGP is that ASes only advertise routing information they intend to adhere. However, ASes can lie about routing information, and advertise connections they do not intend to provide. This has happened in the past, e.g., as a means for censoring [12]. Therefore, the soundness of a given BGP message is not guaranteed. Furthermore, publicly available BGP data is sparse and only covers a small part of the Internet [10, 11, 16]. Consequently, BGP appears to be a suboptimal solution to make inferences about a communication path.

Traceroute. Traceroute is a tool designed to discover the machines participating in a connection. Traceroute can be run from any machine and exploits aspects of the *Internet Protocol* (IP). It is designed to get replies from every server participating in a connection, thus revealing their IP-address in the response. The resulting list of IP-addresses can then be mapped onto *autonomous system numbers* (ASNs). This list then results in the AS path. This approach has problems of its own problems as IP-address to ASN mapping is not necessarily reliable and border routers are known to share IP-addresses between peering ASes as Mao et al. have shown [22].

Furthermore, the question arises, whether the Traceroute path discovered reflects the actual path the connection takes. Paths may change rapidly due to load balancing or be incomplete as not every machine participating in the relay has to respond to traceroute probes.

Contribution. In this thesis we address the question whether Traceroute can be considered reliable for path detection by performing Traceroute measurements from within different ASes to a multitude of targets. We correlate the findings with the corresponding BGP data advertised by the start ASes using the Routeviews projects and RIPE-Ris. To achieve this we implement a tool chain that automates the needed processes. We show a strong correlation between the two data sets as well as the relatively steadiness of the AS paths discovered by either method. Additionally we propose *Route Fluttering*,*Short Lived Routing Problems*, and *Transit ASes* as reasons why Traceroute detected paths cannot always be a perfect match with BGP advertised paths and categorize observed patterns by those reasons. Based on our findings we argue that Traceroute is as a valid path discovery tool.

2 Related Work

There is a multitude of work covering the topic of routing, structure of the Internet and reliability of traceroute or BGP on inferring the observable Internet structure. Furthermore, there is already previous work on the correlation between traceroute and BGP. This chapter covers first work concerning the structure of the Internet and second work relating to the reliability of BGP and traceroute.

Gao published the first of a chain of papers concerning the relationships and connectivity between autonomous systems in the Internet. Gao classifies AS relationships into different categories and presents a solution to infer them from BGP data. Gao uses internal AT&T data and the public the WHOIS service to verify the results [17]. Even though the paper gives insight into the theory of routing policies and relationships between different ASes, it does not approach the topic of package routing predictions, which is the main topic of this thesis. The results of this paper argue strongly toward logical and predictable connectivity and therefore predictable routing paths.

Dimitropoulos et al. build upon the idea of Gao and introduces a heuristics for inferring AS relationships. The authors perform a survey to validate their results. They further compare their results with BGP information and show that BGP does miss a significant quantity of up to 86.2% of true adjacencies [15]. This significantly points out the limitations of BGP in inferring the structure of the Internet on the AS level and the need for a tool that is not based on voluntary public information and company policies.

Anonymous researchers discovered large parts of the Internet Topology in the Internet Census 2012. They infiltrated a large amount of weakly secured embedded devices and ran multiple scans, including Traceroute. The researchers were able to obtain 68 million traceroute records. It is noteworthy though that the means of this discovery approach were unethical and thus the results are not usable for scientific purposes. However, the Internet Census showed that a vast amount of the Internet can be scanned if enough different probe placements can be obtained [1].

Ricardo Oliveira et al. measure how much and how precise the current Internet structure on the AS-level is actually revealed. They establish a ground truth of connectivity and compare it with the inferred topology maps to establish empirical facts about the reliability. As the work is concerned with AS-topology and connectivity, it does not focus on actual routing paths through those networks but shows that BGP data is unreliable for non carrier connections and therefore another approach to gain information about routing and connectivity has to be established [23].

A. Faggiani et al. discuss the potential of traceroute based data in revealing Internet topology information in addition to BGP data. Faggiani et al. point out that the BGP infrastructure alone is only able to cover up to 15.9% of the Internet core. They further show evidence that using traceroute would give a significant improvement on Internet topology and connectivity insight [16]

Zhang et Al. point out problems using traceroute and mapping its result onto ASNs to get ASN paths. The authors show that mapping IP-addresses on ASNs by longest prefix matching is error prone and may yield wrong results in a significant amount of cases. They further propose and present a solution to counter this effect [32]. We consider the presented problems and will address whether they still manifest in current measurements.

Mao et al. address problems of mapping traceroute hops onto AS numbers. The authors present results showing that BGP and traceroute data differ in AS level routes due to Internet Exchange Points, other traceroute detectable anomalies, and an incomplete IP-address to ASN mapping [22]. As this thesis covers the same area of research, we consider the presented problems and discuss their impact on our measurements.

3 Background

In this chapter we explain the necessary background to understand the tool chain that gathers information on the correlation between Traceroute and BGP derived AS paths. The inner workings of the tool chain are described in Chapter 5. This chapter explains BGP in general terms and throughly explains those parts that are utilized by our tool chain. Furthermore, we describe the needed aspects of MRT, which is a format for storing BGP data with its context information. We conclude the chapter by an explanation of the program Traceroute which is also utilized by our tool chain.

3.1 BGP

Routing information is needed by everybody participating in routing a package. A package is routed by continuously forwarding a package until the package reaches its assigned destination. The routing information contain the information for the next step.

BGP is todays main tool for the advertisement of routing information between peering ASes. BGP enables peering ASes to advertise routing information on what IPs they connect to, to their peers. The IPs are encoded in consecutive blocks called prefixes. BGP information is send in messages that are exchanged by peers. There are multiple different types of messages which are used by the BGP protocol. We only focus on *BGP update messages* as the other message types do not contain useful information for us.

3.1.1 BGP Update Messages

BGP messages are in binary format and follow a structure specified in RFC4271, extended in RFC4893 to account for bigger ASN numbers, and by RFC4760 to account for different Network Layer protocols (e.g. IPv6) [6,25,29]. Each *BGP message* is preceded by a header containing a marker, the length of the message, and the type of the message.

A *BGP update message* consists of 5 fields of varying length: Withdrawn Routes Length, *Withdrawn Routes*, Total Path Attribute Length, *Path Attributes*, *Network Layer Reachability Information (NLRI).* The Withdrawn Routes Length and Total Path Attribute Length field determine how big the Withdrawn Routes and Path Attributes fields are respectively. The length of the NLRI field is determined by the overall length of the message minus the already used length [6,29].

An *IP-Prefix (short: prefix)* determines a block of *IPs.* A prefix is specified by an IP-address and the first n significant bits. Any IP-address whose first n bits are identical to the n significant bits of a prefix belongs to the block of IPs defined by that prefix.

A BGP update message advertises new routes in form of prefixes. A BGP update message advertises at most one new route related to multiple prefixes that share the same Path Attributes. The prefixes of an advertised route are either contained in the *NLRI* field of the message or in the *MP_REACH_NLRI* path attribute. Each BGP update message has to contain a single *AS_PATH* attribute. The AS_PATH attribute identifies the autonomous systems the BGP update message has passed through [6,29]. Only the 'best' route is selected for re-advertisement to the peers. A router selects the 'best' route by using different policies which are specified by the administrator.

Routes can be merged. This is done if a router has routes with overlapping prefixes (e.g. *10.0.0.0/30* and *10.0.0.4/30*). In such case both update messages are merged into one that contains the new prefix (*10.0.0.8/30*). The *AS_PATH* attribute is also merged and now consists of a list that only contains the ASN the merging router belongs to. The list is followed by a set that contains all ASNs of both *AS_PATH* attributes of the old routes in no specific order.

A *BGP update message* contains multiple *withdrawn routes* represented by their respective prefix. The prefixes are either listed in the *Withdrawn Routes* field or in the *MP_UNREACH_NLRI* path attribute [6, 29].

3.1.2 MRT

The *Multi-Threaded Routing Toolkit Routing Information Export Format (MRT)* describes a format to export routing information. Thus it provides the means to collect BGP routing information from different routers and save them with their context information for later processing [21]. The relevant *MRT* types used in the later on described tool chain are *TABLE_DUMP_V2* and *BGP4MP*.

TABLE_DUMP_V2. A *BGP Routing Information Base (RIB)* stores all route updates received by a router and thus contains all current routes known to the router. The *TABLE_DUMP_V2 MRT* type is used to save complete *RIB* dumps of each peer of the collector. Rib dumps typically solely contain this data type.

A *TABLE_DUMP_V2* is always started by a *PEER_INDEX_TABLE*. A *PEER_INDEX_TABLE* contains the list of peers of a BGP collector. The *PEER_INDEX_TABLE* is followed by the actual RIB of the peers. A *RIB Entry* sequence is preceded by a header that defines the prefix the *RIB Entries* are related to and the amount of *RIB Entries* that follow.

A *RIB Entry* consists of the *Peer Index*, *Originated Time*, *Attribute Length*, *BGP Attributes* fields. The *Peer Index* references the peer that sent the data, in the *PEER_INDEX_TABLE*. The *Originated Time* contains the time the contained information was received. The *Attribute Length* determines the length of the *BGP Attributes* field. The *BGP Attributes* field contains the *PATH_ATTRIBUTES* of the *BGP update message* relevant for the advertised prefix received by the respective peer [21].

BGP4MP_MESSAGE. The *BGP4MP_MESSAGE* precedes a BGP message received by the peer and is used to store context information. The *BGP4MP_MESSAGE* contains among other fields the *Peer AS Number*, and the *BGP Message*.

The *Peer AS Number* field specifies the AS Number of the peer that sent the contained BGP Message. The *BGP Message* field contains the complete *BGP message* as sent by the peer [21].

A dump usually consists of multiple of such MRT entries and encompasses a certain timeframe chosen by the collector. All BGP messages received by the peers in that timeframe are dumped into the same file using this MRT type.

3.2 Traceroute

Traceroute is a program used to detect participating server in a routing relay. It exploits features of IP to gain responses from participating servers containing their IP-address.

Traceroute uses the *time to live (ttl)* of IP packages to trigger a response from each participating server in a connection. A *ttl* denotes the remaining lifetime of a package in terms of steps. This lifetime ensures that even if the package can never reach its destination, it will not be relayed through the Internet for forever. Each time a package is forwarded to another server the *ttl* is decreased by one [7]. As soon as a package with a *ttl* of zero reaches a server the server notifies the sender of the decay of the package using *ICMP* [8].

Traceroute starts its detection by sending a package to the target with a *ttl* of one and then incrementally increases the *ttl* for the next packages until either the target host is reached or an upper limit of steps is achieved. In theory each participating server will get a package with a *ttl* equal to zero and respond with an *ICMP_TIME_EXCEEDED* message, which will contain its IP-address. In practice not every server has *ICMP* messages enabled and thus *Traceroute* does not necessarily detect every server in a connection. Additionally it is possible to circumvent Traceroute detection by not decrementing the ttl or branching off the traffic with physical means, such as copying the traffic passed through a cable. Both methods ensure that it is impossible for Traceroute to detect that additional parties are involved in a connection. This is different to not sending an ICMP_TIME_EXCEEDED message as there the missing step is recognized.

3.3 IXP

Internet Exchange Points (IXPs) are providers of physical infrastructure to interconnect a large amount of ASes with each other. ASes choose those connections, if available, in favor of direct connections with their peers, as an IXP can be cheaper and faster [18] [20]. Bigger IXPs even provide international data transfer capabilities and thus enable ASes to exchange traffic with a geographical distant AS faster and more reliable than typical connection chain as less parties are involved [27].

Traffic routed via an IXP traverses the routers of the IXP and is thus reflected in Traceroute. However, BGP sessions can be established between the peers connected via the IXP only. Such a connection is called bilateral [26]. This leaves the IXP out of the connection and leads to the BGP path not reflecting the participation of the IXP in traffic

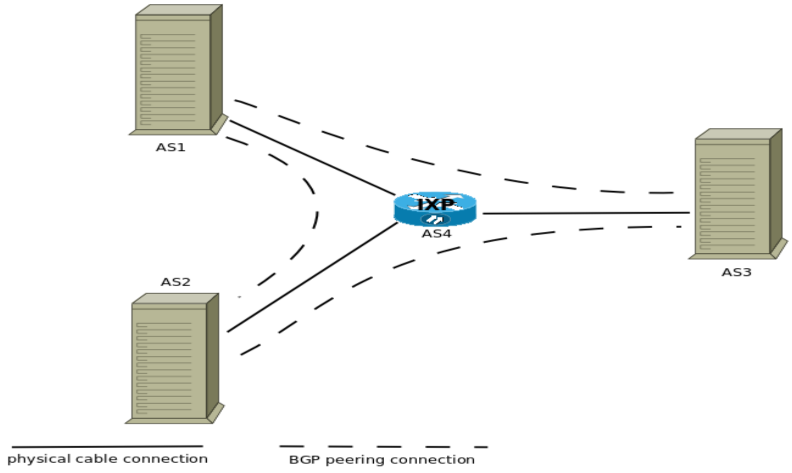

Figure 1: physical AS connection via IXP vs. AS BGP peering

routing (Fig. 1). Such a configuration should reflect in the Traceroute AS path as one additional ASN appearing in between sections of the AS path predicted by BGP.

To the best of our knowledge there is no public list of IXP ASNs, and therefore it is not possible to reliably determine the mismatches caused by IXPs. Mao et al. propose a heuristic detection algorithm to determine possible IXPs in a traceroute path [22]. Owed to the limited scope and width of our measurements, usage of the algorithm is not feasible. The algorithm requires multiple different vantage points to ensure low false positive detections. As we only use two different vantage points with low overall target amount, IXPs would not be reliably detectable. We still list IXPs as a potential reason for mismatches for overall completeness, but studying their influence is out of the scope of the thesis.

4 Expected Patterns of Mismatches

BGP contained paths should, in theory, show the same routing path detected by Traceroute. However, we argue that we have to expect differences due to the distinction between the *exchange of routing information* via BGP and the actual *physical connection and traffic routing* between two peering AS. Therefore, we expect to observe certain patterns in both Traceroute and BGP derived paths. In this chapter we point out three reasons why such differences have to occur. However, we still expect the majority of paths to match and that a high correlation between BGP and Traceroute exists. The intention of this chapter is to explain why a perfect match is not likely and thus Traceroute might be the better choice to detect machine specific routing behavior.

4.1 Route Fluttering

Load balancing is used to distribute the load of handling traffic and connections over different servers. This ensures that each server is able to handle its appointed load. An AS is a network like the Internet, only in a smaller scale. Therefore, choosing different routes potentially sends the traffic to different border routers (Fig. 2). Border routers connect the AS to its peers and different border routers commonly have different peers and therefore different BGP sessions. As a result, the traffic goes different routes if passed through different border routers, resulting in different Traceroute paths in different measurements at different points in time.

We expect this to show in a distinct pattern of differences between the BGP and Traceroute detected paths. We expect that the Traceroute path changes rapidly between two main routes whereas the BGP path remains constant for the same time period. We expect this pattern, as only one border router typically peers with the collector and load balancing potentially leads the traffic to different routers. This necessarily leads to a mismatch if the non peering router with a different routing table forwards the package at one measurement and the peering router at another measurement. Such *fluttering routes* were already theorized and detected as early as 1996 by Paxson [24].

For further analysis, we define a fluttering route as a route which switches multiple times between a matching AS path and a non-matching AS path. As a threshold we chose 3 for-and-back changes as we only want to detect the routes that are clearly fluttering. Routes which are frequently changing its Traceroute derived AS path between paths not matching BGP predicted paths will not be counted. We argue that fluttering cannot be the sole explanation for those mismatches and thus should not be used as the reason. It is likely, that fluttering partially explains those patterns but there have to be other reasons to account for as well. We also do not count paths which present some changes similiar to fluttering, but with a count lower than the threshold. Such occurrences are more likely be explained by a short timed routing problem in the connection leading to using a detour instead of the main route, thus no connection to load balancing.

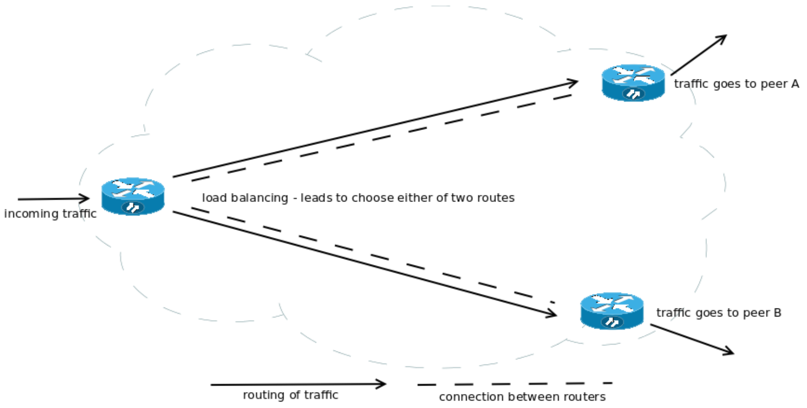
traffic goes to peer A
incoming traffic
load balancing - leads to choose either of two routes
traffic goes to peer B
routing of traffic
connection between routers

Figure 2: Visualization of different routing paths due to AS internal load balancing leading to different border routers

4.2 Short Lived Routing Problems

Routers are based on software that runs on hardware, both representing possible points of failures. Additionally, updates and upgrades have to be installed at some point. Either option leads to a time limited shutdown of the router and thus the router is not reachable. As border routers are routers they eventually experience such a shutdown and are thus unable to be reached by their peers. In such a case a peer will have to account for the change and change its routing behavior by choosing another next step, possibly residing in another AS. This will lead to the Traceroute suddenly detecting a new route. BGP only reflects the new route if the update message that contains the new path has already been propagated to the collector or the downtime did not occur in-between two measurements and the old route was already reinstated. We also speculate that a BGP router does not immediately advertises a change due to non-reachability to its peer but waits to determine if the change is long-term.

We argue that such occurrences should be reflected in measurements by Traceroute detecting a different path than BGP which is only short lived and does not occur regulary and thus should be attributed to route fluttering. For categorization we define any route exhibiting a change in Traceroute but not in BGP as *short lived routing problems* iff

- the mismatch lasts for less than 4 measurement and
- the frequency for the new route is not higher than the threshold for route fluttering

4.3 Transit AS

A transit AS is an intermediate AS between peers that transits the traffic in-between. Usage of transit ASes is is not restricted to different peers. A transit AS can be used if at some location in the AS it is faster to route the traffic through an other AS than using the internal infrastructure. A likely scenario for such an occurrence is a spread of an AS over a vast geographical area, e.g. a split of an AS between two different continents.

Such a split can either be bridged by using IXP infrastructure connecting the two parts of the AS or by using an intermediate AS which simply forwards the traffic from one location to another. If such a forwarding is achieved by configuration of the routing tables in the participating border routers and thus left out of BGP configuration, it will lead to an additional ASN in the Traceroute path but not in the BGP path. This is due to the fact that external configurations are not reflected in the outgoing BGP table.

Even if a *transit AS* is included into the BGP configuration, it is not necessarily reflected in the corresponding Traceroute path. As such an transit AS is used to overcome an area limited efficiency problem it is possible that the border router peering with a collector advertises such solution whereas the border router used by Traceroute does not. This situation would also result in a mismatch.

Both presented characteristics of transit ASes can lead to mismatches. A transit AS connecting the same AS can show in both BGP and Traceroute with a distinct pattern (e.g. {A,B,A}) whereas B is the transit AS). A transit AS connecting two different peers would either be reflected in BGP and thus be no different than any other regular hop or only be reflected in Traceroute if this connection is established by other means than BGP. Such a mismatch would look similar to an IXP but be restricted to very few if not only one distinct AS as the utility of such a setup would be restricted to quite specific circumstances.

We already explained in Chapter 3.3 the problems of detecting IXPs. As we cannot detect IXPs reliable we cannot distinguish between a transit AS, achieved by other means than BGP, and an IXP. Thus we will only focus on transit ASes used for connections in the same AS.

5 Experimental Setup

In this chapter we describe the experimental setup. We start by explaining the tool chain and conclude the chapter by explaining the data points used to gather the data.

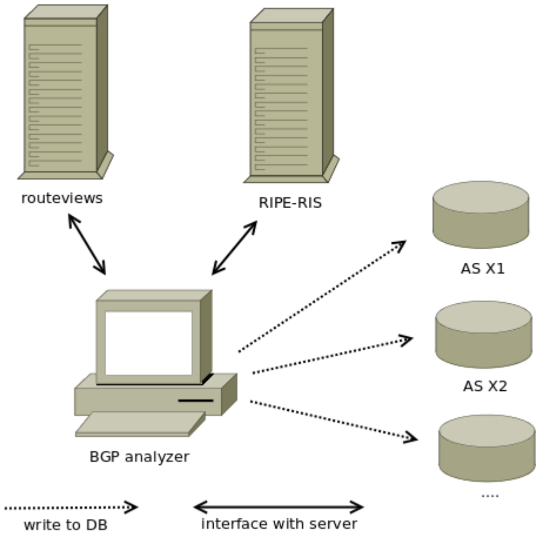

Figure 3: BGP analysis first step

5.1 Tool Chain

The tool chain consists of three tools. The *Mrtreader* and *Tracestat* are used to gather BGP and Traceroute information respectively. The *Analyzer* is used to analyze the gathered data and calculate the relevant data points we are are interested in.

We will address the tools on an abstract level to convey the overall theoretical notion of the underlying implementation. We implemented the described tools using Common-LISP and uploaded the programs to github (https://github.com/simkoc/toolchain).

5.1.1 Mrtreader

The Mrtreader is used to download and parse BGP/MRT data as well as to analyze the retrieved information. Information about the relevant parts of BGP and MRT can be read in Chapter 3.

Download and parsing. To establish a starting point to which future *BGP_UPDATE* messages are applied, a dump containing a *TABLE_DUMP_V2* that is provided by a col-

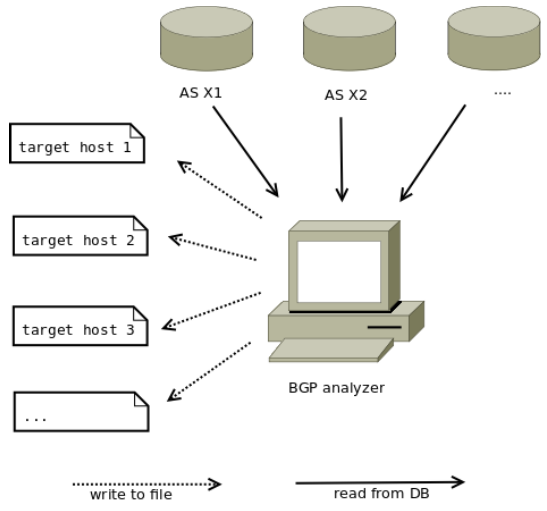

Figure 4: BGP analysis second step

lector peering with an AS of interest, is downloaded and parsed. The creation date of the dump is the start date of the analyzes. The Mrtreader extracts the information about the contained prefixes, their receiving time, and the corresponding *AS_PATHs* and inserts them into a database. Each peer contained in the *PEER_INDEX_TABLE* is assinged its own database. After this step each peer of the collector, who created the dump, has its own database containing all known routes at the time of the dump creation.

Following the processing of the *TABLE_DUMP_V2* all *BGP4MP* containing files provided by the collector which were created within the timespan between the *TABLE_DUMP_V2* and the end of the analyzing period are downloaded. The Mrtreader processes the files consecutively and inserts the information into the respective databases (Fig. 3). Every newly advertised route of a peer is entered into the respective database with its advertisement time. Any withdrawn route is updated with the time of the withdrawal.

Extraction of relevant data. The second functionality of the Mrtreader is to create a list containing the used *AS_PATHs* to a certain target at a certain time from a certain peer. To gather that information we create a routing table. The Mrtreader extracts all prefixes that are advertised but not withdrawn up until the relevant point in time from the database of the respective peer and inserts them into a search tree supporting longest prefix matching. Subsequently the Mrtreader queries the routing table for the relevant IPs and retrieves the corresponding *AS_PATH*. The result is then append to a file respective to the queried IP and peer ASN. Each line in the file contains the relevant point in time, the target IP, a hash of the *AS_PATH* and the *AS_PATH* (Fig. 4).

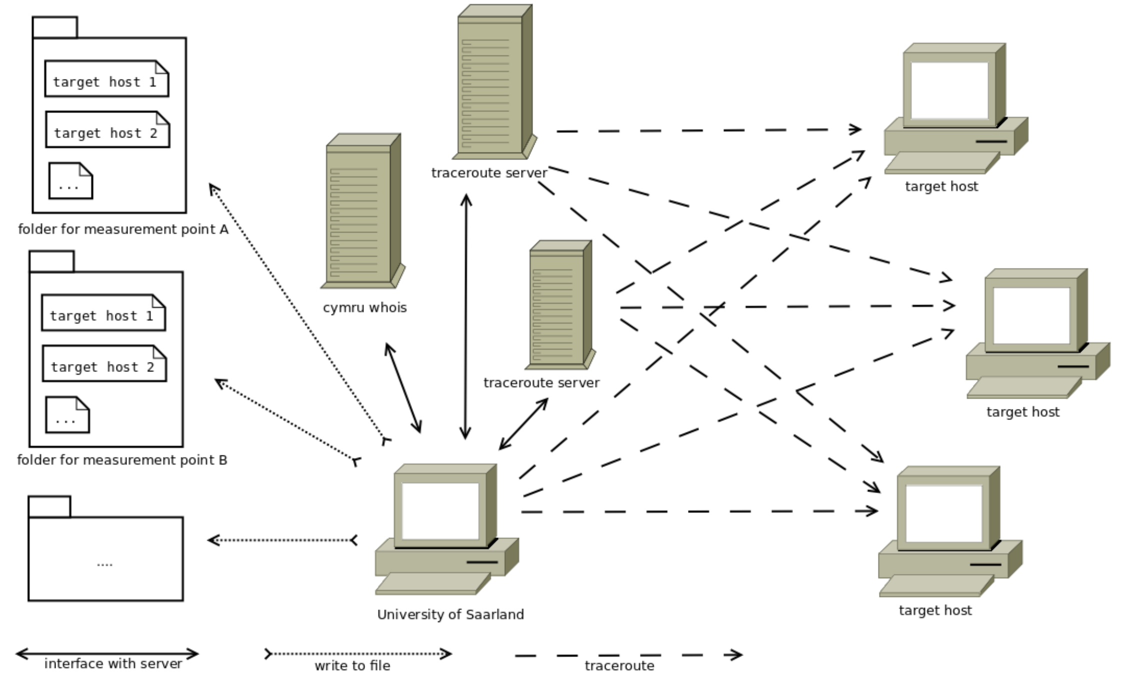

Figure 5: Traceroute

As querying the database and creating the routing table takes a considerable amount of cpu cycles, each datapoint in the file is 60 minutes apart.

5.1.2 Tracestat

Tracestat is used to gather Traceroute paths and translates them into AS paths. Information on Traceroute can be found in Chapter 3. Tracestat is split into two parts. The first part calls Traceroutes placed at different locations. The second part translates and dumps the retrieved information.

First part. The first part calls provided Traceroutes by using predefined interfaces included into Tracestat. It is possible to integrate multiple different Traceroutes into Tracestat and thus enables Tracestat to collect Traceroute paths from multiple different locations. Each Tracestat interface can be called with a single IP and returns a list of IPs representing the path detected by Traceroute. It is possible, that Traceroute returns multiple IP-addresses for the same step, as more than one probe package for each step is sent. In such case Tracestat uses the first IP-addresses of each step.

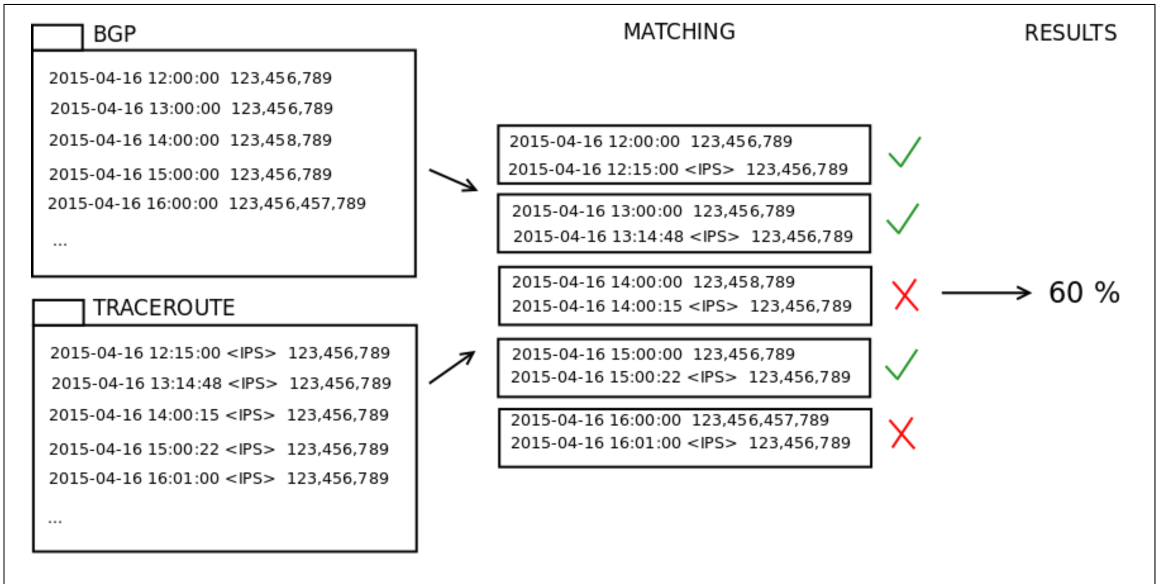
BGP
2015-04-16 12:00:00 123,456,789
2015-04-16 13:00:00 123,456,789
2015-04-16 14:00:00 123,458,789
2015-04-16 15:00:00 123,456,789
2015-04-16 16:00:00 123,456,457,789
...
TRACEROUTE
2015-04-16 12:15:00 <IPS> 123,456,789
2015-04-16 13:14:48 <IPS> 123,456,789
2015-04-16 14:00:15 <IPS> 123,456,789
2015-04-16 15:00:22 <IPS> 123,456,789
2015-04-16 16:01:00 <IPS> 123,456,789
...
MATCHING
2015-04-16 12:00:00 123,456,789
2015-04-16 12:15:00 <IPS> 123,456,789
✓
2015-04-16 13:00:00 123,456,789
2015-04-16 13:14:48 <IPS> 123,456,789
✓
2015-04-16 14:00:00 123,458,789
2015-04-16 14:00:15 <IPS> 123,456,789
✗
2015-04-16 15:00:00 123,456,789
2015-04-16 15:00:22 <IPS> 123,456,789
✓
2015-04-16 16:00:00 123,456,457,789
2015-04-16 16:01:00 <IPS> 123,456,789
✗
RESULTS
60 %

Figure 6: Correlating BGP and Traceroute for a single IP-address

Second part. WHOIS is designed to request information about an IP including the corresponding prefix as well as the owner ASN [5]. Tracestat maps an IP onto an ASN using a server speaking the *WHOIS* protocol. We chose the WHOIS server hosted by Team Cymru as it provides the information as comma seperated values and is thus easy to parse [2]. Tracestat further reduces the resulting ASN paths. It deletes consecutive duplicates of ASNs as well as WHOIS results of IPs the *WHOIS* service was not able to identify, referenced as *NIL* (e.g. *(123,234,NIL,234,345)* becomes *(123,234,345)*). This process leaves paths which are easy to compare as they contain the minimum of relevant data needed. Finally Tracestat dumps the resulting ASN path with its context information into an target IP related file stored in a Traceroute location related folder. Each line of a file contains the timestamp of the corresponding measurement, the target IP, the IP path,the translated ASN path, and the hashes of the paths.

Figure 5 gives a graphical overview of the setup for gathering Traceroute information from multiple different Traceroute locations.

5.1.3 Analyzer

The Analyzer uses the files generated by the Mrtreader and Tracestat to calculate the different intended values. The Analyzer starts by parsing in the previously generated files and makes them accessible for the subsequently described processes. The Analyzer drops any Traceroute measurement which final step did not reach the intended target AS. The

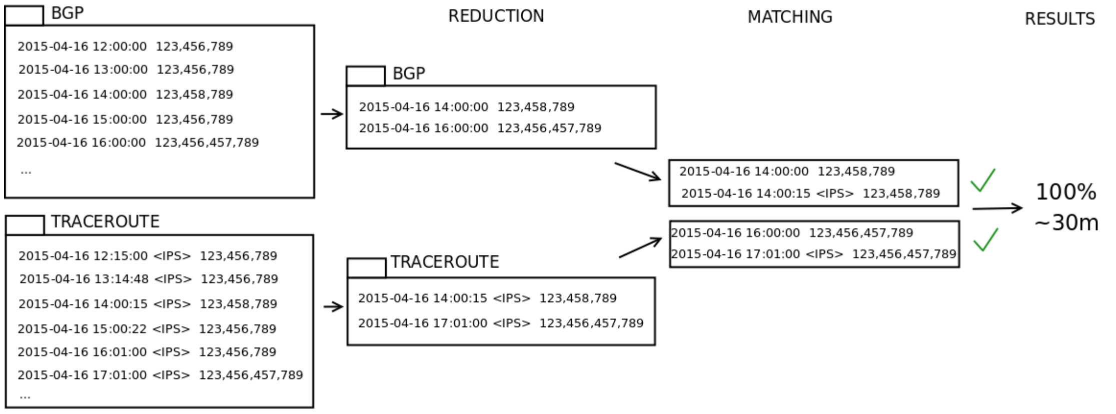
BGP
2015-04-16 12:00:00 123,456,789
2015-04-16 13:00:00 123,456,789
2015-04-16 14:00:00 123,458,789
2015-04-16 15:00:00 123,456,789
2015-04-16 16:00:00 123,456,457,789
...
TRACEROUTE
2015-04-16 12:15:00 <IPS> 123,456,789
2015-04-16 13:14:48 <IPS> 123,456,789
2015-04-16 14:00:15 <IPS> 123,458,789
2015-04-16 15:00:22 <IPS> 123,456,789
2015-04-16 16:01:00 <IPS> 123,456,789
2015-04-16 17:01:00 <IPS> 123,456,457,789
...
REDUCTION
BGP
2015-04-16 14:00:00 123,458,789
2015-04-16 16:00:00 123,456,457,789
TRACEROUTE
2015-04-16 14:00:15 <IPS> 123,458,789
2015-04-16 17:01:00 <IPS> 123,456,457,789
MATCHING
2015-04-16 14:00:00 123,458,789
2015-04-16 14:00:15 <IPS> 123,458,789
2015-04-16 16:00:00 123,456,457,789
2015-04-16 17:01:00 <IPS> 123,456,457,789
✓
✓
RESULTS
100%
~30m

Figure 7: Correlating BGP and Traceroute

Analyzer resolves the expected ASN of the target IP by relying on prefix to ASN mappings provided by the CAIDA [9]. The mapping is provided as a separate file and we suggest using the mapping located in the middle of the timeframe of a measurement.

The main focus of the Analyzer lies on the correlation between the AS level paths obtained by Traceroute and BGP analysis. The Analyzer also calculates the proportion of changes in either Traceroute or BGP paths subsequently also reflected in the sibling data set. As a last bit of information the Analyzer calculates the average time a route stays stable and the influence of length on the possibility of sudden change. Each aspect is addressed separately.

(Strong/Weak) correlation. The Analyzer obtains the correlation between Traceroute and BGP by pairing each measurement of the two respected data sets of the same IP with its closest sibling measurement. Subsequently the Analyzer calculates the proportion of matched and unmatched ASN paths. This process is applied to all obtained data sets and averaged over all results (Fig. 6). For the weak correlation the Analyzer relaxes the exact match restriction by counting the amount of ASNs appearing in both paths in the same order. The Analyzer skips non matching ASNs by performing a look ahead in the path of the sibling to decide whether a given ASN will later on appear in the sibling path or is an additional step not reflected by the measurement.

Change Reflection. To obtain information on how well and fast a change of route in one dataset is reflected in the sibling dataset the Analyzer reduces the data sets to changes only. It only keeps the first data point after a change in an ASN path for further analysis. Consecutively, the Analyzer pairs the measurement points of the sets for the same IP with

$$\bar{x} = \frac{1}{n_x} \sum_{i=1}^{n_x} x_i$$

x_i and y_i denote the respective value in the i-th tuple

$$\bar{y} = \frac{1}{n_y} \sum_{i=1}^{n_y} y_i$$

$$corr(x, y) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 * \sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

Figure 8: formular to calculate correlation coefficient

some restrictions. The Analyzer only pairs two points iff

- the measurement point is chronological after the one to pair with
- the AS paths are the same

Following this the quota of paired and non-paired data points is calculated as well as the average time difference between the matched pairs (Fig. 7). Again the Analyzer subjects all data sets to this process and averages the results over all data sets.

Average Time To Change. The Analyzer calculates the average time it takes before a route changes by reducing the datasets to only reflect the changes that occurred during the measurements. This process is similar to the first step in calculating the change reflection. Subsequently the Analyzer calculates and averages the time difference between two changes. This process is applied to all data sets of the same measurement type and averaged over all results.

Influence of length on change. The Analyzer measures the influence of length on the probability of change by calculating the correlation coefficient between the length of the AS path and the amount of changes in the measurement time frame. The Analyzers traverses all datasets and produces a tuple (x, y) whereas x denotes the average length of the AS path of the dataset and y denotes how many changes happened during the measurement time frame. The Analyzer uses the resulting list of tuples to calculate the correlation coefficient (Fig. 8).

5.2 Data Gathering

In this chapter we explain what measurement points and BGP collectors where used for the *Mrtreader* and *Tracestat*. An in-depth quality analysis of the used data was out of the scope of this thesis.

5.2.1 MRTreader

We found two projects which collect and provide BGP data: the *routeviews project* and the *Réseaux IP Européens Routing Information Service (RIPE-RIS)*. The provided BGP data is contained in MRT data dumps (either *TABLE_DUMP_V2* or *BGP4MP*) that are dumped in different time intervals.

Routeviews project. The *routeviews project* provides access to 13 collectors that have AS as peers who are providing regular BGP updates. It is hosted and maintained by the University of Oregon. The BGP data is available from as early as October 2001 and can be downloaded in 15 minute update dump blocks or 2 hour separated full dumps [11].

RIPE-RIS. *RIPE-RIS* is the second organization and has 13 running collectors spread across the world with substantially different peers. RIPE-RIS provides the data in 5 minute update chunks and 8 hour separated full dumps. Data is available from as early as 2001 [10].

5.2.2 Tracestat

As Tracestat is able to interface with Traceroutes at different locations we had to select ASes which contained an accessible server running Traceroute as well as provided public BGP information. As the University of Saarland is part of the *Deutsches Forschungs Netzwerk* (DFN) that is a peer with a BGP collector, we chose to station one probe on a university server.

For the second probe we searched through the server list of Looking Glass. Looking Glass is a collaboration of different servers providing an online (HTTP) interface to run Traceroute probes from their location and retrieve the information [3]. We chose AS12306 as the other location based on the usability of the interface as well as on the amount of published prefixes. AS12306 belongs to plusline which is a professional hosting service. The Traceroute interface is hosted by the renown German IT-newspaper publisher Heise (http://www.heise.de/netze/tools/Traceroute/). Table 1 lists all available Traceroute probe locations that are in an AS that peers with a collector of either the Routeviews project or RIPE-RIS.

For AS680 we obtained the relevant IPs by downloading the Tor consensus from 13-04-2015. Consecutively we diversified the contained IP-addresses by removing all but one IP having the same first two bytes. With this line of action we ensured that an IP-address will be reachable for Traceroute without having to search the prefix ranges of any given ASN and also secured that multiple different ASes will be targeted. We selected an overall of 200 different IPs for probing.

For AS12306 the possible IPs to probe were more limited due to the limited amount of published prefixes. We extracted all prefixes from the BGP records and probed the first five IPs of each prefix for available servers. We gathered 14 usable IPs for measurements starting

ASN	peer	usable interface	published prefixes
AS680	RIPE-RIS	yes	533926
AS22548	RIPE-RIS	no	540768
AS12306	RIPE-RIS	yes	70
AS33843	RIPE-RIS	no	35
AS2914	RIPE-RIS	no	530662
AS10848	RIPE-RIS	yes	7
AS8816	RIPE-RIS	yes	10

Table 1: table showing ASN,peering project,usability of the interface, and the amount of prefixes published

24-05-2015 and an overall of 24 different IPs for measurements starting at 02-05-2015. We were not able to utilize all 70 published prefixes as some were duplicates and other did not contain a responsive server in the first 5 IPs.

6 Results

In this chapter we describe the measurement results of both measurement periods and give a description of the obtained results. The first measurement period spanned from the 25th to the 30th of April. The second measurement period spanned from the 2nd of Mai to the 16th of Mai. A short overview of the results can be obtained at Table 2. We conclude this segment by comparing the two measurement points and explaining why the observed differences occurred.

	AS680 (25th till 30th)	**AS12360** (25th till 30th)
amount used probe targets	175	14
strong correlation	85.65%	92.62%
BGP Data		
BGP ⟹ Traceroute	42.86%	(no BGP changes)
avg. time to change	5.4h	(no BGP changes)
Traceroute Data		
weak correlation	95.72%	96.41%
Traceroute ⟹ BGP	22.32%	25%
avg. time to change	6.08h	3h
corr. coeff. (IP)	0.43	0.60
corr. coeff. (ASN)	0.17	0.81
	AS680 (2nd till 16th)	**AS12360** (02nd till 16th)
amount used probe targets	177	22
strong correlation	85.77%	90%
BGP Data		
BGP ⟹ Traceroute	71.43%	0%
avg. time to change	14.84	65h
Traceroute Data		
weak correlation	96.12%	93%
Traceroute ⟹ BGP	20.35%	0%
avg. time to change	8.8h	-
corr. coeff. (IP)	0.46	0.31
corr. coeff. (ASN)	0.28	0.29

Table 2: Raw results of the data analysis for the two AS. Amount used probes relates to the amount of IP-addresses finally used for evaluation. ⟹ abbreviates the influence of change in the left data type onto the right one. (IP) and (ASN) stand for IP-path and ASN-path respectively

6.1 AS680

First measurement period. We chose a total of 200 different target IPs of which 175 were usable in the end. 25 proved to be completely unusable as the Traceroute probes never reached the destination and therefore rendered a proper analysis impossible.

In 13 targets we encountered an infrequent mismatch between BGP and Traceroute path. In those cases BGP and Traceroute differentiated at most in 2 measurements and disagreed on one step in the path. We counted 6 targets where BGP consistently added an additional AS and 10 targets in which Traceroute added consistently at least 1 ASN in contrast to BGP. In 5 targets BGP added an additional ASN in at most 5 measurements. A remainder of 4 targets showed additional deviation which did not fit the previously described patterns. A total of 44 targets showed some deviation during the measurement period. The gathered data showed both a high strong and high weak correlation. The strong correlation is at 86% whereas the weak correlation is 96%.

43% of all BGP changes were at some later point reflected in the Traceroute data. 22% of Traceroute changes were reflected in the BGP data at a later measurement. The time between the reflection of a BGP change in the Traceroute data was an average of 0.5h. Traceroute changes were reflected in BGP after an average of 1.12h.

The correlation coefficient of the influence of length on the change quota of the Traceroute data was 0.43 for IP based paths and 0.17 for ASN based paths. If a target experienced changes during our measurements it occurred on average after 5.4h for BGP derived paths and after 6h for Traceroute detected paths.

Second measurement period. We chose the same 200 different target IPs of which 177 were usable in the end. 23 proved to be completely unusable as the Traceroute probes never reached the destination and therefore rendered a proper analysis impossible.

In 15 targets we encountered an infrequent mismatch between BGP and Traceroute path. In those cases BGP and Traceroute differentiated at most in 5 measurements and disagreed on one step in the path. We counted 6 targets where BGP consistently added an additional AS and 11 targets in which Traceroute added consistently, in contrast to BGP, at least 1 ASN and in one case even 2. In 7 targets BGP added an additional ASN in at most 5 measurements. A remainder of 5 targets showed additional deviation which did not fit the previously described patterns. A total of 48 targets showed some deviation during the measurement period. The gathered data showed both a high strong and high weak correlation. The strong correlation is at 86% whereas the weak correlation is 96%.

71% of all BGP changes were at some later point reflected in the Traceroute data. 20% of Traceroute changes were reflected in the BGP data at a later measurement. The time between the reflection of a BGP change in the Traceroute data was an average of 0.5h. Traceroute changes were reflected in BGP after an average of 0.5h.

The correlation coefficient of the influence of length on the change quota of the Traceroute data was 0.46 for IP based paths and 0.28 for ASN based paths. If a target experienced changes during our measurements it occurred on average after 14.44h for BGP derived paths and after 8.89h for Traceroute detected paths.

6.2 AS12306

First measurement period. We chose 14 different target IPs based on the published prefixes of the AS. All results were usable at the end of the measurement period.

We counted only 1 target which did not match Traceroute and BGP paths. During the whole measurement period there it mismatched. Traceroute added one step when compared to the BGP path. Every other target had exactly the same matching BGP and Traceroute predicted paths during the whole measurement period. This resulted in both a high strong and high weak correlation. Strong correlation was at 93% and weak correlation was at 96%.

We observed no BGP changes. The correlation coefficient of the influence of path length on the change quota of the Traceroute data was 0.60 for IP based paths and 0.81 for ASN based paths. If a target experienced changes during our measurements it occurred on average within 2 hours.

Second measurement period. We chose 24 different target IPs based on the published prefixes of the AS. 22 were usable at the end of the measurement period.

We counted only 3 target which did not match Traceroute and BGP paths. During the whole measurement period they mismatched. Traceroute added one step when compared to the BGP path. Two of those three targets had an additional step discovered by Traceroute in at least 50 measurement points. Every other target had exactly the same matching BGP and Traceroute predicted paths during the whole measurement period. This resulted in both a high strong and high weak correlation. Strong correlation was at 90% and weak correlation was at 93%.

The correlation coefficient of the influence of path length on the change quota of the Traceroute data was 0.31 for IP based paths and 0.29 for ASN based paths. If a target experienced changes during our measurements it occurred on average within 65 hours for BGP.

6.3 Comparing Both Measurement Points

In this section we go into details about the differences observed between AS680, AS12306, and between the measurement periods. We will explain why we think those differences occurred and conclude by categorizing AS12306 as a sanity check for the AS680 values for the weak and strong correlation.

Strong and weak correlation. Both datasets yielded quite equal results for the strong and weak correlation correlation. As explained in Chapter 4 we expected that the weak correlation will be significantly stronger than the strong correlation. This was not the case for AS12306. AS12306 showed a strong correlation of 93% and a weak correlation of 96%. We argue that this is due to the small amount of different target used. As only a fraction of them differed the small and strong correlation are expected to be quite equal.

Time to change. The time of change differed significantly between AS680 and AS12306 and between the two measurement periods. Whereas AS680 showed an average time to change of 6h for Traceroute and 5.4h for BGP in the first measurement period and 14.84h for Traceroute and 8.8h for BGP in the second measurement period. AS1230 did not encounter any BGP changes in the first measurement period and the Traceroute changes were 3h apart on average for the first measurement period. For the second measurement period AS12306 encountered one target with two changes which were 65h apart and no target with more than one change in the Traceroute data, thus we were not able to calculate the average time between changes.

AS12306 used a significantly smaller amount of probes than AS680. This lead to targets having more than two changes in either BGP or Traceroute to be a rare event. We therefore argue that the numbers given for AS12306 should not be considered reliable and the AS680 measurements should rather be used.

Reflection of change. The reflection of changes in the other dataset differed greatly between the two datasets as well as between the measurement periods. In the first measurement period 22% of AS680 and 25% of AS12306 targets reflected Traceroute changes in BGP measurements at some later point. 43% of all BGP changes were reflected in Traceroute measurements in AS680. AS12306 experienced no changes in BGP and thus the reflection could not be calculated. In the second measurement period no target in AS12306 reflected changes for either BGP nor Traceroute. AS680 showed higher values for the reflection of BGP changes reflected in Traceroute and a nearly equal value for Traceroute changes reflected in BGP.

We again argue that the differences between the two ASes can be explained by the significant difference in targets. As AS12306 did not show any changes in BGP measurements in the first measurment period and only one in the second we have to assume that neither value of reflection is close to the truth. However, due to the higher amount of targets for AS680 we consider those results to reflect actual behavior. We further argue that the results of the second period are more precise as the timeframe was significantly longer and thus changes could more likely be accounted for.

Correlation of length with change. We observed no strong correlation between the length of a path and the amount of changes it exhibited during our measurement for neither IP nor ASN paths. We argue that this is due to the fact that IP paths are likely to change and thus the length does not have such a high impact. However, all AS paths were short and thus the length did not significantly impact change probabilities. We therefore argue that there is no significant correlation between the length of a path and its changes.

Reflect that we admit at multiple points that AS12306 lacked a significant amount of targets and thus is unlikely to provide exact or reliable results. However, we do consider the results of AS12306 concerning the strong and weak correlation to a good sanity check for the results of AS680. Even though AS12306 lacked a high amount of targets, strong and weak correlation take all measurement points for each target into account and do not rely on rare route changes, thus there are enough measurement points to consider AS12306 to be an indication that the AS680 results are not due to some observation bias or routing exception only exhibited by AS680.

7 Interpretation

In this chapter we categorize the mismatches according to the different patterns introduced in Chapter 4 and the observed patterns. Additionally, we explain the limitations of our approach. We conclude the chapter by arguing that we consider Traceroute to be a valid tool for path detection.

7.1 Observed Patterns

We explain in Chapter 4 that we expect a mismatch between routes and give multiple reasons why such mismatches should occur. In this segment we categorize the observed patterns and mismatches by the previously defined patterns. We conclude by explaining why we do not account for all patterns.

7.1.1 Route Fluttering

During the first measurement period we observed 4 properly fluttering routes for targets of AS680. 2 targets experienced patterns which showed clear evidence of fluttering though none of the main Traceroute paths matched the BGP path. AS12306 contained no fluttering routes.

During the second measurement period we observed 13 properly fluttering routes for targets of AS680. 2 targets experienced patterns which indicate fluttering but none of the main Traceroute paths matched the BGP path. Again AS12306 contained no fluttering routes.

7.1.2 Short Lived Routing Problems

We observed 11 routes to targets of AS680 experiencing patterns we attribute to *Short lived routing problems* during the first measurement period. AS12306 had 0 targets showing such behavior.

We observed 14 routes to targets of AS680 experiencing patterns we attribute to *Short lived routing problems* during the second measurement period. AS12306 had 0 targets showing such behavior.

7.1.3 Transit AS

We observed patterns indicating transit ASes for 1 target in BGP and 11 targets in Traceroute for AS680 during the first measurement period. In 8 of cases this observation coincided with a complete mismatch during the whole measurement period. In 3 cases this pattern only resulted in mismatches when Traceroute switched to that route.

We observed patterns indicating transit ASes for 1 target in BGP and 19 targets in Traceroute for AS680 during the second measurement period. The BGP transit AS mis-

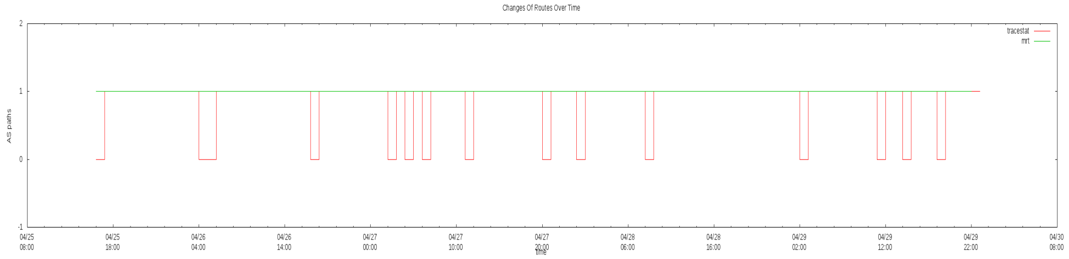

Figure 9: A graph showing route fluttering over the course of the measurements.

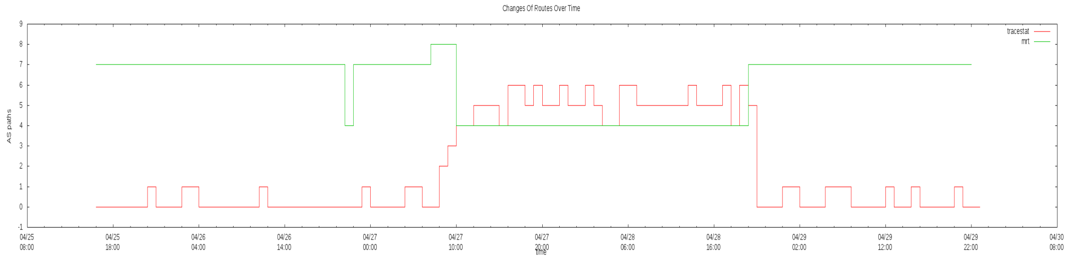

Figure 10: A graph showing route fluttering. But obviously this does not account for all mismatches.

matched coincided with a mismatch for the whole measurement timeframe. In 7 of cases the transit observation in Traceroute coincided with a complete mismatch during the whole measurement period. In 12 cases this pattern only resulted in mismatches when Traceroute switched to that route. AS12306 did not contain any targets containing transits.

7.1.4 Remaining Uncategorized Patterns

In the previous sections we categorized the observed patterns and mismatches into the scenarios described in Chapter 4. In the first measurement period we accounted for 24 of all 44 mismatches in AS680. For the second measurement period we account for 29 of 46 mismatches in AS680. AS12306 encountered fewer patterns and we accounted for 0 of 1 mismatches in the first and 0 of 3 mismatches in the second measurement period.

We were not able to account for all mismatches. However, we argue that this number can be further reduced as soon as a distinction between transit ASes and IXPs can be made. Some of the mismatches showed a constant single step mismatch and thus can either be attribute to IXPs or transit ASes in-between different peers. We further encountered multiple targets showing patterns that do not fit a distinct category and thus can not be properly categorized. We argue that this is due to either a combination of multiple scenarios or a yet unknown new distinct routing phenomenon. We suspend categorizing the remaining mismatches for future work that includes a greater measurement basis.

We conclude that all our predicted patterns occur. However, we are not able to account for all patterns and thus lack a complete framework of patterns to fully explain current routing behavior.

7.2 Limitations of Our Approach

Our evaluation has clear limitations and this thesis should therefore be considered a step towards a consensus rather than a final conclusion on the topic.

The most significant issue is the limited vantage points. This is due to a low availability of public traceroute servers hosted in ASes which also provide their BGP data. It would be possible to improve on this issue by collaborating with different providers. As route prediction is also of great interest to an AS owner a collaboration in which a traceroute probe is hosted in an AS as well as the AS owner peering with a collector provided by the researchers seems realistic. We assume that the main reasons why AS owner do not already peer with public collectors are that such projects are not well known outside the research community, to set up a peering is not worth the time for a AS owner, the owner considers its routing data proprietary and does not want to share with the open public. Stipulating a direct collaboration would lessen or negate such reasons.

The next significant problem is the small amount targets used for different traceroute measurements. Even though the targets were diversified to ensure that different ASes are targeted and therefore redundancies in paths and their influence on the results is minimized only an insignificant amount of the current Internet network of ASes was targeted. We used a public list of tor nodes to ensure availability of the targeted IPs, thus to ensure that a traceroute reaches the targeted AS and the targeted server. Even though this approach is intuitive and simple to implement an organized mapping of available servers throughout the Internet seems feasible in a reasonable amount of time and reasonable resources, at least for IPv4. We considered such approach to be out of scope, and focused on a general tool chain for easy comparison of traceroute and BGP data.

A third issue is the data we used. We did not perform a quality assessment of the data and thus cannot ensure that the data we gathered does not contain a bias which carries over into our results. We consider the risk of a bias to be present as the BGP data collected was gathered from public projects who only collect data from volunteers. Any AS that is not honest with its BGP messages would not peer with such a project as it would significantly increase the risk of being discovered. Our most important placement resided in the *Deutsche Forschungs Netzwerk (DFN)*. The DFN cannot not be considered representative for commercial BGP usage as it is a non profit organization. The second significantly smaller placement was selected from a public list of servers offering Traceroute interfaces. The chosen interface was hosted by a renown German IT-newsite (www.heise.de/netze/tools/traceroute/). As we expect the hoster of the Traceroute to only have a client customer relationship with the provider of the BGP data we consider the risk of a bias to be smaller.

We consider the timeframe of the measurements to be a remaining smaller issue. Our measurements spanned less than a month and one could argue that our time frame is too limited to ensure that the observed patterns are representative. A measurement period of multiple month or years would yield more scientifically sound results. Due to the time required we also considered this to be out of scope for this thesis.

7.3 BGP or Traceroute

We reason by our results that Traceroute should be considered a double-edged sword. It is apparent hat traceroute can show path steps which are either invisible to BGP or simply not in the BGP configuration. We consider this to be path depended as we also have results in which BGP showed additional ASes which were not detected using traceroute. We attribute this to the nature of interconnected machines: Different starts can lead to different paths even if the start resides int the same AS. In our case this means that if a border router was placed differently than the one the Traceroute detection passed a constant mismatch is the result.

We conclude that Traceroute is a valid tool for path prediction for a specific machine. We question the usage of Traceroute as a general means of network topology discovery if measurement positions are abstracted away. This is not to be considered different for BGP. However, we consider Traceroute to be a good tool to establish an overview of all likely paths used if measurements are performed multiple times. We advise against using single measurements as this would lead to missing routes in case of route fluttering which is a common pattern in modern routing behavior. Additionally we want to stress that our conclusion only holds for detection of connection participants not actively evading Traceroute detection as explained in chapter 3.2.

8 Conclusion

In this thesis we presented a tool chain for comparing BGP and Traceroute predicted AS paths. We used the tool chain in a limited study and presented the results.

We showed that the majority of paths in the Internet appear to be stable and match precisely the paths predicted by BGP. We also presented a small but significant amount of measurements where the matching assumption does not hold. However, we are able to explain roughly half of the mismatches. Additionally, most of the mismatches imply that BGP is too static for path prediction, as route fluttering and short time routing problems are behaviors BGP is unable to account for. Furthermore, we argue that the usage of transit ASes is location specific, thus a mismatch due to this reason is specific to the start location of a connection.

We conclude that there is good evidence that Traceroute is a valid tool for path prediction for a specific machine if multiple measurements are performed and the results are only applied for a limited time period. It is necessary to repeat path measurements regularly to account for changes in the routing behavior. We restrict this conclusion to the detection of connection participants to participants not actively avoiding detection, as it is possible to evade Traceroute detection without Traceroute recognizing the missing hops in a connection.

8.1 Future Work

We have convincing results concerning BGP and Traceroute path correlation, but we are aware that a strong correlation between Traceroute and BGP does not necessarily mean that either bears any resemblance towards the real routing of traffic, as they are either trust based (BGP) or exploit a feature of a protocol that was never meant for that pupose. They only present a best effort approach based on current knowledge. We consider any deepening research into routing and Traceroute interesting.

A long time study combined with a collaboration with different ASes could overcome the limitations we discussed in Chapter 7 and would be an important addition to the research already performed. It could further validate the results already established or present a complete novel view on the routing in the Internet. More over, it would reliable quantify the occurrences of the phenomenons theorized in Chapter 4.

As another opportunity for future research, we consider to validate whether Traceroute reflects the paths taken by actual user traffic, e.g. FTP or HTTP. Additionally, it is interesting to validate whether Traceroute-detected paths are a good match for all types of traffic or whether load heavy or constant connections are rerouted outside the scope of the Traceroute detection mechanism. There has already been work on traffic types and their treatment [4] [31] [30]. However, to the best of our knowledge there is no work comparing Traceroute path results with actual paths of different traffic types.

Finally, we would also like to explore whether route fluttering can be linked to the size of a AS. We expect that the size of an AS has a great influence of the amount of border routers and load balancing and thus leads to more a more likely observation of route fluttering. Another assumption we would like to explore and that is linked to the size of an AS, is the amount of different routes leading to the same target. We suspect that this number is also dependent on the size of the AS. Big ASes are more likely to be spaced over a vast geographical area and thus will more likely exhibit transit ASes and different border routers depending on where the measurement was started, either inside or outside the AS.

References

[1] Internet Census 2012. http://internetcensus2012.bitbucket.org/paper.html.

[2] IP to ASN mapping,Team Cymru. http://www.team-cymru.org/IP-ASN-mapping.html.

[3] Looking Glass,www.traceroute.org. http://www.traceroute.org/.

[4] *On Traffic Types and Service Classes in the Internet*, Globecom, San Francisco, CA, USA.

[5] RFC3912,WHOIS. https://tools.ietf.org/html/rfc3912.

[6] RFC4271,BGP.

[7] RFC791,IP. https://www.ietf.org/rfc/rfc791.

[8] RFC792,ICMP. https://tools.ietf.org/html/rfc792.

[9] Routeviews Prefix to AS mappings Dataset (pfx2as) for IPv4 and IPv6.

[10] Réseaux IP Européens Routing Information Service (RIPE-RIS). https://www.ripe.net/data-tools/stats/ris/ris-raw-data.

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

[12] YouTube Hijacking: A RIPE NCC RIS case study. https://www.ripe.net/publications/news/industry-developments/youtube-hijacking-a-ripe-ncc-ris-case-study.

[13] Duncan Campbell. Someboy's Listening. *New Statesman*, August 1998.

[14] S. Chakravarty, M. V. Barbera, G. Portokalidis, M. Polychronakis, and A. D. Keromytis. On the Effectiveness of Traffic Analysis Against Anonymity Networks Using Flow Records. In *Proceedings of the 15th Passive and Active Measurements Conference (PAM '14)*, March 2014.

[15] Xenofontas Dimitropoulos, Dmitri Krioukov, Marina Fomenkov, Bradley Huffaker, Young Hyun, kc claffy, and George Riley. AS Relationships: Inference and Validation. *SIGCOMM Comput. Commun. Rev.*, 37(1):29–40, January 2007.

[16] A. Faggiani, E. Gregori, A. Improta, L. Lenzini, V. Luconi, and L. Sani. A study on traceroute potentiality in revealing the Internet AS-level topology. In *Networking Conference, 2014 IFIP*, pages 1–9, June 2014.

[17] Lixin Gao. On Inferring Autonomous System Relationships in the Internet. *IEEE/ACM Trans. Netw.*, 9(6):733–745, December 2001.

[18] V. Giotsas, S. Zhou, M. Luckie, and k. claffy. Inferring Multilateral Peering. In *ACM SIGCOMM Conference on emerging Networking EXperiments and Technologies (CoNEXT)*, pages 247–258, Aug 2013.

[19] Lauro Poitras Glenn Greenwald, Ewen MacAskill. Edward Snowden: the whistleblower behind the NSA surveillance revelations.

[20] Mike Jansen. Promoting the use of Internet Exchange Points: A Guide to Policy, Management, and Technical Issues. March 2012.

[21] C. Labovitz L. Blunk, M. Karir. RFC6396,Multi-Threaded Routing Toolkit (MRT) Routing Information Export Format. http://tools.ietf.org/html/rfc6396, 2011.

[22] Zhuoqing Morley Mao, Jennifer Rexford, Jia Wang, and Randy H. Katz. Towards an Accurate AS-level Traceroute Tool. In *Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, SIGCOMM '03, pages 365–378, New York, NY, USA, 2003. ACM.

[23] Ricardo Oliveira, Dan Pei, Walter Willinger, Beichuan Zhang, and Lixia Zhang. The (in)Completeness of the Observed Internet AS-level Structure. *IEEE/ACM Trans. Netw.*, 18(1):109–122, February 2010.

[24] Vern Paxson. End-to-end Routing Behavior in the Internet. *SIGCOMM Comput. Commun. Rev.*, 26(4):25–38, August 1996.

[25] E. Chen Q. Vohra. RFC4893,BGP Support for Four-octet AS Number Space. https://www.ietf.org/rfc/rfc4893.txt, 2007.

[26] Philipp Richter, Georgios Smaragdakis, Anja Feldmann, Nikolaos Chatzis, Jan Boettger, and Walter Willinger. Peering at Peerings: On the Role of IXP Route Servers. In *Proceedings of the 2014 Conference on Internet Measurement Conference*, IMC '14, pages 31–44, New York, NY, USA, 2014. ACM.

[27] Patrick S. Ryan and Jason Gerson. A Primer on Internet Exchange Points for Policymakers and Non-Engineers. August 2012.

[28] Gerhard Schmid. on the existence of a global system for the interception of private and commercial communications (ECHELON interception system) (2001/2098(INI)). July 2001.

[29] D. Katz Y. Rekhter T. Bates, R. Chandra. RFC4760,Multiprotocol Extensions for BGP-4. http://tools.ietf.org/html/rfc4760, 2007.

[30] Ying Zhang, Z. Morley Mao, and Ming Zhang. Ascertaining the Reality of Network Neutrality Violation in Backbone ISPs. In *HotNets*. Association for Computing Machinery, Inc., 2008.

[31] Ying Zhang, Z. Morley Mao, and Ming Zhang. Detecting Traffic Differentiation in Backbone ISPs with NetPolice. In *IMC*. Association for Computing Machinery, Inc., 2009.

[32] Yu Zhang, Ricardo Oliveira, Hongli Zhang, and Lixia Zhang. Quantifying the Pitfalls of Traceroute in AS Connectivity Inference. In *Proceedings of the 11th International Conference on Passive and Active Measurement*, PAM'10, pages 91–100, Berlin, Heidelberg, 2010. Springer-Verlag.