Measuring Internet Connectivity Between User Populations Using Active Measurements

Petros Gkigkis

Thesis submitted in partial fulfillment of the requirements for the

Masters' of Science degree in Computer Science and Engineering

University of Crete
School of Sciences and Engineering
Computer Science Department
Voutes University Campus, 700 13 Heraklion, Crete, Greece

Thesis Advisor: Assistant Prof. *Xenofontas Dimitropoulos*

This work has been performed at the University of Crete, School of Sciences and Engineering, Computer Science Department.

The work has been supported by the Foundation for Research and Technology - Hellas (FORTH), Institute of Computer Science (ICS).

University of Crete
Computer Science Department

Measuring Internet Connectivity Between User Populations Using Active Measurements

Thesis submitted by
Petros Gkigkis
in partial fulfillment of the requirements for the
Masters' of Science degree in Computer Science

THESIS APPROVAL

Author: ______________________
Petros Gkigkis

Committee approvals: ______________________
Xenofontas Dimitropoulos
Assistant Professor, Thesis Supervisor

____
Maria Papadopouli
Professor, Committee Member

____
Panagiotis Tsakalidis
Professor, Committee Member

Departmental approval: ______________________
Antonios Argyros
Professor, Director of Graduate Studies

Heraklion, June 2018

Measuring Internet Connectivity Between User Populations Using Active Measurements

Abstract

The Internet is getting better and better at delivering content to end-users; this shift is spearheaded by Internet giants such as Google and Facebook. Nevertheless, a number of applications (e.g., peer-to-peer, Blockchain) rely on user-to-user connections, which raises the –still open– question of how end-users are connected with each other. This is further stressed by the fact that multiple suspicious incidents of path manipulation for user-to-user communications have been reported.

In this work, we use active measurements (i.e., traceroutes between RIPE Atlas vantage points) and publicly available datasets to explore the interconnectivity of the user-facing networks with the largest user populations in any given country. We combine user population per autonomous system (AS) estimates from APNIC with data plane measurements and provide insights into the user-to-user connectivity for 114 countries, over time. In order to study per-country interconnectivity, we construct a framework that stores and processes massive traceroute datasets, making refined results available via an online public API. On a monthly basis, we analyze ∼420K traceroute paths, from ∼3,5K RIPE Atlas probes in ∼2,6K ASes.

We derive statistics and comparisons between countries in terms of: (i) out-of-country vs. in-country paths, (ii) direct connections vs. intermediary networks, (iii) IXP crossings vs. non-IXP crossing paths. We discover among other findings that over time 20% to 50% of the user-to-user connections in Greece cross an IXP; while in the U.S., the fraction of such connections is only ∼3%. We also propose a methodology to infer the transit betweenness of networks in the user-to-user paths. For example, in the U.S. 8% of the user-to-user connections flow through the incumbent provider, consistently in time. Besides, we go beyond eyeball networks, and focus on the differences between the two Internet protocols (IPv4/IPv6) in terms of path lengths, paths staying in or going out of a country, as well as IXP crossings. We observe paths in IPv6 to be shorter than IPv4 for almost all countries. Moreover, in the U.S. the fraction of paths that cross an IXP in IPv6 is ∼50%; two times more than in IPv4. Finally, we evaluate the coverage of RIPE Atlas on user populations around the globe.

Μέτρηση της Διαδικτυακής Σύνδεσης Πληθυσμών Χρηστών με Χρήση Ενεργών Μετρήσεων

Περίληψη

Το Διαδίκτυο βελτιώνεται διαρκώς όσον αφορά την παροχή περιεχομένου στους τελικούς χρήστες. Αυτή η βελτίωση βασίζεται σε μεγάλους παίκτες όπως το Google και το Facebook. Παρ΄όλα αυτά, η εκτεταμένη χρήση εφαρμογών χρήστη-προς-χρήστη (π.χ., peer-to-peer, Blockchain) θέτει το ερώτημα για το πώς οι τελικοί χρήστες συνδέονται μεταξύ τους. Το ερώτημα τονίζεται περαιτέρω από το γεγονός ότι έχουν αναφερθεί πολλαπλά ύποπτα περιστατικά χειραγώγησης μονοπατιών για επικοινωνίες χρήστη προς χρήστη, μεταξύ δικτύων εντός μίας χώρας.

Σε αυτήν την εργασία, χρησιμοποιούμε ενεργές μετρήσεις (δηλαδή traceroutes μεταξύ μετρητικών συσκευών RIPE Atlas) και δημοσίως διαθέσιμα δεδομένα, για να διερευνήσουμε τη διασύνδεση των δικτύων που εξυπηρετούν τους μεγαλύτερους πληθυσμούς χρηστών σε οποιαδήποτε δεδομένη χώρα. Συνδυάζουμε τις εκτιμήσεις του πληθυσμού των χρηστών ανά αυτόνομο σύστημα (AS) από το APNIC, με τις μετρήσεις επιπέδου δεδομένων (traceroutes) και εξάγουμε συμπεράσματα σχετικά με τη συνδεσιμότητα χρηστών για 114 χώρες, στην πάροδο του χρόνου. Προκειμένου να μελετήσουμε τις πτυχές συνδεσιμότητας ανά χώρα, κατασκευάζουμε ένα framework που αποθηκεύει και επεξεργάζεται μαζικά δεδομένα traceroute, καθιστώντας τα επεξεργασμένα αποτελέσματα διαθέσιμα μέσω ενός online δημόσιου API. Σε μηνιαία βάση, αναλύουμε ∼420K traceroute μονοπάτια, από ∼3,5K RIPE Atlas συσκευές σε ∼2,6K αυτόνομα συστήματα.

Εξάγουμε στατιστικά στοιχεία και πραγματοποιούμε συγκρίσεις μεταξύ χωρών όσον αφορά: (i) τα μονοπάτια που εξέρχονται από τη χώρα έναντι εκείνων που παραμένουν εντός των χωρών, (i) τις άμεσες συνδέσεις έναντι των συνδέσεων μέσω ενδιάμεσων δικτύων, (i) τις διελεύσεις μέσω IXP (Internet eXchange Point) έναντι των διελεύσεων εκτός IXP. Μεταξύ των αποτελεσμάτων μας, παρατηρούμε ότι με την πάροδο του χρόνου το 20% έως 50% των συνδέσεων μεταξύ χρηστών στην Ελλάδα διασχίζει ένα IXP, ενώ στις Η.Π.Α., το ποσοστό τέτοιων συνδέσεων είναι μόλις 3%. Επιπλέον, προτείνουμε μια μεθοδολογία για να μετρήσουμε το ποσοστό των συνδέσεων χρηστών που διασχίζουν ενδιάμεσα δίκτυα. Για παράδειγμα, στις Η.Π.Α. το 8% των συνδέσεων μεταξύ χρηστών διέρχεται μέσω του δικτύου ενός μεγάλου παρόχου υπηρεσιών, σταθερά στον χρόνο. Εκτός από τα δίκτυα που καλύπτουν τις μεγαλύτερες πληθυσμιακές ομάδες σε μία χώρα, εξετάζουμε και τις διαφορές μεταξύ των δύο πρωτοκόλλων Διαδικτύου (IPv4 / IPv6) όσον αφορά το μήκος των μονοπατιών, τα μονοπάτια που μένουν εντός ή εξέρχονται από μια χώρα, καθώς και τις διελεύσεις μέσω IXP. Παρατηρούμε ότι τα μονοπάτια του IPv6 είναι μικρότερα ως προς το μήκος από τα αντίστοιχα του IPv4 σε όλες σχεδόν τις χώρες. Επιπλέον, στις ΗΠΑ το ποσοστό μονοπατιών που διασχίζουν ένα IXP στο IPv6 είναι ∼50%, δηλ. δύο φορές μεγαλύτερο από ό,τι στο IPv4. Τέλος, αξιολογούμε την πληθυσμιακή κάλυψη της πλατφόρμας του RIPE Atlas σε όλο τον κόσμο.

Acknowledgements

First of all, I would like to thank my supervisor Prof. Xenofontas Dimitropoulos for showing me his trust and encouragement during my studies as a researcher and as a teaching assistant in his courses and giving me the opportunity to work with him with his great advice and support.

I would also like to express my sincere gratitude to my advisor Dr. Vasileios Kotronis who supported me since day one of my MSc thesis. Thank you for all the technical discussions, our efficient cooperation, exchange of ideas and the time you spent.

I would also like to thank my advisor Emile Aben for giving me the wonderful opportunity to be an intern in RIPE NCC and for all his valuable guidance.

I would also like to thank the members of my dissertation committee, Prof. Maria Papadopouli and Prof. Panagiotis Tsakalidis for conceding to be on my committee.

I would like to acknowledge the Institute of Computer Science (FORTH-ICS) for the scholarship and especially Telecommunications Laboratory for providing me with all the necessary equipment during my studies.

I am also really grateful to Manos Lakiotakis, Lefteris Manassakis for the continuous support, ideas and valuable help and contribution to this work.

I would also like to extend my thanks to my friends and colleagues at FORTH, with whom I have shared many hours of work and joy. Nikoleta, Eftyhia, Antonis, Fotis, Alexandros, Christos, Giorgos, Pavlos, Dimitris, thank you!

Last but not least, I would like to deeply thank my parents, Christos and Kallina, along with my brother Nikos, for their continuous support and love throughout all these years.

The project leading to this application has received funding by the EU Research Council Grant Agreement no. 338402.

May the force be with you...

to my beloved parents...

Contents

Table of Contents **i**

List of Tables **iii**

List of Figures **v**

1 Introduction **1**

1.1 Contributions 2

1.2 Outline of the Thesis 4

2 Datasets and Tools **5**

2.1 RIPE Atlas 5

2.1.1 Probe Features 6

2.1.2 Traceroute Format 6

2.2 APNIC User Estimates per ASN 6

2.3 Geolocation using OpenIPMap 7

2.4 IP-to-AS Mapping 8

2.5 Identifying IXPs 9

2.6 IXP Country Jedi 9

2.6.1 Probe Selection 9

2.6.2 Traceroute execution 10

2.6.3 Analyzing Data 10

3 Eyeball Jedi Framework **13**

3.1 Workflow 13

3.2 Back-End 14

3.2.1 Database Schema 15

3.2.2 Unique Custom Keys for Indexing 18

3.2.2.1 Traceroute ID 18

3.2.2.2 Unique Probe ID 18

3.2.2.3 Unique IP Info ID 19

3.3 Front-End 19

4 Eyeball Connectivity **23**
4.1 Eyeball Networks . . . 23
4.2 IP-path-to-AS-path transformation . . . 24
4.2.1 Rules to transform IP-Path-to-AS-Path . . . 25
4.2.1.1 Limitations of the IP-Path-to-AS-Path transformation . . . 28
4.2.1.2 Alternative Approach: MAP-IT . . . 29
4.3 Eyeball-to-Eyeball Connectivity Matrix . . . 31
4.3.1 Methodology . . . 33
4.3.2 Eyeball-to-Eyeball Matrix in Space . . . 35
4.3.3 Eyeball-to-Eyeball Matrix over Time . . . 36
4.4 Transit Betweenness . . . 42
4.4.1 Methodology to infer transit betweenness of a network . . . 44
4.4.1.1 Symmetric Path Approach . . . 47
4.4.1.2 Asymmetric Path Approach . . . 48
4.4.2 Transit Betweenness over Time . . . 49

5 Further Framework Applications **53**
5.1 Path length . . . 53
5.2 Out-of-Country Paths . . . 55
5.3 IXP Crossing Paths . . . 56

6 Related Work **61**
6.1 IP-path-to-AS-path Mapping . . . 61
6.2 Connectivity between ASes . . . 63

7 Conclusions & Future Work **65**
7.1 Summary and Conclusions . . . 65
7.2 Future Work . . . 66

A Probe Similarity **69**

B Evaluating RIPE Atlas Coverage **73**
B.1 Methodology . . . 73
B.2 Coverage Map . . . 74
B.3 Countries Table . . . 75

Bibliography **77**

List of Tables

4.1 Mapping between IPs and ASes (used for MAP-IT third-party address case) 29

4.2 Mapping between IPs and ASes (used for MAP-IT hidden border address case) 31

A.1 Mapping between the AS Number and the name/holder of the selected Greek networks using RIPEstat. 70

List of Figures

2.1 JSON format of a RIPE Atlas traceroute 7
2.2 Workflow of the IXP Country Jedi Prototype Tool 10

3.1 Workflow of the Eyeball Jedi Framework 14
3.2 Database Schema 15
3.3 Traceroute paths between Greek ASes that passed through North Europe. 20
3.4 Example of `json`-returned object from the API of our framework . 21

4.1 Traceroute 28
4.2 Traceroute Equal paths 28
4.3 Snapshot of eyeball-to-eyeball matrix for Canada (generated on 2017-04-01). 36
4.4 Greece IPv4 in/out of country 37
4.5 Greece IPv6 in/out of country 38
4.6 United Stated of America IPv4 in/out country 39
4.7 Canada IPv4 in/out country 39
4.8 Greece IPv4 IXP crossing 40
4.9 Greece IPv6 IXP crossing 41
4.10 United States of America IXP crossing 41
4.11 Ireland IPv4 IXP crossing 42
4.12 Netherlands IXP crossing 43
4.13 Traceroute measurements between probes {1, 2} of ASX and probes {3, 4} of ASY. 45
4.14 Traceroute measurements between probes {1, 2} of ASX and probe {3} of ASY. 46
4.15 Traceroute measurements between probe {1} of ASX and probe {3} of ASY. 47
4.16 Symmetric Transit Betweenness over time for the United States of America. 50
4.17 Asymmetric Transit Betweenness over time for the United States of America. 51
4.18 Symmetric Transit Betweenness over time for the Netherlands. . . 51
4.19 Asymmetric Transit Betweenness over time for the Netherlands. . 52

5.1 Comparison of IPv4 vs IPv6 average path length over time for various countries 54
5.2 Comparison of IPv4 vs IPv6 % of Out-of-Country paths over time for various countries. 57
5.3 Comparison of IPv4 vs IPv6 IXP crossing paths over time for various countries 59

A.1 Uncovering the interfaces of the border routers between each pair of ASes (used for probe similarity computation) 70
A.2 Probe similarity using Jaccard index 0.5 for AS6799 (OTENET-GR) 71
A.3 Similarity vs. distance between probes of AS6799 72

B.1 IPv4 Public and Private Probes. 74
B.2 RIPE Atlas IPv6 Public and Private Probes. 75
B.3 Table overview of "eyeball coverage" of the RIPE Atlas platform . 76
B.4 Overview of "eyeball coverage" for Greece 76

Chapter 1

Introduction

The overall Internet usage has seen tremendous growth over the last decade. About ~54.4% of the world's population uses the Internet [41]. According to the International Telecommunication Union [19] about 4 billion people (more than half of the world's population) will be online by the end of 2017. However, the majority of the users know little on how the Internet is structured and the sheer complexity of it [38], which results to understanding the Internet as a single abstract entity.

The Internet is a network of networks with distinct modes of technical and political control. Each network entity is called Autonomous System (AS), as it has its own routing policies and strategies. According to an Internet report [42] as of May 2018 about 61K ASes participate in the global Internet routing system. These ASes form transit or peering business relationships with each other, reflected on physical connections between their routing equipment. In the transit relationship, the traditional customer-provider model is applied, where the customer pays the provider to transit its traffic to other ASes with a typically rate-based charge (e.g., based on the 95th percentile approach) [72]. This model results in a transit hierarchy with the ISPs having the largest customer cones at the top. On the contrary, peering relationships express the model of traffic exchange between the users and customers of each network without paying a fee. Moreover, recent research studies discovered that the hierarchical structure of the early days of the Internet has loosened over the years [77].

In this context, the emergence of the Internet Exchange Points (IXPs) enabled medium and small size networks to peer directly with content giants such as Google and Facebook but also with large Content Delivery Networks (CDNs) such as Akamai. The primary objective of these Internet giants is to bring content (and related services) as close as possible to the end-users [52]. By optimizing their connectivity with the user networks they increase the user experience of their services (e.g., reducing latency), while decreasing transit fees. To optimize their Internet connectivity with other ASes, they peer with hundreds of networks either via IXPs [48] or via private interconnections at hundreds of different facilities [64, 59]. As an outcome of all these interconnections between ASes we would expect the Internet to operate as a random game of hot potato routing. However, a research study [4] revealed that only a small portion of autonomous systems carry the majority of the paths for a disproportionate number of routes on the Internet.

Furthermore, a rising tendency of governments around the world trying to legislate and control the Internet by applying censorship on Internet traffic [24, 5, 45], has increased the awareness of the research community to identify possible intermediate networks that could act as points of control [76] for the connectivity of an entire country ecosystem. For example, the AT&T Whistle-Blower's Evidence [21] revealed the case where an ISP cooperated with the U.S. National Security Agency in order to surveil Internet traffic.

Moreover, the interconnection of ASes has become a very active area of research, with the majority of the studies focusing on how networks reach large content providers and their services, as well as on how these routes can be further optimized. Nevertheless, the extended use of user-to-user applications (e.g., peer-to-peer, VOIP, online gaming) and technologies (e.g., Blockchain) raises the following –still open– question. *How is a user serviced by network A in country X connected to another user in network B in the same country?* To address this question we need to focus on how the eyeball networks, *i.e.* the networks that provide Internet access to end-users at the "last mile", are interconnecting with each other. As opposed to content which can be moved around and hosted anywhere in the network [73], end-users usually access the Internet from a limited physical area, typically the area where they reside or work. Therefore, routing IP packets from end-user to end-user is really determined by how the networks serving the users are connected with each other.

In this thesis in order to characterize the user-to-user connectivity in a given country we use the RIPE Atlas measurement platform [31]. The platform has probes deployed inside multiple user networks; these probes can perform active Internet measurements (e.g. traceroutes, pings). Using the platform we explore what RIPE Atlas probes can tell us about the interconnectivity of networks which serve the majority of users in a given country. Our intent is to understand and characterize aspects of user-to-user connectivity in terms of: (i) out-of-country vs. in-country paths, (ii) direct connections vs. intermediary networks, (iii) paths between neighbor country pairs even in light of limited probe coverage within their eyeball ISPs. We believe that such a characterization can help network operators, network researchers and Internet users to discover interesting interconnectivity artifacts and issues [74, 60, 56] within the countries they operate and live, and act upon them.

1.1 Contributions

In this thesis we make the following contributions:

- **Eyeball Jedi Framework - API**
  We propose the Eyeball Jedi framework which can store, enrich with external

datasets and analyze massive traceroute datasets derived from the RIPE Atlas platform. The processed results are available using a publicly available REST API which can be used from network operators as a network debugging tool, but also from researchers working on the field of Internet connectivity using data plane measurements and accounting for how users are connected to each other.

- **Inferring user-to-user connections that**
  - **Go through Internet eXchange Points (IXPs)**
    We characterize the user-to-user connectivity based on the crossing or not of an IXP per AS pair for 114 countries across time. Moreover, we aggregate the results and provide statistics w.r.t. the user-to-user connections that flow through IXPs. We discover that over time, 20% to 50% of the user-to-user connections in Greece cross an IXP; while in the U.S., the fraction of such connections is only ∼3%.
  - **Stay local or go out of country**
    We examine the evolution of the in/out-of-country paths between user-to-user interconnections per AS pair for 114 countries across time. Moreover, we aggregate the results and provide statistics w.r.t. the user-to-user connections that stay in-country or go out of country. We discover in Canada fractions of connections going out of the country ranging from ∼2% to ∼11%, which align with other similar research findings.

- **Quantifying the presence of intermediate networks**
  We propose a methodology that uses data plane measurement results (traceroutes) along with user population per AS estimates and extract quantitative statistics on the "transit betweenness" of the intermediate networks in a country ecosystem. For example, in U.S. 8% of the user-to-user connections flow through the AS7843 (Time Warner Cable Internet LLC), consistently in time.

- **RIPE Atlas Population Coverage**
  To evaluate the RIPE Atlas coverage in user populations per country we created the RIPE Atlas Population Coverage tool. The tool, on a daily basis, estimates the proportion of Internet users across the world situated in networks that can be measured from within RIPE Atlas.

- **Publication to workshop & presentation to RIPE meeting**
  Part of this work, titled ”Characterizing User-to-User Connectivity with RIPE Atlas" has been published at the the Applied Networking Research Workshop (ANRW) 2017 in Prague, Czech Republic [58]. Moreover, early insights from this work have been presented to the Connect Working Group of the RIPE 74 meeting in Budapest, Hungary.

1.2 Outline of the Thesis

The thesis is structured as follows. First, we describe the datasets and the tools that we used to explore the connectivity between user populations (Chapter 2). Then we discuss and introduce a framework which was built in order to analyze the measurement data but also to make the analyzed data available to the community (Chapter 3). After that, we focus on the "Eyeball" connectivity; first, we describe and discuss the methodology to identify these eyeball networks and parse the collected results. Then we describe the methodology to construct the Eyeball-to-Eyeball matrix. After that, we introduce a new metric: the "Transit Betweenness" in order to measure the percentage of user-to-user connections that traverse an intermediate network. Finally, we visualize these results in a meaningful way (Chapter 4). In the next Chapter, we provide additional applications for our framework by studying the path length, the IXP crossings and the paths that went out of country over time but also per IP version to discern the Internet structures of each nation (Chapter 5). Next, we present and discuss the related work (Chapter 6). Finally, we discuss the results, the limitations, and the potential uses of our estimates and our framework (Chapter 7).

Lastly, Appendix A evaluates the current probe selection strategy of the IXP Country Jedi tool. Appendix B presents a tool to explore the proportion of Internet users across the world situated in networks that can be measured from within RIPE Atlas.

Chapter 2

Datasets and Tools

In this Chapter we present the datasets and tools that were used in this work. To make our work reproducible and also to allow other researchers to create new work based on this research, we used only publicly available datasets and open-source tools. This Chapter is structured as follows. First, we describe the RIPE Atlas platform which generated our primary dataset (i.e., traceroute measurements; Section 2.1). Then, we describe the APNIC user estimates per ASN (Section 2.2), which were used to rank the networks –and their contribution in user-to-user communication– based on the estimated number of users they serve on a country level. After that, we describe the datasets that we used to perform IP Geolocation (Section 2.3), IP-to-AS mapping (Section 2.4) and identify whenever a traceroute crosses an IXP or not (Section 2.5). Finally, we describe the IXP Country Jedi prototype tool from which we derived targeted traceroute datasets (Section 2.6).

2.1 RIPE Atlas

RIPE Atlas is one of the largest global Internet Measurement platforms. Its main purpose is providing data on network connectivity and reachability. The platform is supported by thousands of volunteers around the world who host small hardware devices, called probes, in their homes and offices. These probes measure the health of the Internet from all over the world 24 hours a day. The platform is continuously expanding and new probes are connecting all the time. RIPE Atlas was established in 2010 by the RIPE Network Coordination Centre [34]. As of May 2018, it is composed of nearly 25,370 probes and 317 anchors(i.e., probes with advanced hardware and measurement capabilities). Out of the total number of probes, ~10,500 are currently active(online) around the world.

The aggregated data collected by the probes are publicly available to everyone through the RIPE Atlas API [28]. The users who host probes benefit by earning credits on a daily basis. These credits can be used in order to launch their own customized measurements, and gain valuable information about their own network or any other network they want to measure.

In this work, we use Atlas probes to perform traceroute measurements. We explain the main capabilities and features of a probe in section 2.1.1 and the traceroute format used by RIPE Atlas in section 2.1.2.

2.1.1 Probe Features

The backbone of the RIPE Atlas platform is formed by the RIPE Atlas probes. The probes are small hardware devices that actively measure Internet connectivity through `ping`, `traceroute`, `DNS`, `SSL/TLS`, `NTP` and `HTTP` measurements. Probes are able to perform measurements for both IP address protocols (IPv4/IPv6), given that the provider network supports them. The most valuable properties that can be extracted from a probe are the ASN of the network hosting the probe, the prefix where the IP of the probe belongs, and the –user-defined– location of the probe.

2.1.2 Traceroute Format

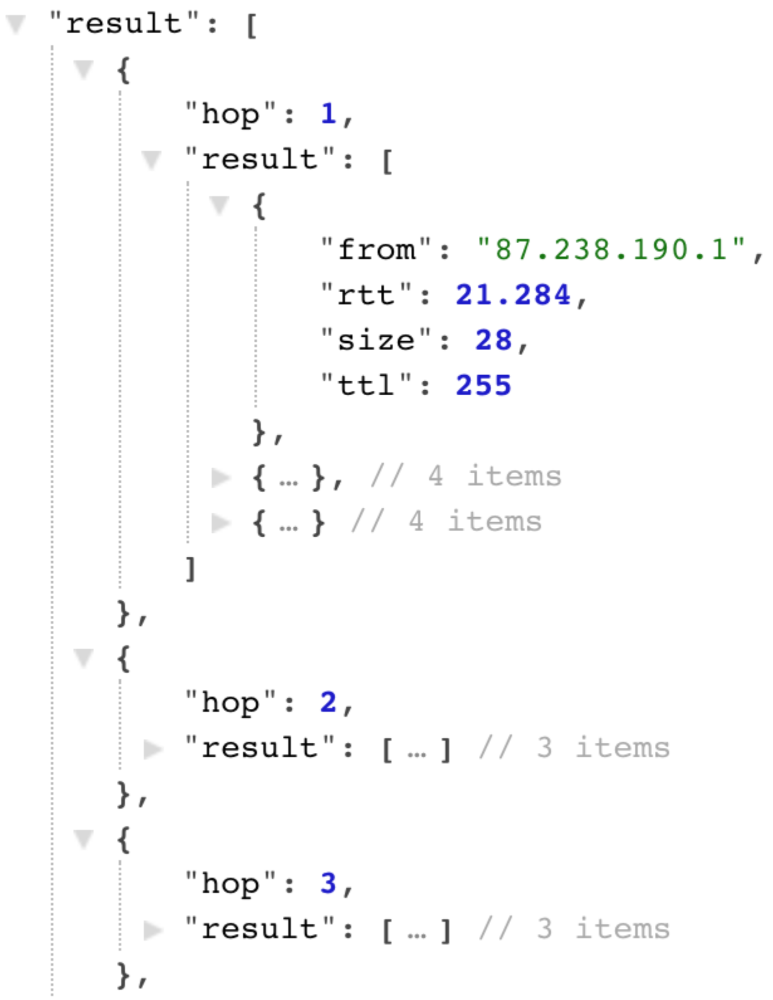
In this work we use used only data plane measurements; specifically traceroutes that were launched using the RIPE Atlas platform. An example of a traceroute result that has been fetched from the RIPE Atlas platform is depicted at Fig:2.1. The results are sorted by their hop number[1] and encapsulated inside the list “result”. For each item of the “result” list an index key ”hop” is provided along with another list called ”result”. The encapsulated ”result” list consists of three or more responses to traceroute probes (3 by default per measurement). Each response item is a dictionary with multiple fields; in our work we focus on the “from” (IPv4 or IPv6 source address in reply), “rtt” (round-trip-time of reply, not present if the response is late) and “ttl” (time-to-live) fields. A more detailed version of the raw data structure of a traceroute is presented in [29].

2.2 APNIC User Estimates per ASN

Measuring the relative size of a network in terms of resources has been done in the past using various techniques. These techniques leverage public information sources such as the number of routed IPs announced by the network[47], as well as the number of transit customers who use the specific network.

However, estimating the size of a network in terms of users or customers, e.g., in case the network is an Internet Service Provider (ISP), is a challenging process [54]. The primary reason is the wide use of NAT (Network Address Translation) inside networks. APNIC [3], the Regional Internet address Registry (RIR) for the Asia-Pacific region, made a first attempt to provide an answer to this question by measuring the number of users inside a network using a methodology based on online advertisements[14]. These measurements provide an estimation of the actual numbers.

[1] Note that the default value of maximum hops in the traceroute implementation of the RIPE Atlas is 32.



```
▼ "result": [
  ▼ {
        "hop": 1,
      ▼ "result": [
          ▼ {
                "from": "87.238.190.1",
                "rtt": 21.284,
                "size": 28,
                "ttl": 255
            },
          ▶ { … }, // 4 items
          ▶ { … } // 4 items
        ]
    },
  ▼ {
        "hop": 2,
      ▶ "result": [ … ] // 3 items
    },
  ▼ {
        "hop": 3,
      ▶ "result": [ … ] // 3 items
    },
```

Figure 2.1: JSON format of a RIPE Atlas traceroute

As an outcome of their study, they created a publicly available report of user populations per Autonomous System (AS)[7]. In our work we use this data; we fetch their published reports on a daily basis. Then, we group the ASes based on their country of origin. We started collecting these reports on July 1, 2017 and we have been collecting them until today.

2.3 Geolocation using OpenIPMap

In the last decade many IP address geolocation services have emerged in the market. The need of this kind of databases is crucial to various domains of the Internet industry, with the main domain being advertising. Advertisers use these data to identify the location of a user and display location-based advertisements.

Today, many companies provide IP geolocation databases and services. The most well-known initiatives are: MaxMind, offering two databases, GeoLite2 [12] and GeoIP2 [22] which are free and paid respectively, Digital Element NetAcuity (mostly known as NetAcuity) [9] and DB11-Lite (mostly known as IP2Location-Lite)[16].

A recent work [57] compared these databases and revealed that they suffer from high inaccuracies when used for IP router interface geolocation. Furthermore, the authors concluded that there is room to improve router geolocation, due to the fact that these databases are mostly good at geolocating IP addresses on the edge of the Internet and not middleboxes such as routers.

OpenIPMap was initially started as a prototype tool developed by Emile Aben[18]. The tool's goal was to provide reliable geolocation data for core Internet infrastructure. The key difference between the OpenIPMap and the other geolocation databases on the market is that OpenIPMap focuses on geolocating Internet infrastracture (such as routers), rather than edge IP addresses. OpenIPMap uses multiple input sources, while utilizing an algorithm based on weighs to extract the IP address location. The tool relies on self-reported data, reverse DNS hostnames, and publicly available geolocation databases. Moreover, it uses a validation technique based on speed-of-light calculations with round-trip times of pings towards the geolocated target. Currently, RIPE NCC has developed a production-level version of OpenIPMap, available at [17, 26]. In our work, to geolocate IP addresses, we used the prototype OpenIPMap tool, since it was available in that time period. In the future, we plan to use the production OpenIPMap version.

2.4 IP-to-AS Mapping

In order to map an IP address to the administrative AS, we used the same IP-to-AS mapping process with the IXP Country Jedi prototype tool (see section: 2.6). The tool performs IP-to-AS mapping using the RIPEStat service [37]. The service offers a REST API query call,[2] which takes as input an IP address and returns a `json`-formatted response with the holder/origin AS of the address.

Specifically, the RIPEStat service uses Internet routing data to derive the mapping between IP addresses and origin holder ASes. The service uses the collected and stored Internet routing data from the Routing Information Service (RIS)[35], which was established in 2001. It is worth mentioning that the RIS project in order to provide a global view of the Internet routing system, collects BGP data from several locations (route collectors) around the globe. Using these RIS data, RIPEStat processes hourly all the routed prefixes for all ASes and then maps prefixes to origin ASes. Moreover, to perform an IP-to-AS lookup, the service searches for the longest routed prefix match and returns its origin AS (as shown in BGP advertisements).

[2] `https://stat.ripe.net/data/prefix-overview/data.json?max_related=0&resource=`

2.5 Identifying IXPs

An Internet exchange point (IX or IXP) is the physical infrastructure through which Internet service providers (ISPs) and content delivery networks (CDNs) exchange Internet traffic between their networks (Autonomous Systems).

Identifying an IXP crossing inside a traceroute path is not a trivial task (we refer the reader to sophisticated IXP identification techniques such as the traiXroute [70]). In our work we use the following simple approach to identify them at scale. On a monthly basis, we collect all the IXP subnets from PeeringDB, and we group them based on the country where each IXP operates.

At this moment, Euro-IX (European Internet Exchange Association) is building a global IXP Database [11]. The objective of this upcoming database will be to keep all IXP related info (e.g. IXP peering subnets and IXP members) up-to-date. We plan to use this dataset as soon as it becomes available, in order to handle cases such as distributed IXPs, or IXPs that operate in multiple countries.

2.6 IXP Country Jedi

IXP Country Jedi is a prototype tool developed by Emile Aben[20]. The goal of the tool is to perform monthly traceroute measurements inside countries and report which paths stay local (inside the same country). It also reports which paths crossed an IXP within a given country. The tool uses RIPE Atlas probes and is publicly available. It is worth noting that there is no state shared between each run of the tool, to avoid transferring any bias across runs, as well as because the measurement probes do not remain stable over time.

In our work, the primary dataset we used to derive the connectivity results on the country level, at scale and in time, starting from July 2015, is produced by the IXP Country Jedi prototype tool. However, this tool uses a specific probe selection methodology to initiate traceroute measurements between probe-hosting ASes in a given country. We next show that this probe selection provides efficient coverage for our own inferences.

2.6.1 Probe Selection

Measuring the connectivity of an AS against an Internet target is not an easy task. A well-known fact is that each AS has its own routing policies and may offer different type of services to its end-users. The diversity of paths between two end-users of the same AS against a target can be unpredictable. IXP Country Jedi studies the connectivity between two ASes using the methodology described in the following:

To validate the probe selection strategy of the tool we examined the path diversity of 5 Greek networks in Appendix A.

Practical probe Selection Strategy. The IXP Country Jedi selects two probes at most per ASN, in a given country. The probe selection methodology relies on two criteria. The first criterion is related to the probe location and the distance from the capital of the given country. The tool selects the closest and farthest probe from each ASN, w.r.t. the capital city, to attempt to provide some level of geo-based path diversity. As a second criterion, it filters probes using the RIPE Atlas system tag 'system-ipv4-stable-1d' and 'system-ipv6-stable-1d' for IPv4 and IPv6 respectively. These tags indicate the status of the probe in the last 24 hours. The status is not only based on probe uptime but also relates to the ratio of successfully completed measurements performed by the probe.

2.6.2 Traceroute execution

The tool runs on a monthly basis for about 114 countries around the globe. Specifically, it launches a measurement campaign on the first day of each month. In order to explore the paths in a given country, it performs full mesh traceroutes between all available RIPE Atlas probe-hosting ASNs. These are probe-to-probe measurements between the probes of AS_X and AS_Y, as chosen by IXP Country Jedi, for each [AS_X,AS_Y] pair. To remove discrepancies stemming from load-balancing and increase the accuracy of the traceroute results, it performs only Paris traceroutes[46].

2.6.3 Analyzing Data

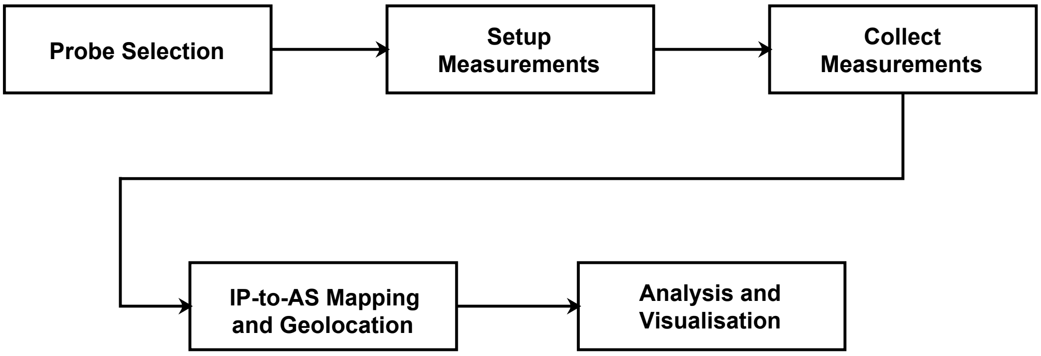


Figure 2.2: Workflow of the IXP Country Jedi Prototype Tool

The IXP Country Jedi workflow is depicted in Fig: 2.2 and consists of the next five steps.

1. **Probe Selection**: The tool selects the available RIPE Atlas probes for a given country according to the methodology we described in Section 2.6.1.

2. **Setup Measurements**: Using the probe list from the previous step, it configures the mesh measurements between all selected probes. The measurements are initialized using the RIPE Atlas API and a list of measurements IDs is generated. The measurement IDs are used to check the status of the

measurements (e.g., “finished”/“unfinished”), and also to collect the actual results.

3. **Collect Measurements**: The measurement results are fetched from the platform. The results usually become available after 5-15 minutes from the initial measurement request.

4. **IP-to-AS and Geolocation**: The IP-to-AS mapping is performed using RIPEStat data, as described in Section:2.4. Along with the IP-to-AS mapping, IP geolocation is applied using the OpenIPMap tool (described in Section:2.3).

5. **Analysis and Visualisation**: Then the enriched results are analyzed, in order to explore path properties such as whether the path crossed an –in-country– IXP or went out of the country.

   To easily explore the results, a set of visualizations is generated. Some of these visualizations are the following:

   - **GeoPath**: It compares the paths in IPv4 and IPv6, by indicating where the paths went in order to reach the final destination.
   - **AS Graph**: The output is a graph of all ASNs that have been discovered in the measurement.
   - **IXP Country**: The IXP Country is a tabular structure in which rows and columns correspond to different probes, used as sources and destinations respectively. This structure allows to view the traceroute paths that went out of country or stayed local, and also if they crossed an IXP.
   - **RTT Mesh**: Is similar to the IXP Country tabular structure but this one focuses on the Round Trip Time (RTT) between the probes.
   - **Per ASN report**: Creates a report per each ASN in the measurement and tries to display the most interesting paths to other ASes in the same country that host RIPE Atlas probes. It focuses on paths that went out of country and paths that crossed intermediate ASes.

Chapter 3

Eyeball Jedi Framework

In this Chapter, we propose and describe the Eyeball Jedi framework that we constructed in order to “mine”, at a fine granularity, the already collected data of IXP Country Jedi. The primary functionality of the proposed framework is to collect, index and enrich traceroute data. Furthermore, it provides an online REST API to allow network operators and other researchers to apply useful queries on the data. This framework is the basis on which we run other applications (related to inter-domain routes), that we describe in Chapters 4 and 5.

This Chapter is structured as follows. First, we analyze the workflow of the proposed framework, from raw traceroute data to processed –usable– path-related results (Section 3.1). Then, we describe the back-end and the techniques that we use to store and index data (Section 3.2). Finally, we discuss the framework’s front-end, and provide a list of currently supported API calls (Section 3.3).

3.1 Workflow

In this section we present the workflow of the implemented framework which is depicted in Fig: 3.1. The first step in the workflow parses the measurement `json` file of each IXP Country Jedi run for a given country and a given date. This file includes all the RIPE Atlas measurement IDs in IPv4 and IPv6. Then, using these measurement IDs and the RIPE Atlas Cousteau API [30], it fetches the traceroute data. After that, the traceroute results are parsed and meta-data tags are generated to store alongside the initial data in the database.

As a next step, the stored results can be enriched using the meta info from the IXP Country Jedi. The enrichment process is generic and can be performed using additional sources, such as IXP identification via traiXroute [70], IP-to-AS mapping via TraceMON [43], and a variety of different IP geolocation databases [9, 12, 16, 22].

Finally, an online REST API is available to retrieve the data. The API supports filtering queries based on the traceroute properties such as “retrieve data per AS and address family (v4/v6)”.

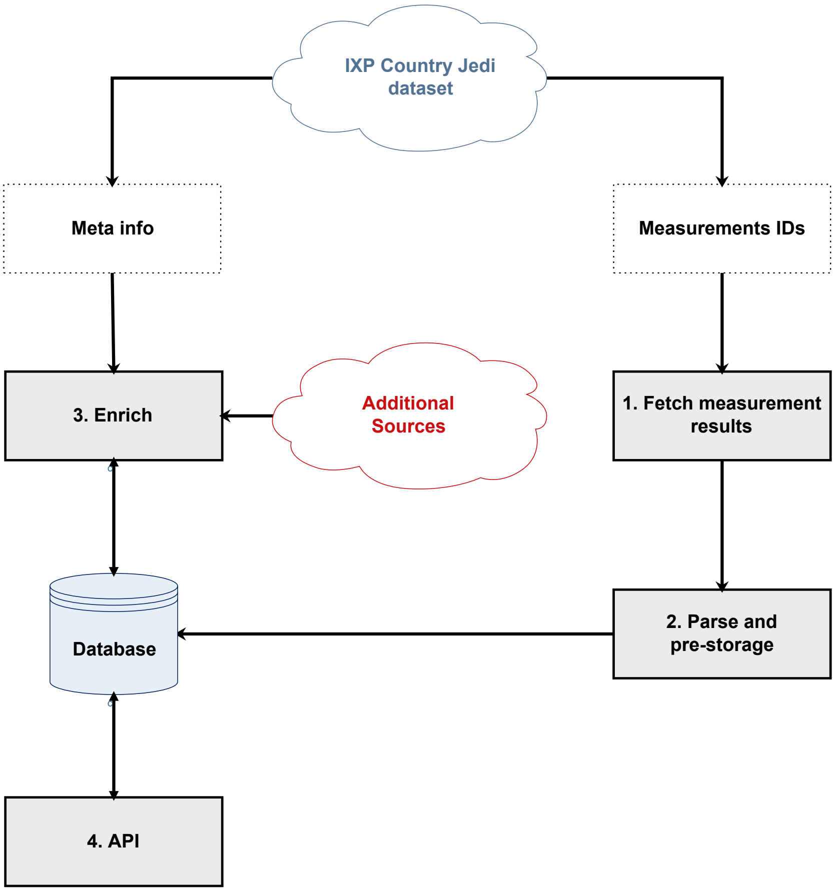


Figure 3.1: Workflow of the Eyeball Jedi Framework

3.2 Back-End

To be able to process and analyze the data in a fast, easy to use and maintainable environment, we selected Django [10] as our primary back-end solution. The main reason behind this selection was that it is a high-level Python Web framework, which will allow anyone to easily contribute to our code-base in the future. Furthermore, the Django project maintains some of the most active developer communities, within which future maintenance and development are almost guaranteed.

Another significant decision had to do with data storing and indexing mechanisms. From the beginning, we required that our framework should be able to perform a range of different data analysis tasks. In order to keep with this approach we selected PostgreSQL [27] (also known as Postgres), which is an object-relational database management system (ORDBMS) with emphasis on extensibility. Using an object-relational database provides a fast and reliable way to analyze the traceroute data.

3.2.1 Database Schema

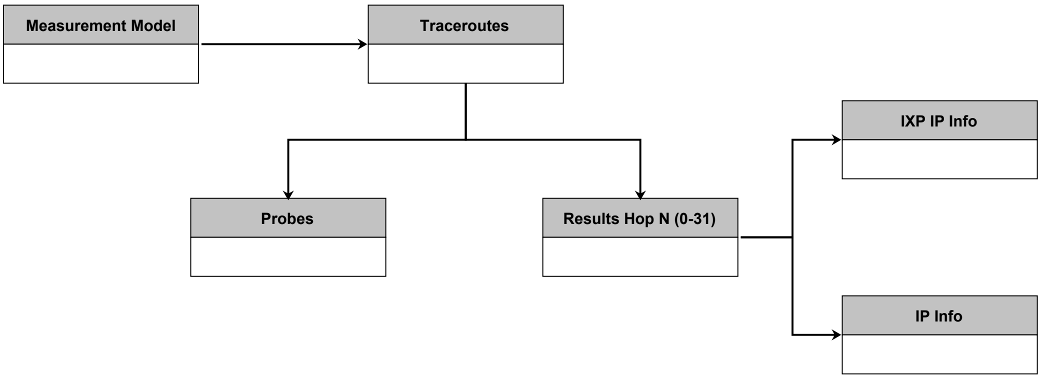


Figure 3.2: Database Schema

Designing a database to store traceroute results was a challenging process; we had to find the “right” data structures and abstractions to handle million of results quickly at scale. As we already mentioned, we selected PostgreSQL to achieve this. Although PostgreSQL is one of the most optimized DB solutions in terms of performance, the nature of traceroute data makes them hard to store in an ORDBM system. While designing the database schema, we had to make a couple of significant design choices that would allow us to join and respond to queries with acceptable performance.

In the rest of this section we explain and discuss some of the built database tables along with their most important key attributes. An abstraction of the database schema is depicted in Fig. 3.2.

- **Measurement Model Table**: at this table we store the RIPE Atlas measurement IDs that were imported in the database.
  - **ID**: Primary key of the ”measurement” table.
  - **Ripe Atlas Measurement ID**: The RIPE Atlas measurement ID.
  - **Af**: Address Family (IPv4/IPv6).
  - **Timestamp**: Start time of the measurement.
- **Traceroutes Table**: is one of the most significant tables of the database. The table includes meta info regarding each traceroute result, along with index pointers, in order to be able to rebuild the actual traceroute by joining other tables.
  - **ID**: Primary key of the ”traceroutes” table.
  - **Source Probe**: The probe ID from which the traceroute was initialized.

  - **Source IP**: The public source IP of the probe.
  - **Destination IP**: The destination IP of the traceroute.
  - **Source ASN**: The ASN of the probe from which the traceroute was intialized.
  - **Destination ASN**: The destination ASN of the traceroute.
  - **Af**: Address Family (IPv4/IPv6).
  - **Source Country**: The country code where the source probe belongs.
  - **Destination Country**: The country code where the destination IP belongs.
  - **Unique Probe ID**: Custom key to track probe changes over time. We describe in detail this key in Subsection: 3.2.2.2.
  - **Unique Traceroute ID**: Custom key to identify each traceroute result. We describe in detail this key in Subsection: 3.2.2.1.
  - **Number of hops**: The number of hops of the traceroute.
  - **Out of country**: Flag to mark if at least one of the traceroute hops was geolocated to a country different from source and destination.
  - **IXP detected**: Flag to mark if we observed an IXP crossing in –at least one of– the traceroute hops.

- **Hop N**$\in [0, 31]$ **Results Table**: To parallelize and speedup the queries on our database, we constructed 32 parallel tables to store the traceroute results. We remind the reader that each traceroute result can contain at most 32 hops, as this is the default value of maximum hops in the traceroute implementation of RIPE Atlas. Each hop result is stored in a table, based on its order of appearance in the traceroute path. Not all tables are used for every single traceroute. For example, assuming that we have a traceroute result with 8 hops, we will use only the first 8 tables to store the results.

  - **ID**: Primary key of the "Hop N$\in [0, 31]$ Results" table.
  - **From x where x** $\in [0, 2]$: For each hop we store the 3 IPs that responded.
  - **RTT x where x** $\in [0, 2]$: For each hop we store the 3 Round-Trip Time values of each packet.
  - **Unique Traceroute ID**: Custom key to identify each traceroute result. We describe in detail this key at Subsection: 3.2.2.1
  - **IXP IP Info**: If an IP in this hop crossed an IXP, this field provides a key, so that we find the IXP in the IXP table.
  - **From x where x** $\in [0, 2]$ **IP Info**: For each of the 3 stored IPs we store an IP identification, in order to be able to retrieve meta info from multiple sources.

- **IP Info Table**: at this table we store the IP enrichment meta info.
  - **ID**: Primary key of the ”IP Info” table.
  - **Unique IP info key**: Custom key to retrieve meta info for an IP. We describe in detail this key in Subsection: 3.2.2.3.
  - **IP**: The actual IP address.
  - **Longitude**: The longitude coordinate.
  - **Latitude**: The latitude coordinate.
  - **Hostname**: The hostname of the IP extracted from the reverse DNS lookup mechanism.
  - **Country Code**: The country where the IP is geolocated.
  - **ASN**: The ASN in which the IP was mapped.
  - **External Source**: The data source from which this meta info was derived.
- **Probes Table**: at this table we store information for every probe that contributed to traceroute results during the measurement campaign(s).
  - **ID**: Custom key to track the ”Probes” over time inside the table. We describe in detail this key in Subsection: 3.2.2.2.
  - **Atlas Probe ID**: RIPE Atlas probe ID.
  - **Probe description**: Probe description, as extracted from RIPE Atlas.
  - **Longitude**: Longitude coordinates.
  - **Latitude**: Latitude coordinates.
  - **ASN in IPv4**: IPv4 ASN, where the probe is hosted.
  - **ASN in IPv6**: IPv6 ASN, where the probe is hosted.
  - **IPv4 Address**: Public IP(IPv4) of the probe.
  - **IPv6 Address**: Public IP(IPv6) of the probe.
  - **Country Code**: Country where each probe is hosted, in ISO 3166-1 alpha-2 format[8].
  - **Dates**: A list of IXP Country Jedi runs in which this probe has participated. In order to append a new entry to this list, the properties that compose the unique Probe ID should be immutable. If at least one property changes, a new key will be formed, and a new entry at the table will be added.
- **IXP Info Table**: at this table we store the IXP name and prefixes that were imported in our framework. Using this table we are able to perform identification of IXP crossings inside the traceroute paths.

- **ID**: Primary key of the "IXP Info" table.
- **IXP name**: The name of the IXP.
- **Subnets list**: List of IXP subnets.
- **Country Code**: The country where the IXP operates.

3.2.2 Unique Custom Keys for Indexing

To easily extract the results, but also track any changes over time for the stored data, we had to create a set of custom keys. These custom keys ensure the uniqueness of the stored results, as well as our capability of monitoring their changes over time. A complete list of these keys, along with a detailed explanation follows.

3.2.2.1 Traceroute ID

RIPE Atlas does not offer any unique traceroute identifier, while aggregates measurement results via measurement IDs. As a consequence, to be able to identify and track each single traceroute result, we constructed a unique identifier per result (the "Traceroute ID"). It is consisted of 4 distinct traceroute properties.

Format of the Traceroute ID:

$$\{Measurement\ ID\}\ -\{Probe\ ID\}\ -\{Timestamp\}\ -\{Dest\ Addr\} \tag{3.1}$$

$$\begin{gathered} where: \\ Measurement\ ID:\ RIPE\ Atlas\ measurement\ ID \\ Probe\ ID:\ RIPE\ Atlas\ probe\ ID \\ Timestamp:\ Start\ time\ of\ the\ traceroute \\ Dest\ Addr:\ Destination\ IP\ address \end{gathered}$$

3.2.2.2 Unique Probe ID

RIPE Atlas uses a probe ID to refer to each probe. However, the probe AS and location may frequently change. To track down those changes in time, we constructed a unique probe identifier which consists of 4 properties that characterize each probe.

Format of the Probe ID:

$$\begin{gathered}\{ASN_{v4}\} \ -\{ASN_{v6}\} \ -\{Latitude\} \ -\{Longitude\} \qquad (3.2)\\ where:\\ ASN_{v4}: \ ASN \ where \ the \ IPv4 \ address \ of \ the \ probe \ resolved\\ ASN_{v6}: \ ASN \ where \ the \ IPv6 \ address \ of \ the \ probe \ resolved\\ Latitude: \ Latitude \ coordinates \ of \ the \ probe\\ Longitude: \ Longitude \ coordinates \ of \ the \ probe\end{gathered}$$

The constructed Probe ID key ensures the historic track of a probe over time in our framework.

3.2.2.3 Unique IP Info ID

The stored traceroute hop IP results can be enriched using various external sources. To achieve this functionality, we created a key based on the meta info of each IP address accompanying the respective data source.

Format of the IP Info ID:

$$\begin{gathered}\{IP\} \ -\{ASN\} \ -\{Country\} \ -\{Latitude\} \ -\{Longitude\} \qquad (3.3)\\ where:\\ IP: \ The \ IP \ address\\ ASN: \ The \ ASN \ where \ the \ IP \ address \ resolve\\ Country: \ The \ Country \ code \ where \ the \ IP \ address \ geolocated\\ Latitude: \ The \ latitude \ coordinate \ of \ the \ probe\\ Longitude: \ The \ longitude \ coordinate \ of \ the \ probe\end{gathered}$$

This specified key format ensures that the IP meta info changes can be tracked over time.

3.3 Front-End

Building a framework that will allow network operators and researchers to easily access and query the stored traceroute data was an incentive from the early start of the Master thesis project. In order to fulfill this requirement we created a public REST API using the Django Rest Framework [1].

The API can be used by developers in order to build new applications but also by researchers to retrieve, explore and analyze path-related data. As an example, a network administrator can use it to fetch traceroutes and employ it as a Network debugging tool between his network and other networks in the same country over time. Attributes such as paths that crossed a foreign country or paths that went through an IXP or not, can be extracted with a single API call for a given AS. An example of paths between Greek ASes that went through Europe instead of staying inside Greece is depicted in Fig: 3.3. This strange routing policy was discovered using our framework API; the geolocation of the paths was performed using the OpenIPMap tool [17].

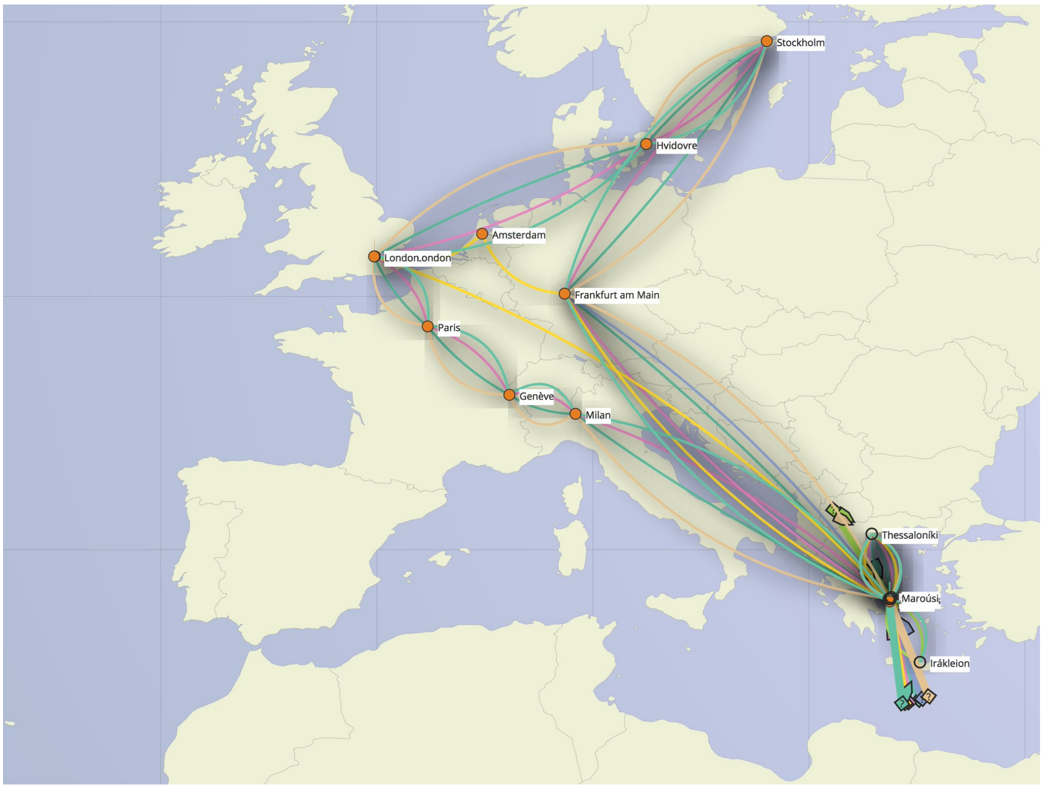


Figure 3.3: Traceroute paths between Greek ASes that passed through North Europe.

The API is highly extensible and new API calls can be added on demand. An indicative list of some interesting calls that are currently supported is the following.

API Calls

The root url of the API is `https://www.eyeball-jedi.net/api/v1/results`

1. **Fetch Traceroute by ID**

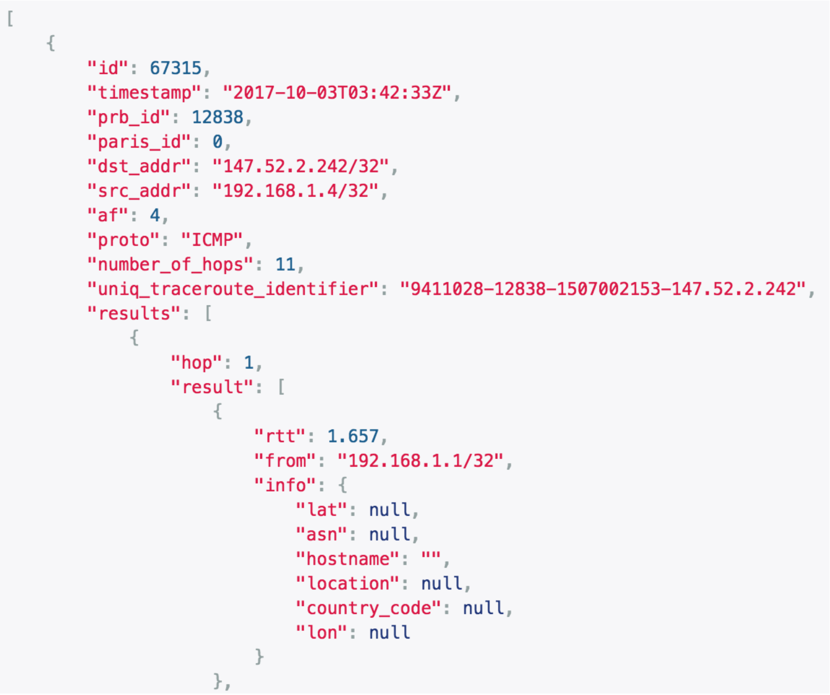
   The Framework assigns a unique ID for each traceroute it stores inside the database; using this unique ID we can fetch the traceroute details. This API call is responsible to fetch the data from the database and serve them to the requester in a `json` format.

   An example call is the following:
   `*/traceroute/id/<traceroute_id>`

2. **Fetch Traceroutes using filtering**

   To easily explore the stored data we created an API call that filters out results based on the ASN, the Address Family (v4/v6) and the date.

   - **Retrieving traceroute results for a pair of ASes**
     `*/cc/<country_code>/src_asn/<ASN>/dst_asn/<ASN>/af/<v4|v6>`
   - **Retrieving data in/out of country paths**
     `*/cc/<country_code>/asn/<ASN>/incountry/<true|false>/af/<v4|v6>`
   - **Retrieving data IXP crossing paths**
     `*/cc/<country_code>/asn/<ASN>/ixp/<true|false>/af/<v4|v6>`
   - **Retrieving all data for a given AS**
     `*/cc/<country_code>/asn/<ASN>/af/<v4|v6>`



```
[
    {
        "id": 67315,
        "timestamp": "2017-10-03T03:42:33Z",
        "prb_id": 12838,
        "paris_id": 0,
        "dst_addr": "147.52.2.242/32",
        "src_addr": "192.168.1.4/32",
        "af": 4,
        "proto": "ICMP",
        "number_of_hops": 11,
        "uniq_traceroute_identifier": "9411028-12838-1507002153-147.52.2.242",
        "results": [
            {
                "hop": 1,
                "result": [
                    {
                        "rtt": 1.657,
                        "from": "192.168.1.1/32",
                        "info": {
                            "lat": null,
                            "asn": null,
                            "hostname": "",
                            "location": null,
                            "country_code": null,
                            "lon": null
                        }
                    },
```

Figure 3.4: Example of `json`-returned object from the API of our framework

To use the API, a detailed documentation of the supported calls is provided at the following website: `https://www.eyeball-jedi.net/api/docs/`.

Chapter 4

Eyeball Connectivity

In this Chapter we study the connectivity between eyeball networks. First, we discuss how we can characterize a network as eyeball or non-eyeball (Section 4.1). Then, we describe the methodology that we used in order to parse the traceroute results and extract the AS-level path from the IP-level route (4.2). After that, we present two different approaches to study the connectivity of the networks inside a country. The first approach aims to provide an insight into the peering policies between the eyeball networks that operate in a given country. To this end, we describe the methodology that we used to construct the Eyeball-to-Eyeball connectivity matrix (Section 4.3). The second approach focuses on the networks that act as intermediate (e.g., transit) between two networks in a country. We propose a methodology to rank these networks by using the transit betweenness metric (Section 4.4). Finally, we evaluate the two approaches and provide insights for the eyeball ecosystem of a set of countries.

4.1 Eyeball Networks

In our work, with the term "eyeball" network we refer to the user-facing networks with the largest user populations in any given country. Typically, an eyeball network is an Internet Service Provider (ISP) which offers Internet access to customers such as home users or small enterprises. To classify a network as an eyeball or non-eyeball, we apply a threshold on the fraction of the users of the total population that this network serves. As a source of the fraction of users per AS we use the APNIC estimates (see Section 2.2). A range of different thresholds can be applied to make this classification, such as 1.0% or 10.0% of the total Internet user population (on the country level). In our study, we use as a threshold a conservative threshold of 1.0% since this allows us to study the country-level eyeball ecosystems at a fine granularity. This method typically represents a majority of Internet users (if we consider all country-wide eyeball AS with over 1% coverage), on average covering 90.5% of end- users per country, though there are outliers such as Russia with only 29.3% coverage due to a highly fragmented eyeball ecosystem. It worth mentioning that since the used threshold is not authoritative on our behalf, we make it tunable for the user. In the visualizations and results which are available through the online portal [40], the user can actually specify the threshold to characterize networks as “eyeballs”.

Despite the potential use of different thresholds to distinguish eyeball from non-eyeball networks, the number of the “eyeball” networks in a country is closely related with the location and size of the country. We can identify three types of country categories:

1. Countries with a large incumbent network/ISP; usually this network owns more than 50% of the market share and is often the national ISP of the country.

2. Countries with a couple of networks where each one owns more than 10% of the market share.

3. Countries where very few networks own more than 10%.

An interesting case is Russia, where the eyeball ecosystem looks highly diverse and fragmented. In this case, many of the networks we would need in order to cover 95% of the market fall below a 0.1% market share threshold. For anything under this threshold, one would need 343 ASNs to cover the 95% of market share.

To infer the “eyeball” networks in a country we used the user population per ASN estimates from APNIC. We describe this dataset in detail in Section 2.2. APNIC updates these estimates on a daily basis; percentages may change over time. In this work, to avoid any errors and outliers and to provide a consistent basis to estimate the user population on the AS level we used the following methodology.

On a daily basis, we started fetching the estimates for all countries and ASes from APNIC. Then, we selected a time window of one month between 15 January 2018 to 15 February 2018 and calculated the average percentage values for each country and AS. For all methodologies, calculations and results in this thesis we use this dataset.

4.2 IP-path-to-AS-path transformation

To identify a network as intermediate between two ASes we analyze the traceroute paths and then we extract the AS path. As a first step to extract the AS path we need to perform IP-to-AS mapping for each IP in the traceroute. For this mapping we used the dataset and methodology that we described in Section 2.4. However, mapping a traceroute result from the IP level to the AS level is not a trivial task. The main challenges behind this mapping/transformation are how the non-responding hops or unknown hops should be handled but also the limitations of the traceroute tool itself. Routers may respond with different interfaces than the inbound or use third party addresses to respond to the probing packets.

4.2.1 Rules to transform IP-Path-to-AS-Path

To study the eyeball connectivity at the AS level in a clean and consistent IP-to-AS transformation we created and applied the following set of rules.

1. **Unknown/Non-responding Hops**

   This set of rules focus on the non-responding/unknown hops of a traceroute result. We define a set of sub-rules that match patterns of AS-level paths.

   (a) In case we have only one unknown hop in the traceroute path, we assume that this hop belongs either to the previous AS or to the next AS. In case of a single non-responding/unknown hop between distinct ASes we do not assume that a hidden AS is present in the path, since the working assumption is that a packet needs to traverse at least 2 routers (and the respective interfaces) at the same AS to enter and exit this AS.

   $$AS_1 \rightarrow * \rightarrow AS_1 \rightarrow AS_2 \tag{4.1}$$

   $$AS_1 \rightarrow * \rightarrow AS_2 \tag{4.2}$$

   Example of this pattern match is depicted at 4.1 and 4.2. For both of the two cases we extract as AS path $AS_1 \rightarrow AS_2$.

   (b) If we have more than one unknown hops in the traceroute path, and these hops are between hops that resolved to the same AS, we assume that these hops belong to the same AS, since the working assumption is that AS-level loops, forbidden in the control plane (BGP) should not be reflected on the data plane (traceroutes).

   $$AS_1 \rightarrow AS_1 \rightarrow * \rightarrow * \rightarrow AS_1 \rightarrow AS_2 \tag{4.3}$$

   Example of this pattern match is depicted at 4.3. For this traceroute we extract as AS path $AS_1 \rightarrow AS_2$.

   (c) If we have more than one unknown hops in a row and these hops are between hops that are resolved to different ASes we can not make any claims. Two consecutive non-responding/unknown hops may signal the traversal of an entirely new AS which we do not know about.

   $$AS_1 \rightarrow * \rightarrow * \rightarrow AS_2 \tag{4.4}$$

   Example of this pattern match is depicted at 4.4. At this case we report $AS_1 \rightarrow AS_* \rightarrow AS_2$.

2. **Hops with IPs that resolved to different ASes**

   With this set of rules we examine the cases where multiple IPs have been found on a hop, belonging to different ASes.

   (a) In case we found a hop with multiple IPs that belong to the previous and the next AS, we assume that this hop belongs to one of the two ASes.

   $$AS_1 \rightarrow \{AS_1, AS_2\} \rightarrow AS_2 \tag{4.5}$$

   Example of this pattern match is depicted at 4.5. In this case, we report as AS path $AS_1 \rightarrow AS_2$.

   (b) In case we found a hop with multiple IPs and the previous hop or the next hop is a non-responding/unknown hop, we examine if the IPs of the hop belong to the previous or next AS. If yes, we assume that this hop belongs to one of the two ASes.

   $$AS_1 \rightarrow * \rightarrow \{AS_1, AS_2\} \rightarrow AS_2 \tag{4.6}$$

   $$AS_1 \rightarrow \{AS_1, AS_2\} \rightarrow * \rightarrow AS_2 \tag{4.7}$$

   Example of this pattern match is depicted at 4.6 and 4.7. For both of these cases we report as AS path $AS_1 \rightarrow AS_2$.

   (c) In case we found a hop with multiple IPs and the previous hop or the next hop is a non-responding/unknown hop, we examine if the IPs in the hop belong to the same AS which is either the previous or the next AS.

   $$AS_1 \rightarrow \{AS_1, AS_2\} \rightarrow * \rightarrow AS_1 \tag{4.8}$$

   Example of this pattern match is depicted at 4.8. For both of these cases we report as AS path AS_1.

   This rule also applies to the following cases:

   $$AS_1 \rightarrow * \rightarrow * \rightarrow \{AS_1, AS_2\} \rightarrow AS_2 \tag{4.9}$$

   $$AS_1 \rightarrow \{AS_1, AS_2\} \rightarrow * \rightarrow * \rightarrow AS_2 \tag{4.10}$$

   At 4.9 and 4.10 we have more than one non-responding/unknown hop either before or after the hop with the multiple IPs. However, the IPs of the hop belong to the previous or next AS so for both of the cases we report as AS path $AS_1 \rightarrow AS_2$.

$$AS_1 \rightarrow * \rightarrow \{AS_1, AS_2\} \rightarrow \{AS_1, AS_2\} \rightarrow AS_2 \tag{4.11}$$

$$AS_1 \rightarrow \{AS_1, AS_2\} \rightarrow \{AS_1, AS_2\} \rightarrow * \rightarrow AS_2 \tag{4.12}$$

At 4.11 and 4.12 we have more than one hop with multiple IPs in a row along with a non-responding/unknown hop. If the IPs of these hops resolve to the previous or next AS we report as AS path $AS_1 \rightarrow AS_2$.

(d) In case we have a hop with multiple IPs and one or more non-responding/-unknown hop in a row, we examine it the IPs of the hop belong to the previous or to the next AS.

$$AS_1 \rightarrow \{AS_1, AS_2\} \rightarrow * \rightarrow AS_3 \tag{4.13}$$

$$AS_1 \rightarrow \{AS_1, AS_2\} \rightarrow * \rightarrow * \rightarrow AS_3 \tag{4.14}$$

Example of this pattern match is depicted at 4.13 and 4.14. For both of the two cases we report as AS path $AS_1 \rightarrow AS_* \rightarrow AS_3$.

(e) In case we have hops with multiple IPs in a row and at least one of the hop IPs belong either to the previous or to the next AS, we examine if there is at least one IP in each hop that resolves to the same AS.

$$AS_1 \rightarrow AS_1, AS_2 \rightarrow AS_1, AS_3 \rightarrow AS_2 \tag{4.15}$$

Example of this pattern match is depicted at 4.15 where we report as AS path $AS_1 \rightarrow AS_2$.

As we already mentioned, we created this set of rules in order to map IP-level paths to AS-level paths based on a consistent IP-to-AS transformation. However, even with this approach, some traceroute limitations such as routers responding with a different interface than the inbound, are not addressed. In Section 4.2.1.1 we list the known traceroute limitations and in Section 4.2.1.2 we evaluate the use of an advanced alternative approach to perform the IP-path-to-AS-path mapping.

4.2.1.1 Limitations of the IP-Path-to-AS-Path transformation

Traceroute is one of oldest network diagnostic tools for displaying the route (path) between two hosts. Although the tool has been used for years by network operators and researchers, there are some well-known limitations which we should take into account when using it. First of all, the tool does not discover paths at the router level, but at the interface level. Furthermore, when routers do not respond to probes the tool fails to get a response from the specific hop. Finally, the tool in the presence of traffic load balancing may indicate a path that does not exist. In our work, to address the load balancing effect, we employ the Paris traceroute approach [46].

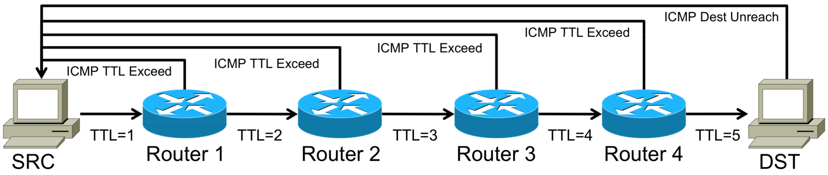


Figure 4.1: Traceroute

A traceroute path example is depicted at Fig: 4.1. The path consists of four routers between the SRC and DST hosts. In the beginning, traceroute sends a probe packet with TTL equals to one; the TTL value one will force the first router on the path to respond with an "ICMP TTL exceeded" message. The tool uses this ICMP error message and marks the source address of the packet as the IP of router one. Then, it increases the TTL by one to seek a response from the next router on the path. The probing stops either when the DST address responds (with an "ICMP destination unreachable" message) or in the case of multiple hops in a row not responding to the probes, indicating filtering of the probes in-path. Note that the default number for packet probes sent for per-hop discovery are 3 (thus the multiple responses per hop).

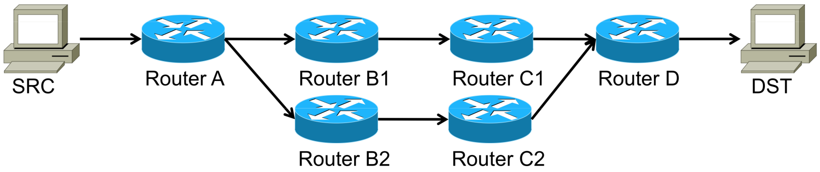


Figure 4.2: Traceroute Equal paths

As we already mentioned we used paris traceroute to address the load balancing effect in the path. However, even with the use of paris traceroute we found cases where packet responses in a hop were generated with different source IPs. Interestingly, there were cases where this set of different IPs after IP-to-AS mapping results to a set of different ASes. Moreover, an example of a load balancing using equal paths is depicted in Fig: 4.2. In the Figure we observe that the router

Address	Administrative AS
IP_{1}	AS_{1}
IP_{2}	AS_{4}
IP_{3}	AS_{2}
IP_{4}	AS_{3}

Table 4.1: Mapping between IPs and ASes (used for MAP-IT third-party address case)

A performs load balancing by splitting the traffic sent toward the DST across two different equal-cost paths.

Although the load-balancing behavior is absolutely reasonable and can explain the difference in the responding interfaces, the RFC1812 dictates that the source address of an ICMP error packet must be the same with the corresponded outgoing interface of the ICMP reply, and not the address of the interface on which the packet triggering the error was received. The authors of [68] found that traceroute output can lead to inaccurately reconstruct the route by overestimating the load balancers along the path toward the destination.

Finally, we remind to the reader that the traceroute measurements are subject to the constraints of the routers they visit. Possible constraints can be packet dropping, silently forwarding packets without altering and decreasing the TTL but also modifying the TTL in unpredictable ways [23].

4.2.1.2 Alternative Approach: MAP-IT

As an alternative and more sophisticated approach to transform the IP-level paths to AS-level paths we used the Multipass Accurate Passive Inferences from Traceroute (MAP-IT) algorithm [69]. The algorithm focuses on inferring the exact interface addresses between the point-to-point inter-AS links. Following the inference of the point-to-point links across AS borders, we can discover and extract the exact AS-level paths.

In our work, the MAP-IT algorithm can help us remove false positive and false negative results in the process of transformation of the IP-level path to the corresponding AS-level path. We distinguish two cases where MAP-IT overcomes and improves our IP-path-to-AS-path transformation.

1. **The router responds with a third-party address**

   As we already described in section 4.2.1.1 the traceroute tool discovers the interfaces of the routers and not the routers themselves. Due to this limitation a traceroute path may include an IP response from a router interface that uses a third-party address. One of the improvements that the MAP-IT algorithm can offer to our work is the case where a router may respond with an address from a third-party AS that has an active peering connection with the expected AS. We will provide an example of a path where our IP-path-to-AS-path methodology falsely identifies such a third-party AS in the path, while MAP-IT filters out this AS from the path.

An example of an IP-level path extracted from a traceroute is depicted at 4.16. The path consists of four IPs which respectively are the interfaces of four routers of the path. Using the IP-to-AS mapping we extract the mapping between IPs and ASes as the Table 4.1 depicts.

$$Traceroute: \ IP_1 \rightarrow IP_2 \rightarrow IP_3 \rightarrow IP_4 \tag{4.16}$$

We consider as the true level AS-path the path that 4.17 shows.

$$True\ AS-level\ path: \ AS_1 \rightarrow AS_1 \rightarrow AS_2 \rightarrow AS_3 \tag{4.17}$$

With the naive approach of IP-path-to-AS-path, we map the hop of the IP_2 to AS_4 using the IP-to-AS dataset. However, we wrongly infer the AS_4 in the path due to the fact that the router in the second hop responded with a third-party interface IP that belongs to AS_4 as 4.18 shows.

$$\begin{gathered} IP-to-AS\ naive\ path: \ AS_1 \rightarrow AS_4 \rightarrow AS_2 \rightarrow AS_3 \\ (IP_2\ belongs\ to\ third\ party\ AS_4) \end{gathered} \tag{4.18}$$

On the other hand, the MAP-IT algorithm detects that the IP_2 on the second hop is from an interface of a router that belongs to AS_1 (4.19) and not to AS_4 as we infer with the naive IP-to-AS.

$$MAP-IT\ path: \ AS_1 \rightarrow AS_1 \rightarrow AS_2 \rightarrow AS_3 \tag{4.19}$$

2. **Hidden border**:

   The second case where MAP-IT algorithm improves our naive IP-to-AS approach is when a hidden border router exists between a pair of ASes. In this case the naive IP-to-AS approach reports false negatives as it misses to detect the AS.

   An example of an IP-level path extracted from a traceroute is depicted at 4.20. Using the IP-to-AS dataset, the corresponding mapping between IPs and ASes is shown in Table 4.2.

$$Traceroute: \ IP_1 \rightarrow IP_2 \tag{4.20}$$

Address	Administrative AS
IP_{1}	AS_{1}
IP_{2}	AS_{1}

Table 4.2: Mapping between IPs and ASes (used for MAP-IT hidden border address case)

We consider as the true level AS-path the path that 4.17 shows.

$$True\ AS - level\ path:\ AS_1 \rightarrow AS_2 \tag{4.21}$$

Using the naive IP-to-AS approach we calculate as AS path the one of 4.22. In this case we fail to discover the AS_2 in the path.

$$\begin{gathered} IP - to - AS\ naive\ path:\ AS_1 \rightarrow AS_1 \\ (since\ IP_2\ belongs\ to\ AS_1) \end{gathered} \tag{4.22}$$

However, again the MAP-IT algorithm detects that the IP_1 on the second hop is from an interface of a router that belong to AS_2 as 4.23 depicts and not to AS_1 as we infer with the naive IP-to-AS approach.

$$MAP - IT\ path:\ AS_1 \rightarrow AS_2 \tag{4.23}$$

To evaluate and compare our IP-path-to-AS-path approach against the MAP-IT approach we created the Algorithm: 1. We discovered that the MAP-IT performs worse in terms of mapping IP pairs to AS-level pairs in with respect to coverage and requires a lot of manual tuning to achieve a decent performance based on the input data. Moreover, the MAP-IT algorithm needs as input hundreds of traceroute results to efficiently detect the point-to-point inter AS-links. It also takes as input lists of IXP prefixes, AS-to-organization and IP-to-AS mappings. These factors make difficult the application of MAP-IT on the old collected data, since there are no publicly available data to evaluate if it is worth using it by validating its inferences. Therefore, in this work, we used only our IP-path-to-AS path approach. However, we plan in the future to create a hybrid approach by combining MAP-IT with our methodology.

4.3 Eyeball-to-Eyeball Connectivity Matrix

The IXP Country Jedi tool (see Section 2.6) performs on a monthly basis, full-mesh probe-to-probe measurements between all the probe-hosting ASes of a given

Algorithm 1 Naive IP-path-to-AS-path Mapping vs MAP-IT

```
procedure PathComparison(traceroute_hops)//The traceroute hop results
    match_true = 0 //MAP-IT and Naive IP-path-to-AS-path agree
    match_false = 0 //MAP-IT and Naive IP-path-to-AS-path disagree

    for hop in traceroute_hops do
        ips_list_curr = get_ips(hop)//Get the IPs from the hop
        ips_list_next = get_ips(hop + 1)
        as_set_curr = get_naive_as_set(ips_list_current)
        as_set_next = get_naive_as_set(ips_list_next)

        for ip in ips_list_next do
            mapit_as_pair = get_mapit_as_pair(ip)

            if mapit_as_pair is None then
                match_false += 1 //We can not map this pair of border IPs
                continue

            mapit_asx = mapit_as_pair[0]
            mapit_asy = mapit_as_pair[1]

            if mapit_asx in as_set_curr and mapit_asy in as_set_next  then
                match_true += 1
            else if mapit_asy in as_set_curr and mapit_asx in as_set_next  then
                match_true += 1
            else
                match_false += 1

    return [match_true, match_false] //Number of matched and unmatched
```

country. Then, the tool analyzes the traceroute data and creates a set of visualizations to provide insights on the data. One of this visualizations is the peering matrix of all the measured ASes in a country. The peering matrix aims to provide useful insights to network operators in order to explore how their network reach other networks in the same country. The generated data are also useful to the research community as they express a snapshot of the connectivity of a country's AS ecosystem.

Although the IXP Country Jedi peering matrix can provide useful insights on the connectivity between ASes on a country level, it can not provide any insights on the connectivity between user populations. To provide such insights we used the APNIC user population per ASN estimates in order to create a similar peering matrix including only the eyeball networks for a given country, while weighing the contributions of each eyeball pair in terms of user connection paths. The methodology that was used to construct this eyeball to eyeball peering matrix is described in Section 4.3.1. In this thesis, we focus on three properties that can characterize the connectivity of the user populations in a country. We explore the fraction of user-to-user connection paths that stayed local in the country or go out of the country, the fraction of the paths that cross an IXP or not but also the fraction of the paths where the connection between two eyeball ASes was direct (between the 2 peering eyeballs) or indirect (over an intermediate e.g., transit provider).

4.3.1 Methodology

The first step in order to construct the Eyeball-to-Eyeball connectivity matrix relies on the identification of the eyeball networks in a given country. The identification process as we already described in section 4.1 is based on the estimated fraction of users that are served from a network in a country. These user population per ASN estimates are derived from the APNIC dataset (see section 2.2). Furthermore, the process to characterize a network as eyeball or non-eyeball is described in section 4.1.

In Chapter 3 we introduced a framework responsible to collect and store the monthly IXP Country Jedi tool measurement data. Using the API of this framework we retrieve the already enriched traceroute results for all networks that we characterized as eyeballs. We support both IPv4 and IPv6 address family protocols and we group the results using them. The retrieved traceroute results are already enriched with geolocation and IP-to-AS info for each hop IP but also analyzed with respect to IXP crossings. The framework uses the OpenIPMap (see Section 2.3) tool to geolocate the IP addresses, the IP-to-AS mapping (see Section 2.4) to map an IP to an AS and data from the PeeringDB to identify the IXP crossings (see Section 2.5) respectively.

The collected data between a pair of ASes (i.e., AS_X, AS_Y) may include more than one traceroute results. This is possible due to the fact that the IXP Country Jedi tool selects at maximum two probes per AS in a country to perform probe to probe traceroute measurements. So, it is reasonable for the number of traceroute results between two ASes to depend on the number of probes that were available in these two networks at the times of the measurement. We list all these different cases in detail in Section 4.4.1. In case of more than one traceroute results for the path $AS_X \rightarrow AS_Y$, we examine all the available traceroute paths for this pair and inspect if the same property (i.e., path goes out of country) holds to all the paths from $AS_X \rightarrow AS_Y$ before we make any claims. In case of differences between the paths for a given property we characterize the connectivity from $AS_X \rightarrow AS_Y$ pair as inconsistent (with respect to that property). Using this approach, and assuming that the selected probes can capture the full path diversity of an AS and accurately represent the local –user– market, we can estimate the percentage of user-to-user connections that stay in (or go out of) the country, cross an IXP or not as well as whether they are direct or not.

To characterize the AS path from the AS_X towards to the AS_Y as direct or indirect we use the IP-path-to-AS-path methodology that we described in Section 4.2. The transformation of each traceroute result to an AS-level path allows us to detect any intermediate ASes in the path. If the path includes at least one intermediate we characterize the connection as indirect or else we consider it as direct. In the case of any IXPS in the path between the source and destination ASes we still consider the path as direct, since IXPs provide a direct layer-2 switching fabric between two peering networks. Finally, in case of differences between the traceroute results we consider as inconsistent the direct/indirect property for the $AS_X \rightarrow AS_Y$ connection, similar to the other properties.

Our main objective was to create a similar tabular structure as the IXP Country Jedi peering matrix focusing only on the eyeball networks. In this new matrix, similar to the IXP Country Jedi, the rows and the columns correspond to different eyeball networks (AS), which are used as sources and destinations respectively. However, unlike the IXP Country Jedi peering matrix where all boxes are equally sized, in the new matrix the resulting boxes are sized according to the APNIC estimations of the coverage of Internet users per AS (see Section 2.2).

With such a structure, we can calculate metrics related to the user population that interconnects via (direct or not) paths within or outside a country. The basic metric we use is represented by the area of the displayed boxes, which corresponds to a product of coverage percentages. By dividing such areas with the total area, we can calculate percentages of user-to-user connections with certain characteristics. More specifically, we calculate the percentages of user-to-user connections using the Eq. 4.24.

$$user - to - user(AS_X \rightarrow AS_Y) = \frac{ASX_X * ASX_Y}{Total} \tag{4.24}$$

where AS_X, AS_Y are the percentages of the user population per ASN estimates from APNIC and Total is the square area of the matrix.

Moreover, the colors of the boxes correspond to different types of interconnectivity information, such as out-of-country, in-country, IXP crossing or not while red borders mark indirect AS-level connections. An example of this AS-to-AS matrix for Canada for April 2017 is presented in Section 4.3.2. In Section 4.3.3, we take a look at how properties such as paths in/out of country and IXP crossings are evolving across time per country.

4.3.2 Eyeball-to-Eyeball Matrix in Space

In this section we evaluate the methodology of the previous Section 4.3.1 and present the Eyeball-to-Eyeball matrix for eyeball networks. Using the data from the IXP Country Jedi we can construct this matrix for about 114 countries around the globe across time, starting from September 2015.

An example of the Eyeball-to-Eyeball matrix for Canada in April 2017 is depicted in Fig. 4.3. In this example we use as a threshold to consider a network as eyeball the value of 1%, since this allows us to study the country-level Canada eyeball ecosystem at a fine granularity. In this figure we evaluate the Canada ecosystem against two different properties. We examine the paths that go in/out of country and the direct/indirect connectivity between the ASes. The colors in the figure are mapped as follows. The in-country paths are colored as green and the out-of-country paths as orange. The eyeball networks without RIPE Atlas probe coverage are colored as light grey and the in-/out-of-country inconsistencies between probes are black. Finally, the red borders of the boxes mark indirect AS-level eyeball connections and the blue borders of the boxes mark direct/indirect inconsistencies.

As the figure depiction is based on a threshold of 1% to filter the eyeball networks, 16 ASes were marked as eyeballs. By aggregating the user population estimates of these 16 ASes we cover the 84.5% of Internet users in Canada. Moreover, the cumulative area of user connections, seemingly served via in-country paths, is 47.1%. Only 9% are indirect (with ≥ 1 intermediaries) on the AS-level. 3.1% leave the country. Moreover, 18.1% suffer from lack of RIPE Atlas probe coverage, and only 3.2% exhibit inconsistencies with respect to achieving consensus on whether traffic actually leaves the country or remains within it. Finally, the 28.6% is the rest of the area not examined by the AS-to-AS matrix. We note that the asymmetries displayed in the matrix may stem either from inference errors or from interesting ISP policy differentiation per traffic direction, something we plan to investigate further in future work.

Finally, we have created and made publicly available an online version of this AS-to-AS matrix visualization for the eyeball networks of about 114 countries. In the online visualization, the user can specify the threshold of the networks that can be consider as eyeball. Moreover, the user can select the date of the snapshot (month/year) but also the address family protocol (IPv4/IPv6). At this moment, the available properties that can be explored through this visualization in order to characterize the user-to-user connections is the in/out-of-country paths, the IXP

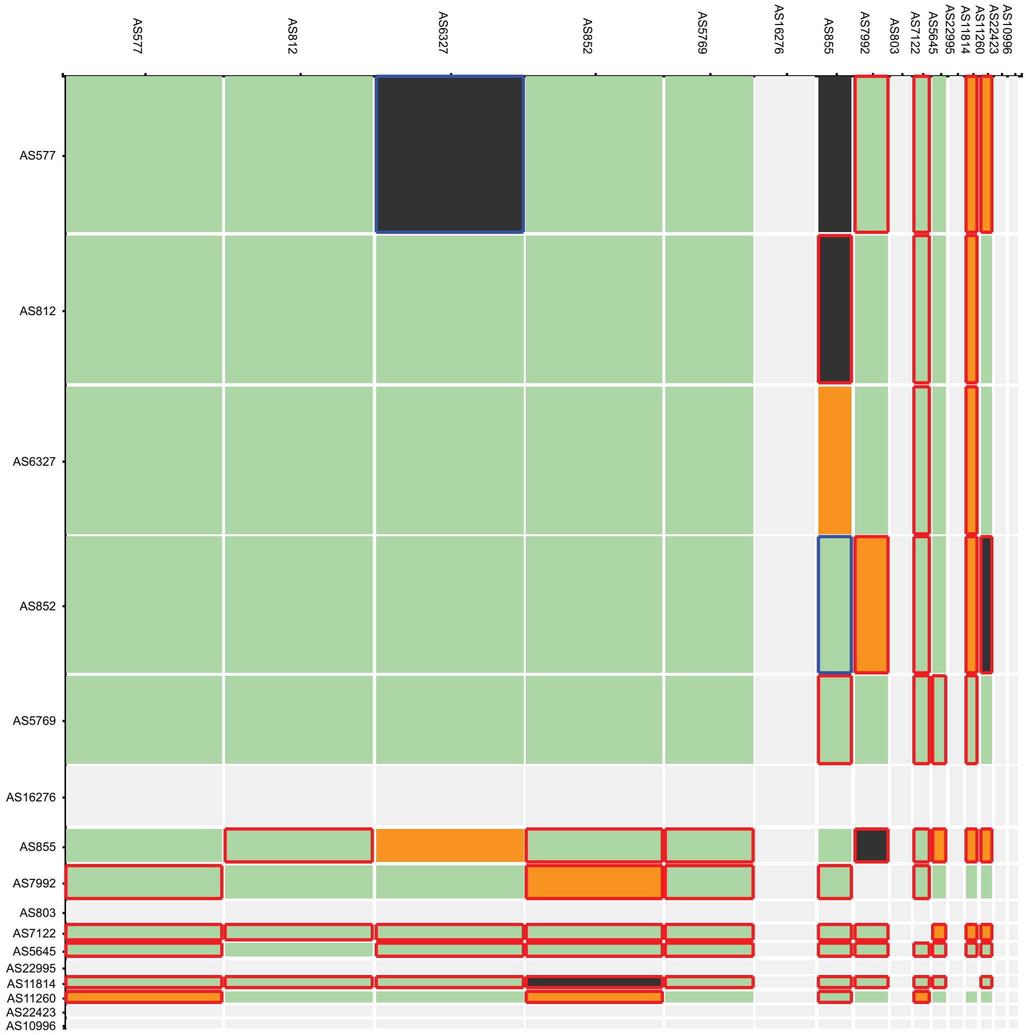


Figure 4.3: Snapshot of eyeball-to-eyeball matrix for Canada (generated on 2017-04-01).

crossing or not and the direct/indirect paths between a pair of ASes. The online AS-to-AS matrix visualization along with statistics are provided at the following website: `https://www.eyeball-jedi.net`.

4.3.3 Eyeball-to-Eyeball Matrix over Time

In the previous section we presented an example of the Eyeball-to-Eyeball matrix for Canada for April 2017. Although this matrix can provide useful insights in order to characterize the connectivity of users inside a country, it only reveals and visualizes data of a monthly snapshot. To get useful insights of how the connectivity between the eyeball ASes of a country evolves across time, we used the monthly generated statistics of the Eyeball-to-Eyeball matrix. In this section we examine how metrics such as the in/out of country and IXP crossing paths evolve over time in a set of countries. As a starting point to explore such kinds of connectivity evolution, we set September 2015 (first run of the IXP Country Jedi) and we provide data until February 2018. We remind that to consider a network as eyeball we used as threshold 1% of the total user population.

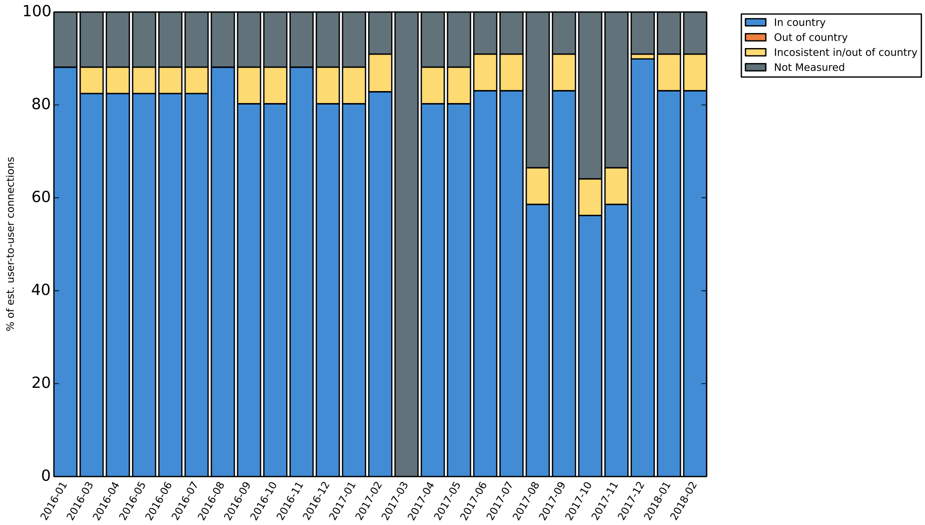


Figure 4.4: Greece IPv4 in/out of country

In/out-of-country user-to-user connections over time

At first, we examine the evolution of the in/out-of-country paths between eyeball networks across time. We compare the fraction of estimated user-to-user connections in IPv4 against the fraction in IPv6. Fig. 4.4 depicts the evolution of the fraction of estimated user-to-user connections w.r.t. the in/out-of-country metric for Greece in IPv4. The blue color represents the fraction of connections that stay in-country (inside Greece), the orange the out-of-country and the yellow the inconsistent ones. Our first observation is that using the RIPE Atlas platform and a threshold of 1% we can measure across time about 90% of the total user population of Greece. We observe that about 80% of the total user-to-user connections in IPv4 stay local across time. Moreover, we observe that about 5-10% of the user connections appear to be inconsistent over time. We found the root cause of these inconsistencies to be a probe in a large ISP that was connected through tunneling to an AS that operates in United States. The current probe methodology can not filter out probes with such behavior and we plan in the future to exclude them from our measurements. In 2017-03, it is clearly visible that we don't have any data as the the monthly run of the IXP Country Jedi tool failed. Finally, we observe a drop in the coverage for three different months. This may be the result of a probe instability (probe going off-line) in an eyeball AS that hosts only one probe.

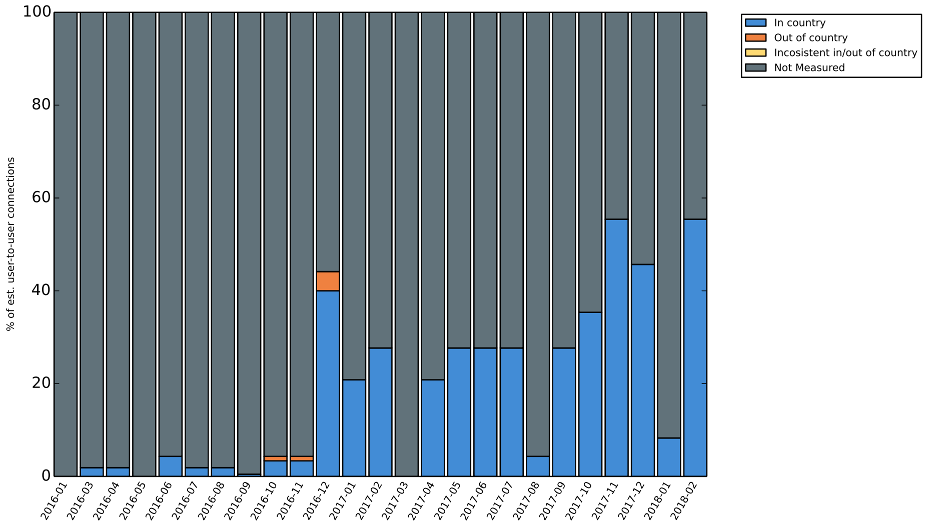


Figure 4.5: Greece IPv6 in/out of country

In Fig. 4.5 we depict the evolution of the fraction of user-to-user connections w.r.t. the in/out of country metric for Greece in IPv6. The colors follow the same pattern that we previously described. The first observation is that the estimated fraction of user population coverage is significantly lower than the one in IPv4. This is a clear evidence of of low IPv6 adoption inside the eyeball networks. In 2016-12, we observe a significant increase in the user population coverage that we measure. Interestingly enough, this chronically matches with the wide launch of the IPv6 protocol to the home user subscribers of the largest Greek ISP (COSMOTE) according to [6] which amounts to ∼45% of the Greek market-share. We observe that also in IPv6 the majority of fraction of user-to-user connections stay inside the country across time. We previously examined the Greek eyeball ecosystem across time and found that almost all user-to-user connections between Greek ASes stay local in country. This is good news for the Greek users as their packets stay in country and not traverse foreign countries that may perform packet inspecting. Also this is an indication of well-connectedness and peering between the ASes of Greece.

Next, we present the Eyeball-to-Eyeball matrix across time for in/out-of-country paths for the United States and Canada. We selected these two countries as in the past, [74] revealed that traceroute paths starting from Canadian ASes targeting Canadian destinations often crossed suspicious Internet monitoring locations in the United States, a phenomenon which they named "boomerang routing".

In Fig. 4.6 we depict the estimated fraction of user-to-user connections for the United States in IPv4 across time. We observe that the % measured user-to-user

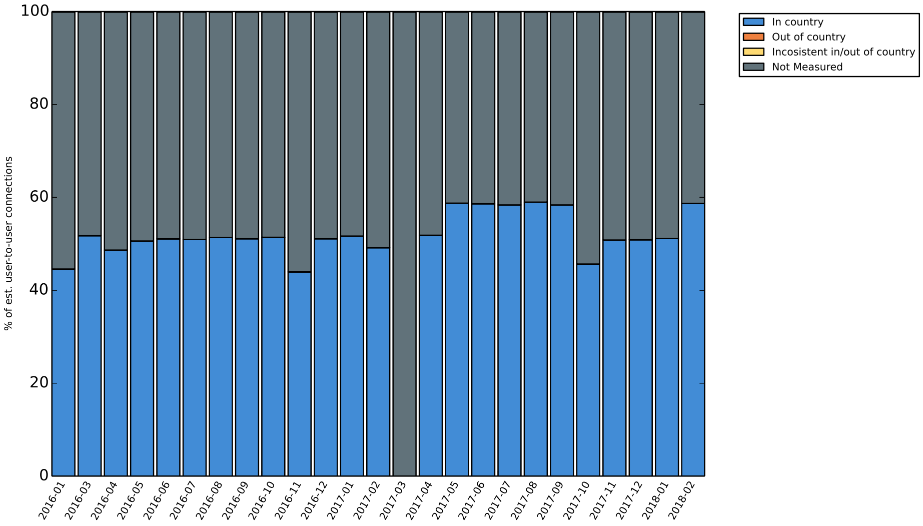


Figure 4.6: United Stated of America IPv4 in/out country

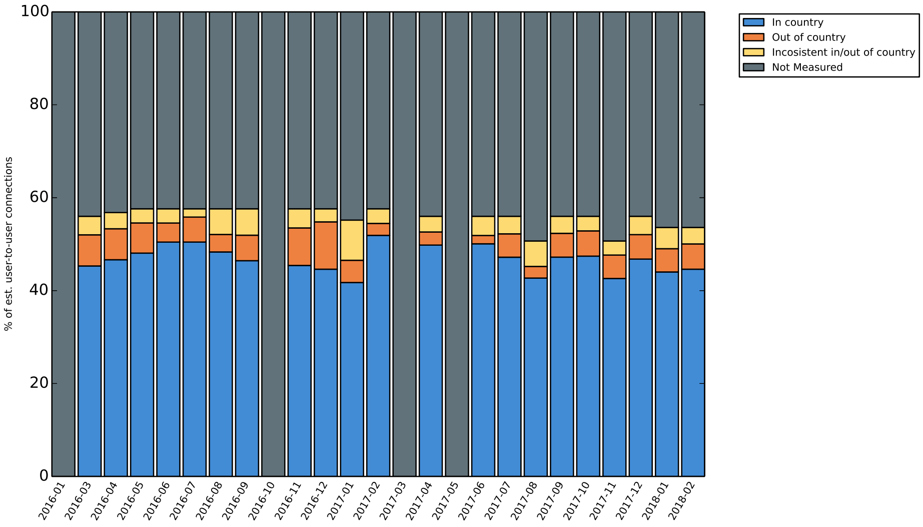


Figure 4.7: Canada IPv4 in/out country

connections range from ∼40% to ∼60% across time. Moreover, we observe that all user-to-user connections stay local and there are no out-of-country or inconsistent connections that we can see. On the other hand, Fig. 4.7 depicts the fraction of user-to-user connections for Canada in IPv4. Interestingly the % measured user-to-user connections is close to the United States findings and is about ∼60%. However, we can clearly observe the existence of out-of-country paths and inconsistencies across time. About ∼45% of the connections stay local (in-country) and the out-of-country vary over time up-and-down from ∼2% to ∼11%. This is an interesting finding in comparison with the findings of the boomerang routing study which found roughly one quarter of the Canada-to-Canada traceroute paths to traverse the United States.

IXP crossings of user-to-user connections over time

Next, we examine the fraction of estimated fraction of user-to-user connections that cross an IXP that operates in the same country (as the eyeball networks) across time but also per address family. Fig. 4.8 depicts the estimated user-to-user connections that cross or not an IXP for Greece in IPv4. The blue color represents the fraction of estimated user-to-user connections that did not traverse an IXP (inside Greece), orange the IXP-crossing connections and the yellow the inconsistent ones. A first observation is that the no-IXP-crossing connections stay almost stable over time to about 35%. Interestingly the IXP-crossing connections range from 20% to 50% across time. We remind that the only known IXP that operates in Greece is the GR-IX [13], so we can safely assume that the IXP-crossing connections are traversing this IXP. Moreover, the inconsistencies between paths of the same AS pair range from 3% to 25% which is an indication of path diversity between the paths from an AS_X against an AS_Y.

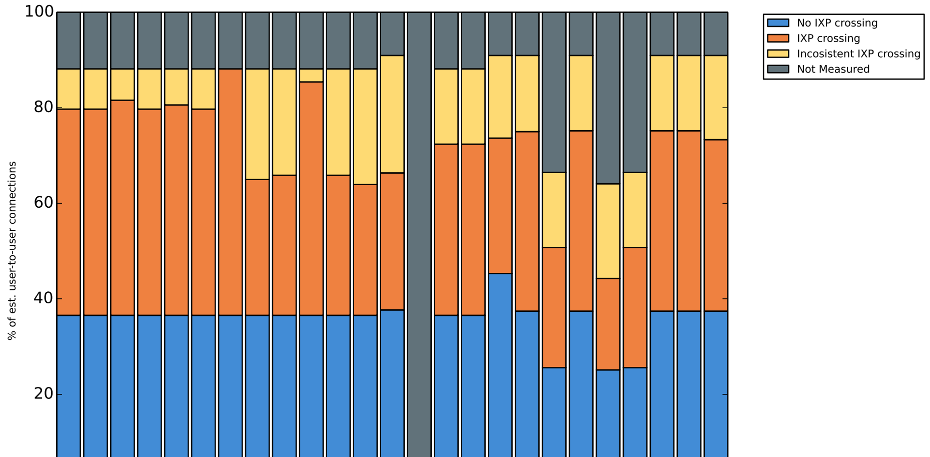


Figure 4.8: Greece IPv4 IXP crossing

In Fig. 4.9 we depict the estimated fraction of user-to-user connections that flow or not through an IXP in IPv6 for Greece. A first observation is that as the IPv6 coverage increases due to the IPv6 deployment, the fraction of IXP-crossing IPv6 paths also increases. This is an indication that networks seem to prefer also the IXP peering model to peer with other networks in IPv6.

Furthermore, we present the Eyeball-to-Eyeball matrix across time for paths

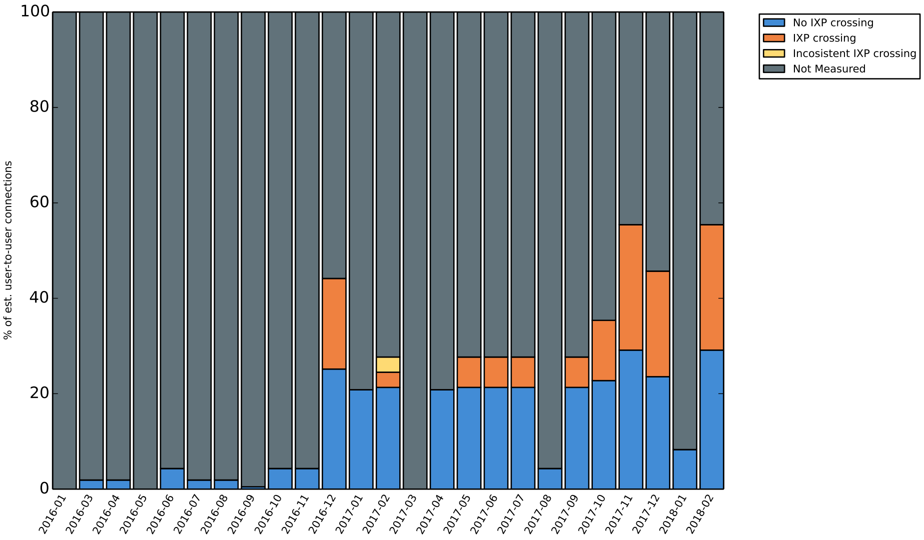


Figure 4.9: Greece IPv6 IXP crossing

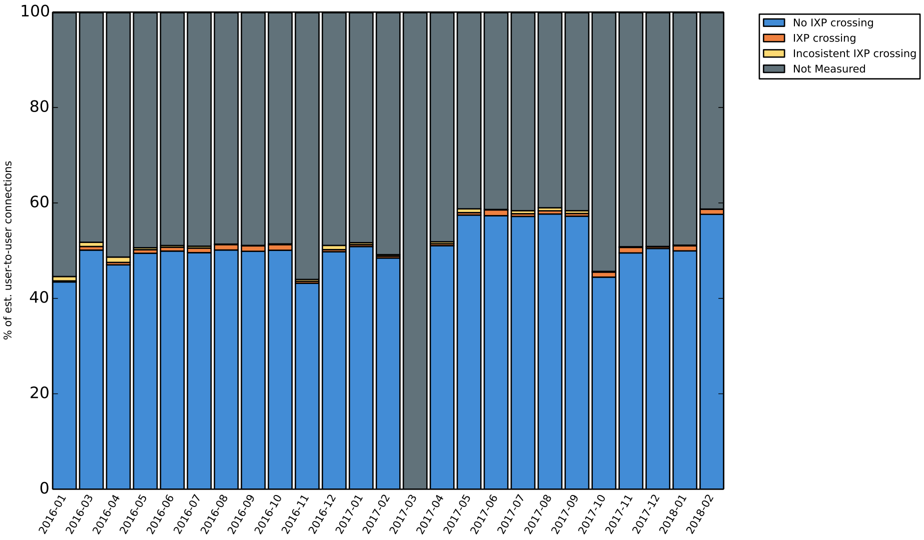


Figure 4.10: United States of America IXP crossing

crossing an IXP or not for the United States, Ireland and Netherlands in IPv4. In Fig. 4.10 we depict the estimated fraction of user-to-user connections in IPv4 for the United States of America. The first clear observation that we can make is that almost all connections are not crossing an IXP, but are setup through other types of –private or public– peering. Moreover, the fraction of no-IXP-crossing paths stays stable over time.

In addition to the United States of America, the eyeball network ecosystem

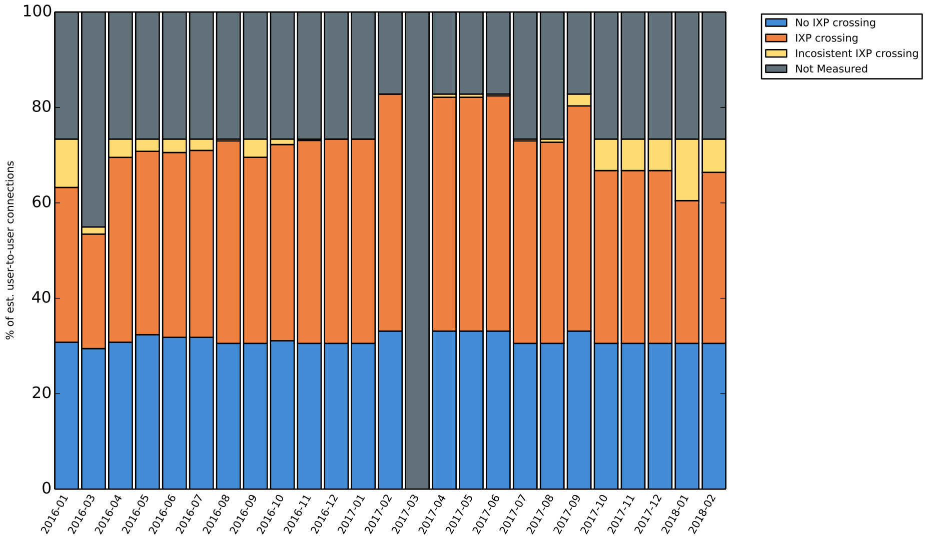


Figure 4.11: Ireland IPv4 IXP crossing

of Ireland and Netherlands depends and uses IXPs, as a significant fraction of the estimated user-to-user connections pass through them. Fig. 4.11 depicts the estimated fraction of user-to-user connections for Ireland in IPv4, traversing –or not– an IXP that operates in the same country. We observe that the higher fraction of connections pass through an IXP across time. The IXP crossing fraction ranges from 20% to 50%, while the no-IXP crossing fraction remains almost stable at about 30%. The high fraction of connections passing through an IXP for Ireland can be explained, as Ireland is the base of operations of the INEX [15] IXP. Moreover, Fig. 4.12 depicts the fraction of estimated user-to-user connections for Netherlands in IPv4. We also observe a high fraction of connections passing through an IXP but in contrast to Ireland, the fraction of IXP -crossing connections do not surpass the fraction of no-IXP-crossing connections. The high fraction of IXP-crossing connections can be attributed to the presence of two of the largest (in terms of member base sizes) IXPs worldwide in Netherlands. More specifically, the AMS-IX [2] and [25] operate there.

Finally, the stability of the results across time, the correspondence of the observed results to facts related to how each country-level eyeball ecosystem operates, and the fact that the IXP Country Jedi does not keep any state between each monthly measurement (allowing for independent repeatable hypothesis tests), provide a good indication of the usefulness of our methodology.

4.4 Transit Betweenness

The Internet is a global network consisting of thousands of interconnected computer networks. Each network applies its own routing strategies and policies [50]

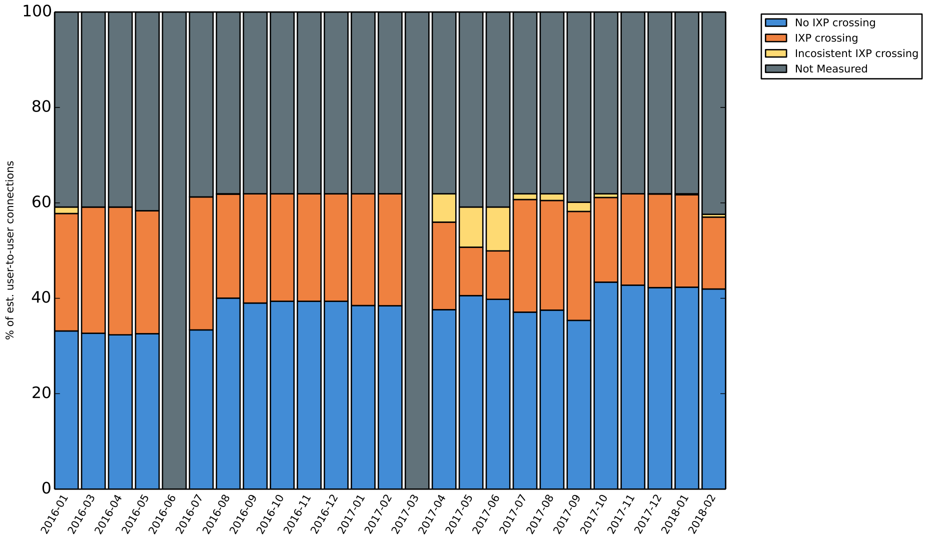


Figure 4.12: Netherlands IXP crossing

which results to a complex set of routing paths, involving several networks. So, a common question that naturally arises in the field of Internet routing is: *which are the intermediate networks that a packet will traverse, when sent from point A to point B?*. Assuming that the packet starts from the network A and heads towards to the network B it is possible to traverse zero (e.g., in case of direct A-B peering) or more (e.g., A's and B's upstream providers) intermediate networks [51], including IXPs. Moreover, a research study [4] revealed that only a small portion of ASes carry the traffic for a disproportionate number of routes on the Internet. Therefore, the answer to the original question can be extracted using two different types of measurements. The first option is to conduct control plane measurements. For this type, researchers rely on BGP data, in order to extract the AS path between two ASes. However, the result of this type of measurement may be incomplete [75]. The second option is to conduct data plane measurements. This type of measurement can be performed using tools such as `ping` and `traceroute`. However, data plane measurements also suffer from limitations and may result to discovery of false or incomplete paths.

In this work, in order to identify the intermediate networks between a pair of ASes we only rely on results generated by data plane measurements based on the abundant data plane data we have at our disposal. More specifically, we use the proposed framework (see Section 3) to extract and analyze traceroute data. Moreover, we combine these data with the APNIC population estimates that we described at Section 2.2. In contrast to other studies we do not only identify the networks between a pair of ASes in a country, but we also rank these intermediate networks based on the fraction of user-to-user paths that pass trough them. Our goal is to provide insights of how many user-to-user connections out of a total set of possible user-to-user connections would pass through each network inside a country. This can help researchers study metrics such as censorship, path diversity and routing resilience on a country level. To measure this "flow" of user-to-user connections we introduce the *transit betweenness* custom metric; we describe the methodology for its inference/calculation in Section 4.4.1. The possible values of transit betweenness for a network (AS/IXP) can range from 0% to 100%.

4.4.1 Methodology to infer transit betweenness of a network

As we mentioned earlier (see section 2) we use the generated data of the IXP Country Jedi prototype tool. Due to the fact that the tool does not keep any state between each run, the number of ASes and number of probes vary between each run. Studying a dataset where vantage points and number of measurements change over time is a challenging process. To overcome these challenges and measure the transit betweenness inside a country with a consistent and reliable approach we do the following.

As a first step, using the proposed framework (see Section 3) we retrieve all the available traceroute data of a given month. We filter out the source and destination ASes where the APNIC estimates fail to provide any information. By studying the paths between ASes that are covered by the APNIC dataset we can calculate the fraction of user-to-user paths that flow through each one of the intermediate networks.

Next, we analyze the traceroute paths between a pair of ASes in a country. To identify the intermediate networks, we transform each traceroute from an IP-level path to an AS-level path using the IP-path-to-AS-path transformation that we described in Section 4.2. Assuming a $[AS_X, AS_Y]$ pair of source/destination ASes, we consider two different approaches in order to calculate the transit betweenness of any intermediate networks between them; the *symmetric* and the *asymmetric* approach (described in detail in Section 4.4.1.1 and 4.4.1.2 respectively).

The *symmetric* approach relies on the assumption that between the $probe_1$ and $probe_3$ which belong to AS_X and AS_Y respectively, any observed intermediate network may "control" the bidirectional communication between the two ASes, irrespectively of whether it is present only in the direction $AS_X \rightarrow AS_Y$ or $AS_Y \rightarrow AS_X$ (on the respective probe-to-probe paths). In contrast, the *asymmetric* approach takes into account the presence of intermediate networks *separately per direction.* Both these approaches are useful to understand whether the direction of the communication plays an important role in what intermediate networks are observable, depending of course on whether the probe-to-probe paths are symmetric or not.

Before proceeding to the description of the symmetric approach in section 4.4.1.1 and the asymmetric approach in section 4.4.1.2 we need to explain how we weigh each traceroute result between a pair of ASes, in order to extract information on the fractions of user-to-user connections that flow on this path. We remind the reader that the IXP Country Jedi tool selects at maximum two probes per AS and performs probe to probe measurements between all ASes in the country. In this thesis, we make the working assumption that by using two probes we can expose the full path diversity of an AS (w.r.t. paths towards other ASes in the country). To measure the paths between two ASes we perform traceroute measurements from each of the selected probes of the source AS towards the selected probes of the destination AS. Using max two probes per AS generates up to 8 traceroute results (considering the direction of the traceroute) between the pair of $[AS_X, AS_Y]$. We assume that these 8 results expose the full transit betweenness of an intermediate network between a pair of ASes. However, using two probes per AS is not always possible as there are many ASes hosting a single probe. Additionally, it is always possible for a probe to fail during the measurement process. We distinguish three different cases which are related to the number of probes that were used to measure the path. We introduce a penalty for associated missing traceroute paths on which we refer as the "Unknown betweenness".

Measurement cases based on number of probes

- **Two probes on both ASes**

  In the first case (Fig. 4.13), we select and use two probes ($\{1,2\}$) from the AS_X and two probes ($\{3,4\}$) from the AS_Y and perform full mesh probe-to-probe traceroute measurements.

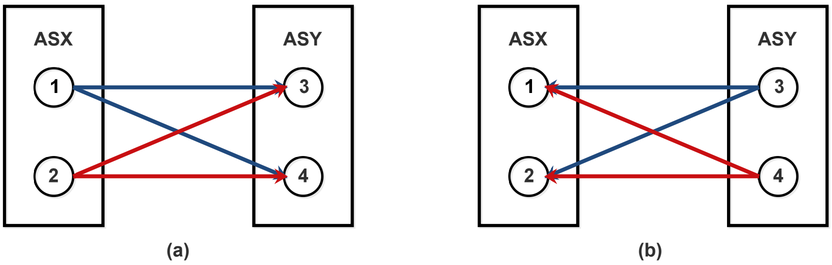


Figure 4.13: Traceroute measurements between probes {1, 2} of ASX and probes {3, 4} of ASY.

  Fig. 4.13 (a) depicts the four traceroute measurements from probes of AS_X towards the probes of AS_Y. The initiated traceroute measurements from probe $\{1\}$ are marked with a blue arrow while the measurements from $\{2\}$ are marked with a red arrow.

  Fig.4.13 (b) depicts the four traceroute measurement from probes of AS_Y towards the probes of the AS_X. The initiated traceroute measurements from probe $\{3\}$ are marked with a blue arrow while the measurements from $\{4\}$ are marked with a red arrow.

We end up with four traceroute results for the path of $AS_X \rightarrow AS_Y$ and four traceroute results for the path of $AS_X \leftarrow AS_Y$. In total eight traceroute results have been produced for this pair of ASes.

- **Two probes on ASX and one probe in ASY**

  In the second case as the Fig. 4.14 depicts, we select and use two probes ($\{1, 2\}$) from AS_X and a single probe ($\{3\}$) from AS_Y in order to perform full mesh probe-to-probe traceroute measurements.

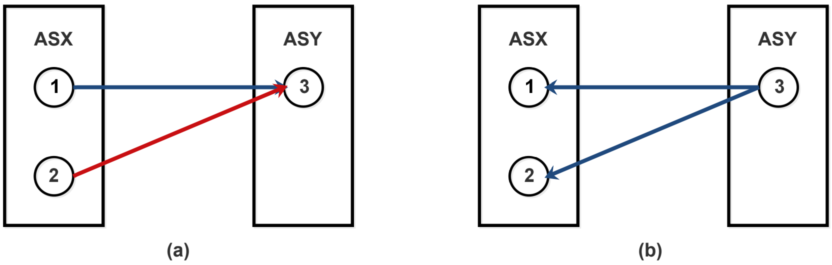


Figure 4.14: Traceroute measurements between probes {1, 2} of ASX and probe {3} of ASY.

  Fig.4.14 (a) depicts two traceroute measurements from probes of the AS_X towards the single probe of AS_Y. The initiated traceroute measurement from probe $\{1\}$ is marked with the blue arrow, while the measurement from $\{2\}$ is marked with the red arrow.

  Fig.4.14 (b) depicts two traceroute measurements from the single probe of the AS_Y towards the two probe of AS_X. The initiated traceroute measurements from probe $\{3\}$ are marked with the blue arrow.

  We end up with two traceroute results for the path of $AS_X \rightarrow AS_Y$ and two traceroute results for the path of $AS_X \leftarrow AS_Y$. In total four traceroute results have been produced for this pair of ASes.

- **One probe to both ASes**

  In the last case as the Fig. 4.15 depicts, we use one probe ($\{1\}$) from both ASes in order to perform full mesh probe-to-probe traceroute measurements.

  Fig.4.14 (a) depicts the single traceroute measurement from the probe of the AS_X towards the single probe of AS_Y. The initiated traceroute measurement from probe $\{1\}$ is marked with the blue arrow.

  Fig.4.14 (b) depicts the single traceroute measurement from the probe of the AS_Y towards the probe of AS_X. The initiated traceroute measurement from probe $\{3\}$ is marked with the blue arrow.

  In this last case we end up with one traceroute result for the path of $AS_X \rightarrow AS_Y$ and one traceroute result for the path of $AS_X \leftarrow AS_Y$. In total two traceroute results have been produced for this pair of ASes.

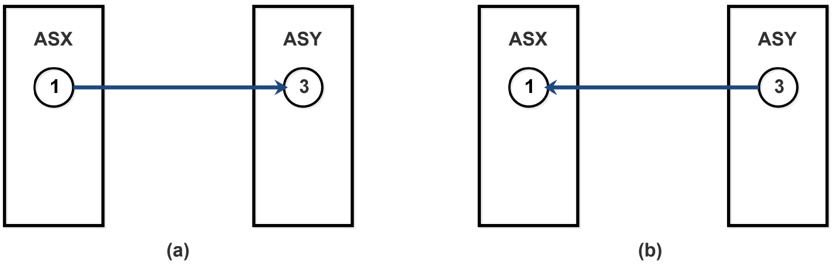


Figure 4.15: Traceroute measurements between probe {1} of ASX and probe {3} of ASY.

4.4.1.1 Symmetric Path Approach

In the symmetric approach, to calculate the transit betweenness of an intermediate network, we assume that between the $probe_1$ and $probe_3$ which belong to the AS_X and AS_Y respectively, the directions $probe_1 \rightarrow probe_3$ and $probe_1 \leftarrow probe_3$ do not affect the presence of intermediate ASes w.r.t. the bidirectional communication. More specifically, if we identify a network as intermediate in the path from $\{AS_X - probe_1\} \rightarrow \{AS_Y - probe_3\}$ we consider that this network is also intermediate in the path of $\{AS_X - probe_1\} \leftarrow \{AS_Y - probe_3\}$.

To calculate the transit betweenness of an intermediate AS using this approach we need to match the traceroute paths between each pair of probes. We remind that in this work we assume that the maximum path diversity between a pair of $[AS_X, AS_Y]$ ASes can be measured using two probes per AS. An example of a full mesh probe to probe measurement is depicted in Fig. 4.13. The figure depicts 8 traceroute paths (arrows) in total. We group these results based on the following 4 probe pairs $[1,3], [1,4], [2,3], [2,4]$.

We then weigh all the probe pairs according to the number of traceroute paths they include. If a probe pair $[probe_1, probe_3]$ includes two traceroute paths (i.e. $\{AS_X - probe_1\} \rightarrow \{AS_Y - probe_3\}$ and $\{AS_X - probe_1\} \leftarrow \{AS_Y - probe_3\}$) we weigh it with 0.25 (2/8 paths). If it consists only of one traceroute path we weight it with 0.125 (1/8 paths).

As 4.25 shows, the Symmetric Transit Betweenness (STB) of an intermediate AS_Z is highly related to the number of probe pairs on which the AS was identified.

$$STB_{ASX \leftrightarrow ASY}(AS_W) = \sum_{probe\ pairs} Weight(probe\ pair) * (EST_{ASX} * EST_{ASY}) \tag{4.25}$$

where EST_{ASX}, EST_{ASY} are the percentages of the user population estimates from APNIC for AS_X and AS_Y respectively, and the Weight of probe pair can be either 0.25 or 0.125, as described in 4.26.

$$Weight(probe\ pair) = \begin{cases} 0.125, & \text{if 1 } traceroute\ result \\ 0.25, & \text{if 2 } traceroute\ results \end{cases} \tag{4.26}$$

Finally, we add up the weight of all probe pairs; if it is less than 1, we apply 4.27 to calculate the Unknown Symmetric Transit Betweenness ($USTB$) between $AS_X \leftrightarrow AS_Y$.

$$USTB(AS_X \leftrightarrow AS_Y) = (1 - \sum_{probe\ pairs} Weight(probe\ pair)) * (EST_{ASX} * EST_{ASY}) \tag{4.27}$$

where EST_{ASX}, EST_{ASY} are the percentages of the user population estimates from APNIC for AS_X and AS_Y respectively, and Weight is as described above.

To calculate the corresponding aggregate metrics we define the following:

$$STB_{all}(AS_W) = \sum_{ASX \leftrightarrow ASY\ where\ AS_W\ is\ present} STB_{ASX \leftrightarrow ASY}(AS_W) \tag{4.28}$$

$$USTB_{all} = \sum_{ASX \leftrightarrow ASY\ where\ paths\ are\ missing} USTB(AS_X \leftrightarrow AS_Y) \tag{4.29}$$

4.4.1.2 Asymmetric Path Approach

In the asymmetric approach, to calculate the transit betweenness of an intermediate network, we assume that the traceroute paths between the probes of the AS_X and AS_Y are independent. Moreover, using the assumption that we need two probes per AS to expose the path diversity we need 8 traceroute results to give the maximum transit betweenness to an intermediate AS. As a result we equally rate each of the eight traceroute results with a weight of 0.125.

Then, for each intermediate network in the traceroute path we apply 4.30 to calculate the Asymmetric Transit Betweenness (ATB) of it.

$$ATB_{ASX \leftrightarrow ASY}(AS_W) = \sum_{found\ ASX \leftrightarrow ASY\ paths} 0.125 * (EST_{ASX} * EST_{ASY}) \tag{4.30}$$

where EST_{ASX}, EST_{ASY} are the percentages of the user population estimates from APNIC for AS_X and AS_Y respectively.

Finally, in cases where we have less than 8 traceroute results available for the pair of ASes we apply 4.31 in order to calculate the Unknown Asymmetric Transit Betweenness ($UATB$) for this pair of ASes.

$$UATB(AS_X \leftrightarrow AS_Y) = (8 - |found\ ASX \leftrightarrow ASY\ paths|) * 0.125 * (EST_{ASX} * EST_{ASY}) \tag{4.31}$$

where EST_{ASX}, EST_{ASY} are the percentages of the user population estimates from APNIC for AS_X and AS_Y respectively, and *found* is the number of available traceroute paths for the $[AS_X, AS_Y]$ pair of ASes.

To calculate the corresponding aggregate metrics we define the following:

$$ATB_{all}(AS_W) = \sum_{ASX \leftrightarrow ASY\ where\ AS_W\ is\ present} ATB_{ASX \leftrightarrow ASY}(AS_W) \tag{4.32}$$

$$UATB_{all} = \sum_{ASX \leftrightarrow ASY\ where\ paths\ are\ missing} UATB(AS_X \leftrightarrow AS_Y) \tag{4.33}$$

4.4.2 Transit Betweenness over Time

In this section, we apply the methodology of Section 4.4.1 and present the Transit Betweenness using the symmetric and asymmetric approach across time for the United States and Netherlands. We use the data of the IXP Country Jedi to construct this Transit Betweenness evolution across time for about 114 countries around the globe, starting from March 2016[1]. Moreover, we support both address family protocols (v4/v6). It is worth noting that the transit betweenness metric for all ASes doesn't sum up to 100% since we can have multiple intermediate networks in the path between two eyeballs; these networks will all receive the same transit betweenness value with respect to this eyeball pair.

United States is the first country that we examine. Fig. 4.16 depicts the transit betweenness for all networks that appeared at least one time in the top 5 networks of a month using the symmetric approach in IPv4. We can see that the network with the highest betweenness across time is AS7843 (Time Warner Cable Internet LLC) with a betweenness of ∼8% of the total user-to-user connections. The "not-coverage" from RIPE Atlas betweenness is ∼17% and the Unknown betweenness is ∼19%. Moreover, Fig. 4.17 depicts the transit betweenness using the asymmetric approach in IPv4. As we can see the asymmetric transit betweenness of the networks is decreased in contrast to the symmetric approach. This is a clear indication of routing asymmetries for the eyeball-to-eyeball communications within the country; an intermediate network may be present only in one direction of a path, resulting to lower asymmetric betweeness than the symmetric case where we consider its presence irrespectively of the direction. Moreover, we observe that the unknown betweenness in the two approaches significantly differ, as the symmetric one is more sensitive to missing traceroute data.

[1] We have data from September of 2015 but we chose to plot a smaller time window for demonstration purposes.

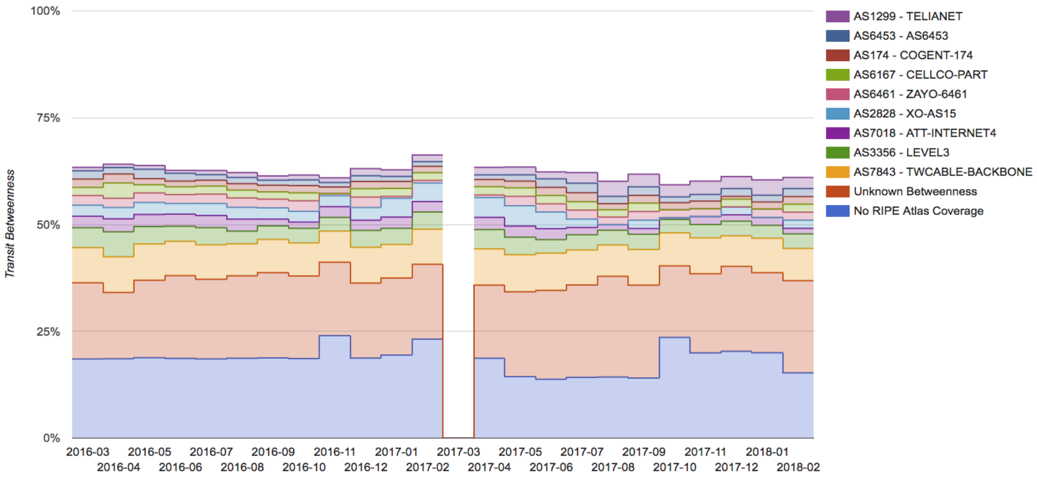


Figure 4.16: Symmetric Transit Betweenness over time for the United States of America.

The next country ecosystem on which we study transit betweenness of intermediate networks is Netherlands. Fig 4.18 depicts the transit betweenness of the intermediate networks that appeared at least one time in the top 5 networks on a month as the ones with the highest transit betweenness using the symmetric approach. The first observation is related with the presence of two major IXPs, AMS-IX[2] and NL-IX[25]. The transit betweenness of the AMS-IX at March 2016 is about ∼10% and across time tends to decrease, as in February 2018 it is about ∼3%. On the other hand, NL-IX increases its transit betweenness over time. Moreover, we observe a significant increase for the AS6830 (Liberty Global Operations B.V.) as it moves from about ∼0.01% to about ∼9% but also for AS286 (KPN) which increases its transit betweenness from ∼5% to ∼11%.

Furthermore, Fig. 4.19 depicts the transit betweenness using the asymmetric approach in IPv4. For the AS6830 (Liberty Global Operations B.V.) we observe the same increment as we observed in the symmetric approach. However, we observe that the for NL-IX and AS286 (KPN) the increase of the transit betweenness that we observed is not present using this approach. Again this seems to be related with routing asymmetries. It is worth noting that on June 2017 we observe a significant drop in transit betweenness of almost all intermediate networks for both of the two approaches. This is due to a failed measurement run of the IXP Country Jedi tool.

Finally, an online version (for 114 countries) of the Transit Betweenness for the symmetric and asymmetric approach where the user can select the address family protocol (v4/v6), but also the number of top transit networks to report per month is available at: `https://www.eyeball-jedi.net`.

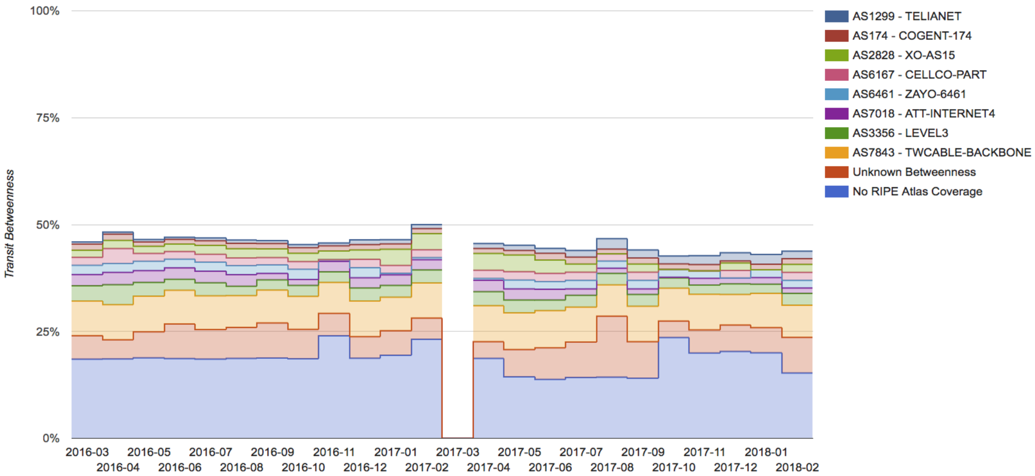


Figure 4.17: Asymmetric Transit Betweenness over time for the United States of America.

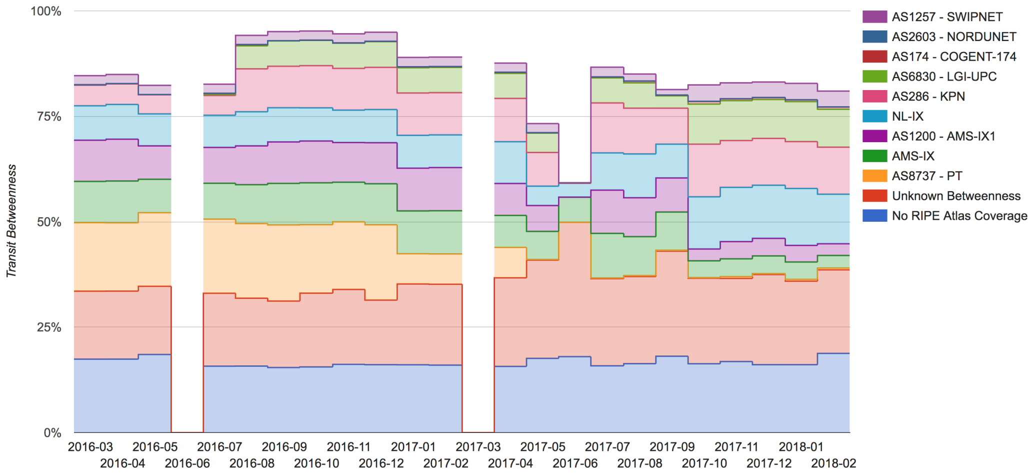


Figure 4.18: Symmetric Transit Betweenness over time for the Netherlands.

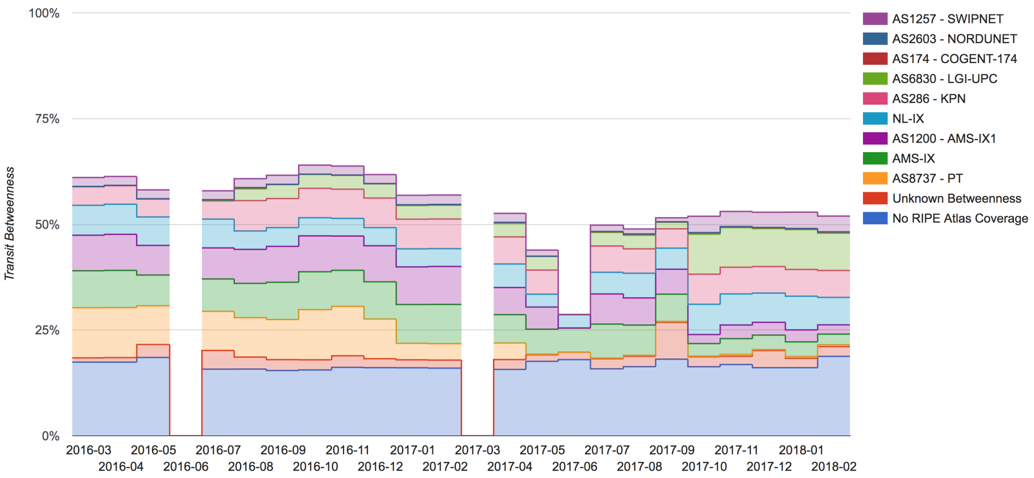


Figure 4.19: Asymmetric Transit Betweenness over time for the Netherlands.

Chapter 5

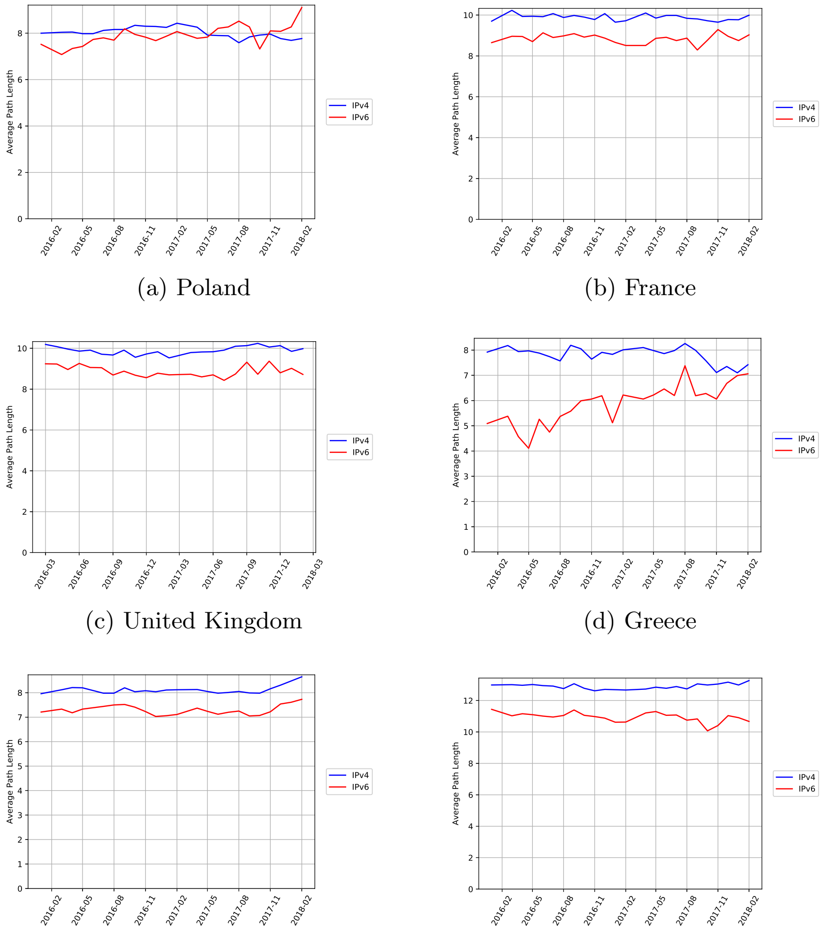
Further Framework Applications

In Chapter 3 we proposed a framework to store, enrich and easily explore and mine/analyze the traceroute data of the IXP Country Jedi tool (see section 2.6). We provided a detailed explanation of an application of this framework on studying eyeball-to-eyeball connectivity within a given country in 4. In this Chapter, we use the proposed framework to further demonstrate its capabilities on performing data analysis. The network ecosystem of a country can be characterized by multiple metrics and traceroute path attributes. In this chapter we focus on three different metrics. We study these metrics over time but also we compare the metrics per address family protocol (IPv4/IPv6). In Section 5.1 we examine the average path length of the traceroute paths inside a given country. After that, in Section 5.2 we explore the fraction of paths that did not stayed local in the country but went out of country. Finally, in Section 5.3 we look into the fraction of paths that crossed an IXP that operates in the same country.

5.1 Path length

One of the most valuable kinds of information that can be extracted from a traceroute path is its path length. The length in most of the cases represents the number of routers that a packet needs to traverse between a pair of two hosts. However, as we described in Section 4.2.1.1 traceroute measurements are subject to router constraints. Possible router constraints that can affect the path length of a measurement include the router silently forwarding packets without altering and decreasing the TTL but also modifying the TTL in unpredictable ways. These traceroute measurement constraints are beyond the scope of this thesis.

Typically, a short (hop-wise) path length facilitates the quick transfer of information. It also can be considered as indicator of good connectivity between a pair of ASes or a whole network (e.g., eyeball) ecosystem. Using the API of the framework we analyzed all the collected data for all measured ASes per month and per country. After that, we extracted the average path length on IPv4 and IPv6 over time. We consider as invalid the traceroute paths that did not reach the destination IP and we excluded them from our results.



(a) Poland

(b) France

(c) United Kingdom

(d) Greece

(e) Netherlands

(f) United States

Figure 5.1: Comparison of IPv4 vs IPv6 average path length over time for various countries

Fig: 5.1 depicts the average path length over time but also per IP version (blue for IPv4 and red for IPv6) for six different countries. The listed countries are (a) Poland, (b) France, (c) United Kingdom, (d) Greece, (e) Netherlands and (f) United States. We observe that the average path length using IPv6 is shorter in 5 of the 6 countries, while in Poland it seems to follow and surpass IPv4.

The observed results can lead us only to assumptions (and educated guesses) as we do not have ground truth data to validate any possible claims. We can assume that IPv6 traceroute paths, experience on average shorter path lengths due to the phenomenon of tunneling IPv6 over IPv4 (according to the RFC 4213

[71]) when a part of network does not support dual stack routing (IPv4 and IPv6) but supports only IPv4. In this case the IPv6 packets are encapsulated into IPv4 packets and are routed using an IPv4 header. As a result, the TTL value of IPv6 is not decreasing and IPv6 routers in the path are not discovered.

Another assumption is that due to the lower traffic volume of IPv6, IPv6 paths are subject to less complex/sophisticated/advanced traffic engineered and packets are routed in a simpler topology. Furthermore, the ASes that adopt IPv6 are changing their network equipment in order to support dual stack routing. The new hardware is capable of providing higher throughput in router-to-router links. This may lead network administrator to create simpler topologies, as they can handle higher traffic volumes using less links to perform load balancing.

5.2 Out-of-Country Paths

In the last decade, studies [45, 55, 74] revealed that traffic between two ASes of the same country often traverses one or more countries before reaching the destination AS. One of the most well-known cases is the Canadian case [74] where researchers discovered a phenomenon which they named as ”boomerang routing” whereby Internet transmissions originating and terminating in Canada are routinely routed through the United States. This phenomenon can occur due to a set of possible reasons such as network misconfiguration and more economical routing through a transit that operates abroad; however, they can also happen on purpose.

The typical business relationship between a pair of ASes can be either transit or peering. The transit relationship expresses the traditional customer-provider model. The customer pays the provider to transit his traffic towards other networks with a typically rate-based charge (usually based on the 95th percentile approach) [72]. On the other hand, Internet peering is a business relationship where two ASes agree to provide access to each other customers and exchange Internet traffic without paying a fee (e.g., “settlement-free” peering). One of the primary objectives of the IXP Country Jedi prototype tool (see Section 2.6) was to allow network administrators to easily see how their network reaches the other networks in the same country on a monthly basis. Using the tool they can derive data to find possible misconfigurations in their network but also to examine their peering strategies against other ASes.

One of the primary causes that can explain the out-of-country paths can be economical reasons, more specifically the transit agreements between ASes of the same country. If the source and destination AS do not peer either by a private peering agreement or through an IXP they are obligated to hand over the traffic to one or more third-party ASes. These ASes will act as transit providers to deliver the traffic to the destination AS.

As an example, we have the following ASes AS_X, AS_Y, AS_Z which operate in country $Country_A$ and AS_W that operates in $Country_B$. AS_X does not peer with AS_Z so the only way to successfully forward traffic to this network is by using a transit provider. AS_Y and AS_W peer with both AS_X and AS_Z. However, AS_X will choose as transit provider the cheapest option from the set of available transit providers AS_Y and AS_W. In case the cheapest is AS_W, traffic may be routed through AS_W in $Country_B$ and then back to $Country_A$ to reach the destination AS_Z.

Another factor that can make paths between ASes of the same country leave the country is network misconfiguration. This type of errors can affect and create anomalies in the connectivity of a significant portion of ASes. Routing on the Internet between the ASes is performed using the Border Gateway Protocol (BGP). Using the BGP protocol the ASes exchange routing information and routers build their routing tables. In recent years many cases where a network administrator made a misconfiguration and announced –using the BGP protocol– a wrong routing path to other ASes have occurred. In most of these cases the outcome was traffic anomalies but also packet loss. However, there are cases where an AS may on purpose hijack paths of other ASes to gather their traffic. These types of misconfigurations may affect routing paths and force packets to traverse ASes out of country (BGP prefix hijack events).

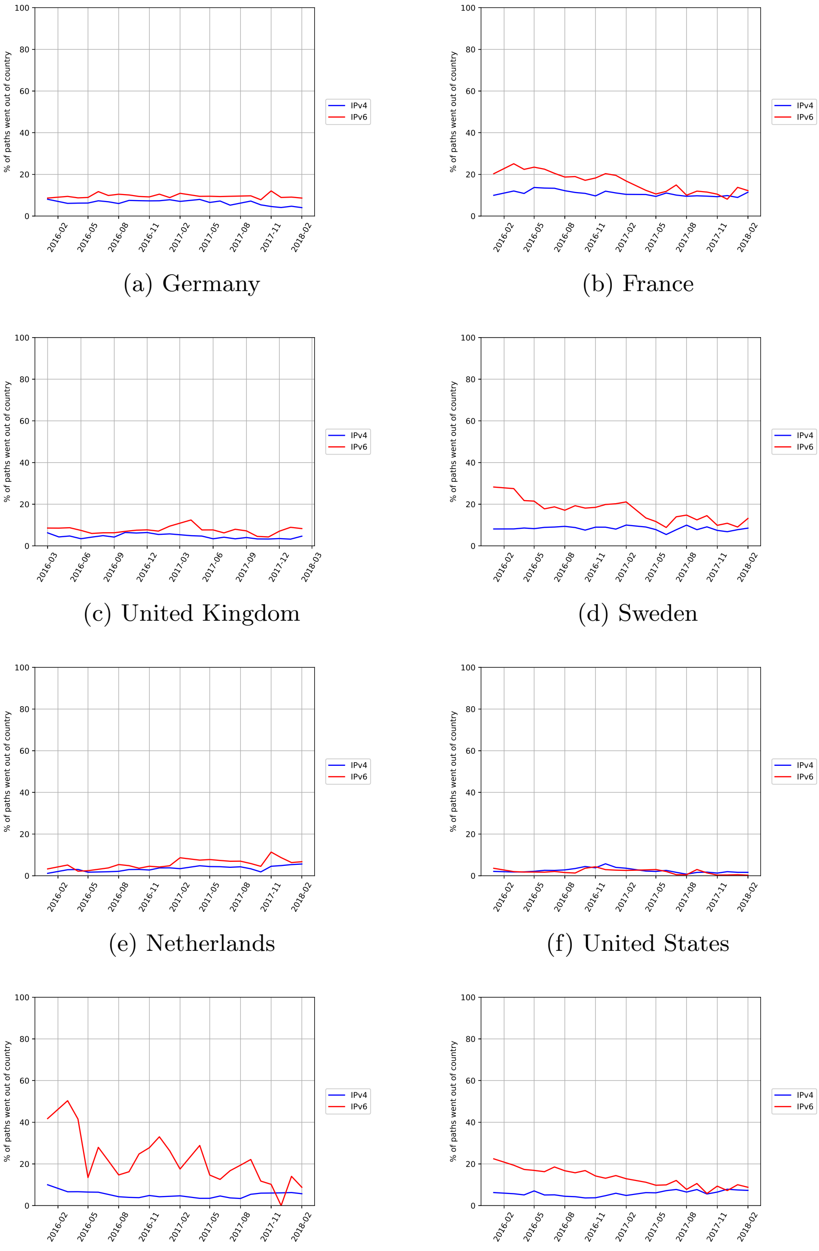
Finally, packets may traverse another country on purpose. When a packet traverses a country, it is subject to the laws of this country. Additionally, the packets could be subject to inspection, filtering and snooping. Moreover, a recent work [45] revealed the case where traffic is routed to a foreign country due to interdependent legal and technical loopholes that the US intelligence community could use to circumvent constitutional and statutory safeguards (applying to the country of origin) to inspect the traffic.

Figure 5.2 depicts the average fraction of paths between all ASes in eight countries that went of country. The listed countries are (a) Germany, (b) France, (c) United Kingdom, (d) Sweden, (e) Netherlands, (f) United States, (g) Spain and (h) Italy. An observation that we can easily make for all the eight countries is that the IPv6 seems to start matching IPv4 across time.

However, the observed results can lead us only to assumptions as we do not have any ground truth data to validate any possible claims. An observation pertaining to the IPv4 and IPv6 paths is that the IPv6 paths have the highest fraction of paths going out of country. This may be the outcome of poor IPv6 adoption of the ASes in the country. If the transit providers or the already established peering links between ASes in the country support only IPv4, then networks are forced to route the IPv6 paths with a different policy. For example, they may offload IPv6 traffic to global transit providers that can eventually route the IPv6 traffic to the destination AS.

5.3 IXP Crossing Paths

IXPs (Internet Exchange Points) have become a very popular place for ASes and CDNs (Content Delivery Networks) to peer with other ASes. Peering in IXPs can



(a) Germany

(b) France

(c) United Kingdom

(d) Sweden

(e) Netherlands

(f) United States

(g) Spain

(h) Italy

Figure 5.2: Comparison of IPv4 vs IPv6 % of Out-of-Country paths over time for various countries.

have significant advantages to the networks such as reducing the cost of transit fees, improving their latency against other ASes but also reach over direct connections large CDN networks. Furthermore, as we already discussed previously (see section 5.2), IXPs can play a key role on the decrease of out-of-country paths between

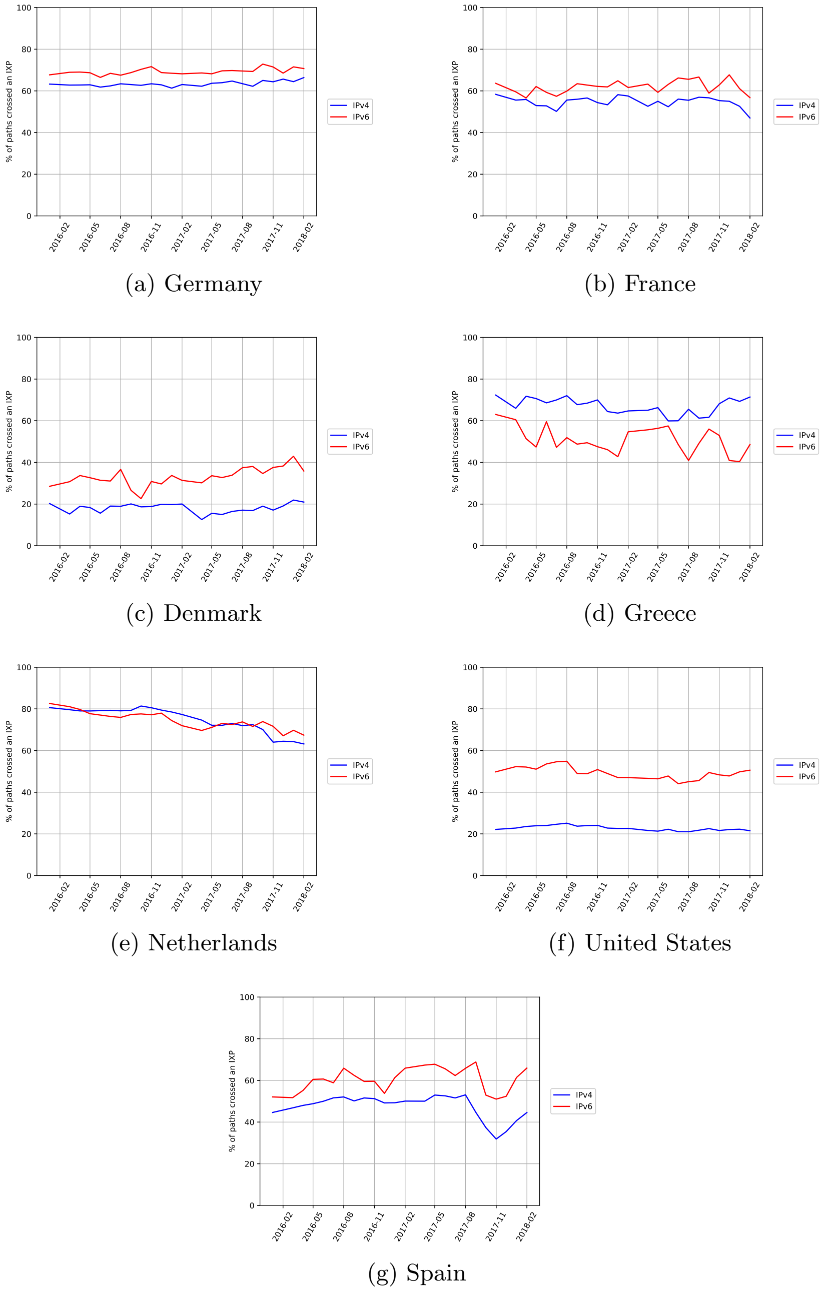
ASes of the same country.

In this section, we examine the fraction of paths between ASes of the same country that cross an IXP that operates in the same country. The IXP crossing identification is performed using the methodology that we describe in Section 2.5. The correctness of the results is subject to the limitations of the IXP crossing inference and we plan in the future to use more complex IXP identification techniques such as the traIXroute tool [70].

Fig. 5.3 depicts the fraction of paths that went through an IXP that operates in the same country over time in IPv4 and IPv6. The listed countries are (a) Germany, (b) France, (c) Denmark, (d) Greece, (e) Netherlands, (f) United States and (g) Spain. An interesting observation is that the IPv6 paths seem to cross more frequently an IXP than the IPv4 paths in Germany, France, United States and Spain. However, the observed results may be biased due to the fact that networks with IPv6 deployment are usually major ISPs that peer aggressively with other networks and need large pools of IP addresses to serve their customers.

On the other hand, Greece is an interesting case where we observe that IPv4 paths have a higher fraction of IXP crossings than the IPv6 ones. The only Greek IXP at this moment in Greece is the GR-IX (Greek Internet Exchange) [13] which is operated by the national research and academic network.

In the Netherlands, the fractions of IPv4 and IPv6 paths seem to match (i.e., converge) over time; we further observe a tendency of decrease over time.



(a) Germany

(b) France

(c) Denmark

(d) Greece

(e) Netherlands

(f) United States

(g) Spain

Figure 5.3: Comparison of IPv4 vs IPv6 IXP crossing paths over time for various countries

Chapter 6

Related Work

This chapter provides the related work and background information that is required to support our work. The content is based on the literature survey that I carried out during the first and the second year of my MSc. We categorize the related work into two different sectors. First, we discuss works that analyze and map traceroute paths from the IP level to the AS level in order to discover the AS-path from a traceroute result. Then, we discuss related research studies on the connectivity between a set of ASes. The goal of this chapter is to help the reader understand the differences in our approach from other related studies.

6.1 IP-path-to-AS-path Mapping

The process of extracting an AS path from a traceroute path has been an active area of research for more than two decades. The first step towards the AS-level path discovery, is mapping IP addresses to routers (mostly known as alias resolution). This step is crucial to identify and characterize inter-domain connections between ASes as routers may respond using different interfaces each time they are being probed. Although the research community has made important steps to successfully map IPs to routers, the problem is not entirely solved. [62] studied the performance of the existing alias resolution techniques and implementations and found that every alias resolution technique which was tested has strengths and weaknesses. Later, the MIDAR [63] implemented an extremely precise ID comparison test by integrating multiple probing methods from multiple vantage points to perform alias resolution.

The next step towards the extraction of an AS-level path from a traceroute IP-level path, is to map each observed IP address to the administrative AS that announces the longest prefix containing this address. To perform this mapping BGP data are frequently used. However, [78] found that this approach can lead to significant false AS-link inferences, due to the fact that a large number of routers respond to traceroute probes with source IPs from a different network (i.e. a neighbor AS, resulting to the presence of a third-party address in the traceroute).

The majority of the Internet mapping studies have derived the network structure, at the AS-level, from a limited number of vantage points of either traceroute or BGP data sources. Although the methodologies and techniques to infer the network topology are continuously improved, the number of vantage points is significantly outpaced by the rapid growth of the Internet. For this reason, [51] proposed a technique to increase the network topology exposure of traceroute measurement results by using active measurements from P2P networks. They calculate the AS paths using BGP data and combine the methodology of [67] with a filtering methodology to avoid false positive reports of ASes. Moreover, in their approach they remove paths that cross IXPs to avoid reporting a wrong AS path, and then they use the IP-to-AS mapping provided by Team Cymru [39], which incorporates both publicly available and private BGP information.

Moreover, Zhuoqing M. et al. [67] proposed a methodology to create an accurate AS-level traceroute tool by using data derived from traceroute results along with BGP data from publicly available route collectors. Their study revealed that about 10% of the traceroute paths have one or more hops that do not map to a unique AS number, and around 15% of the traceroute AS paths have an AS loop. To increase and improve how IP addresses of network infrastructure map to the ASes that operate the infrastructure, they combined the traceroute results with reverse DNS lookups, BGP routing tables, and BGP update messages. Later, they proposed [66] a new algorithm to improve the inaccurate IP-to-AS mappings. The new algorithm can reduce the initial mismatch ratio of 15% between BGP and traceroute AS paths to 5% while changing only 2.9% of the assignments in the initial IP-to-AS mappings.

The authors of [65] focused on tackling the challenge of accurate inferring inter-AS links at the granularity of individual border routers between a network with at least one vantage point (that can perform active targeted traceroutes) and directly connected networks, with high precision. Concurrently, Alexander M. et al. [69] proposed the MAP-IT algorithm which can infer the router ownership on a traceroute path using multi-pass inference mechanisms. The algorithm is capable of inferring the exact interface addresses used for point-to-point inter-AS links, as well as the specific ASes involved.

In this thesis, after understanding in depth the related work and the challenges of accurately transforming an IP-level path to an AS-level path, we developed a simple and scalable IP-path-to-AS-path rule-based transformation technique. Our technique, in order to perform the IP-to-AS mapping, employs the approach of mapping each observed IP address to the administrative AS that announces –in BGP– the longest prefix containing this address, as a first step. As the literature discovered [78] the IP-to-AS mapping may be incorrect because routers may respond to probe packets with third party interfaces. Moreover, in some cases it may be impossible to extract an AS, as some IPs fail to map to an AS (private/not announced IPs). To overcome all these challenges we created a set of filtering rules. The rules on this set are derived from findings of other research works such as [67, 68]. We know that our technique is simpler than other more sophisticated and complex techniques which may also use control plane data (BGP AS-paths) to find AS paths. We attempted to improve our technique using the state-of-the-art MAP-IT algorithm [69] to accurate identify the AS borders. However, the algorithm requires various input sources, which makes the process of applying it on old datasets challenging. In contrast, the current rule-based technique to perform the IP-path-to-AS path transformation requires only BGP data for IP-to-AS mappings; this kind of data are publicly available across time from the *routeviews* project [44]. We describe our IP-path-to-AS-path technique along with the set of rules that we use in Section 4.2.

6.2 Connectivity between ASes

In the last decade, measuring and analyzing the interconnection of ASes has become a very active area of research. The growth of interest for research on this field has been driven by a set of different events; governments around the world trying to legislate and control the Internet by applying censorship on Internet traffic, discovery of cases of suspicious "boomerang routing" (see next), and monitoring of users' data. Moreover, recent events such as the AT&T Whistle-Blower case, where a former engineer of the AT&T reveled that the U.S. National Security Agency was surveilling Internet traffic at a major U.S. Internet backbone network [21], attract the attention of researchers. Moreover, the [45] article described interdependent legal and technical loopholes that potentially allowed the US intelligence community to circumvent constitutional and statutory safeguards for Americans. The article also described how modern Internet protocols can be manipulated to deliberately divert American's traffic abroad where traffic can then be collected under a more permissive legal regime.

To provide insights on potential traffic manipulation, [74] analyzed traceroute paths between Canadian ASes and discovered a phenomenon which they named "boomerang routing" whereby Internet transmissions originating and terminating in Canada are routinely routed through the United States. They pointed out that these Canadian packets traveling through the United States could be subject to U.S. law and could expose Canadians to potential U.S. surveillance activities. However, they analyzed a small number of traceroute paths (about ∼25,000) for a specific time period and only for Canada. The importance of in-country traffic staying local inspired [53] which started as a research initiative in order to map and analyze the routes Canadian data packets take across the Internet backbone. In addition to these two works, in our work on a monthly basis we analyze in-country traceroute paths between all ASes that are covered by the RIPE Atlas platform for about 114 countries. Furthermore, in our analysis we use the OpenIPMap [17] tool to geolocate routers and avoid using commercial IP geolocation databases, as these database suffer from high inaccuracies when it comes to the geolocation of Internet infrastructure. We describe the geolocation dataset of our work in Section 2.3.

The authors of [49] highlighted the non-uniformity of interconnected country ecosystems. They combine real world data (e.g. number of IXPs in a country) along with data from BGP tables (e.g. number of ASes that announce prefixes in the country) and calculate network metrics (e.g. betweenness centrality, clustering coefficient) for 25 countries. In addiction, in our work we use traceroute data and infer the transit betweenness (taking into account also the user populations of the in-country eyeball networks) of ASes for 114 country ecosystems around the globe.

Hal et. al [76] proposed a methodology for mapping national networks of ASes in order to identify small sets of ASes that could possibly act as points of control for a country. They discovered that across countries only a few autonomous systems act as points of control. To identify these sets of ASes they used AS relationship data derived from BGP path announcements collected by the routeviews [44] project, and analyzed the IP address allocation to ASes. More specifically, they calculated the points of control for each country as the minimum set of autonomous systems necessary to connect to 90% of the IP addresses in the country. Their methodology could not be applied to U.S. due to insufficient data of the active IP addresses in ASes that operate in United States, as well as countries with small numbers of IP addresses. In our work, we can infer these “networks in the middle” using traceroute measurements but also infer user-to-user connections that flow through them. Our methodology can be applied to any country where the RIPE Atlas platform provides sufficient coverage.

Finally, [55] measured the country-level paths towards popular domains and characterized transnational routing detours. They found that traffic from a country targeting the same country is often traversing known surveillance states. In their work, they examined only 5 countries, and they analyzed only paths in IPv4. In our work we support both address family protocols (v4/v6) and we analyze 114 countries. Moreover, to geolocate the IP addresses they used the MaxMind database which suffers from high inaccuracies in geolocating Internet infrastructure. In this work we use the more accurate OpenIPMap [17] tool to geolocate routers (see Section 2.3).

Chapter 7

Conclusions & Future Work

In this chapter, we first summarize what we have done and what we have learned while doing this thesis in section 7.1. We conclude with an overview of future work associated with this thesis in section 7.2.

7.1 Summary and Conclusions

In this thesis, we began with the motivation to characterize the connectivity of the "*eyeball*" networks, i.e. user-facing networks with the largest user populations, at large scale (114 countries) and across time. At first, we propose and describe the Eyeball Jedi framework architecture which collects, enriches, analyzes and provides access through a publicly available API to millions of traceroute results generated by the IXP Country Jedi prototype tool starting from September 2015. The framework provides an online API available to researchers and network operators to build on top of this path-related applications.

Using the proposed framework we explore to what extent we can provide insights on country connectivity of all ASes, covered by RIPE Atlas, in a given country. To provide quantitative insights on the user-to-user connectivity we use the population per AS estimates from APNIC. On top of these estimates we build the Eyeball-to-Eyeball matrix which provide useful insights on the connectivity of the eyeball networks. We infer the number of user-to-user connections that stay local or leave the country, crossing or not an IXP and are direct or indirect (involving at least on intermediate AS). Moreover, we analyze the monthly results across time to study the evolution of country ecosystems regarding these properties.

For example, we discover that over time 20% to 50% of the user-to-user connections in Greece cross an IXP (GR-IX); in Ireland this fraction ranges from 30% to 50% and involves INEX IXP. In addition, the user-to-user connections crossing IXPs in United States is only ~3% across time. This is an indication of the wide variety of the peering ecosystems across countries.

Moreover, we examine the fraction of user-to-user connections that stay local in the country or go out of country; we discover that in Canada it varies up-and-down from ~2% to ~11%. This is an interesting finding as our results match other studies that focused only on the Canadian ecosystem[53, 74]. This provides confidence in our data plane approach.

Furthermore, we propose a methodology to examine the transit betweenness of the intermediate networks between in-country networks and provide quantitative insights on the fraction of user-to-user connections that flow through them. We employ both the symmetric and asymmetric approach to take into account presence of intermediates in different directions. In the U.S. we see that the incumbent provider (Time Warner Cable) accounts for approximately 8% of the user-to-user connections as an intermediate network consistently in time. For Netherlands the rise of the NL-IX IXP over time; however, this rise is accompanied with potential routing asymmetries.

In addition, we conducted a mini validation of the probe selection strategy of the IXP Country Jedi using a large ISP in Greece, and we evaluated the coverage of the RIPE Atlas on user populations across the globe.

Besides the aforementioned quantitative insights we learned also that accurately exposing the path diversity of large eyeball networks requires several (more than two) probes; this is because the number of probes should ideally match the user population distribution that they represent. The latter dynamic approach would reduce both the unknown betweenness as well the sensitivity of the methodology to missing or problematic traceroute data. We understand that transforming accurately an IP-level path to an AS-level path is a challenging process and requires a lot of manual effort to be done at scale. We believe that the proposed framework and applications will be valuable for researchers working on the field of Internet connectivity using data plane measurements and accounting for how users are connected to each other. Moreover, it can also be used as a network debugging tool for network operators. For this purpose we make our code, measurements and online API available to the community.

Finally, part of this work has been already published at the Applied Networking Research Workshop (ANRW) 2017 in Prague, Czech Republic [58]. Moreover, insights from this work has been presented to the Connect Working Group of the RIPE 74 meeting in Budapest, Hungary.

7.2 Future Work

We see the following research and engineering directions as interesting future work.

- **Probe Selection**
  With this thesis we made a first attempt to explore the diversity of paths between two RIPE Atlas probes of the same AS against a specific target. Although path diversity can be unpredictable, and we found that the probe selection methodology which the IXP Country Jedi tool uses (selects two probes per ASN -closest and farthest- to the capital) is not optimal in terms

of uncovering the full path diversity, it is a scalable and practical method to examine the path diversity. Moreover, we found probes with weird behavior such as accessing the Internet through a tunnel and not from the origin AS; we need to apply filtering for such probes to avoid distortion of our results. We plan in the future work to assess if this assumption holds in practice. Finally, we see room for potential research on improving the probe selection strategy, in order to expose the full path diversity of an AS.

- **IP-path-to-AS-path Mapping**
  The current IP-path-to-AS-path uses a simple IP-to-AS mapping and a set of rules to transform IP-level paths into AS-level paths. In the context of this work, we tried more sophisticated approaches such as MAP-IT to improve our methodology. However, that approach requires a lot of tuning to provide useful results. We plan in the future to expand our methodology to combine data plane and control plane measurements such as AS paths to increase the accuracy of the generated AS paths.

- **Transit Betweenness Methodology**
  The current betweenness methodology suffers from missing data. One of the challenges that we need to rethink and address is the penalty that we apply in the transit betweenness metric. For small or non-eyeball networks this penalty distorts the transit betweenness percentages of their intermediate networks in a small fraction. However, for large eyeball networks, applying the penalty may distort in a significant way the results of the percentages of the other networks.

- **Framework and Portal**
  The current version of the framework and the portal is in the stage of prototype tool. We plan to continue working on them and later this year to open source the code of the tool. Moreover, we plan to increase the number of external sources that the framework can use to enrich the traceroute data such as the production OpenIPMap, use MAP-IT to enrich data but also tools such as the traIXroute tool to improve the IXP crossing identification process. Finally, we intend to improve the system under the hood to achieve better performance (especially in regarding memory- and time-intensive queries) and scalability.

- **Comparison of user-to-user vs. user-to-content paths**
  It is a well known fact that the CDN traffic is highly optimized, which raises the question how the user-to-user paths, compare with the CDN paths. Are they also getting better (more optimized) or do they lag behind? We plan in the future to compare the user-to-user against the user-to-content paths and answer such questions from different points of view (e.g., latency, throughput, loss, etc.).

- **Validation of our findings**
  One of our main objectives as future work is to validate the findings of this thesis. Finally, with the portal website we will try to crowd source our findings and seek ground truth data from the network operators themselves.

- **Neutralizing biases**
  Like any other measurement-based research work out there, the results presented in this thesis are accompanied with some biases and limitations. For example, RIPE Atlas probes are usually hosted by network enthusiasts; their deployment is not uniform within the user populations of the different countries. Probe selection is thus affected by such a placement bias. Moreover, we have already mentioned the limitations of IP-path-to-AS-path transformation, which becomes especially hard when applied at scale and over time. Apart from that, traceroute datasets may be incomplete, thus giving us only a partial view of a connectivity ecosystem. Finally, we note that APNIC, while being the only known source of user population estimates per AS per country, is subject to several sources of bias related to their measurement methodology and estimation of ISP market shares. In future work, we plan to address these biases separately; in this thesis we made a best effort to account for them where feasible.

Appendix A

Probe Similarity

The main limitation of the probe selection methodology (see Section 2.6.1) is that we do not have any ground truth data to compare against. A recent work [61] tried to look into the probe similarity and define a probe similarity metric. However, their similarity metric focuses on topological similarity without taking into account categorical properties such as the geographic region and the AS that the probes belong to.

To provide useful insights regarding the probe selection strategy, we examined how similar are probes which are hosted in the same AS. The key idea to compare how similar are two probes was that if both of the two probes use the same border router to exit their origin AS, then the rest of the path will be the same. Therefore, e2e path diversity is dictated by the border router diversity. Furthermore, we also examined if the probes use the same border router of the next AS. We constructed a set of IPs that each probe revealed in a specified time period and compared them using the Jaccard index.

Greece was the first country that we looked for probe similarity. The country was selected for a couple of reasons: (i) an acceptable number of Atlas probes are hosted in Greek eyeball networks, the diversity between the probe location (mainland/islands) is noteworthy, and most importantly we had knowledge of the Greek network ecosystem. The top 5 Greek ASes in terms of number of probes being hosted were 'AS6799', 'AS1241', 'AS3329', 'AS25472', 'AS6866'. The aggregate number of active probes being hosted on those ASes during that period was 51 probes. A complete mapping of each AS number to the name and holder organization using the RIPEstat service can be found at Table: A.1.

The dataset that we used to infer the probe similarity includes all the RIPE Atlas public IPv4 traceroute results from 01/02/2017 to 28/02/2017 that were initialized from probes hosted in any of the ASes of the Table A.1. To remove any inconsistencies between active and inactive probes we filter them based on the number of traceroute measurements they run per day. We kept only the probes that initialized at least 120 public measurement per day for the selected time period. This resulted to a decrease on the number of probes from 51 to 37. In

AS Number	Name and holder of this ASN
6799	OTENET-GR Ote SA (Hellenic Telecommunications Organisation)
1241	FORTHNET-GR Forthnet
3329	HOL-GR Vodafone-Panafon Hellenic Telecommunications Company SA
25472	WIND-AS Wind Hellas Telecommunications SA
6866	CYTA-NETWORK Cyprus Telecommunications Authority

Table A.1: Mapping between the AS Number and the name/holder of the selected Greek networks using RIPEstat.

total, more than 2 million traceroute results were analyzed from these probes.

To infer the probe similarity we applied the following methodology.

As a first step, IP-to-AS mapping was applied to the collected traceroute results. To perform the IP-to-AS mapping we used the daily dumps of all routed prefixes from the RIPE RRC16 Collector. The results were grouped per day and mapped using the dump file of that day. The dump files can be found at the public RIS Raw data ftp directory. [36].

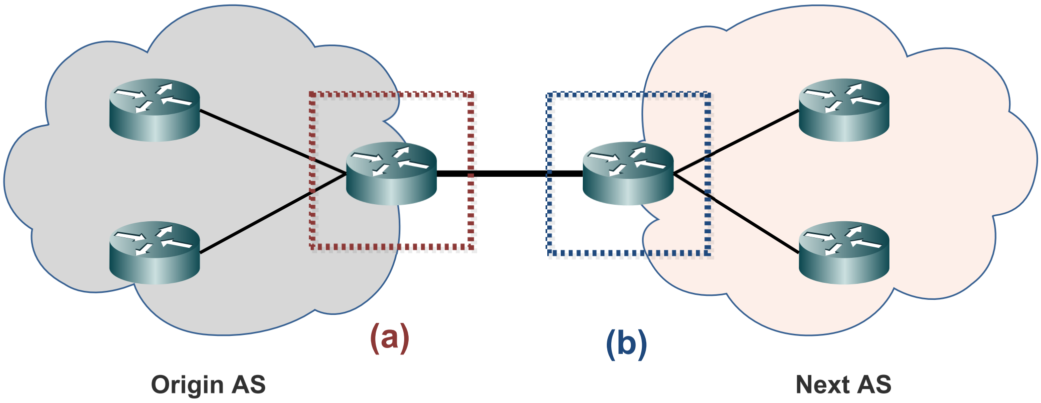


Figure A.1: Uncovering the interfaces of the border routers between each pair of ASes (used for probe similarity computation)

The next step of our methodology was to filter out tuples of border router IPs between the origin AS and the next AS. We removed tuples in which we had at least one unknown hop. Figure: A.1 depicts the connection between the origin and the next AS. For sake of simplicity for the visualization we used only one set of border routers between the two ASes. Furthermore, the first part of the tuple

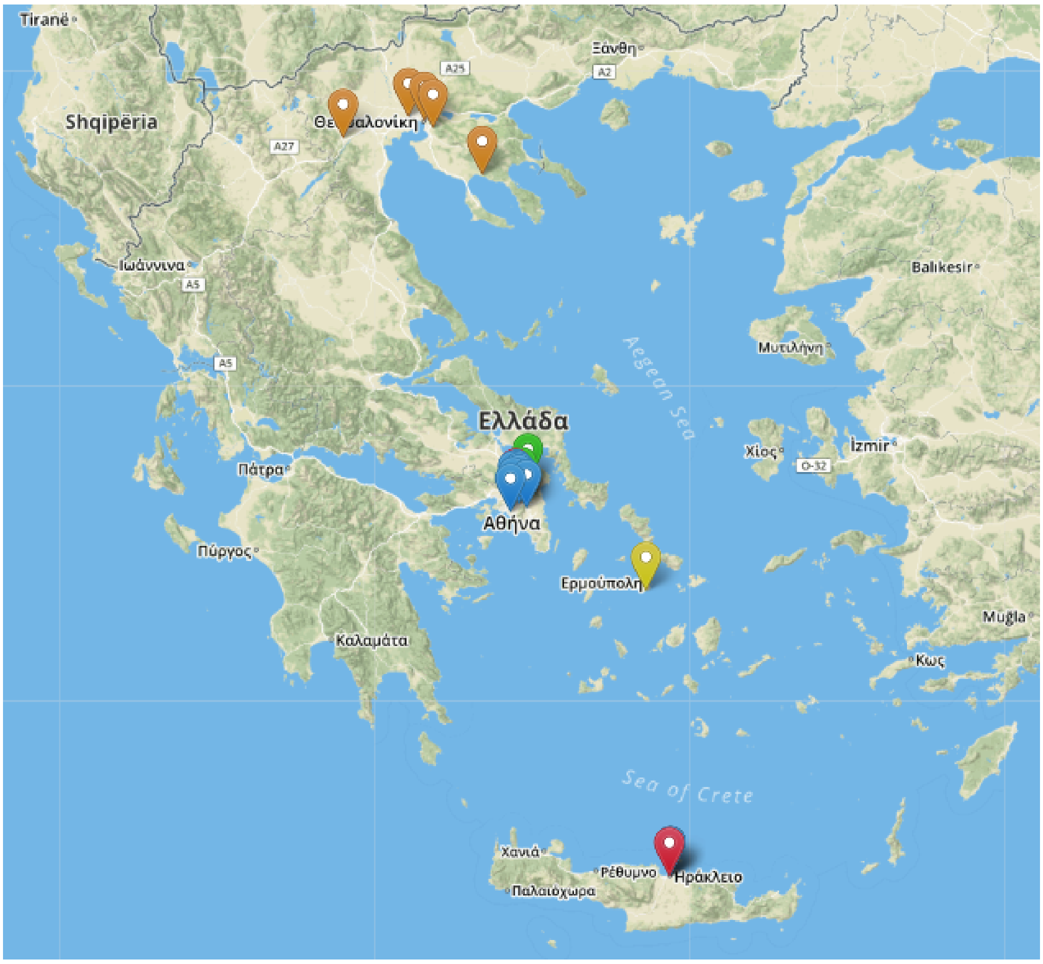


Figure A.2: Probe similarity using Jaccard index 0.5 for AS6799 (OTENET-GR)

includes the IP responses from the interfaces of router (a) and the second one from the router (b).

Finally, we ended up with a list of IP pair tuples per probe. We group the probe results per probe and AS and then apply a set of different thresholds for Jaccard similarity index.

An example of grouping the probes of AS6799 (OTENET-GR) using Jaccard similarity with threshold 0.6 is depicted at Fig: A.2. We observe that multiple groups of probes were formed even for probes hosted in the same city. However, we see that the AS6799 use a geographical policy routing between the North and South part of the country. The differences between probes on the same city provide evidence of the existence of different routing policies between customers, even on a city-level.

Finally, to test the current probe selection strategy; we correlated –using the Pearson correlation coefficient– the distance between a pair of probes with the

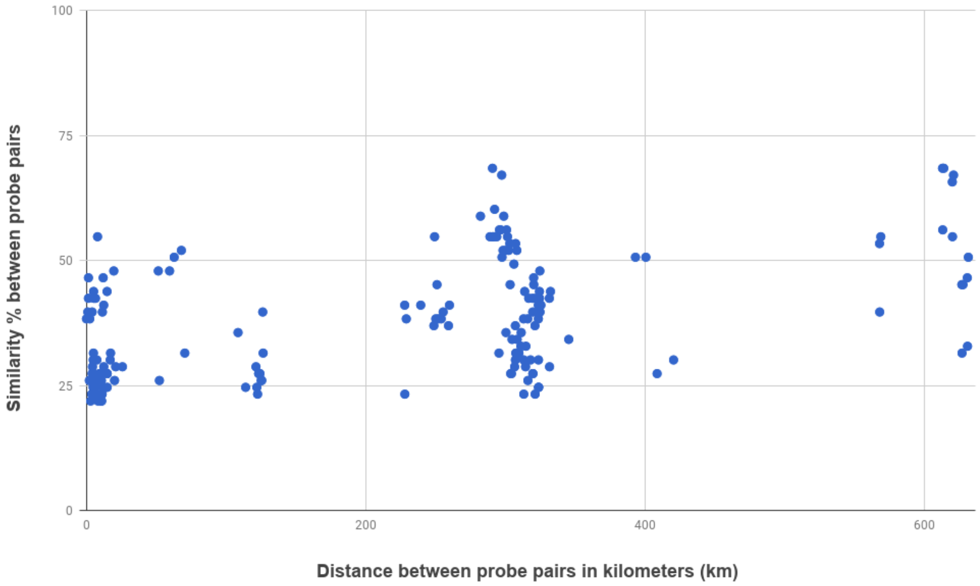


Figure A.3: Similarity vs. distance between probes of AS6799

Jaccard similarity index of the respective border IP sets. An example for AS6799 is depicted at Figure: A.3. The Pearson coefficient between the probe distance and Jaccard similarity is weak and negative with a value of -0.528. The negative correlation indicates that as one of the variables increases, the other tends to decrease, and vice versa.

These results indicate that while the default IXP Country Jedi probe selection strategy (see Section 2.6.1) is not optimal in terms of uncovering full path diversity, it is a scalable and practical method to examine the path diversity.

Appendix B

Evaluating RIPE Atlas Coverage

The RIPE Atlas platform is continuously expanding and new probes are deployed daily on networks that the platform did not cover before. Thousands of probes are spread across multiple countries and different ASes. Based on the official statistics from the platform [15] RIPE Atlas covers 177 countries in 5 continents which results to 90.3% of the total countries and 100% of continents. In terms of networks, RIPE Atlas covers more than 3622 ASNs in IPv4 and 1397 ASNs in IPv6 which results to 6% and 9.3% of the total ASNs respectively. Without any doubt, RIPE Atlas is one of the most expansive measurement platforms covering a significant number of countries and networks.

A question that frequently arises between users of the platform is to what percent of population the covered ASNs map. To estimate the proportion of Internet users per country connected to networks that contain active RIPE Atlas probes, we created the RIPE Atlas Population Coverage tool [32]. This is useful to:

- Quickly access the number of networks across the world that are reached by RIPE Atlas.
- Determine where more RIPE Atlas probes are needed by spotting gaps in RIPE Atlas coverage at a glimpse.

B.1 Methodology

First of all, in order to discover the size of an AS (autonomous system) in terms of users we use the APNIC estimates (see Section 2.2). On a daily basis, we fetch the APNIC estimates for 249 countries and apply certain thresholds in order to consider a network as an eyeball network. Per country we use the inferred ISP market shares to estimate which networks (ASes) are the dominant players up to a cumulative fraction of 95% of the Internet users in that country. In the portal, we do not consider an ISP as a major eyeball network if it has less then 0.1% of market share. Sometimes this last restriction causes us to report on less then 95% of the market.

An interesting outlier is Russia, where the eyeball ecosystem looks highly diverse and fragmented. In this case, many of the networks we would need in order to cover 95% of the market fall below our 0.1% market share threshold. For anything under this threshold, one would need 343 ASNs to cover the 95% of market share.

Finally, we combine these estimates with the daily RIPE Atlas active probes [33] and we assume that if at least one active probe exist inside an AS, then the user population of the AS is covered.

B.2 Coverage Map

To provide an easy way to view the RIPE Atlas coverage we created the Coverage Map.

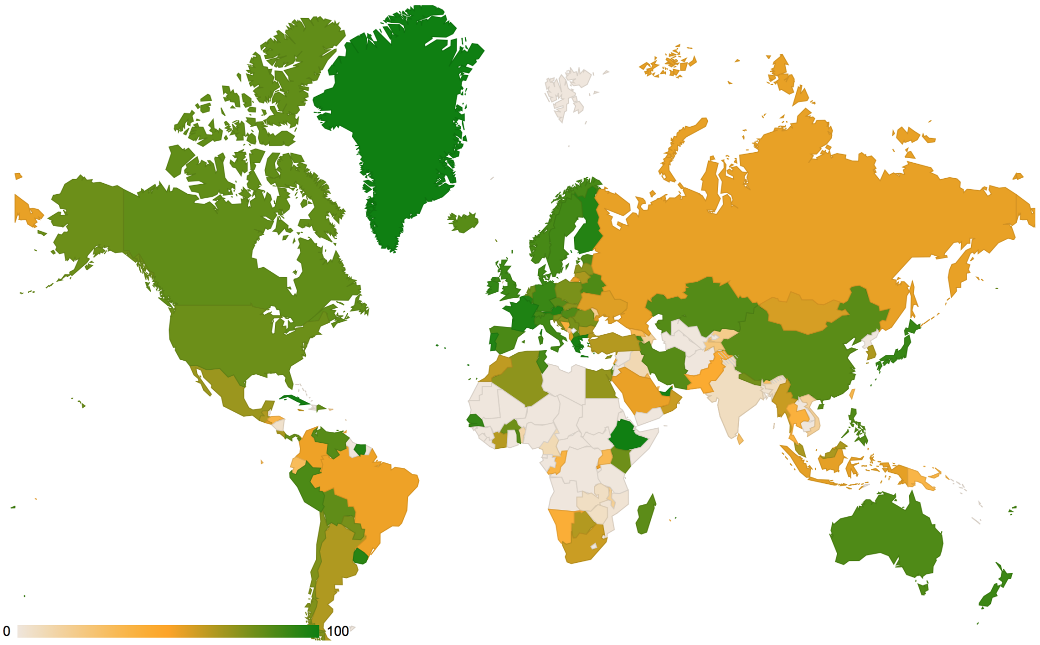


Figure B.1: IPv4 Public and Private Probes.

Fig. B.1 depicts an example of the world map showing the coverage using IPv4 public and private RIPE Atlas probes. On that map we can see that the RIPE Atlas coverage on eyeball networks in Africa is none or very low. RIPE Atlas ambassadors in the Africa region can use the portal to identify the missing networks and target them to deploy new probes and increase the diversity and coverage of the RIPE Atlas network. As an example, in Nigeria RIPE Atlas only cover an estimated 1.3% of the end-user market.

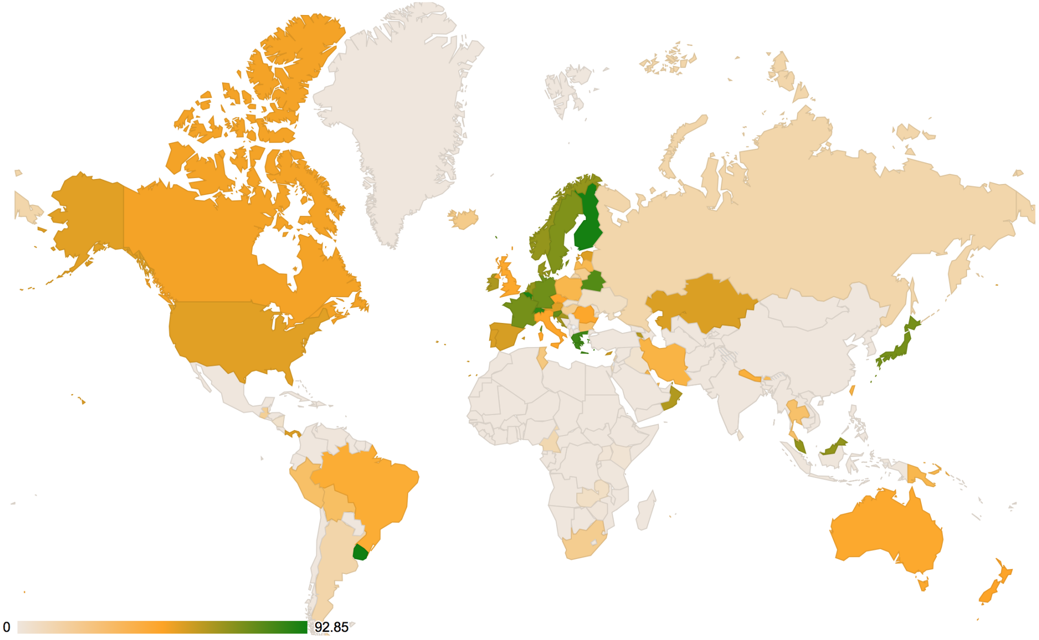


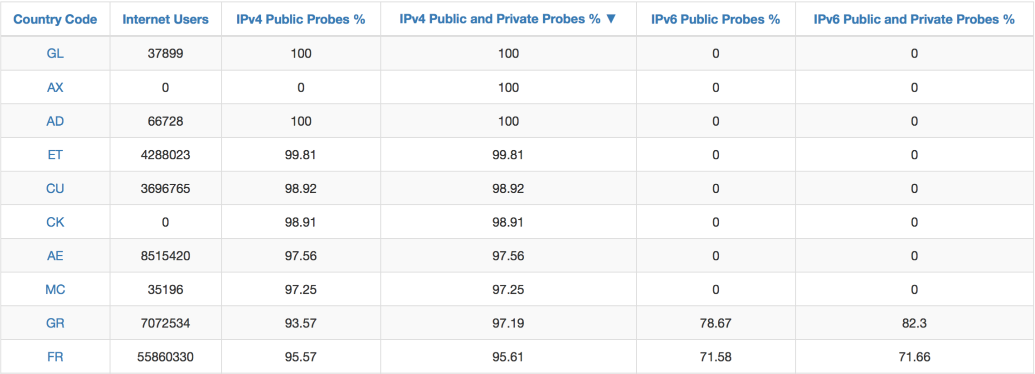
Figure B.2: RIPE Atlas IPv6 Public and Private Probes.

Fig. B.2 depicts a world map showing coverage using IPv6 public and private RIPE Atlas probes. We can observe that the IPv6 coverage of eyeball networks is significantly lower in comparison with IPv4 support. Interestingly, Uruguay, Finland and Belgium seem to perform very well, with more than 90% IPv6 coverage of their eyeball networks.

B.3 Countries Table

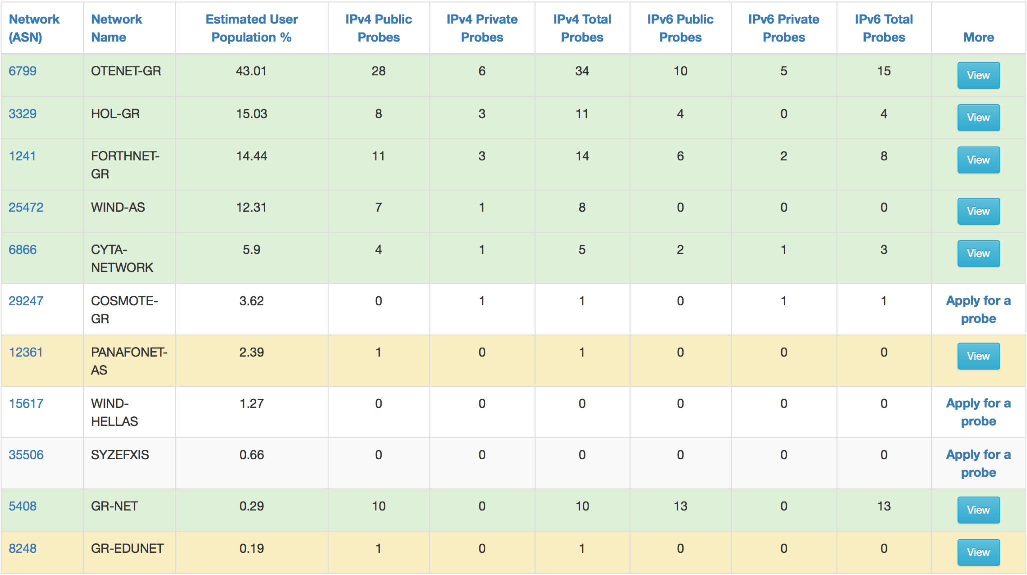
To access the data in a easy to explore form, we created a table where the user can find all countries in (by ISO 3166 two-letter country code [8]) along with estimates of how many end-user market networks are covered by RIPE Atlas probes. We present these estimates per address family protocol (v4/v6), but also per probe type (public/private). An example of the table is depicted in Fig. B.3.

Moreover, we created a per country view that lists the networks with the largest estimated market share, together with the number of probes that are deployed in each network. Rows are coloured green if three or more IPv4 capable public probes are in that network; i.e. if RIPE Atlas have some redundancy in sources from that particular network. If RIPE Atlas have one or two probes, the row is coloured yellow. Rows without colour represent networks where it would potentially be interesting to deploy new RIPE Atlas probes. This is particularly interesting for RIPE NCC staff distributing RIPE Atlas probes, and also for RIPE Atlas ambassadors to see where to focus attention for probe distribution. An example for Greece is is depicted in Fig. B.4.



Country Code	Internet Users	IPv4 Public Probes %	IPv4 Public and Private Probes % ▼	IPv6 Public Probes %	IPv6 Public and Private Probes %
GL	37899	100	100	0	0
AX	0	0	100	0	0
AD	66728	100	100	0	0
ET	4288023	99.81	99.81	0	0
CU	3696765	98.92	98.92	0	0
CK	0	98.91	98.91	0	0
AE	8515420	97.56	97.56	0	0
MC	35196	97.25	97.25	0	0
GR	7072534	93.57	97.19	78.67	82.3
FR	55860330	95.57	95.61	71.58	71.66

Figure B.3: Table overview of ”eyeball coverage” of the RIPE Atlas platform



Network (ASN)	Network Name	Estimated User Population %	IPv4 Public Probes	IPv4 Private Probes	IPv4 Total Probes	IPv6 Public Probes	IPv6 Private Probes	IPv6 Total Probes	More
6799	OTENET-GR	43.01	28	6	34	10	5	15	View
3329	HOL-GR	15.03	8	3	11	4	0	4	View
1241	FORTHNET-GR	14.44	11	3	14	6	2	8	View
25472	WIND-AS	12.31	7	1	8	0	0	0	View
6866	CYTA-NETWORK	5.9	4	1	5	2	1	3	View
29247	COSMOTE-GR	3.62	0	1	1	0	1	1	Apply for a probe
12361	PANAFONET-AS	2.39	1	0	1	0	0	0	View
15617	WIND-HELLAS	1.27	0	0	0	0	0	0	Apply for a probe
35506	SYZEFXIS	0.66	0	0	0	0	0	0	Apply for a probe
5408	GR-NET	0.29	10	0	10	13	0	13	View
8248	GR-EDUNET	0.19	1	0	1	0	0	0	View

Figure B.4: Overview of ”eyeball coverage” for Greece

Bibliography

[1] Django REST framework is a powerful and flexible toolkit for building Web APIs. http://www.django-rest-framework.org.

[2] AMS-IX - Amsterdam Internet Exchange. https://ams-ix.net/.

[3] APNIC Asia-Pacific Network Information Centre. http://www.apnic.net.

[4] B. Huffaker and k. Claffy. "IPv4 & IPv6 Internet Topology Map January 2009." Cooperative Association for Internet Data Analysis. 2009. http://www.caida.org/research/topology/as_core_network/pics/ascoreipv4-ipv6.200903_poster_1250x850.png.

[5] China has launched another crackdown on the Internet — but it's different this time. https://www.cnbc.com/2017/10/26/china-internet-censorship-new-crackdowns-and-rules-are-here-to-stay.html.

[6] "Cosmote introduces IPv6" Internet Society 16 October 2016. https://www.internetsociety.org/blog/2016/10/cosmote-introduces-ipv6/.

[7] Costumer Populations per ASN APNIC report. https://stats.labs.apnic.net/aspop.

[8] Country Codes - ISO 3166. https://www.iso.org/obp/ui/.

[9] Digital Envoy. 2016. Digital Element NetAcuity databases. https://www.digitalelement.com/netacuity.

[10] Django [Computer Software]. (2013). Retrieved from https://djangoproject.com. https://www.djangoproject.com.

[11] Euro-IX - Internet eXchange Point Database (IXPDB)., howpublished = https://www.euro-ix.net/en/ixpdb/ixpdb/.

[12] GeoLite2 is the free IP geolocation database offered by MaxMind. https://dev.maxmind.com/geoip/geoip2/geolite2/.

[13] GR-IX - Greek Internet Exchange. https://www.gr-ix.gr.

[14] How Big is that Network? APNIC user population estimates. https://labs.apnic.net/?p=526.

[15] INEX is the Internet peering point for the island of Ireland. https://www.inex.ie.

[16] IP2Location. 2017. IP2Location Databases. http://lite.ip2location.com/.

[17] IPMAP - A Collaborative Approach to Mapping Internet Infrastructure.

[18] IPMAP - Prototype Github respository.

[19] ITU: Telecommunications development sector. http://www.itu.int.

[20] IXP Country Jedi Prototype Github respository. https://github.com/emileaben/ixp-country-jedi.

[21] M. Klein "AT&T Whistle-Blower's Evidence." Wired 17 May 2006. https://www.wired.com/2006/05/att-whistle-blowers-evidence.

[22] MaxMind - GeoIP2 is the paid and most accurate IP geolocation database that MaxMind offers. https://www.maxmind.com/en/geoip2-databases.

[23] MPLS fundamentals: Forwarding labeled packets. http://www.ciscopress.com/articles/article.asp?p=680824&seqNum=4.

[24] Net Neutrality is about Government Control of the Internet. https://fee.org/articles/net-neutrality-is-about-government-control-of-the-internet/.

[25] NL-IX - the neutral internet exchange. https://www.nl-ix.net.

[26] OpenIPmap is the RIPE NCC tool for mapping core Internet infrastructure. https://openipmap.ripe.net.

[27] PostgreSQL is a powerful, open source object-relational database system. https://www.postgresql.org.

[28] REST RIPE Atlas API. https://atlas.ripe.net/docs/api/v2/.

[29] RIPE Atlas - Raw data structure documentation. https://stat.ripe.net/.

[30] RIPE Atlas Cousteau - A python wrapper around RIPE ATLAS API. https://github.com/RIPE-NCC/ripe-atlas-cousteau.

[31] RIPE Atlas Measurement platform. https://atlas.ripe.net/.

[32] RIPE Atlas Population Coverage. http://sg-pub.ripe.net/petros/population_coverage.

[33] RIPE Atlas Probe Archive. `https://ftp.ripe.net/ripe/atlas/probes/archive/`.

[34] RIPE NCC - The Regional Internet Registry for Europe, the Middle East and parts of Central Asia. `https://www.ripe.net`.

[35] RIPE RIS - Routing Information Service. `https://ris.ripe.net`.

[36] RIPE RIS Raw Data Collector 16. `http://data.ris.ripe.net/rrc16/`.

[37] RIPEstat — Internet Measurements and Analysis. `https://atlas.ripe.net/docs/api/v2/`.

[38] "Study reveals how much people understand Internet" - phys.org. `https://phys.org/news/2014-11-reveals-people-internet.html`.

[39] Team Cymru "IP-to-ASN mapping as a service". `https://www.team-cymru.com/IP-ASN-mapping.html`.

[40] The Eyeball Jedi portal and API. `https://www.eyeball-jedi.net`.

[41] The Internet big picture world Internet users and 2015 population stats., howpublished = `http://www.internetworldstats.com/stats.htm`.

[42] Tony Bates; Philip Smith; Geoff Huston. "CIDR report". `http://www.cidr-report.org/as2.0/`.

[43] TraceMON - Network Debugging Made Easy. `https://labs.ripe.net/Members/massimo_candela/tracemon-traceroute-visualisation-network-debugging-tool`.

[44] University of Oregon: Route Views Project., howpublished = `http://www.routeviews.org/`.

[45] Axel Arnbak and Sharon Goldberg. Loopholes for circumventing the constitution: Unrestricted bulk surveillance on americans by collecting network traffic abroad. *Mich. Telecomm. & Tech. L. Rev.*, 21:317, 2014.

[46] Brice Augustin, Xavier Cuvellier, Benjamin Orgogozo, Fabien Viger, Timur Friedman, Matthieu Latapy, Clémence Magnien, and Renata Teixeira. Avoiding traceroute anomalies with paris traceroute. In *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*, pages 153–158. ACM, 2006.

[47] Suso Benitez-Baleato, Nils B Weidmann, Petros Gigis, Xenofontas Dimitropoulos, Eduard Glatz, and Brian Trammell. Transparent estimation of internet penetration from network observations. In *International Conference on Passive and Active Network Measurement*, pages 220–231. Springer, 2015.

[48] Timm Böttger, Felix Cuadrado, Gareth Tyson, Ignacio Castro, and Steve Uhlig. A hypergiant's view of the internet. *ACM SIGCOMM Computer Communication Review*, 47(1), 2017.

[49] Görkem Çakmak and Mehmet N Aydin. A country-specific analysis on internet interconnection ecosystems. In *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2017 9th International Congress on*, pages 232–237. IEEE, 2017.

[50] Gorkem Cakmak and Henna Suomi. A comparison of isp and mno interconnection models. In *Telecommunications (ICT), 2014 21st International Conference on*, pages 431–436. IEEE, 2014.

[51] Kai Chen, David R. Choffnes, Rahul Potharaju, Yan Chen, Fabián E. Bustamante, Dan Pei, and Yao Zhao. Where the sidewalk ends: Extending the internet as graph using traceroutesfrom p2p users. *IEEE Transactions on Computers*, 63:1021–1036, 2009.

[52] Yi-Ching Chiu, Brandon Schlinker, Abhishek Balaji Radhakrishnan, Ethan Katz-Bassett, and Ramesh Govindan. Are we one hop away from a better internet? In *Proceedings of the 2015 Internet Measurement Conference*, pages 523–529. ACM, 2015.

[53] Andrew Clement. Ixmaps—tracking your personal data through the nsa's warrantless wiretapping sites. In *Technology and Society (ISTAS), 2013 IEEE International Symposium on*, pages 216–223. IEEE, 2013.

[54] Alberto Dainotti, Karyn Benson, Alistair King, Bradley Huffaker, Eduard Glatz, Xenofontas Dimitropoulos, Philipp Richter, Alessandro Finamore, and Alex C Snoeren. Lost in space: improving inference of ipv4 address space utilization. *IEEE Journal on Selected Areas in Communications*, 34(6):1862–1876, 2016.

[55] Anne Edmundson, Roya Ensafi, Nick Feamster, and Jennifer Rexford. Characterizing and avoiding routing detours through surveillance states. *arXiv preprint arXiv:1605.07685*, 2016.

[56] Rodérick Fanou, Pierre Francois, and Emile Aben. On the diversity of interdomain routing in africa. In *International Conference on Passive and Active Network Measurement*, pages 41–54. Springer, 2015.

[57] M. Gharaibeh, A. Shah, B. Huffaker, H. Zhang, R. Ensafi, and C. Papadopoulos. A Look at Router Geolocation in Public and Commercial Databases. In *Internet Measurement Conference (IMC)*, Nov 2017.

[58] Petros Gigis, Vasileios Kotronis, Emile Aben, Stephen D Strowes, and Xenofontas Dimitropoulos. Characterizing user-to-user connectivity with ripe atlas. In *Proceedings of the Applied Networking Research Workshop*, pages 4–6. ACM, 2017.

[59] Vasileios Giotsas, Christoph Dietzel, Georgios Smaragdakis, Anja Feldmann, Arthur Berger, and Emile Aben. Detecting peering infrastructure outages in the wild. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, pages 446–459. ACM, 2017.

[60] Arpit Gupta, Matt Calder, Nick Feamster, Marshini Chetty, Enrico Calandro, and Ethan Katz-Bassett. Peering at the internet's frontier: A first look at isp interconnectivity in africa. In *International Conference on Passive and Active Network Measurement*, pages 204–213. Springer, 2014.

[61] Thomas Holterbach, Emile Aben, Cristel Pelsser, Randy Bush, and Laurent Vanbever. Measurement vantage point selection using a similarity metric. In *Proceedings of the Applied Networking Research Workshop*, pages 1–3. ACM, 2017.

[62] Ken Keys. Internet-scale ip alias resolution techniques. *ACM SIGCOMM Computer Communication Review*, 40(1):50–55, 2010.

[63] Ken Keys, Young Hyun, Matthew Luckie, and Kim Claffy. Internet-scale ipv4 alias resolution with midar. *IEEE/ACM Transactions on Networking*, 21(2):383–399, 2013.

[64] Aemen Lodhi, Natalie Larson, Amogh Dhamdhere, Constantine Dovrolis, et al. Using peeringdb to understand the peering ecosystem. *ACM SIGCOMM Computer Communication Review*, 44(2):20–27, 2014.

[65] Matthew Luckie, Amogh Dhamdhere, Bradley Huffaker, David Clark, et al. Bdrmap: Inference of borders between ip networks. In *Proceedings of the 2016 Internet Measurement Conference*, pages 381–396. ACM, 2016.

[66] Zhuoqing Morley Mao, David Johnson, Jennifer Rexford, Jia Wang, and Randy Katz. Scalable and accurate identification of as-level forwarding paths. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 3, pages 1605–1615. IEEE, 2004.

[67] 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*, pages 365–378. ACM, 2003.

[68] Pietro Marchetta, Valerio Persico, Antonio Pescapé, and Ethan Katz-Bassett. Don't trust traceroute (completely). In *Proceedings of the 2013 workshop on Student workhop*, pages 5–8. ACM, 2013.

[69] Alexander Marder and Jonathan M. Smith. Map-it: Multipass accurate passive inferences from traceroute. In *Internet Measurement Conference*, 2016.

[70] George Nomikos and Xenofontas Dimitropoulos. traixroute: Detecting ixps in traceroute paths. In *International Conference on Passive and Active Network Measurement*, pages 346–358. Springer, 2016.

[71] Erik Nordmark and Robert E. Gilligan. Basic Transition Mechanisms for IPv6 Hosts and Routers. RFC 4213, RFC Editor, July 2005.

[72] William B Norton. *The Internet peering playbook: connecting to the core of the Internet*. DrPeering Press, 2011.

[73] Erik Nygren, Ramesh K Sitaraman, and Jennifer Sun. The Akamai network: a platform for high-performance Internet applications. *ACM SIGOPS Operating Systems Review*, 44(3):2–19, 2010.

[74] Jonathan A Obar and Andrew Clement. Internet surveillance and boomerang routing: A call for canadian network sovereignty. 2013.

[75] Ricardo Oliveira, Dan Pei, Walter Willinger, Beichuan Zhang, and Lixia Zhang. The (in) completeness of the observed internet as-level structure. *IEEE/ACM Transactions on Networking (ToN)*, 18(1):109–122, 2010.

[76] Hal Roberts, David Larochelle, Rob Faris, and John Palfrey. Mapping local internet control. In *Computer Communications Workshop (Hyannis, CA, 2011), IEEE*, 2011.

[77] Dennis Weller, Bill Woodcock, et al. Internet traffic exchange: Market developments and policy challenges. Technical report, OECD Publishing, 2013.

[78] Yu Zhang, Ricardo Oliveira, Hongli Zhang, and Lixia Zhang. Quantifying the pitfalls of traceroute in as connectivity inference. In *International Conference on Passive and Active Network Measurement*, pages 91–100. Springer, 2010.