UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

LUCAS FERNANDO MÜLLER

Improving the Accuracy of Spoofed Traffic Inference in Inter-Domain Traffic

Thesis presented in partial fulfillment
of the requirements for the degree of
Doctor of Computer Science

Advisor: Prof. Dr. Marinho Barcellos

Porto Alegre
February 2020

CIP – CATALOGING-IN-PUBLICATION

Müller, Lucas Fernando

Improving the Accuracy of Spoofed Traffic Inference in Inter-Domain Traffic / Lucas Fernando Müller. – Porto Alegre: PPGC da UFRGS, 2020.

146 f.: il.

Thesis (Ph.D.) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2020. Advisor: Marinho Barcellos.

1. Stability. 2. Spoofing. 3. Security. 4. Customer cone. 5. Inter-domain routing. I. Barcellos, Marinho. II. Título.

"Science is a way of life. Science is a perspective. Science is the process that takes us from confusion to understanding in a manner that's precise, predictive and reliable – a transformation, for those lucky enough to experience it, that is empowering and emotional."

— Brian Randolph Greene,

– professor of physics and mathematics at

Columbia University, NY, U.S.A

"It is not the critic who counts; not the man who points out how the strong man stumbles, or where the doer of deeds could have done them better. The credit belongs to the man who is actually in the arena, whose face is marred by dust and sweat and blood; who strives valiantly; who errs, who comes short again and again, because there is no effort without error and shortcoming; but who does actually strive to do the deeds; who knows great enthusiasms, the great devotions; who spends himself in a worthy cause; who at the best knows in the end the triumph of high achievement, and who at the worst, if he fails, at least fails while daring greatly, so that his place shall never be with those cold and timid souls who neither know victory nor defeat."

— Theodore Roosevelt, 26th president of U.S.A

– "Man in the Arena" April 23, 1910.

ACKNOWLEDGEMENTS

"Gratus animus est una virtus non solum maxima, sed etiam mater virtutum omnium reliquarum."
(A thankful heart is not only the greatest virtue, but the parent of all the other virtues).
Marcus Tullius Cicero, In: Oratio Pro Cnæo Plancio, XXXIII

First of all, I am especially thankful to my advisor, Marinho Barcellos, for his guidance in the past years. Thank you for always believing in my potential and for invariably encouraging me to move forward. Marinho provided continuous encouragement, support, and patience throughout this work, as well as constant feedback, that was essential to my development as a researcher. I would also like to thank him for all the exchange of ideas and research endeavors during all those years. Thank you so much for everything!

This thesis would not have been possible without the help and influence of many special people, who played a significant role in my formation. I am also beyond grateful to Kc Claffy (CAIDA/UC San Diego), Matthew Luckie (University of Waikato, New Zealand), and Bradley Huffaker (CAIDA/UC San Diego), all of which co-advised this thesis. I feel very fortunate for the chance I had to work with them at the Center for Applied Internet Data Analysis (CAIDA), San Diego/California – the leading research group in Internet data analysis. It helped me greatly to become a better researcher and develop a deep appreciation for CAIDA's rigorous scientific approach to Internet measurements. My deepest thanks, I truly appreciate you all!

Each of my advisors had a unique and critical contribution to my success, and I am grateful for that. I would also like to thank Renata Cruz Teixeira (INRIA Paris), Lisandro Granville (INF/UFRGS), Artur Ziviani (LLNC, Brazil), and Jéferson Nobre (INF/UFRGS), who agreed to be part of my doctoral examination committee for reviewing this thesis.

I am very thankful to IX.br and NIC.br teams – the directors, coordinators, infrastructure operators, and members. Leandro Bertholdo (IX.br/RNP, nowadays at UTwente, Amsterdam), Bruno Lorensi (IX.br), Cesar Loureiro (RNP), Julio Sirota (IX.br), Milton Kashiwakura (IX.br), Demi Getschko (NIC.br), for their excellent support, feedback, and discussions that allowed this work to be possible.

Special thanks go to my close collaborators and people who helped me and encouraged me throughout this Ph.D.: Rodrigo Oliveira, Matheus Lehmann, Miguel C. Neves, Pedro de Botelho Marcos, Sergio Gutierrez, Fabricio Mazolla, Leandro Bertholdo, Bruno Lorensi, Kolby Kothmann, Daniela Casas Velasco, Raphael Ozawa, Bruno Brendler,

Shuai Hao, Roderick Fanou, Marco Chiesa, Marco Canini, Christoph Dietzel, Pradeeban Kathiravelu, Diogo Mattos, Rejane Frozza, Daniela Peranconi. I am grateful to each one of them! I also need to express my particular gratitude to Rodrigo Oliveira and Miguel C. Neves, who helped me on this journey many times – thanks a lot for everything!

I am also grateful to Brazil's National Network for Education and Research (RNP) for our collaboration on the IPÊ-Analytics research project. It enabled me to improve my views on the challenges of managing large and distributed backbone networks. Many thanks to everyone who was involved in the IPÊ-Analytics project development, especially Alex Moura, Marcos Schwarz, Iara Machado, and Eduardo Moraes Sathler.

During these remarkable years, I have had the privilege to work with and met many exceptional colleagues, researchers, and professors. More importantly, I have made many friends for life. My sincere gratitude to all members of the Computer Networks Group of INF/UFRGS, each student and professor whom I had the pleasure of interacting with. I was super lucky to share the lab 208 at INF/UFRGS with Rodrigo Oliveira, Matheus Lehmann, Miguel C. Neves, Pedro de Botelho Marcos, Sergio Gutierrez, Fabricio Mazolla, Rodrigo Mansilha, Rodolfo Antunes, Daniel Marcon, Lucas Leal, and Rodrigo Dal Ri. I was also very fortunate to share special moments with CAIDA's family: Kc, Bradley Huffaker, Matthew Luckie, Josh Polterock, Dan Andersen, Alberto Dainotti, Ken Keys, Shuai Hao, Roderick Fanou, Mingwei Zhang, Amogh Dhamdhere, David Clark, Alistair King, Ramakrishna Padmanabhan, Ricky Mok, Paul Hick, Alex Ma, Marina Fomenkov, Philipp Winter, and Young Hyun.

Last but not least, I would like to send my warmest thanks to my family. I would like to thank my parents Ângela and Rony, and brother Tomás for the unconditional support and example of determination and perseverance they have always been for me. Thank you for always being there for me. Their unconditional love and affection always inspire me to see the best in people and give me the strength to pursue my goals.

Finally, my acknowledgments to Microsoft Azure, CNPq, and CAPES for the projects financed and CAPES for the doctoral scholarship I received during the period (CAPES Finance Code 001). My special thanks also to PPGC/UFRGS and CAIDA/UC San Diego for all the support during my Ph.D.

ABSTRACT

Ascertaining that a network will forward spoofed traffic usually requires an active probing vantage point in that network, effectively preventing a comprehensive view of this global Internet vulnerability. We argue that broader visibility into the spoofing problem may lie in the capability to infer lack of Source Address Validation (SAV) compliance from large, heavily aggregated Internet traffic data, such as traffic observable at Internet Exchange Points (IXPs). The key idea is to use IXPs as observatories to detect spoofed packets, by leveraging Autonomous System (AS) topology knowledge extracted from Border Gateway Protocol (BGP) data to infer which source addresses should legitimately appear across parts of the IXP switch fabric. In this thesis, we demonstrate that the existing literature does not capture several fundamental challenges to this approach, including noise in BGP data sources, heuristic AS relationship inference, and idiosyncrasies in IXP interconnectivity fabrics. We propose Spoofer-IX, a novel methodology to navigate these challenges, leveraging *Customer Cone* semantics of AS relationships to guide precise classification of inter-domain traffic as In-cone, Out-of-cone (*spoofed*), Unverifiable, Bogon, and Unassigned. We apply our methodology on extensive data analysis using real traffic data from two distinct IXPs in Brazil, a mid-size and a large-size infrastructure. In the mid-size IXP with more than 200 members, we find an upper bound volume of Out-of-cone traffic to be more than an order of magnitude less than the previous method inferred on the same data, revealing the practical importance of Customer Cone semantics in such analysis. We also found no significant improvement in deployment of SAV in networks using the mid-size IXP between 2017 and 2019. In hopes that our methods and tools generalize to use by other IXPs who want to avoid use of their infrastructure for launching spoofed-source DoS attacks, we explore the feasibility of scaling the system to larger and more diverse IXP infrastructures. To promote this goal, and broad replicability of our results, we make the source code of Spoofer-IX publicly available. This thesis illustrates the subtleties of scientific assessments of operational Internet infrastructure, and the need for a community focus on reproducing and repeating previous methods.

Keywords: Stability. spoofing. security. customer cone. inter-domain routing.

Aprimorando a Precisão da Inferência de Tráfego Spoofing na Troca de Tráfego Inter-domínio

RESUMO

A constatação de que uma rede encaminhará tráfego falsificado geralmente requer um ponto de vantagem ativo de medição nessa rede, impedindo efetivamente uma visão abrangente dessa vulnerabilidade global da Internet. Isto posto, argumentamos que uma visibilidade mais ampla do problema de spoofing pode estar na capacidade de inferir a falta de conformidade com as práticas de Source Address Validation (SAV) a partir de dados de tráfego da Internet altamente agregados, como o tráfego observável nos Internet Exchange Points (IXPs). A ideia chave é usar IXPs como observatórios para detectar pacotes falsificados, aproveitando o conhecimento da topologia de sistemas autônomos extraído dos dados do protocolo BGP para inferir quais endereços de origem devem aparecer legitimamente nas comunicações através da infra-estrutura de um IXP. Nesta tese, demonstramos que a literatura existente não captura diversos desafios fundamentais para essa abordagem, incluindo ruído em fontes de dados BGP, inferência heurística de relacionamento de sistemas autônomos e características específicas de interconectividade nas infraestruturas de IXPs. Propomos o Spoofer-IX, uma nova metodologia para superar esses desafios, utilizando a semântica do Customer Cone de relacionamento de sistemas autônomos para guiar com precisão a classificação de tráfego inter-domínio como In-cone, Out-of-cone (*spoofed*), Unverifiable, Bogon, e Unassigned. Aplicamos nossa metodologia em análises extensivas sobre dados reais de tráfego de dois IXPs distintos no Brasil, uma infraestrutura de médio porte e outra de grande porte. No IXP de tamanho médio, com mais de 200 membros, encontramos um limite superior do volume de tráfego Out-of-cone uma ordem de magnitude menor que o método anterior inferiu sob os mesmos dados, revelando a importância prática da semântica do Customer Cone em tal análise. Além disso, não encontramos melhorias significativas na implantação do Source Address Validation (SAV) em redes usando o IXP de tamanho médio entre 2017 e 2019. Na esperança de que nossos métodos e ferramentas sejam aplicáveis para uso por outros IXPs que desejam evitar o uso de sua infraestrutura para iniciar ataques de negação de serviço através de pacotes de origem falsificada, exploramos a viabilidade de escalar o sistema para infraestruturas IXP maiores e mais diversas. Para promover esse objetivo e a ampla replicabilidade de nossos resultados, disponibilizamos publicamente o código fonte do Spoofer-IX. Esta tese

ilustra as sutilezas das avaliações científicas da infraestrutura operacional da Internet e a necessidade de um foco da comunidade na reprodução e repetição de métodos anteriores.

Palavras-chave: Segurança, spoofing, Customer Cone, estabilidade, roteamento, inter-domínio.

LIST OF ABBREVIATIONS AND ACRONYMS

ACL	Access Control List
AFRINIC	African Network Information Centre
APNIC	Asia Pacific Network Information Centre
ARIN	American Registry for Internet Numbers
AS	Autonomous System
ASN	Autonomous System Number
BCP	Best Current Practice
BGP	Border Gateway Protocol
c2p	Customer-to-Provider
CAIDA	Center for Applied Internet Data Analysis
CDC	Context-Dependent Classification
CDN	Content Distribution Network
CF	Colocation Facility
CIC	Context-Independent Classification
CP	Content Provider
DDoS	Distributed Denial of Service
DNS	Domain Name System
DoS	Denial of Service
FC	Full Cone
GEANT	pan-European Network for Research and Education
GRE	Generic Routing Encapsulation
HTTP	Hypertext Transfer Protocol

IANA	Internet Assigned Numbers Authority
IBR	Internet Background Radiation
ICMP	Internet Control Message Protocol
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocols
IP	Internet Protocol
IPIP	IP over IP encapsulation
IRR	Internet Routing Registry
ISOC	Internet Society
ISP	Internet Service Provider
IXP	Internet Exchange Point
LACNIC	Latin America and Caribbean Network Information Centre
LIR	Local Internet Registry
MANRS	Mutually Agreed Norms for Routing Security
NIR	National Internet Registry
NREN	National Research and Education Network
OSPF	Open Shortest Path First
p2p	Peer-to-Peer
PLCC	Prefix-Level Customer Cone
PoP	Point of Presence
PPCC	Provider/Peer-Observed Customer Cone
RFC	Request for Comments
RIPE NCC	Réseaux IP Européens Network Coordination Centre
RIPE RIS	RIPE Routing Information Service

RIR	Regional Internet Registry
RNP	Rede Nacional de Ensino e Pesquisa
ROA	Route Origin Authorizations
RPKI	Resource Certificate Public Key Infrastructure
RV	RouteViews project
s2s	Sibling-to-Sibling
SAV	Source Address Validation
TCP	Transmission Control Protocol
ToR	Type of Relationship
TP	Transit Provider
UDP	User Datagram Protocol
VLAN	Virtual Local Area Network
VP	Vantage Point
VRRP	Virtual Router Redundancy Protocol

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

TABLE OF CONTENTS

1 INTRODUCTION

We organize this chapter into five parts. First, we explain the design goals and deficiencies of two of the most fundamental elements of the Internet: Internet Protocol (IP) and Border Gateway Protocol (BGP) (§1.1). Then, we link the previous discussion to introduce the problem definition and scope (§1.2) that lead us to study and tackle spoofed traffic on the Internet, the main topic of this thesis. Following, we describe our objectives (§1.3), and we give an overview of the contributions (§1.4) achieved with this work to the current state-of-the-art. Finally, we detail how the remainder of the thesis is organized (§1.5).

1.1 Thirty Years in the Evolution of the Internet Ecosystem

Nowadays, the Internet is composed by close to 70k independent networks (APNIC, 2020; NRO, 2020), also known as Autonomous Systems (ASes). An AS can be an Internet Service Provider (ISP), a Transit Provider (TP), a Content Provider (CP), or a smaller organization such a university or a corporation that autonomously administers a domain of connected IP prefixes. Packets within an AS are routed according to a set of metrics and Interior Gateway Protocols (IGPs) that are determined by each AS operator separately and can differ significantly between ASes (HAWKINSON; BATES, 1996). Each AS owns only a subset of the IP address space and typically covers a limited geographical area, which means that end-to-end traffic often needs to traverse multiple AS domains (often Transit Providers) before reaching its destination.

In inter-domain routing, BGP is used as the de-facto protocol for the exchange of reachability information at the boundary of ASes (MAUCH; SNIJDERS; HANKINS, 2017). Before starting to exchange traffic, two ASes need first to establish a physical connection and agree to a contractual relationship that determines the economic and technical aspects of their connectivity (MARCOS et al., 2018). Business relationships between ASes can be broadly classified into two types: Customer-to-Provider (c2p) (or transit) and Peer-to-Peer (p2p) (GAO, 2001; DHAMDHERE; DOVROLIS, 2010; LUCKIE et al., 2013). In a c2p relationship, the customer pays the provider for traffic sent between the two ASes. In return, the customer gains access to all ASes reached by the provider, including those which the provider reaches through its own providers. In a p2p relationship, the peering ASes gain access to each other's customers, typically without either AS paying

the other. Peering ASes have a financial incentive to engage in a settlement-free peering relationship, instead of relaying in a paid provider to carry their traffic. In this context, neither ASes have a customer role.

BGP provides flexibility via policy-based routing, that is, the ability to express routing policies on how reachability information is propagated, allowing AS operators to enforce their contractual agreements and implement complex traffic engineering techniques for load and cost balancing (ANWAR et al., 2015). As a result of policy-based routing, inter-domain traffic does not necessarily follow the shortest path between two ASes. Policy-based routing has been one of the initial design aspects of BGP aiming to enable AS operators to chose which routes will be accepted, which will be preferred and which will be propagated to their neighbors (FLACH et al., 2016). Internet inter-domain routing is a collaborative effort between ASes. ASes negotiate contractual agreements to define their business relations and impose technical restrictions on traffic exchange. On the Internet, connectivity does not imply traffic reachability, which is fundamentally determined by the business relationships between ASes (KATZ-BASSETT et al., 2008; FAYAZ et al., 2016).

The Internet Protocol (IP) provides a simple abstraction for communication over the Internet, identifying hosts by, in theory, globally uniquely addresses. This allows data to cross heterogeneous networks and reach the intended destination. Despite the simplicity of IP and BGP, both were designed with an implicit premise of *mutual trust* between the users (POSTEL, 1981; REKHTER; LI, 1994). Because the major development of these protocols took place in the late 1980s and early 1990s, *virtually no security mechanism were foreseen in these protocols because this was considered unnecessary* (TIMBERG, 2015). It would have added a non-negligible overhead on systems that had low computing power. As a result, the IP and BGP protocols are simple and reliable, but mostly unsecured (BELLOVIN, 2004).

Although the successful operation of all the elements just described and the overwhelming success of the Internet there has been an explosion of security threats. Arguably the greatest architectural vulnerability in the TCP/IP protocol suite as designed is the fact that it provides no explicit mechanism to prevent packets with forged headers from traversing the network (BELLOVIN, 1989). Due to the destination based packet forwarding scheme of the current Internet, routers deliver IP packets without checking the validity of the packets' source addresses, enabling malicious parties to leverage such ability, known as **IP spoofing**, forging or "spoofing" the source address of IP packets. Networks that forward spoofed source IP addresses in packets are a cybersecurity risk on the global

Internet, because they enable attacks such as spoofed Denial of Service (DoS) that are operationally infeasible to trace back to the actual source.

Recognizing that lack of *Source Address Validation (SAV)* is fundamentally an architectural limitation (MORRIS, 1985; BELLOVIN, 1989), the Internet Engineering Task Force (IETF) introduced Best Current Practices (BCPs) recommending that networks block the forwarding of packets with spoofed source addresses (FERGUSON; SENIE, 2000; BAKER; SAVOLA, 2004). Compliance with this practice faces misaligned incentives i.e., protects the *rest* of the Internet from attacks being sourced from the network that must pay a non-trivial cost for deploying and accurately maintaining the filters. Thus, despite many attempts to improve SAV deployment, some of the most damaging DoS attacks in the Internet continues to use IP spoofing as a primary attack vector for large-scale DoS attacks (SCHEID, 2016; MORALES, 2018; KLABA, 2018; NETSCOUT, 2019), and these attacks continue to increase in prevalence and intensity (JONKER et al., 2017; KOTTLER, 2018; SHANI, 2019).

Identifying networks that do not filter spoofed packets is critical to global network infrastructure protection, because it provides a focus for remediation and policy interventions (LUCKIE et al., 2019; ISOC, 2019). However, identification of these networks is challenging at Internet scale. The definitive method requires an active probing vantage point in each network being tested, to see if a spoofed packet successfully traverses the network (BEVERLY et al., 2009; CAIDA, 2018c). Since there are approximately 790K independently routed prefixes from almost 70K ASes on the Internet in 2019 (APNIC, 2020; NRO, 2020), this method has limited feasibility for a comprehensive assessment of Internet spoofing.

1.2 Problem Definition and Scope

In this thesis, we approach the issue of inter-domain networks which lack compliance of SAV best practices. More specifically, we devise a methodology to obtain a broader visibility into the spoofing problem, which we argue may lie in the capability to infer lack of SAV compliance from large, heavily aggregated Internet traffic data, such as traffic observable at Internet Exchange Points (IXPs). Most ASes connect to an IXP to exchange traffic between their customers, i.e., via peering relationships where neither AS pays the other for transit. For these ASes, legitimate source addresses in packets will belong to direct or indirect customers of the AS sending the packets across the IXP fabric

to their peers.

However, inferring SAV deployment at an IXP is remarkably challenging, far more so than has been captured in the literature, due to a combination of operational complexities that characterize today's interconnection ecosystem. First, determining which source addresses are valid in packets arriving at a given port of an IXP switch fabric is challenging, because there is no registry of which addresses networks should forward; in practice, we must heuristically infer valid source addresses. Second, while the original role of IXPs was to promote peering between ASes, networks now also use IXPs to obtain IP transit services from a provider (AGER et al., 2012), and we have found evidence of organizations joining their sibling network ASes across an IXP. For ASes offering transit across the IXP, and for sibling networks, it is infeasible to infer invalid source addresses from IXP traffic data – the set of valid addresses is potentially the entire address space. Third, while IXPs may be thought of as a single switching fabric, in practice IXPs and resellers offer complex services, including remote peering, layer-2 transport, and virtualized segmenting of traffic into multiple Virtual Local Area Networks (VLANs). These interconnection practices occur below and are thus not visible to the IP layer or in the BGP protocol. Accurately inferring SAV deployment at an IXP requires navigating all of these aspects. In this thesis, we describe a methodology that does so, with focus on IPv4 traffic exclusively, as it accounts for more than 95% of the traffic observed at the vantage points of interest (IX.br, 2020; AMS-IX, 2020a; DE-CIX, 2020).

1.3 Objectives

Our main objective is to propose a methodology for the analysis and accurate classification of spoofed traffic in the inter-domain level using heavily aggregated Internet traffic data, obtaining a broader visibility into the spoofing problem. We base our methodology on the hypothesis that *"one can tell the presumably legitimate packets from those potentially spoofed by observing their IP source address, as well as the direction of those packets"*.

Based on the hypothesis above, we need to deal with the Internet topology incompleteness due to the lack of readily available connectivity data. The key idea is to map the inter-domain topology to be able to construct and maintain lists of valid source addresses, per AS, which varies through time according to their established relationships and traffic engineering policies. These lists specify exactly which source addresses should legitimately appear in packets at the observation point in a given time window, as well as

the direction of those packets guiding the precise classification of inter-domain traffic into distinct traffic categories. Given the hypothesis stated, the following questions guide this thesis:

- Can one devise a methodology to augment our capacity to observe and remediate security properties of the Internet, more specifically the lack of traffic filtering (SAV), helping to improve the global cybersecurity? And, if so, how?
- When performing the deployment of the methodology, what will be the vantage points used, and the restrictions imposed by them on the application of the methodology?
- Considering the definition of the valid source address space for each AS, and the continually changing nature of the routing policies accordingly with the network conditions, how to deal with the dynamics of the inter-domain routing?
- Given the organization diversity of networks, as well as their business practices, to what extent the network infrastructure complexities will affect the inter-domain traffic analysis?

In order to answer the research questions posed and assess the effectiveness of our design, we partnered with real network infrastructures and evaluated our methodology by means of measurements using real data. In the various stages of this research, our methodology was analyzed using data obtained from two distinct IXPs from the IX.br ecosystem (IX.br, 2020), spanning traces from three distinct years (2017, 2018, and 2019) [1]. Even though some epistemological challenges remain, the results obtained show that it is possible to obtain a more comprehensive view of this global Internet vulnerability. The results also show that the existing literature does not capture several fundamental challenges to this approach, including noise in BGP data sources, heuristic AS relationship inference, and idiosyncrasies in IXP interconnectivity fabrics. Lastly, we are aware of the importance of being able to replicate scientific results, so all the work reported in this thesis was conducted with replicability in mind.

[1] The first preliminary data analyses were performed in 2016 with a mid-size IXP, focused on understanding the problem, as well as a presentation (MULLER et al., 2016) during the IX Forum 10, part of the Internet Infrastructure Week in Brazil organized by NIC.br and CGI.br.

1.4 Contributions

Our methodology provides new data that improves the understanding of spoofed traffic and, if used by ASes connected to IXPs, can aid increase the Internet infrastructure resilience to attacks. We believe that ultimately our methodology will be integrated into expert system capabilities rather than be amenable to complete layer-3 automation due to networks specific knowledge requirements. More specifically, in designing Spoofer-IX and performing an extensive data analysis, we make the following contributions:

1. We provide a detailed analysis of the methodological challenges and their implications on building IP spoofing detection at IXPs, extended with a comprehensive analysis of previous work.
2. We design and developed Spoofer-IX, a novel methodology for the purposes of accurately inferring spoofed traffic and the absence of SAV in AS members of IXPs, which allows network operators to measure, fix, and support filtering.
3. We apply our tool to extensively study the situation at a medium-sized IXP in Brazil across more than 200 ASes, considering two periods, two years apart, showing the opportunities to fix and improve the deployments of SAV. Interactions with network operators support our findings, together with our analysis.
4. We assess the deployment of Spoofer-IX to more complex IXP architectures and to distinct networks. We partnered with a larger IXP with over one thousand members to assess the scalability of our methodology and implementation. We explored practical application and generalizability of our Spoofer-IX methodology and how networks could independently adopt Spoofer-IX to detect and eliminate spoofed traffic.
5. We describe and publish our code to promote further work. Commercial and privacy sensitivities prevent sharing of traffic data that would enable directly reproducibility of much work in the field of Internet security. But in the interest of *replicability*, we publicly release our code (MULLER et al., 2019b) so that other researchers and IXPs can use it to improve our collective ability to measure and expand deployment of SAV filtering.

1.5 Overview and Structure

The rest of this thesis is organized as follows.

- Chapter 2 presents all the fundamentals of this thesis. The reader who is familiar with the concepts may choose to skip this chapter entirely, or use it as a quick reference.
- Chapter 3 discusses the state-of-the-art in dealing with the Spoofing Problem and the algorithms for the inference of AS relationships and Customer Cones.
- Chapter 4 analyzes the methodological challenges and their implications for applying BGP-based SAV inference methods to modern IXP connectivity fabrics.
- Chapter 5 introduces a novel methodology for the inference of spoofed traffic in inter-domain traffic crossing IXPs. First, we provide an overview of the Spoofer-IX. Then we discuss in detail its two fundamental stages, the related algorithms, and the methodology implementation.
- Chapter 6 describes the datasets employed in our extensive analyses and how they were collected. The analyses are based on data from the third-largest IXP in Brazil, with more than 200 member ASes connected at the IXP switching fabric. We discuss the traffic classification results, including a comparison against the state-of-the-art. We report our validation efforts regarding the inferences made and the results obtained.
- Chapter 7 discusses how to scale the traffic analysis to more complex network infrastructures, even beyond IXPs. We present results obtained in collaboration with three Colocation Facility (CF) that constitute part of a second large IXP, also in Brazil.
- Chapter 8 summarizes the contributions and our key findings, and we present prospective directions for future research.
- The Appendix A includes a list of publications that show the results achieved in the development of this thesis. This list also contains the main collaborations carried out in papers that focus on related subjects, as well as public presentations, research projects, research grants, and (co-)supervised final BSc student papers.

2 BACKGROUND

In this chapter, we present the background context of this thesis in five sections. The first section (§2.1) describes the Spoofing Problem, the attacks it enables, and the common strategies employed to execute these attacks. The second section (§2.2) explains how the IP address space is organized and allocated to organizations. The third section (§2.3) reflects on the IP address space organization, explaining the distinct IP addresses categories. The fourth section (§2.4) presents the definition of the distinct business relationship classes between Autonomous Systems (ASes), how they directly impact in the definition of valid traffic in the Internet and also introduces the definition of the Customer Cone model. Lastly, the fifth section (§2.5) illustrates what makes IXPs a great vantage point to explore Internet security proprieties.

2.1 Framing the Spoofing Problem

The Internet architecture provides no explicit mechanism to prevent packets with forged headers from traversing the network. Due to the destination based packet forwarding scheme of the current Internet, routers deliver IP packets without checking the validity of the packets' source addresses. Figure 2.1 illustrates this vulnerability that exists in the IP Protocol v4/v6 and allows end users (e.g., user D) to send IP packets with fake source addresses (e.g., to an application server B), i.e., the addresses that are not assigned to them, which is known as **IP spoofing**.

Figure 2.1: IP Spoofing Problem Overview. Internet Protocol (IP) does not include built-in source address validation. As a consequence, end-hosts in Autonomous Systems (ASes) can forge IP packet header information.

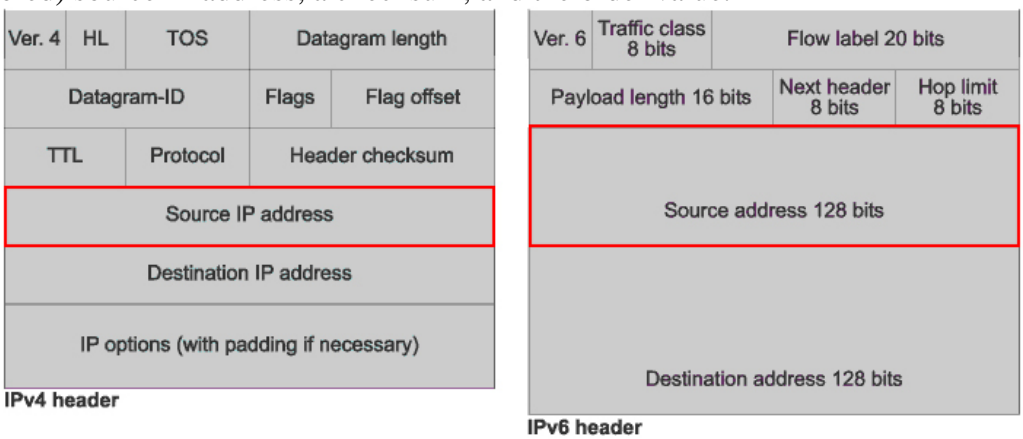
Source: by author (2019).

Malicious parties can leverage the ability to forge or "spoof" the source address of IP packets to mount various attacks. With the increasing size of the Internet (CAIDA,

2019b), we have seen a growing number of attacks that take advantage of the network's large scale. Below we discuss four general types of such attacks (RYBA et al., 2015; ZARGAR; JOSHI; TIPPER, 2013):

- *Distributed Denial of Service (DDoS)*, in which collections of hundreds or thousands of compromised machines are coordinated to simultaneously send floods of bogus traffic toward a target, completely overwhelming the target's resources, or the resources of the target's network;
- *Self-propagating Malicious Code, or Worms*, which compromises hundreds of thousands of Internet hosts in a matter of a small timeframe (usually in the scale of hours) allowing mass control to further coordinated network attacks (US-CERT, 2018);
- *Attacks on Internet Essential Services* refer to those which attempt to subvert the key components of the Internet's underlying infrastructure (e.g., Domain Name System (DNS), BGP routing);
- *Attacks on Large-scale Services* take advantage of publicly accessible Internet services, e.g., web servers, Content Distribution Network (CDN), online game servers. Attackers make use of their large deployment scale to open a new front of highly automated attacks exploiting the lack of network security configurations and outdated software versions.

Figure 2.2: IP v4 and v6 headers. An attacker illicitly impersonates another machine by manipulating IP packets. IP Spoofing involves modifying the packet header with a forged (spoofed) source IP address, a checksum, and the order value.



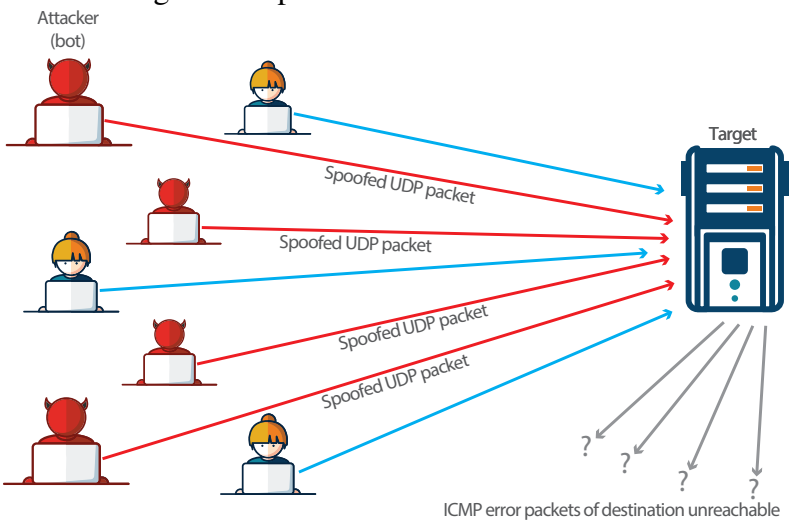
Adapted from: (CISCO, 2006).

These kind of large-scale Internet attacks are usually challenging to counter because of the difficulties in tracing them back (in some cases even impossible) or deploying widespread defensive measures. The diversity of exploits attests both to the continued threat of spoofing-based attacks as well as the ability to spoof on the Internet.

Following, we illuminate two attack strategies based on the method employed to set the IP source address of the packets (depicted in Figure 2.2) on such attacks: random and selective source addresses.

Random Spoofing. The strategy behind the attacks carried out using this pattern employ a wide range of source IP addresses leveraging the whole IP Address Space. This results in flood attacks in which a large number of packets of a given protocol or a combination of protocols (e.g., User Datagram Protocol (UDP), Hypertext Transfer Protocol (HTTP), or Internet Control Message Protocol (ICMP)), are sent to a target with the aim of overwhelming that device's (or targeted network) ability to process and respond, depleting all its resources. Figure 2.3 shows a type of Distributed Denial of Service (DDoS) attack using such a pattern. An attacker (or a botnet) will spoof the source IP address of the UDP packets, impeding the attacker's true location from being exposed and potentially saturated with the response packets from the targeted server. As a result of the target server utilizing resources to check and then respond to each received UDP packet, the target's resources can become quickly exhausted when a large flood of UDP packets are received, resulting in denial-of-service to normal traffic (blue in Figure 2.3).

Figure 2.3: UDP flood attack – Random spoofing. Employs a wide and random range of source IP addresses during attack operation.

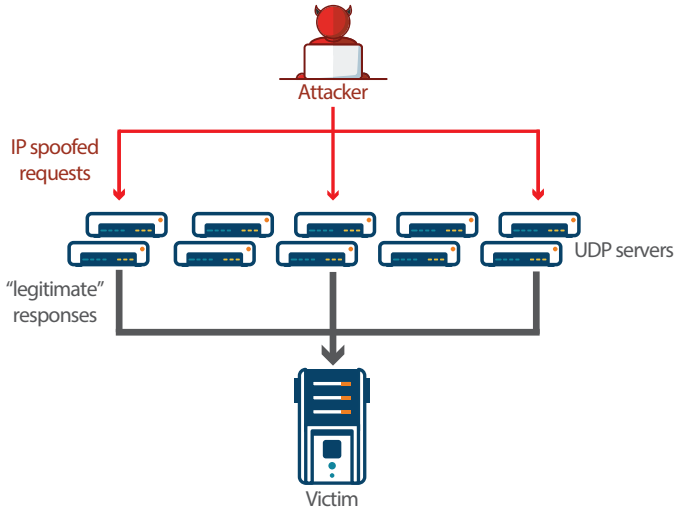


Source: by author (2019).

Selective Spoofing. This pattern is normally employed when there is a specific target to attack, for example, it can be a device, a network service, or an entire facility/infrastructure. Differently from random spoofing, it has a set of specific requirements. First, it requires a protocol vulnerable to reflection/amplification. Second, a list of reflectors – i.e., servers that support the vulnerable protocol. Third and last, the victim IP address or prefix. The attacks using this pattern then require selective spoofing of source IP addresses

of victims. Figure 2.4 illustrates how it works. The attacker sends fake UDP requests, which contains spoofed source IP address; however, the attacker set the victim's IP address in the source IP address field. The spoofed packets traverse the Internet and eventually are delivered to the reflector servers. The reflector server receives the fake packet and sends the response in good faith. The response, though, is directed to the victim. The victim will end up receiving a large volume of response packets it had never requested. The responses delivered to victim might be larger than the spoofed requests (hence *amplification*).

Figure 2.4: Reflection attacks (often called "amplification attacks") – Selective spoofing. Relies on reflectors, servers or Internet services that present vulnerable protocols.



Source: by author (2019).

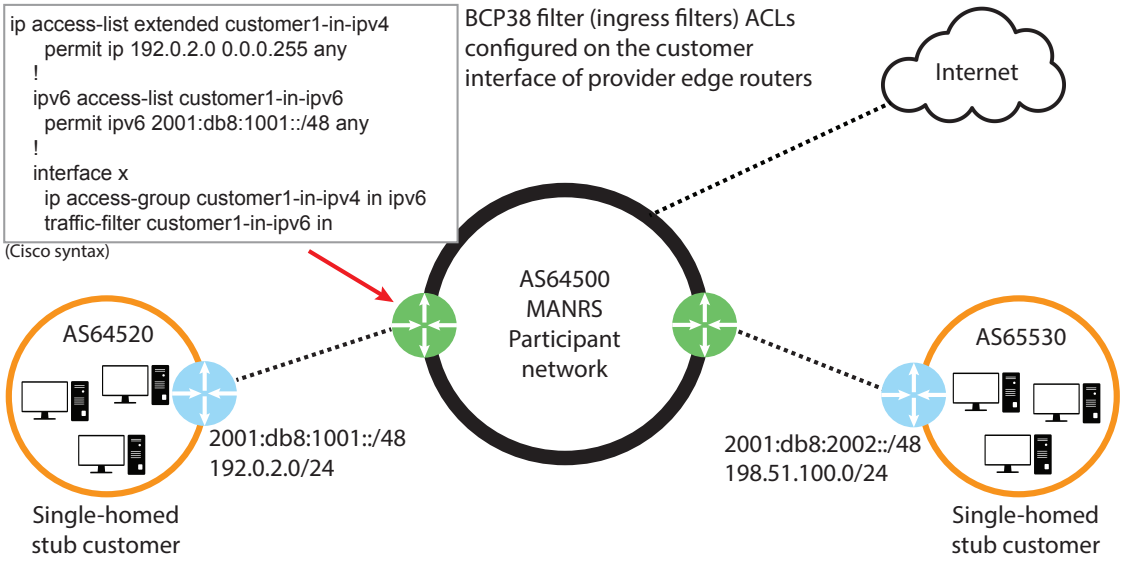
Moreover, worth to say that the multi-vector pattern, i.e., the combination of both strategies, is also common nowadays (MAJKOWSKI, 2018b; VERISIGN, 2018).

2.2 Source Address Validation (SAV)

The first initiative to tackle the spoofing problem came from the IETF community in the early 90s. At that time the IETF introduced the Best Common Practices (BCPs) 38 and 84 (FERGUSON; SENIE, 2000; BAKER; SAVOLA, 2004) – to protect against source spoofing, frequently referred to as Source Address Validation (SAV). In these documents, Ferguson et al. (2000) and Baker et al. (2004) described how to prevent inter-domain spoofing by using Ingress filtering. Ingress filtering is a technique that verifies which prefix of an IP source address routes to the network from which the packet was received. Network operators implement SAV by using *ingress filters* in routers, as illustrated in Figure 2.5, dropping packets with source addresses outside the locally valid address space before they

enter the global Internet. The key insight in ingress filtering is simplicity: the decision of whether to accept or to reject an IP source address can be made solely based on the information available from routing protocols. In practice, this is achieved by deploying Access Control Lists (ACLs) that only allow traffic with source IP addresses covered by specific prefixes to enter the network. The router maintains a continuously updated list of all prefixes for which it is allowed to accept traffic on a certain interface, from a specific peer. Traffic with IP addresses that are not covered by these prefixes will be dropped before entering the network. In Figure 2.5, the optimized ACL strategy would be to place an explicit permit filter on the customer interface. Explicit permit filters permit specific address ranges and then deny all else. For example, if the operator's customer (AS64520) is allocated 192.0.2.0/24, the BCP 38 ACL applied at the edge router of AS64500 would permit all source addresses from 192.0.2.0/24 and then deny all packets whose source address is not 192.0.2.0/24.

Figure 2.5: Networks should implement anti-spoofing techniques know as SAV on the customer interface of provider edge routes. Allow only packets with source IP addresses from the customer's networks (2001:db8:1001::/48, 2001:db8:2002::/48, 192.0.2.0/24, 198.51.100.0/24)



Adapted from Mutually Agreed Norms for Routing Security (MARNS) (ISOC, 2020a).

Unfortunately, operators responsible by the network's management do not cooperate, and many ASes till to this day do not implement the cited BCPs. They claim that the action of installing filters costs money and that their personnel is not capable of installing those filters (MCCONACHIE, 2014). Therefore it makes economic sense for these network operators not to install filters – *"no one is attacking my network, that is someone else's problem!"* (MCCONACHIE, 2014). The profile of the situation fits in the classic Tragedy

of the Commons problem (HARDIN, 1968) when we have a shared-resource system where individual users acting independently according to their self-interest behave contrary to the common good of all users by depleting or spoiling that resource through their collective action.

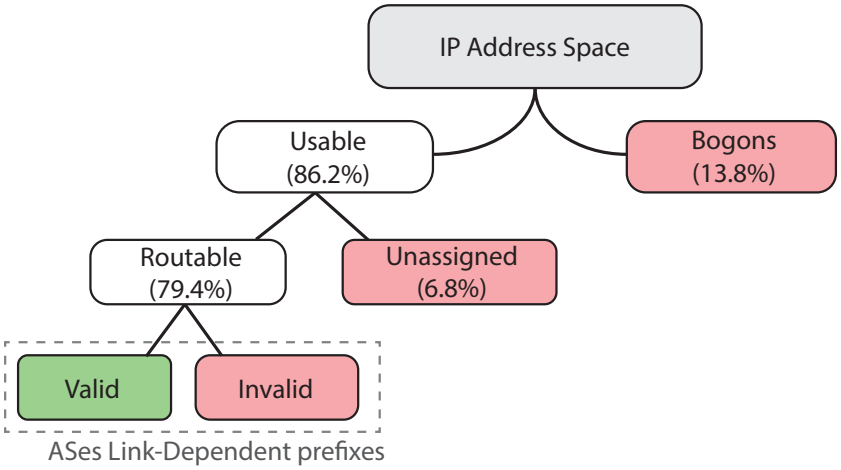
Even though other distinct initiatives appeared along the decades proposing new ways to tackle the Spoofing problem (see more in Chapter 3), the SAV filtering (BCPs 38 and 84 (FERGUSON; SENIE, 2000; BAKER; SAVOLA, 2004)) continues to be the best viable solution (ISOC, 2020b). However, not all networks care of implementing and keeping such configurations in place. Given the situation of lack of compliance with SAV, a joint community effort led by Internet Society (ISOC) with MANRS project (ISOC, 2020b; ISOC, 2018) and also endorsed by the Cybersecurity Tech Accord (Tech Accord, 2018) is tackling the problem in distinct ways. They are educating operators worldwide of the importance of SAV, helping the operators to identify networks lacking SAV, and working to establish policies to incentivize operators to adopt such practices.

2.3 Address Space Fundamentals

The IP address space is divided into multiple blocks by distinct institutions before being assigned to end users (e.g., ISPs, corporations, or academic institutions). The assignation hierarchy works as follows. The Internet Assigned Numbers Authority (IANA) is in charge of distributing /8 prefixes to Regional Internet Registry (RIR). There are five RIRs (AFRINIC, APNIC, ARIN, LACNIC, RIPE NCC), each responsible for a different geographical area. In turn, RIRs allocate IP address space to Local Internet Registry (LIR), such as ISPs, large companies, etc. LIRs distribute IP address blocks at the local level, i.e., to end users. That said, Figure 2.6 illustrates how the IP address space is exploited in the Spoofing Problem.

The global BGP routing table now contains over 790k distinct IPv4 prefixes (APNIC, 2020). These include invalid prefixes that should not be globally announced, such as the **bogons** (IETF reserved or Martians) and **unassigned** IP space. While **bogon** prefixes (MOSKOWITZ et al., 1996; WEIL et al., 2013; COTTON et al., 2013) can be used in specific cases, such as private networks, loopback interfaces, some are reserved for future use. The **unassigned** prefixes have been allocated to an RIR, but not assigned by that RIR to an end-user (e.g., an ISP). IANA keeps a list of allocated, and reserved prefixes (IANA, 2018b; IANA, 2018c) and each RIR has its own list of prefixes assigned

Figure 2.6: IP Address Space organization, reflecting on the usage for inter-domain traffic exchange. Highlighted in red are the categories which enable the occurrence of spoofed traffic in the wild, and in green is the only set of valid IP addresses that should be used. Percentages relative to the total IPv4 Address Space in April, 2019 (IANA, 2018b).



Source: by author (2019).

to end-users (NRO, 2020). In the routable address space, we have the *dynamic prefixes* one AS should expect when exchanging traffic. However, different from the previous cases (i.e., bogon and unassigned), there is no global registry of routable prefixes per AS; in practice, we must heuristically infer these addresses (using the concept of cones - details in §2.4). Therefore, from the perspective of an AS, these prefixes can be invalid or valid. **Invalid** when packets are sent without respecting interconnection business agreements, i.e., using prefixes that do not belong to the networks (or the ASes involved do not have agreed to use in) exchanging the traffic, or **Valid** when the packets exchanged by the networks are in accordance with its established neighboring network relationships. Next, we discuss concepts that help in the definition of the set of prefixes that are link-dependent (§2.4).

2.4 AS Relationships and Customer Cones

Internet inter-domain routing is a collaborative effort between ASes. ASes negotiate contractual agreements to define their business relations and impose technical restrictions on traffic exchange (MARCOS et al., 2018). On the Internet, connectivity does not imply traffic reachability, which is fundamentally determined by the business relationships between ASes (KATZ-BASSETT et al., 2008; FAYAZ et al., 2016). The AS business relationships are coarsely divided into three primary classes – customer-provider (c2p, p2c), peering (p2p) and sibling (s2s) (LUCKIE et al., 2013). Figure 2.7 illustrates them and their subtleties.

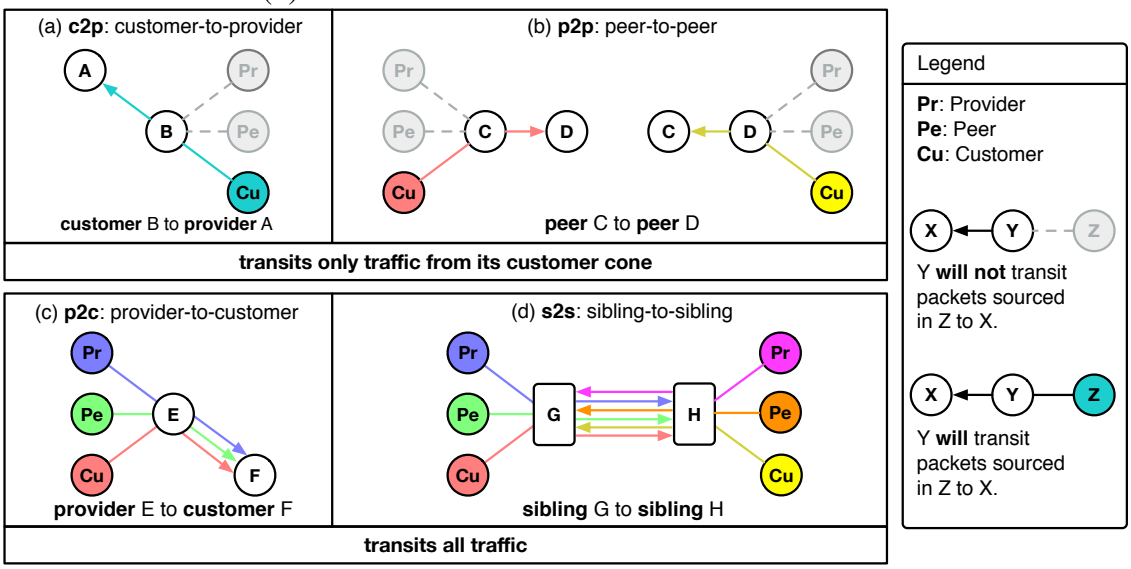
1. **Transit relationship, which includes Customer-to-Provider (c2p) and Provider-to-Customer (p2c)**. It is established when an AS (customer) pays a better-connected AS (provider) to transit traffic with the Internet. The providers act as a gateway to the rest of the Internet. An AS can have multiple providers for purposes of resilience and load balancing. Such ASes are called *multihomed*.
2. **Peering relationship (p2p)**. Peer-to-Peer (p2p) relationships allow two ASes to freely exchange traffic between themselves and their customers as a means to avoid the cost of sending traffic through a provider. These interconnections can be settlement-free or paid, depending on who benefits the most from the agreement, e.g., traffic imbalance or route diversity from larger ISP.
3. **Sibling relationship (s2s)**. Sibling-to-Sibling (s2s) relationships represent the case where a single organization may own and operate multiple ASes, and may transit packets received from any source. This case enables the establishment of s2s relationships, in which the links connect two or more ASes that belong to the same administrative entity (CAIDA, 2018b) without any cost or routing limitations.

The dynamics of the AS ecosystem are determined both by external factors (e.g., the state of the global economy or the popularity of new Internet applications) and by complex incentives and objectives of each AS. Specifically, ASes attempt to either optimize their utility or financial gains by dynamically changing, directly or indirectly, the ASes they interact with (DHAMDHERE; DOVROLIS, 2011a). Consider the following examples. For a transit provider, the objective may be to maximize its profit, and it may approach this goal through competitive pricing and selective peering. On the other hand, if we consider a Content Provider, the objective may be to have highly reliable Internet access and minimal transit expenses, and to achieve that, it may pursue through aggressive multihoming and open peering policy.

Aligned with such dynamics, a fundamental model to explore the Internet interconnection ecosystem is the **Customer Cone**. This model results from the AS relationship graph [1] allowing the comparison across ASes, as well as the definition of the expected source addresses one should look forward to see in valid traffic. The Customer Cone constraints which source IP addresses one should see in valid inter-domain traffic transiting from a Customer to its provider, or between peers. Figure 2.7 illustrates the subtleties: an AS in a c2p or p2p relationship with another AS should only send packets with a source address from within its customer cone – respectively, (a) and (b) in Figure 2.7. In contrast,

[1] When discussing the State-Of-The-Art in Chapter 3 - §3.2, we provide more formal details.

Figure 2.7: The Customer Cone constrains the set of source addresses expected in valid inter-domain traffic transiting an AS behaving rationally in a c2p or p2p relationship. In the c2p relationship shown in (a), B transits traffic from its customers to A, but not its peers and providers. Similarly, in the p2p relationship shown in (b), C only transits traffic from its customers to D (likewise, from D to C). However, as shown in (c), the p2c relationship does not constrain the source addresses transited by E to F, and neither does the s2s relationship between G and H in (d).



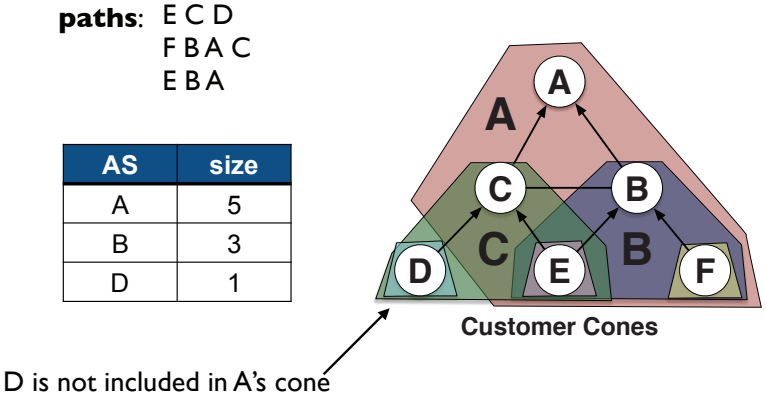
a link between a provider to its customer or between two siblings may forward packets with *any* routed source address – (c) and (d) in Figure 2.7.

More precisely, a Customer Cone is defined as the set of ASes that a given AS can reach using its customer (*p2c*) links. This includes the direct (customers directly connected to the AS) and indirect customers (customers of direct customers, semi-recursively). Looking specifically at the **AS-level Customer Cone**, we define an AS *A's* AS-level Customer Cone as the AS *A* itself plus all the ASes that can be reached from A following only *p2c* links in BGP paths observed. In other words, *A's* Customer Cone contains A, plus *A's* customers, plus its customers' customers, and so on (LUCKIE et al., 2013; DIMITROPOULOS et al., 2007).

Figure 2.8 illustrates the concept (CAIDA, 2018a). With an input set of three BGP AS paths [2], extracted from publicly available sources (RIPE, 2018; ROUTEVIEWS, 2018), have been inferred the relationship between ASes and the Customer Cone for each AS. The size of the Customer Cone of an AS reflects the number of other elements (ASes, IPv4 prefixes, or IPv4 addresses) found in its set. An AS in the Customer Cone is assumed to

[2] A BGP AS path is the sequence of autonomous systems that network packets traverse to get to a specified router. AS numbers are assembled in a sequence that is read from right to left. For example, for a packet to reach a destination using a route with an AS path 5 4 3 2 1, the packet first traverses AS 1 and so on until it reaches AS 5. In this case, AS 5 is the last AS before the packet destination.

Figure 2.8: Customer Cone Organization. An AS Customer Cone contains the set of ASes we observe the AS announce to its peers or providers. In practice, this is the set of ASes it can reach through its customers.



Adapted from: (Huffaker, B., 2018; CAIDA, 2018a)

pay, directly or indirectly, for transit, and provides a coarse metric of the size or influence of an AS in the routing system. The example depicts several AS Customer Cones: ASes D, E, and F all sit at the bottom of the hierarchy, and so, only have a single AS in their cone. Both C and B tie with 3 ASes. Note that B and C both have E in their respective cones. A is ranked at the top of the hierarchy with 5 ASes in its Customer Cone (note that D is not included in A's cone given the input BGP AS paths observed).

The size of the AS Customer Cone can also reveal its importance in the Internet's capital and governance structure (LUCKIE et al., 2013; LODHI et al., 2015). According to the number and type of links, an AS can be categorized as Tier-1, Tier-2, Content Provider (CP), Internet Service Provider (ISP), and stubs. At the top of this hierarchy are the Tier-1 networks, which do not pay for transit to upstream providers at all; instead they peer with each other to provide connectivity to all destinations in the Internet. Tier-2 networks are also large ASes, mainly providing IP transit to other ASes, but not for free. The Content Providers (CP) are the global networks that focus primarily in transit traffic between content generators and end-users. To achieve low cost and low end-to-end delay they seek to establish *p2p* relationships, but they also have providers for redundancy and fail-over purposes. ISPs can have both customers and providers, and usually, their coverage is at the national or regional level. The ISPs provide Internet connectivity to stub ASes or to end-hosts. The ISPs that exclusively operate as access providers for end-hosts are called eyeballs (MA et al., 2008). At the bottom of the hierarchy are stub[3] ASes which do not have their own customers and pay providers to reach all destinations in the Internet (e.g., universities, research networks).

[3] A stub Autonomous System is an AS that is connected to only one other AS.

2.5 IXPs as Observatories

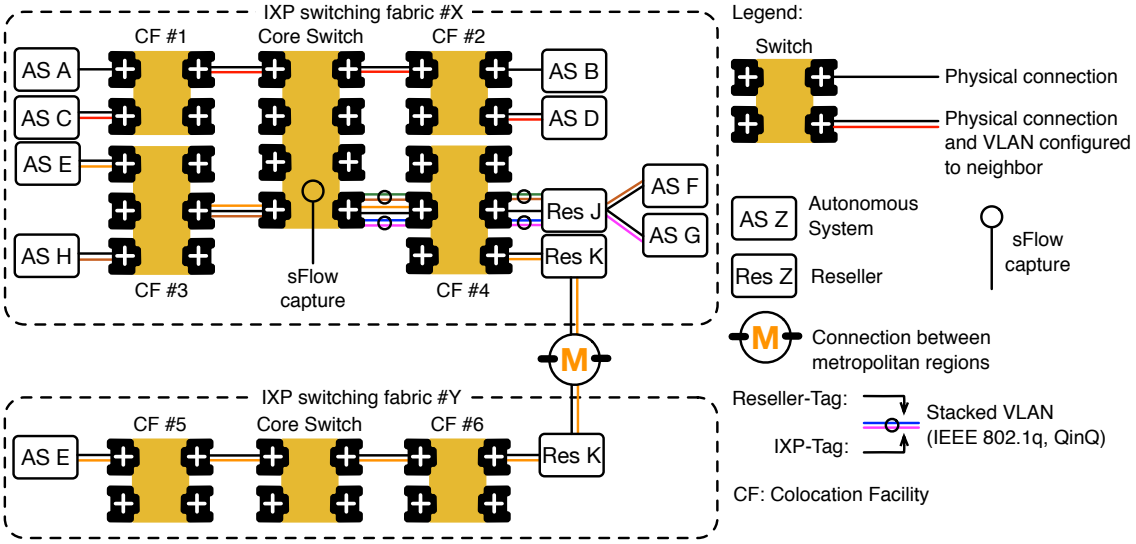
Internet Exchange Points (IXPs) and their co-located facilities act as key enablers for inter-domain connectivity worldwide. In fact, they have an increasingly central role on the Internet dynamics as a whole (AGER et al., 2012; CHATZIS et al., 2013). IXP infrastructures carry traffic in the range of Terabits per second (IX.br, 2020; DE-CIX, 2018; AMS-IX, 2018), their membership is constantly growing (KLöTI et al., 2016; LODHI et al., 2014), and support hundreds of thousands of peerings (WOODCOCK; FRIGINO, 2016).

The combination of the aforementioned properties with the dense connectivity, and increasing AS presence at IXPs worldwide (AGER et al., 2012; BRITO et al., 2016) make them strategic vantage points to investigate and deploy new methodologies aiming to increase the overall Internet infrastructure security worldwide. Following, we explain IXPs' organization and highlight characteristics of these network infrastructures.

IXPs are attractive vantage points to observe signals of SAV deployment, as hundreds of ASes may be present at a single logical location. The IXP operator assigns each member a unique IP address from a prefix controlled by the operator. The member assigns the address to its router interface connected to the IXP, which is used to establish BGP routing with other members. When a member AS's router transmits a packet across the Ethernet switching fabric, the source and destination MAC addresses in the Ethernet frame uniquely identify the AS pair exchanging the packet, and its direction.

Figure 2.9 illustrates the architecture of many modern IXPs (Euro-IX, 2019b; IX.br, 2020; DE-CIX, 2020; AMS-IX, 2020b; LINX, 2020; IX.br Forum 12, 2019). The figure contains two separate IXPs and their switching fabrics #X and #Y, with a core switch for each IXP. While some IXPs may consist of a single core switch where participants interconnect, operators achieve the scale of modern large IXPs by placing switches at distinct physical colocation facilities, any of which can serve as an IXP attachment point. The figure shows that the switches are adjacent, but in practice colocation facilities are usually in different buildings (GIOTSAS et al., 2015b; MOTAMEDI et al., 2019). IXP operators often use sFlow (P. Phaal, S. Panchen, and N. McKee, 2001) or NetFlow (CLAISE, 2004) to collect traffic flow statistics. A comprehensive view of all traffic from all services at the IXP would require flow data captured from all switches in the switching fabric, as traffic between participants at a single colocation facility will not travel to the core switch.

Figure 2.9: Illustration of the architecture of modern IXPs. They typically construct a switching fabric using a core switch that interconnects other switches located in remote Colocation Facilities. ASes typically connect to a switch located in a Colocation Facility, and can form bilateral peering relationships with neighbors. These ASes may request a VLAN to isolate their traffic from other members at the IXP. Resellers can provide services such as remote peering and layer-2 transport.



Participants can exchange traffic directly across the switching fabric in a bilateral session. In Figure 2.9, ASes A and B exchange traffic directly. However, modern IXPs often use VLANs to provide logical isolation between different types of interconnection (CHATZIS; SMARAGDAKIS; FELDMANN, 2013; Euro-IX, 2019a). For example, an IXP may provide a route server, but only offer that route server on a specific VLAN. Similarly, traffic between two participants may be sufficiently sensitive or high volume that members request a VLAN from the IXP to isolate their communications (DE-CIX, 2019; AMS-IX, 2019b; LINX, 2019b). In Figure 2.9, ASes C and D exchange traffic in their own isolated VLAN.

To foster IXP growth and enable more networks to interconnect, IXPs have supported resellers, which provide value-added services at an IXP, such as remote peering and layer-2 transport (CASTRO et al., 2014; NOMIKOS et al., 2018; MEGAPORT, 2019; IX Reach, 2019). A reseller provides remote peering services so that an AS which is not physically present at a colocation facility can still reach other members at the IXP, without the AS incurring colocation facility fees or port charges from the IXP operator. These resellers require some cooperation with the IXP, e.g., (LINX, 2019a; AMS-IX, 2019a). The IXP assigns the remote peers any VLAN tags they require to participate at the exchange as local members do.

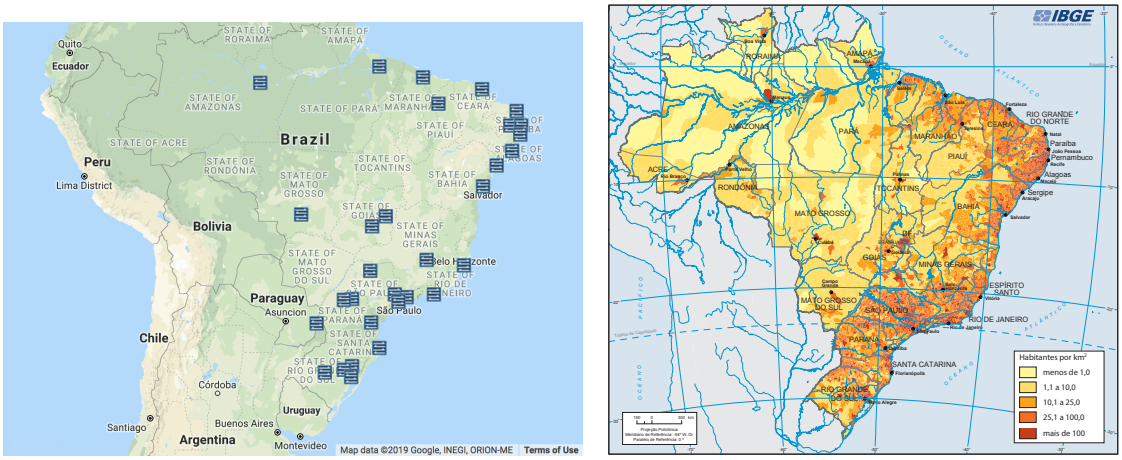
An IXP may use different technical approaches to support remote peering providers (CASTRO et al., 2014; NOMIKOS et al., 2018; IX.br Forum 12, 2019). A reseller can bridge Ethernet networks so that the MAC address of the customer router's interface will uniquely identify the origin of traffic in the peering fabric. A second approach is for a reseller to push a tag (reseller-tag) to uniquely identify their specific customer AS to the IXP, so that the MAC address of the Ethernet frame corresponds to the reseller's router. Figure 2.9 illustrates this second approach, where reseller J allows customer ASes F and G to reach other members. When the reseller transmits these packets into the IXP, the reseller also pushes a tag (reseller-tag) to uniquely identify their specific customer AS to the IXP. The IXP bridges traffic into the IXP switching fabric by removing the outer-most reseller-tag while keeping the IXP-tag. In Figure 2.9, the sFlow tap sees the IXP-tag and the MAC address of the reseller, which uniquely identifies the specific AS who sent the packet.

A reseller can also provide remote peering to members colocated at one IXP who want to reach members in a different IXP, reaching the reseller across the peering fabric. Figure 2.9 shows a more complicated example, where AS E bridges its network between metropolitan regions using the services of a reseller (K) present at both IXPs.

Overview of the IX.br Ecosystem

In this work, we apply our method and analyses taking into consideration the largest IXP ecosystem in Latin America – the **Brazilian IX.br Ecosystem** (IX.br, 2020).

Figure 2.10: Brazilian IX.br Ecosystem Overview.



(a) A total of 31 IXPs distributed in the Brazilian national territory. Blue square-box markers shows the existing IXPs.

(b) Map of the Population Distribution in Brazil from the latest official Population Censuses.

Adapted from: (IX.br, 2020; IBGE, 2013).

The IX.br ecosystem comprehends a total of 31 independent IXPs and 110 colo-

cation facilities (colos), with current peak traffic in the scale of more than 8 Tbps in aggregate (IX.br, 2020). These IXPs are located at strategic cities along the country as seen in Figure 2.10(a). Back in 2016, they were a total of 26 different locations and 102 colos, a significant increase in an already large ecosystem. The presence and location of such infrastructures were driven essentially by the size of the population in each city (as can be seen in Figure 2.10(b), economic growth, local companies interest, and the development of landing points from international submarine cables (TELEGEOGRAPHY, 2019; ITU, 2019).

According to the latest data available (as of December 2019), there are 7401 ASes registered in Brazil (Registro.br, 2020; Sirota, J., 2019; Sirota, J., 2018), representing $\sim$73% of total ASNs assignments (10171 ASNs) from LACNIC region (NRO, 2019). Data from the IX.br coordinator team (IX.br, 2020) show that around 2.5k unique ASes ($\sim$34% of the total Brazilian ASes) are already connected to, at least, one of the IXPs in the Brazilian Ecosystem.

3 RELATED WORK

In this chapter, we first discuss the most recent advances in measuring deployment of SAV (§3.1). Afterward, we revisit the most prominent approaches regarding both, the AS relationships inference algorithms and the cone construction inference methods (§3.2). These algorithms will be used to define the set of IP addresses a given AS should expect to receive traffic from and can generate traffic to (we present the methodology details in Chapters 4 and 5). We complete the chapter highlighting the goal of our proposal (§3.3).

3.1 Measuring Deployment of SAV

Many academic research efforts have described techniques, and even alternative architectures, to promote deployment of SAV (Chapter 2 - §2.2) (DUAN; YUAN; CHANDRASHEKAR, 2006; YAAR; PERRIG; SONG, 2006; LIU et al., 2008; LIU; BI; VASILAKOS, 2014). Other studies suggest improvements to protocols to mitigate the impact of spoofed packets. Rossow identified and studied protocols prone to amplification attacks (ROSSOW, 2014). Kührer et al. suggest approaches that help to reduce the number NTP servers vulnerable to amplification by 92% (KüHRER et al., 2014). Zhu et al. (ZHU et al., 2015) suggest connection-oriented DNS to prevent the exploitation of open DNS resolvers, e.g., for amplification attacks (KüHRER et al., 2015).

Fewer efforts have tried to empirically measure SAV compliance for networks attached to the global Internet. In 2005, Beverly, *et al.* developed a client-server technique to allow users to test networks to which they are currently attached (BEVERLY; BAUER, 2005), and operationalized a platform to track trends over time (BEVERLY et al., 2009; CAIDA, 2018c). However, this active measurement method relies on users downloading and running client software, inducing a sparse data collection, which may not adequately capture or represent the full dimensionality of the problem. To overcome this requirement for a vantage point in every network, over the last few years researchers have investigated opportunistic creative techniques to infer lack of SAV in other macroscopic Internet datasets.

In 2013, Dainotti *et al.* described a technique to identify spoofed traffic in unsolicited Internet Background Radiation (IBR) traffic (DAINOTTI et al., 2013; DAINOTTI et al., 2016), based on the assumption that unrouted address should not appear as source addresses in legitimate packets. It attempts to remove spoofed traffic preventing it from

corrupting the signal otherwise extracted from IBR traffic. This method cannot identify networks sending the spoofed traffic, and thus failing to help deploy SAV.

In 2017, Lone *et al.* reported a technique to infer evidence of spoofed traffic in massive traceroute archives, based on the knowledge that an edge network should not appear to provide transit in a traceroute path (LONE et al., 2017). This method is also limited by whatever appears in the traceroute archives, as well as by the inconsistent addressing conventions used in traceroute implementations (MARDER et al., 2018). In 2018, Lone *et al.* experimented to boost CAIDA's Spoofer project deployment with the help of crowdsourcing marketplaces (Lone et al., 2018), which worked well while recruiting and remunerating workers.

Most closely related to our study, in 2017 Lichtblau *et al.* used a large European IXP as a vantage point for inferring which networks at the IXP had not deployed SAV (LICHTBLAU et al., 2017). For each member at the IXP, their method infers a set of IP prefixes containing addresses that may legitimately appear in the source field of IP packets crossing an IXP. They infer that a member AS that sends a packet into the IXP switching fabric with a source address outside of those prefixes has not deployed SAV. They argued against using AS relationships and AS Customer Cones which they claimed did not address asymmetric routing. However, their method did not consider ASes forming Customer-Provider or Sibling relationships at the IXP, where all routed addresses may be legitimate source addresses in IP packets crossing an IXP – (c) and (d) in §2.4, Figure 2.7. In these cases, there is no way to infer SAV deployment across these links at the IXP.

3.2 AS Relationships and Customer Cones Inferences

Internet studies demand knowledge on the relationships between ASes. More specifically, in our method, knowledge of the business relationships between ASes is essential to define the set of IP addresses a given AS should expect to receive traffic from and can generate traffic to in a given link. However, most ASes try to hide their business relations. Service providers consider the policy details of their business relationships as proprietary information and do not generally make them public (LUCKIE et al., 2013; OLIVEIRA et al., 2010; GAO, 2001). Therefore, Internet researchers have to rely on indirect AS relationship inference algorithms in order to build a picture of Internet business structure. In the last years, researchers have introduced a number of algorithms to infer the AS relationships and a few cone construction inference methods. Following, we review

some of the most prominent research on the field. We start the section by discussing efforts aimed to infer AS relationships. Then, we review studies which proposed cone construction methods. In this section, we follow a chronological order of publications.

AS Relationships Inferences. Gao's seminal work (GAO, 2001) inspired many researchers to seek approaches to inferring ISP business relationships using information from publicly BGP routing tables. Gao used the concept of valid paths as the basis for her inference heuristic and identified the top provider in a given path based on AS degree (the number of ASes connected to a given AS). Her solution relies on the assumption that BGP paths are hierarchical, or valley-free, i.e., a customer route can be exported to any neighbor, but a route from a peer or a provider can only be exported to customers. The valley-free rule describes a typical AS path, i.e., aims to prevent an AS from providing free transit either to their providers or peers. Most reachable paths which are valid for traffic routing are valley-free, as they serve the business interest of ASes – to minimize operation cost and maximize revenue.

After Gao's work, several studies have been developed to infer AS relationships, therefore, in the following, we will highlight the ones widely cited in the Internet research community. Subramanian *et al.* (SUBRAMANIAN et al., 2002) relaxed the problem by not inferring sibling links and provided a mathematical formulation based on the concept of valid paths. They formalized Gao's heuristic into the Type of Relationship (ToR), a combinatorial optimization problem. Assuming maximization of the number of valid paths as a natural objective, they formulated the AS relationship inference problem as: given an undirected graph G derived from a set of BGP paths P, assign the edge type (*c2p* or *p2p*) to every edge in G such that the total number of valid paths in P is maximized. They conjectured that the ToR problem is NP-complete and developed a heuristic-based solution. Following, Di Battista *et al.* (BATTISTA; PATRIGNANI; PIZZONIA, 2003) proved that the ToR is indeed NP-complete. More importantly for practical purposes, they demonstrated that *p2p* links cannot be inferred in the ToR problem formulation and developed mathematically rigorous approximate solutions to the ToR problem but inferred only *c2p* and *p2c* links. Still in line with previous approaches, Dimitropoulos *et al.* (DIMITROPOULOS et al., 2007) proposed a solution by reducing the multi-objective optimization problem to the MAX2SAT problem (a boolean algebra problem). However, MAX2SAT is NP-hard and their implementation does not scale for recent AS graphs.

At that time, Xia and Gao (XIA; GAO, 2004) first introduced techniques to evaluate the accuracy of the existing algorithms. They showed that neither (SUBRAMANIAN et al.,

2002) nor (GAO, 2001) technique offers a solution to the problem of reliable identification of *p2p* links due to their low accuracy. Given the results, Zhang *et al.* (ZHANG et al., 2005) propose a way to incorporate other sources than BGP tables, including Route Servers, Looking Glasses, and Internet Routing Registries (IRR) to compile a topology. Their algorithm starts with a set of ASes inferred to be in the Tier-1 clique, then infers links seen by these ASes to be *p2c*; all other links are *p2p*. Despite significant progress, much remains unknown in terms of the quality of the inferred AS connectivity due to the lack of ground-truth data for validation of the results obtained.

Researchers then performed analysis to quantify the incompleteness of the observed AS-level connectivity as seen by the commonly-used vantage points/datasets (COHEN; RAZ, 2006; OLIVEIRA et al., 2010). Their findings showed that none of the available topology discovery methodologies are able to capture the complete inter-domain topology. The incompleteness of the resulted topologies is mainly a result of policy-based routing that restricts the propagation of certain link types (e.g., peering and backup relationships). Given its importance to the field of Internet research, Luckie *el al.* (LUCKIE et al., 2013) revisited the science of AS relationship inference and gave particular attention to validation. They presented and validated to an unprecedented level a new algorithm for inferring AS relationship using publicly available BGP data (RIPE, 2018; ROUTEVIEWS, 2018). The authors assembled the largest source of validation data for AS relationship inferences to date, validating 34.6% of their relationship inferences, finding the c2p and p2p inferences to be 99.6% and 98.7% accurate, respectively.

Cone Construction Inferences. As introduced previously in §2.4, to explore the Internet interconnection ecosystem, its dynamics, as well as to support distinct Internet analysis, the concept of Customer Cones was proposed (DIMITROPOULOS et al., 2007). However, only in 2013 the concept gained more strength, when Luckie *el al.* (LUCKIE et al., 2013) used their new AS relationship inferences results (discussed above) to propose distinct methods to build the Customer Cone of each AS. The reason of multiple methods to infer the customer cone of a given AS is due to ambiguities inherent in BGP data analysis. The authors of that paper used the customer cone as a metric of influence to study top selected ASes and the Internet topology flattening effect. Nowadays, they release periodic results to the community through an online system (and an API) called AS-Rank (CAIDA, 2018a).

The results achieved with the Customer Cone model have been proven stable, being applied to the study of the Internet, from Internet topology mapping (LUCKIE et al.,

2014; MARDER et al., 2018; GIOTSAS et al., 2015b), inter-domain routing policies (ANWAR et al., 2015) to Internet topology evolution (DHAMDHERE; DOVROLIS, 2010; DHAMDHERE; DOVROLIS, 2011b). But most recently, Lichtblau *et al.* (LICHTBLAU et al., 2017) (recall §3.1) proposed the Full Cone algorithm to build new cones. They argued against using AS relationships and AS Customer Cones which they claimed did not address asymmetric routing. The authors then proposed an algorithm to build cones without considering the AS relationships classes (Chapter 2 - §2.4) and they use their resulting cones to spoofing traffic detection. Following, in the next chapter, we discuss in details both cone constructions, i.e., Full Cone and Customer Cone, and their application to spoofed traffic classification.

3.3 Summary

Information presented thus far, on Sections 3.1 and 3.2, are summarized in Tables 3.1 and 3.2 considering only the most related approaches to ours. As one can observe from the state-of-the-art, few efforts (BEVERLY; BAUER, 2005; LICHTBLAU et al., 2017) have tried to empirically measure SAV compliance for networks attached to the global Internet (Table 3.1). For years the Spoofer Project (BEVERLY; BAUER, 2005; LUCKIE et al., 2019) was the only methodology proposed and developed to help identify precisely networks lacking SAV compliance. Even though, the project kept growing, adding support to IPv6 checking and handling some NAT scenarios (CAIDA, 2018c), it still suffers from sparse visibility on the global Internet due to the huge challenge that is to obtain the collaboration from networks to install and run their spoofer client periodically.

Most closely related to our study is (LICHTBLAU et al., 2017), as discussed previously. Despite their potentialities, their proposal does not tackle the problem properly. They aim to minimize false spoofing detections instead of dealing with the network complexities that exist in shared peering infrastructures (Chapter 4 - §4.2). To this end, they proposed and used the Full Cone inference method. The purpose of it is to infer the valid address space per AS, assuming as valid all BGP announcements and updates, besides the decision to argue against and ignore the distinct types of AS relationships (Chapter 4 - §4.1). As a result, their decisions lead to drastically impact the precision of the resulting classification inferences (Chapter 6 - §6.3 and §6.6). In addition, there is no validation of their results. The authors did not released an official publicly shared codebase to enable research reproducibility.

In this thesis, we argue that broader visibility into the Spoofing problem lies in the capability to infer lack of SAV compliance from strategic vantage points, such as at Internet Exchange Points. Our proposal emphasizes the importance of the details underlying the cone construction method to detect not only spoofed traffic at an IXP, but the member of the IXP network sending it. We advocate that using public BGP information and aggregated traffic flow data to infer lack of SAV requires:

1. deep understanding of idiosyncrasies in IXP interconnectivity fabrics;
2. filtering of unverifiable traffic flows;
3. sanitizing input (BGP) data to filter erroneous or unverifiable paths;
4. identifying and using AS relationships to infer Customer Cones, and
5. tracking the legitimate source address space of each sending-receiving pair of ASes using the IXP, including accounting for routing dynamics, e.g., traffic engineering, policy changes, or asymmetric routing.

We show that these measures are key to accurately inferring spoofed traffic crossing the switching fabric of an IXP. Our approach classifies traffic traces (e.g., NetFlow (CLAISE, 2004), Sflow (P. Phaal, S. Panchen, and N. McKee, 2001)) leveraging information of the IXP peering fabric infrastructure and our constructed filters to reliably isolate spoofed traffic, and metrics to detect atypical traffic behaviors. Additionally, we make our proposal reproducible by enabling other researchers or network operators to obtain the same results under different conditions using our developed artifacts.

Table 3.1: Summary of related approaches to network SAV compliance identification.

Authors	Method	Objective	Target	IP Version	Base Technique	Address Space Definitions	Deployment	Results Visibility	Traffic Sanitization	Handle Network Complexity	Reproducibility
CAIDA Spoofer Project (BEVERLY; BAUER, 2005) (CAIDA, 2018c)	Active	Accuracy	Any network connected to the Internet	v4 and v6	Craft spoofed test packets	CAIDA reserved prefixes	On each AS being checked	Sparse	Not designed to perform traffic flow analyses	Handle NAT client scenarios	No
Full Cone and Spoofing Classification (LICHTBLAU et al., 2017)	Passive	Minimize false spoofing detections	Members connected to IXPs	v4	Inferences	Bogon prefixes, Full Cone inference algorithm	At an IXP	Localized	Does not filter unverifiable traffic	No	No
Spoofer-IX (this work, composed by the Prefix-Level Customer Cone and a Spoofing Classification Pipeline)	Passive	Accuracy	Members connected to IXPs, also enables Colocation Facilities (and more broadly all networks at the inter-domain)	v4	Inferences	Bogon prefixes, Full Bogons list, Prefix-Level Customer Cone inference algorithm	It offers different choices: per IXP, or Colocation Facility, or per individual switch at the peering infrastructure under analysis	Localized	Defines and filter a set of unverifiable flows	Yes, distinct cases (e.g., Remote Peering, VLANs, p2c, Transport Provider, siblings).	Yes

Source: by author (2019).

Table 3.2: Summary of related approaches to the IP Address Space inferences per AS.

Authors	Objective	AS-Path Sanitization	AS Relationship Inference	Maintaining Address Spaces	Asymmetric Routing	Traffic Engineering
CAIDA Customer Cone / ASRank (LUCKIE et al., 2013), (CAIDA, 2018a)	Metric of influence to study top selected ASes and the Internet topology flattening effect	Filters erroneous information	Correctly infers relationships – Provider-to-Customer, Customer-to-Provider, and Peer-to-Peer relationships.	Generate on a monthly basis, consuming, one RIB daily for the first-5-days. Produces inferred relationships between AS pairs and provider/peer observed AS cones.	yes	no
Full Cone (LICHTBLAU et al., 2017)	Inference of valid address space per AS in the Internet	Assumes all announcements and updates are valid	Assumes all relationships are equal, and bidirectional, i.e., all ASes share all prefixes they can reach with all, peers, customers or providers	Update information on a weekly basis, consuming 9-days of all available RIBs, and update files from public BGP collectors	no	no
Prefix-Level Customer Cone (this work, part of Spoofer-IX)	Inference of valid address space per AS in the Internet	Filters erroneous information	Leverages the AS inference algorithm and the Provider/Peer Observed Customer Cone from (LUCKIE et al., 2013) to build a novel Prefix-level Customer Cone.	Update information on a weekly basis – 7-days, consuming one RIB only a day from public BGP collectors	yes	yes

Source: by author (2019).

4 TACKLING METHODOLOGICAL CHALLENGES

In the previous chapter, we discussed the most prominent solutions to the spoofing problem, as well as the AS Relationships and Cone Construction Inferences algorithms used to build the set of valid IP address per AS. In this chapter, we describe the core of our methodology in the context of two complex groups of challenges to inferring spoofed traffic in IXP traffic data. The first one (§4.1) is determining which addresses are valid source addresses in traffic transiting a given neighbor AS, i.e., packets with a source address that are *In-cone* (potentially valid) for that AS. An incomplete set of valid addresses could yield false inferences of failure to deploy SAV, for that should a valid address appear in the observed packets but not be in the In-cone set, i.e., it will be *Out-of-cone* (potentially spoofed) for that AS. The second group of challenges (§4.2) is related to the analytical implications of modern IXP interconnection practices that can prevent the visibility of both topology and traffic. These practices complicate the analysis of which ASes exchanged traffic and their routing relationship. Once addressed these challenges, we describe in the next chapter how these different results fit together in our methodology that deals with IXP specifics.

4.1 Subtleties in Cone Construction

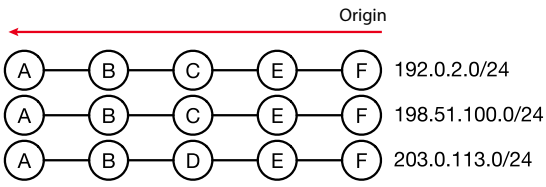
Inferring the set of valid source addresses for packets traveling from a specific AS to a specific adjacent AS at an IXP requires investigating a multidimensional parameter space. Precision in this process is crucial. Mistakenly excluding valid addresses allows certain valid packets to be classified as spoofing, therefore incorrectly inferring a given AS as not SAV compliant (which we call a false positive). Conversely, including invalid source addresses in the cone definitions allows spoofed packets to be misclassified as valid, therefore failing to detect lack of SAV compliance (false negatives).

As mentioned in Chapter 1, there is no global registry that contains ground truth on which addresses are valid source addresses for packets transited by an AS; instead, we must infer them from the best available public BGP routing data sources (RIPE, 2018; ROUTEVIEWS, 2018; PCH, 2020), even though these sources may contain spurious announcements (LUCKIE, 2014). We explore the two approaches for inferring valid source addresses reported in the literature (Chapter 3 - §3.2): the Full Cone (FC) (LICHTBLAU et al., 2017) and the Customer Cone (CC) (LUCKIE et al., 2013).

4.1.1 Full Cone

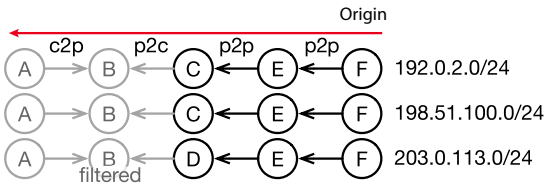
The Full Cone, used in (LICHTBLAU et al., 2017), is the more permissive of the two construction methods (Chapter 3 - §3.3, Table 3.2). Aiming to minimize false positives, Lichtblau *et al.* (2017) chose to *"not distinguish between peering/sibling, customer-provider and provider-customer links. Rather, whenever [the algorithm sees] two neighboring ASes on an AS path, [the algorithm] presumes a directed link between the two, where the left AS is considered upstream of the right AS."* The resulting cone for an AS includes every prefix that contains that AS in the BGP route's AS path, for all routes observed by public route collectors in snapshots (RIBs) and updates during the measurement period.

Figure 4.1: Comparison of construction of Full Cone and Customer Cone given identical set of three BGP AS paths at top. The red line indicates how a BGP AS Path is read (the last AS, the "origin" AS, "owns" the prefix).



ASes	Full Cones:
A, B, E, F	192.0.2.0/24, 198.51.100.0/24, 203.0.113.0/24
C	192.0.2.0/24, 198.51.100.0/24
D	203.0.113.0/24

(a) Example Full Cones (§4.1.1) for six ASes given these BGP paths. The Full Cone for an AS includes every prefix that contains that AS in the path for all routes observed by public route collectors, regardless of the underlying relationships.



ASes	AS-Level Customer Cones:
A, B, F	(contain only themselves with these paths)
C	C, E, F
D	D, E, F
E	E, F

ASes	Prefix-Level Customer Cones:
A, B	(empty with these paths)
C, D, E, F	192.0.2.0/24, 198.51.100.0/24, 203.0.113.0/24

(b) Example Customer Cones (§4.1.2) for six ASes using the same BGP paths from Figure 4.1(a). In Customer Cone construction, we annotate each AS link with a c2p, p2c, or p2p relationship before inferring the Prefix-level Customer Cone. With this specific set of paths AS B is filtered out of the process (the PPCC cone construction uses routes observed from its providers and peers), and AS A has no customers or peers considering only these BGP paths.

Source: by author (2019).

Lichtblau *et al.* (2017) acknowledge that this method intentionally sacrifice specificity, i.e., inflating the address space considered legitimate for each AS pair, in the interest of avoiding false positives, i.e., avoiding mistakenly attributing a failure to deploy SAV. Using this method, a stub AS that provides a public BGP view containing all prefixes it

received from its peers and providers will have *all* of these prefixes included in its Full Cone, i.e, the entire routed address space will be deemed valid.

Figure 4.1 illustrates the complexities of building the cones. We compare both methods, FC and CC, using the same set of three BGP AS paths extracted from BGP route announcements. In Figure 4.1(a) we show the results of the Full Cones for six ASes (A to F); assuming A is a stub AS and a customer of B, all three prefixes would be included in A's Full Cone even though no system in A should originate packets with those source addresses. Next, we discuss how the Customer Cone construction work.

4.1.2 Customer Cone

The Customer Cone is the more restrictive of the two construction methods. It takes into account the semantics of AS relationships. As described in Chapter 2 - §2.4, the AS-level Customer Cone defines the set of ASes reachable using customer links from the AS, including the AS itself (LUCKIE et al., 2013). We use the Provider/Peer-Observed Customer Cone (PPCC) algorithm defined in (LUCKIE et al., 2013) to build an AS-level Customer Cone. Using the paths in Figure 4.1(b), the PPCC method constructs the cone of AS C using routes observed from its providers and peers. The PPCC method accommodates hybrid relationships, where an AS may not propagate all of its customer routes to all of its peers and providers. Customer Cone inference critically relies on accurate routing relationship inferences; a customer link incorrectly inferred to be a peer link will result in address space that the provider AS transits being incorrectly excluded from its customer cone. Figure 4.1(b) illustrates the AS-level Customer Cones for the same ASes and paths as Figure 4.1(a), with link annotations to identify the inferred routing relationships between ASes. However, an AS-level Customer Cone is not sufficient to define the set of valid source addresses in traffic transiting a given neighbor AS.

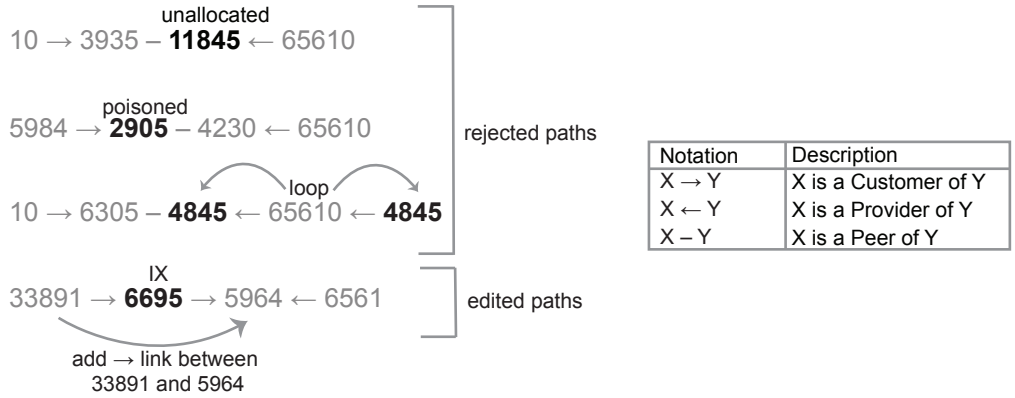
Once we have the AS-level Customer Cone for C, we transform it into its corresponding Prefix-Level Cone by including all prefixes originated by ASes in the AS-level Customer Cone for C during the same observation window. This *novel Prefix-Level Cone Construction* accommodates traffic engineering practices, where an AS may announce different prefixes through different providers (e.g., segmenting the address space in smaller prefixes and announcing them separately), but forward traffic from within these prefixes according to the best route to the destination. To illustrate, in Figure 4.1(b), we include 203.0.113.0/24 in C's Prefix-Level Customer Cone, even though that prefix is not observed

in any BGP paths involving C, because F is in C's customer cone. Importantly, we do not include these three prefixes in A's customer cone, because A has no customers. Although this is not depicted in the figure, we also combine the Prefix-Level Customer Cones of siblings, because a sibling C may transit packets from the customer cone of any of C's siblings to C's peers or providers.

4.1.3 Filtering and Sanitizing AS Paths

As we mentioned before, precision in the process of constructing the cones is crucial. The BGP-based collection infrastructure used to obtain AS-level topology data (ROUTEVIEWS, 2018; RIPE, 2018) suffers from artifacts induced by misconfigurations (i.e., reserved or unallocated ASes), poisoned paths (i.e., AS loops or non-adjacent Tier-1 ASes), and prepended IXP route server ASes all of which hinder the AS-relationship inference results. Figure 4.2 exemplifies some cases of BGP AS paths containing these artifacts and indicates the type of processing made on each case, i.e., reject or edit the AS path. We use the method from (LUCKIE et al., 2013) to incorporate steps to remove such artifacts.

Figure 4.2: Examples of BGP AS paths containing artifacts. An Invalid path implies a link (and thus relationship) between two ASes, where in reality neither may exist.



Source: by author (2019).

Table 4.1 shows, as an example of a cone construction, the results of the path sanitization process we used to construct our cones. These results reflect the data sanitization across all BGP monitors used to compute a dataset within a five days sample (1-5 April 2017), i.e., the results are a sum of the number of records discarded or cleaned across monitors. This process removed approximately 24K AS paths and 342K /24 prefixes, due to poisoned AS Paths and/or the existence of reserved/unallocated ASes which makes the AS Path invalid. Moreover, we edited 259K AS Paths, removing IXP ASes

inadvertently or deliberately prepended. To better understand the case of IXP ASes it is worth to remember that ASes often establish p2p relationships over the shared switching fabric provided by IXPs. To facilitate dense peering connectivity, IXPs provide BGP Route Servers (RICHTER et al., 2014) over which ASes establish many-to-many (multilateral) interconnections. Route Servers typically have their own ASN, but according to best practices it should be filtered-out from the AS path since the Route Server does not participate in the routing decision process (JASINSKA et al., 2016). However, for debugging reasons, some IXP members append the Route Server ASN in the BGP path. We sanitize BGP paths to remove Route Server ASNs since essentially the peering links are between the IXP members, and not between the IXP and ASes. Note that the removal of Route Server ASNs does not discard the prefixes, although affects inferred relationships, and consequently, observed prefixes by the ASes when we build the cones. The Full Cone method (LICHTBLAU et al., 2017) *did not filter any paths*, thus it would classify as *In-Cone* (valid) any traffic using addresses in any invalid paths.

Table 4.1: Filters applied to AS paths, and number of paths and /24 address blocks (contained in /24 or shorter prefixes) filtered, while building Customer Cone (RIPE RIS and RV data, 1-5 April 2017).

Filtering and Sanitizing Operations over AS Paths	**AS Paths**	**Prefixes (/24)**
Poisoning: paths with AS loops or non-adjacent Tier-1 ASes	11,089	13,253
Reserved/Unallocated: paths with reserved or unallocated ASes	12,517	329,223
Prepended IXP ASes: edit paths to remove Route Server ASNs/IXP ASNs	258,335	—

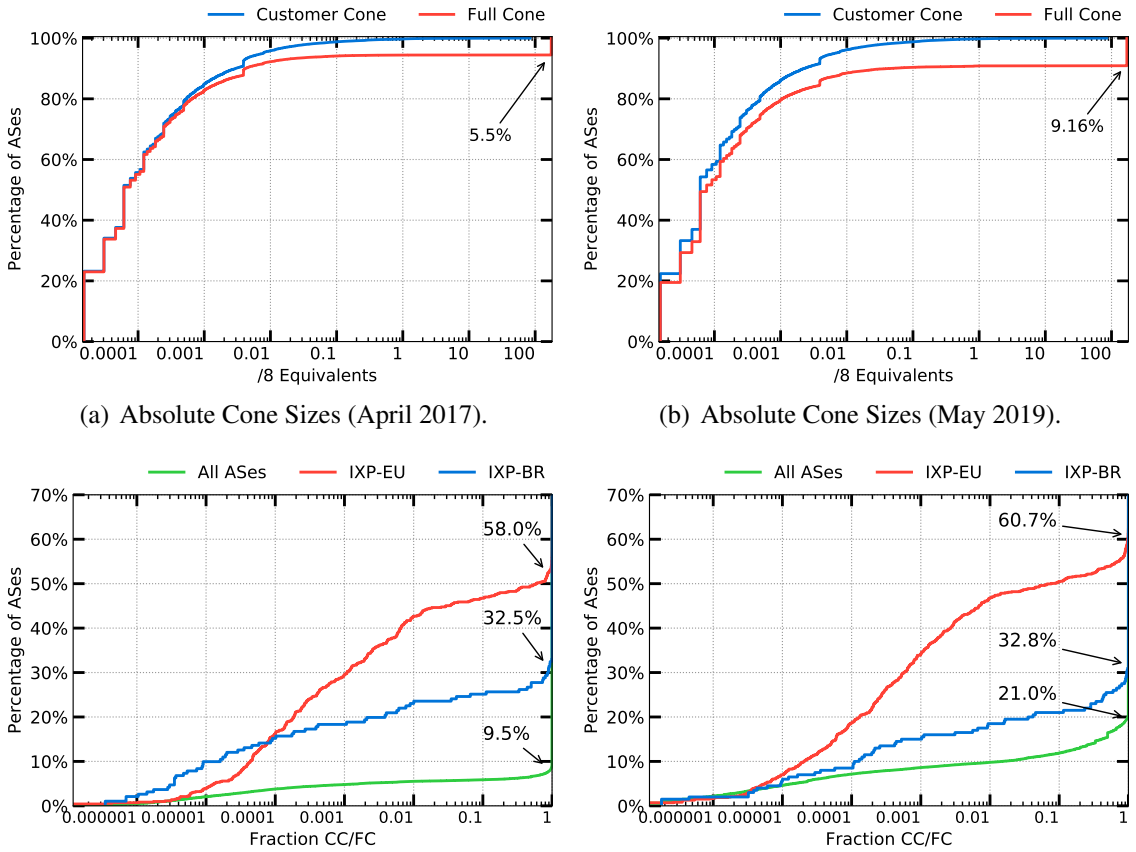
Source: by author (2019).

4.1.4 Impact of the Cone Construction Method

Figure 4.3 shows how the choice of cone construction method impacts the inference of valid address space per AS. Figures 4.3(a) and (b) present the results for all ASes in the Internet, the X-axis shows the absolute number of /8 equivalent prefixes and the Y-axis show the percentage of ASes. In addition, Figures 4.3(c) and (d) include specific analysis of members from two IXPs, a large European IXP used in (LICHTBLAU et al., 2017) and the mid-size IXP-BR in our study; the X-axis shows the relative cone size between our cone (Customer Cone) and the Full Cone and in the Y-axis the percentage of ASes. These plots were computed using traffic and BGP data from April 2017 and May 2019 (see Chapter 6 - §6.1 for further detail on the datasets we used).

For example, in Figure 4.3(a) we show that 5.5% of all ASes in the Internet had

Figure 4.3: The cone construction approach significantly impacts the source addresses each method will consider valid. The results in (a) and (b) refer to all ASes in the Internet, while (c) and (d) include specific analysis of members from two IXPs (EU and BR).



(a) Absolute Cone Sizes (April 2017).

(b) Absolute Cone Sizes (May 2019).

(c) Relative Cone Sizes (April 2017).

(d) Relative Cone Sizes (May 2019).

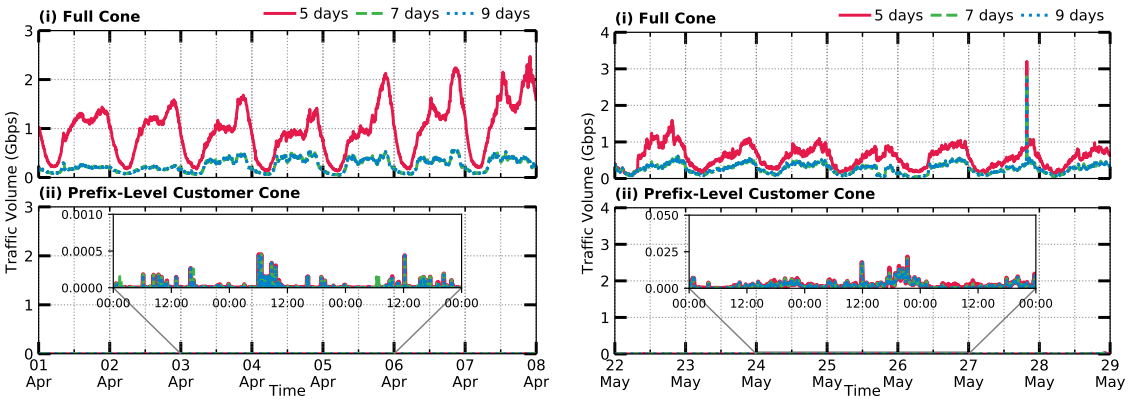
Source: by author (2019).

the equivalent of all routed IPv4 address space (175 /8 equivalents) in their Full Cone in April 2017. In Figure 4.3(c) we show that for 90.5% of ASes, the Full Cone and Customer Cone were congruent (included the same addresses), but 58% of IXP-EU member ASes had Full Cones covering more addresses than the Customer Cone, with 42% having a list 100 times larger in the Full Cone than the Customer Cone. Note that we also looked two years later using data from May 2019, and the differences increased with the pace of the IPv4 allocation space depletion, as can be seen in Figures 4.3(b) and 4.3(d).

The disparity of the results of cone sizes for all ASes compared to those at the IXP is because while over 80% of the Internet's ASes are stubs (ROUGHAN et al., 2011), i.e., do not provide transit, these are less likely to peer at an IXP (AGER et al., 2012). Further, IXPs are popular places to operate public route collectors because the collector can obtain BGP routing views from multiple ASes at a single place. Therefore, those ASes at an IXP that provide a routing view will have all of the prefixes they announce in routes to the collector, including those from their peers and providers, in their Full Cone.

Figure 4.4 show how the cone construction methods are affected by the choice of BGP observation window (COMARELA; GüRSUN; CROVELLA, 2013) on the inference of Out-of-cone traffic at our mid-size IXP in Brazil in two complete distinct periods, April 2017 (Figure 4.4(a)) and May 2019 (Figure 4.4(b)) contrasting the Full Cone and the Prefix-Level Customer Cone. This effect is because of the FC's permissive nature, which exposes the cone inference to announcements across the whole Internet. Note that the spike that appears in Figure 4.4(b)(i) is valid traffic incorrectly classified by FC as spoofed. This portion of the traffic is unverifiable and can not be classified by any method [1]. Next in §4.2, we follow this discussion explaining the traffic visibility challenges.

Figure 4.4: The inferred Out-of-cone traffic volume for the Full Cone (FC) is sensitive to changing BGP observation window sizes in the construction of the cone, while Prefix-level Customer Cone (PLCC) is not. In Figure 4.4(a)(i) and Figure 4.4(b)(i) while the 7 and 9 day lines are almost identical (overlap), the 5-day line contains an order of magnitude more traffic because the set of valid addresses for each AS is smaller. This contrasts with the PLCC – Figure 4.4(a)(ii) and Figure 4.4(b)(ii), where Out-of-cone traffic is robust to changes in the BGP table input window.



(a) Week-1, April 2017. (b) Week-4, May 2019.

Source: by author (2019).

4.2 Topology and Traffic Visibility

While the original role of IXPs was to promote peering between ASes physically present and connected to a switching fabric, in practice IXP services have become more complex. For example, many networks now obtain transit services from a provider at the IXP (AGER et al., 2012), or an organization can connect its sibling networks using the IXP switching fabric. IXPs may also offer services such as remote peering and layer-2 transport,

[1] Further, we explain more details in Chapter 5 and present analyses in Chapter 6 - §6.3 regarding the unverifiable traffic. The spike identified by FC is composed of 84% Transport Provider traffic and 16% Provider-to-Customer traffic.

as well as virtualized segmenting of traffic into multiple virtual LANs (VLANs). These services present three challenges to accurate inference of SAV deployment. Following, we discuss three main hurdles.

First, the semantics of the AS relationship between two IXP members impacts whether the customer cone can serve to constrain inference of valid source address space. That is, when inferring whether the source IP address of a packet is spoofed when crossing an IXP, we need to consider the AS relationship between the two IXP members exchanging traffic. As discussed in Chapter §2 - §2.4, a provider may forward packets with a source address from any routed prefix in the Internet to their customer, and a sibling may forward packets from the provider of one sibling to the customer of another sibling. In these cases, we cannot apply a cone of valid addresses to infer the SAV policy of the transmitting member. We can only make this inference when that member has a peering or transit relationship with (i.e., receives transit service from) another member. The authors of (LICHTBLAU et al., 2017) did not consider the routing relationship between a pair of ASes when evaluating the source address of a packet sent by a member at the IXP. In contrast to prior work, we consider the routing relationship between the two IXP member ASes exchanging traffic when evaluating the source address of a packet crossing the IXP.

Second, there is a set of three traffic visibility impediments. First, as discussed in Chapter §2 - §2.5, traffic between members connected to the same switch will stay within the switch. In a distributed switching fabric, observing all member traffic requires traffic capture from all switches. Second, ASes that establish private interconnections with other ASes at the same Colocation Facility; their traffic exchange does not use the core IXP switching fabric. Third, to infer SAV policy of an IXP member, we require hosts in the cone of the IXP member to attempt to send spoofed packets to hosts they would reach across the IXP. Because most ASes in the Internet peer at an IXP, only destinations in the customer cone of the receiving AS could receive that packet, i.e., the victim or the amplifier must be reached via the IXP. Because most customer cones are small (Figure 4.3, where only 5% of ASes have more than 0.006% of the routed address space in their customer cone) the chance of a victim or amplifier also being reached via a peering relationship at the IXP is small; a victim or amplifier is more likely to be reached via a transit relationship at the IXP.

Lastly, shared use of IXP ports creates attribution challenges. While the IXP can supply the AS number for a given port, with the associated Ethernet MAC address, that port does not necessarily uniquely identify the sending AS when a reseller uses the port to

provide layer-2 transport, in cases of remote peering and port resale (§2.5), or when the port connects to another exchange. Prior work has illustrated measurement challenges of inferring remote peering (CASTRO et al., 2014; NOMIKOS et al., 2018). In this thesis, the IXP provided us the reseller and IXP tags they used to bridge remote peers. This IXP-specific knowledge exemplifies why we believe a Customer-Cone-based approach to SAV inference will ultimately be integrated into expert system capabilities rather than be amenable to complete layer-3 automation.

5 SPOOFER-IX METHODOLOGY

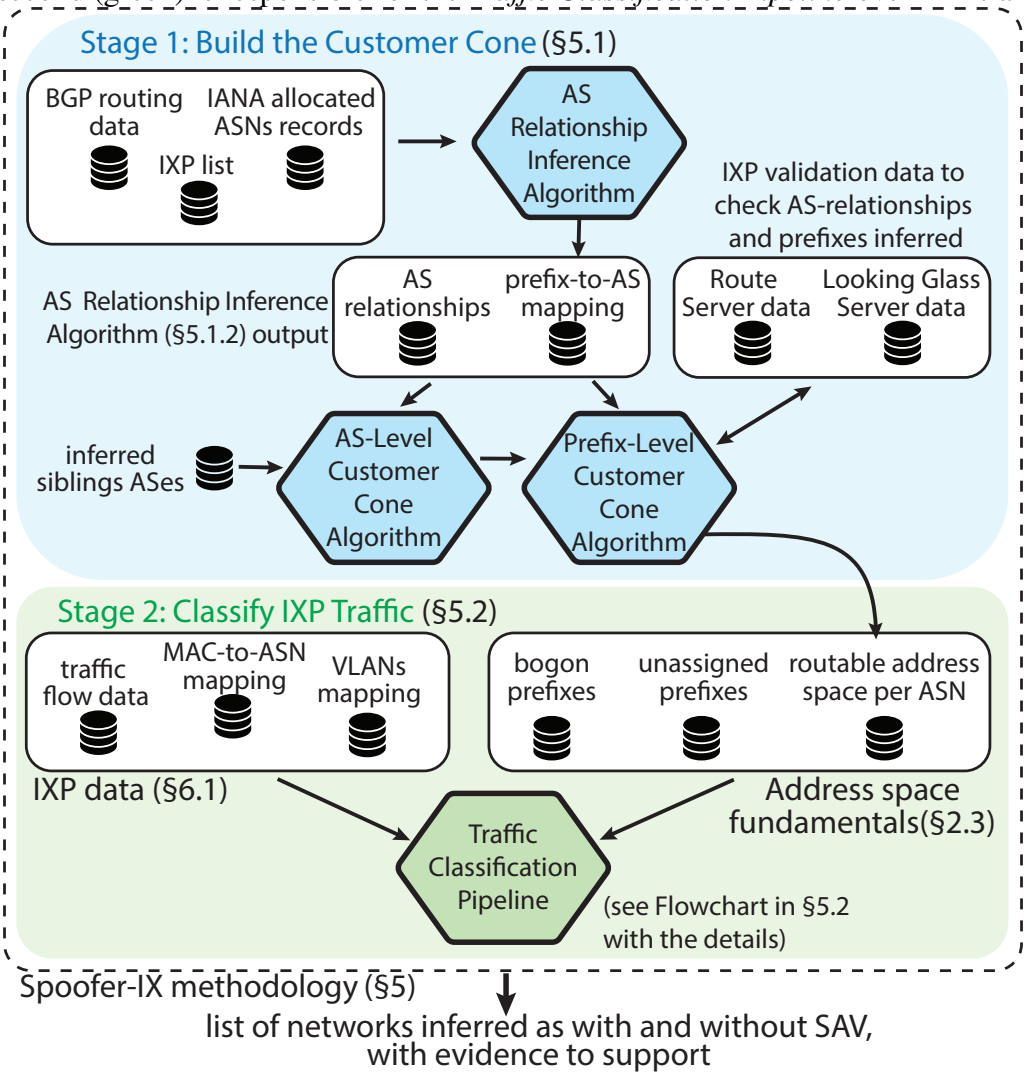
The Customer Cone construction method and the approaches to tackle traffic visibility impediments, both described in Chapter 4, underpin our traffic classification method – how we infer invalid source addresses (presumably spoofed) in packets crossing an IXP, and the ASes responsible for transmitting them.

In this chapter, we describe how these pieces fit together in our methodology implementation, which relies on IXP passive traffic measurements and topological information, i.e., BGP data and IXP switching fabric forwarding databases. Our methodology, illustrated in Figure 5.1, has two key stages. Section 5.1 describes the first stage with the strategy responsible for building an accurate *Prefix-Level Customer Cone* from BGP public data. The second stage, in Section 5.2, focuses on the explanation of our Traffic Classification Pipeline over IXP traffic. In Section 5.3, we explain the set of tools developed as part of Spoofer-IX. Finally, in Section 5.4 we close the chapter with some considerations.

5.1 Stage 1: Build the Customer Cone

This section presents the methodology to build the *Prefix-Level Customer Cone (PLCC)*, in which the goal is to infer the valid IPv4 address space per AS aiming to detect spoofed traffic. PLCC leverages the methodology proposed by Luckie *et al.* (LUCKIE et al., 2013) (Chapter 3 - §3.2) to infer AS relationships and the AS-Level Customer Cone. First, for our purposes we set the AS Relationship Inference Algorithm to receive and process more days of public BGP data as input, which is fundamental to capture a stable view of the inter-domain routing system (recall Chapter 4 - §4.1). Second, we propose PLCC, a new cone construction method to take into account the routing dynamics, i.e., traffic engineering and policy changes. For completeness, we introduce in this chapter the different sources of publicly available topology data we use in the process of building the cones and their differences. Moreover, we present how the algorithms work while we explain our approach.

Figure 5.1: Spoofer-IX Inference Methodology Overview divided into two stages. The first (blue) builds an accurate *Prefix-Level Customer Cone* from public BGP data, while the second (green) is responsible for the *Traffic Classification Pipeline* over IXP traffic.



Source: by author (2019).

5.1.1 Data Sources

The AS Relationship inference and cone construction algorithms rely on one main data source: public BGP data (RIPE, 2018; ROUTEVIEWS, 2018) being actively collected all over the globe. It also takes as input the list of IANA allocated ASNs to RIRs and organizations, as well as a list identifying IXP ASNs worldwide. Moreover, other sources of routing information, like the Route Servers (RICHTER et al., 2014) and Looking Glass Servers (GIOTSAS; DHAMDHERE; CLAFFY, 2016), were used as part of our inference results validation processes. The Internet Routing Registry (IRR) could not be used as a source because, at the time of writing, the corresponding infrastructure in Brazil was still not fully operational (IX.br, 2018; NLNETLABS, 2018).

Public BGP Data. BGP tables can provide information on AS-level connectivity. Access to BGP routing tables can be obtained mainly through BGP Looking Glass Servers, Route Servers (usually available at IXPs), or projects with distributed BGP monitors (RIPE, 2018; ROUTEVIEWS, 2018). Worth to note that Looking Glass Servers, in most cases, offer access to a limited number of routers so one cannot extract the full routing table [1] of an AS, while the Route Servers provide full route tables. BGP monitors offer the most complete BGP data by peering with backbone ASes and by collecting full BGP tables, and also, update messages (ORSINI et al., 2016).

We leverage the BGP monitors from RouteViews project (RV) (ROUTEVIEWS, 2018) and RIPE Routing Information Service (RIPE RIS) (RIPE, 2018) to process the BGP paths derived from routing table snapshots. Both RouteViews and RIPE RIS have deployed 24 monitors each (at the time of writing) around the world that continuously collect BGP tables and BGP update messages from hundreds of different backbone routers. The motivation for ASes to offer access to their backbone routers is to understand how the global routing system views their prefixes. The projects differ in the ways they collect data [2]; it is necessary to take them into account when processing data.

IANA Allocated ASNs. In order to identify valid AS numbers assigned to organizations and RIRs, we used IANA's list of AS assignments (IANA, 2018a). This list indicates which ASNs are assigned, as well as the ranges of ASNs not assigned and the reserved ones by IETF.

IXP ASes list. We collect a list of ASNs used by IXPs offering Route Server services in order to sanitize BGP paths. We remove Route Server ASNs from AS-Paths since essentially the peering links are between the IXP members, and not between the IXP and ASes. To extract such list, we query PeeringDB (PeeringDB, 2019) for networks of type "Route Server" and extract the ASN.

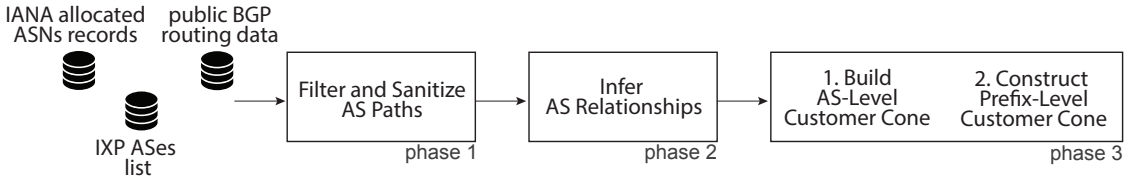
[1] A full routing table is a table which contain all the routes the BGP neighbor is aware of (advertise more than 400,000 prefixes (ORSINI et al., 2016)), while a partial table is a table filtered so that only some specific routes are exchanged.

[2] *RIPE RIS*: offers RIB snapshots every 8h, 3 RIBs per day, UPDATE snapshots every 5min, 288 files per day; *RouteViews*: offers RIB snapshots every 2h, 12 RIBs per day, UPDATE snapshots every 15min, 96 files per day. *PCH*: offer 1 RIB snapshot per day in CISCO txt dump format, UPDATE snapshots every 1 min, 1440 files per day with $\approx$1h delay to publish the data.

5.1.2 Prefix-Level Customer Cone Inference Method

This stage has four main steps, as depicted in Figure 5.2. The first step is to filter and sanitize the AS Paths extracted from the public BGP data. With the data sanitized, the second step is to run the AS Relationship Inference algorithm. Next, with the inferred ASes relationships, we build the AS-Level Customer Cones. We then use the results as input to the fourth and last step, construct our Prefix-Level Customer Cone through the Provider/Peer-Observed Customer Cone (PPCC) algorithm. Following, we discuss more details on each one of these steps.

Figure 5.2: Methodology overview to build the Prefix-Level Customer Cone.



Source: by author (2019).

Phase 1: Filter and Sanitize AS Paths. To avoid incorrectly identifying non-existent links, we discard paths with artifacts. We use the method from (LUCKIE et al., 2013) (recall Chapter 4 - §4.1.3). We filter out paths with AS loops, i.e., where an ASN appears more than once and is separated by at least one other ASN. Such paths are an indication of poisoning. We also discard paths with non-adjacent Tier-1 ASes, i.e., after inferred the clique, we remove paths where any two ASes in the clique are separated by an AS that is not in the clique. Moreover, we filter out paths containing reserved/unassigned ASes (IANA, 2018a) and discard paths to prefixes longer than /24 or shorter than /8, as there is a consensus to not propagate them in the inter-domain routing system.

Phase 2: Infer AS Relationships. We use the AS Paths from Phase 1 to derive AS relationships on a weekly basis. We use as input to the algorithm 7-days of public BGP data, being one RIB file per day, instead of the 5-days practice from Luckie *et al.* (LUCKIE et al., 2013). We have defined this number after having performed a series of evaluations with public BGP data and our traffic flow traces (as discussed in Chapter 4 - §4.1). We were seeking to balance the number of files being processed and the update frequency of the data input that could lead us to the best inference results, taking into consideration the current periodic changes (COMARELA; GüRSUN; CROVELLA, 2013) seen in inter-domain routing system (e.g., due to traffic engineering or routing policy updates).

For completeness and ease of reference, Algorithm 5.1 shows each high-level step

Algorithm 5.1: AS Relationship Inference Algorithm, adapted from (LUCKIE et al., 2013).

Input: AS paths, Allocated ASNs, IXP ASes list
1 Discard or sanitize paths with artifacts;
2 Sort ASes in decreasing order of computed transit degree, then node degree;
3 Infer a transit-free clique (i.e., Tier-1) ASes at top of AS hierarchy and label the links between every pair of ASes in the clique as *p2p* links;
4 Discard poisoned paths;
5 Visit ASes in order of the ranking in (2), and label a link as *c2p* if its previous link in a BGP path is composed of two clique members, or if its previous link in a BGP path is already labeled as *c2p*;
6 Infer *c2p* relationships from VPs inferred not to be announcing provider routes;
7 Infer *c2p* relationships for ASes where the customer has a larger transit degree;
8 Infer customers for ASes with no providers;
9 Infer *c2p* relationships between stub ASes and clique ASes;
10 Infer *c2p* relationships where adjacent links have no relationship inferred;
11 Infer remaining links left as *p2p* relationships;

in the AS Relationship Inference technique. This algorithm uses AS node[3] and transit degrees[4] as metrics of AS-level connectivity and applies heuristics to annotate each link with either a transit (*c2p, p2c*) or peering (*p2p*) relationship (lines 5 – 11). In the former case, a customer buys access to routes that reach the global Internet. In the latter, two ASes share route to their networks (including their customers' networks), sometimes without either AS paying the other (§2.4). For a detailed description of the algorithm, as well as the validation of its results we recommend checking (LUCKIE et al., 2013).

The state-of-the-art AS relationship inference algorithm makes three generally accepted assumptions: 1) there is a clique of large transit providers at the top of the hierarchy, 2) most customers purchase transit in order to be globally reachable, and 3) there are no cycles of p2c links. It uses Algorithm 5.2 to infer the transit-free clique of the Internet, i.e., the clique of Tier-1 ASes at the top of the Internet AS hierarchy. To extract the Internet clique, it leverages as input the same public BGP AS Paths and applies upon them a set heuristics based on AS transit degree and AS Path triples (adjacent pairs of links). The set of ASes resulting from this technique has routes to all other networks on the Internet through customer or peering links without the need to pay for transit.

Phase 3: Construct the Prefix-Level Customer Cone. An AS's *Prefix-Level Customer Cone (PLCC)* is the set of prefixes that the AS can reach through its customer links. Conceptually, constructing this cone is the most complicated part of the methodology, and where mistakes can impact its accuracy. We construct a Prefix-Level Customer Cone using

[3] AS node: number of neighbors an AS has.
[4] AS transit degree: number of unique neighbors that appear on either side of an AS in adjacent links.

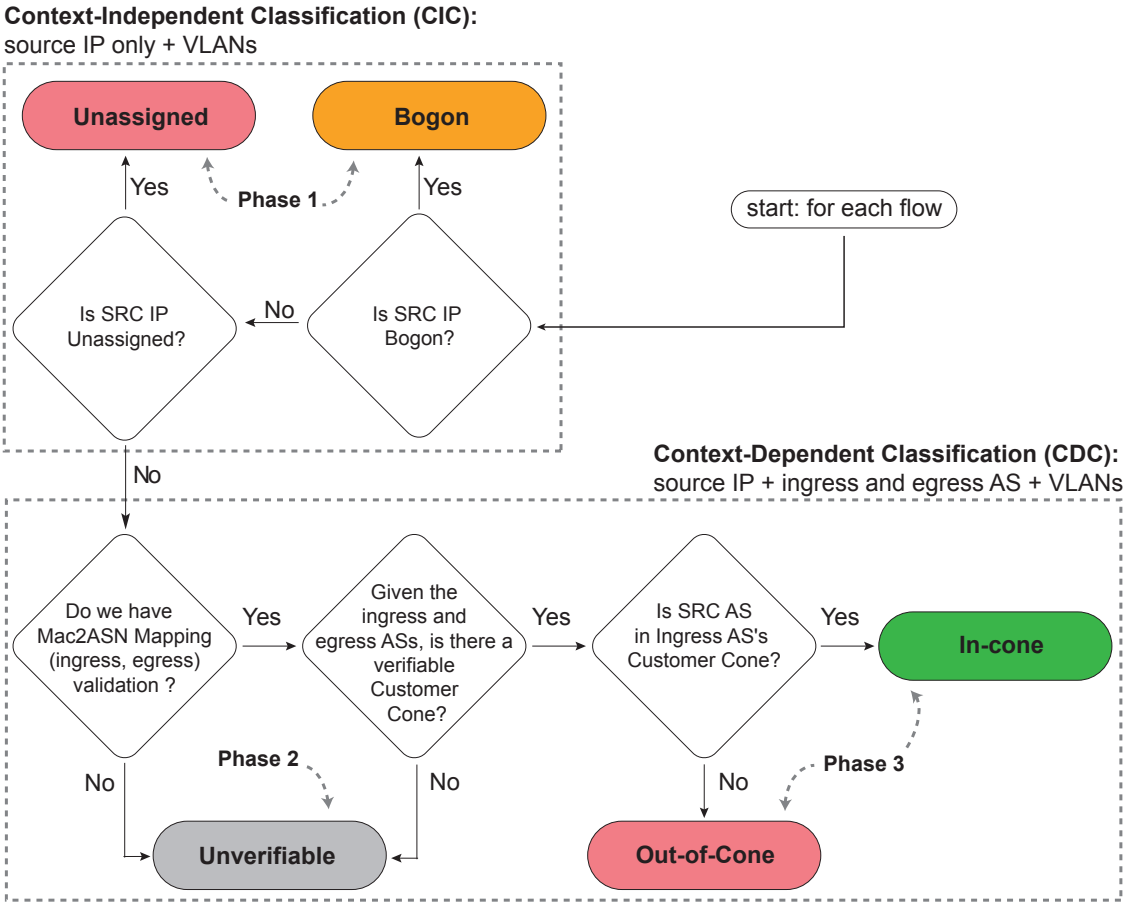
Algorithm 5.2: Internet Clique ASes Inference Algorithm, adapted from (LUCKIE et al., 2013).

Input: AS paths

1 Find the top 10 ASes by transit degree;
2 If there are three consecutive members (X-Y-Z) in the top 10 ASes showing up in paths, and there are more than 5 ASes downstream from X Y Z (to make sure that the paths containing three consecutive members are not poisoned), disconnect the edge between X and Z even though X and Z are connected in some paths;
3 Find the largest clique in terms of transit degree sum among the top 10 ASes, denoted as C;
4 Visit the rest ASes top to down by transit degree, add an AS Z to C if Z has links with all members in C;
5 Similar to Step 2: If there are three consecutive members (X-Y-Z) in C showing up in paths, and there are more than 5 ASes downstream from X Y Z, disconnect the edge between X and Z;
6 Find the largest clique in C in terms of transit degree sum as the final inferred clique;

the method we described in Chapter 4 - §4.1.2. We divide this phase into two parts, as illustrated in Figure 5.2. First, we use the *Provider/Peer-Observed Customer Cone (PPCC)* algorithm defined in (LUCKIE et al., 2013) to build an AS-level Customer Cone. Take Figure 4.1(b) as an example. The PPCC method constructs the cone of AS C using routes observed from providers and peers of C. This accommodates hybrid relationships, where an AS may not send all of its customer routes to all of its peers and providers. Second, once we have the AS-level Customer Cone for C, we transform it into its corresponding Prefix-level cone by including all prefixes originated in public BGP data by ASes in the AS-level Customer Cone for C during the same BGP observation window. This process accommodates traffic engineering, where an AS may announce different prefixes through different providers, but forward traffic from within these prefixes according to the best route to the destination. Figure 4.1(b) shows that we include 203.0.113.0/24 in C's Prefix-Level Customer Cone, even though that prefix is not observed in any BGP paths involving C, because F is in C's Customer Cone. Note these are the same paths we used to illustrate the Full Cone construction (Figure 4.1(a)), except we have annotated links to identify the inferred routing relationships between ASes. Importantly, we do not include these three prefixes in A's Customer Cone, because A has no customers. We also combine the Prefix-Level Customer Cones of siblings, because a sibling C may transit packets from the Customer Cone of any of C's siblings to C's peers or providers.

Figure 5.3: Flowchart showing our traffic classification pipeline (Stage 2 of methodology).



Source: by author (2019).

5.2 Stage 2: Classify IXP Traffic

The first stage of the methodology was about constructing the correct Customer Cone for each AS. The second, described in this section, is related to traffic classification between *In-Cone* or *Out-of-Cone*. In Figure 5.1, it is featured in green at the bottom. This stage has three phases, illustrated in Figure 5.3: (i) filtering out traffic with source IP addresses that match a static list of bogon or dynamic unassigned addresses; (ii) filtering unverifiable packets; and (iii) filtering based on the inferred Prefix-Level Customer Cone. The first phase is independent of any routing semantics, so we call it the *Context-Independent Classification (CIC)*. The second and third phases take into account the ingress and egress ASes for the monitored link, the routing relationship between them, and the Prefix-Level Customer Cone of the ingress AS. We call them the *Context-Dependent Classification (CDC)*.

Our traffic classification method tags each flow, based on its source IP address,

into one of five categories: *Bogon*, *Unassigned*, *Unverifiable*, *Out-of-Cone*, and *In-Cone*. The existence of any invalid traffic (Unassigned in the first step or Out-of-Cone in the second step), illustrated in red in Figure 5.3, is evidence that the ingress AS has failed to deploy SAV. At the same time, Bogon is in orange because it requires analysis of the traffic properties since we have identified cases where the ASes have been using these prefixes together with layer-4 tunneling protocols (e.g., IPIP, GRE) within the IXP switching fabric infrastructure.

Phase 1: Filter Bogon and Unassigned Addresses.

We first classify traffic with *Bogon* and *Unassigned* source IP addresses, according to Team Cymru (Team CYMRU, 2018b) (see dataset details in Chapter 6 - §6.1). Networks sending packets with unassigned source IP addresses are unlikely to have implemented SAV correctly since the most obvious implementation blocks traffic from such addresses because they are not routed, therefore they have no feasible return path. This phase is independent of any routing semantics, unlike the subsequent two phases, which consider the sending and receiving ASes for the monitored link, the routing relationship between them, and the Prefix-Level Customer Cone of the sending AS.

Phase 2: Filter Unverifiable Packets.

This phase classifies traffic flows as suitable to inference of spoofing (recall discussions from Chapter 4) using the Customer Cone, marking unsuitable traffic as *Unverifiable*. Verifiable traffic must satisfy all of the following:

1. It must have a valid MAC-to-ASN mapping for both the sending and receiving MAC addresses.
2. It must not have a known router IP address in the source IP address of the packet. Such a source IP address could be from any interface on the router, which might be assigned by an AS whose address space is not in the customer cone of the router's owner.
3. It must not have a known IP address of the IXP LAN prefix. These prefixes are assigned to the IXP ASN and should not be publicly announced, but sometimes member ASes mistakenly announce them.
4. It must not have a source MAC address from a remote peer or layer-2 transport provider.
5. It must not have a source MAC address from a known provider or sibling of the receiving AS.

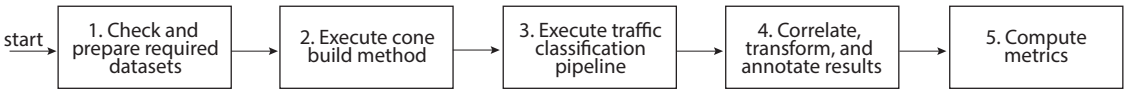
Phase 3: Classify Packets with Customer Cone.

The remaining traffic has a valid MAC-to-ASN mapping, and is either transmitted by a customer of a transit provider at the IXP, or by a peer of another AS at the IXP. If a relationship was not visible in BGP, then we assume the traffic between these members was p2p and use the cones to classify the traffic exchanged. For these transmitting ASes, we classify traffic as *In-Cone* or *Out-of-Cone* using the Prefix-Level Customer Cone (henceforth *Customer Cone* or *CC*) created in the previous stage. A packet whose source IP belongs to the sending AS's Customer Cone address space is classified as *In-Cone*. Otherwise, the packet is classified as *Out-of-Cone*.

5.3 Using Spoofer-IX Implementation

We developed Spoofer-IX as a set of tools to enable the use of our inference methodology by other researchers and network infrastructures, fostering replicability of experimental results. Figure 5.4 depicts the implementation of Spoofer-IX in five steps. Each step contains its tools and interfaces with the subsequent operations. A full-run of Spoofer-IX is comprised of all five steps, which can be employed to distinct network infrastructures (see details in Chapter 7). However, as discussed in Chapter 4, precise knowledge about the network topology and interconnections is required to obtain robust inferences from Spoofer-IX. In the following, we present the implementation details of each step.

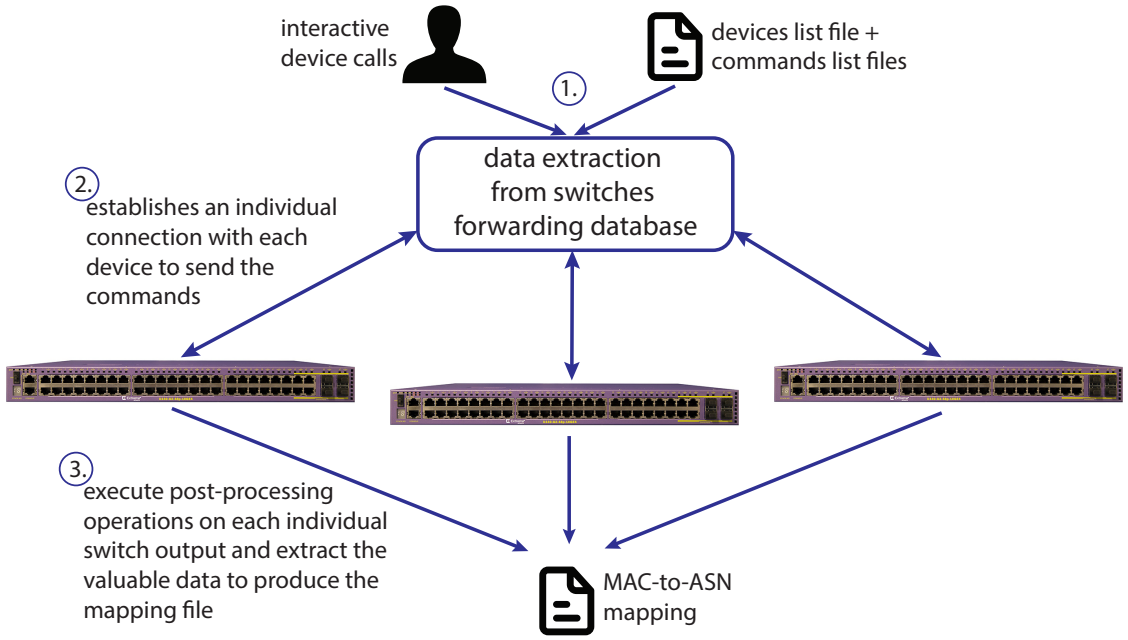
Figure 5.4: Overview of the global steps comprising the analysis methodology.



Source: by author (2019).

Step 1. Check and prepare required datasets (Chapter 6 - §6.1). To obtain accurate results, it is important to align the time windows of the datasets. We provide helper scripts to download, prepare, and optimize datasets. We make available helper scripts (using the Python Scrapy library (SCRAPINGHUB, 2019)) that download and process the BGP routing data files from public BGP route collectors (ROUTEVIEWS, 2018; RIPE, 2018) to build the cones. In addition, we also provide helper scripts to automate topology information extraction from various switch manufacturers using Python's

Figure 5.5: Topology data extraction from switches to create the MAC-to-ASN mapping (Chapter 6 - §6.1).



Source: by author (2019).

Netmiko (BYERS, 2019) and Google's TextFSM (GOOGLE, 2019) libraries. As illustrated in Figure 5.5, the process of extracting topology data from switches has three key phases. The first phase is the definition of the specific set of local commands that should be executed in each device (varies by manufacturers and device model). The next phase requires the establishment of a connection with each device to send the commands. We provide two alternative methods to send the commands, one via interactive calls with a given switch or in batch for a pre-defined list of switches, each running a pre-defined set of commands from input files. The third and last phase involves the application of post-processing operations over each device output answers to generate the required output mapping files.

Step 2. Execute cone construction in three phases (stage 1, §5.1). The codebase of this stage is written in Perl. This step starts with the filtering and sanitization of AS Paths from the previously downloaded BGP data files. Then, proceeds with the execution of the AS Relationships inference algorithm. Lastly, the construction of the Prefix-Level Customer Cone.

Step 3. Execute traffic classification pipeline (stage 2, §5.2). The core implementation of the pipeline and the next steps (4 and 5) were developed in Python, and some additional Bash helper scripts are used to automate parameterized execution. This step saves classification results to disk in Apache Avro (APACHE, 2019) format for use in the

next step.

Step 4. Correlate, transform, and annotate results. Using the classification results and correlation datasets (e.g., MAC-to-AS mapping, prefix-to-ASN mapping), we proceed to create new intermediate files with data enriched with additional information, useful to compute distinct metrics. The results are then forwarded to a set of transformation processes that will group the traffic data in distinct ways (e.g., time, IP address, prefix), compute unique records across a given time range, and verify differences between parametrized time bins (e.g., 5-min, 15-min, 1-hour, etc).

Step 5. Compute metrics. Use data created in the previous step to look for atypical network events, which could hint to network attacks. We implement two distinct metrics to assess the IPv4 address activity and behavior over time: Activity and Churn, and Spatio-Temporal Properties in active IP addresses (RICHTER et al., 2016; DAINOTTI et al., 2013). The former allows measuring the volatility of address activity over time, while the latter captures aggregated properties of active IPs seen in each time-window.

We provide two modes to setup the environment: automatic and personalized setups. The automatic setup is based on a Bash helper script to install and configure most of the dependencies of the project (e.g., NFDUMP (HAAG, 2019), Apache Avro (APACHE, 2019), RIPE NCC BGPdump (RIPE NCC, 2019)), enabling its use out-of-the-box, e.g., on a fresh Linux Ubuntu server. All the steps were developed and made available with multiprocessing support. The source code and the documentation are available online at (MULLER et al., 2019b).

5.4 Considerations

This chapter presented a new methodology using IXPs as observatories to infer spoofed packets and networks that leak them in the Internet. We dealt with operational complexities that characterize today's interconnection ecosystem, the noise inherent in public BGP data sources, and heuristic AS relationship inferences.

The next chapter presents the results we got when we applied our methodology to traces from two distinct years – 2017 and 2019 – from a mid-size IXP with $\approx$200 members and a peak traffic volume of 200Gbps. The accurate inferences shed new light on the deep subtleties of scientific assessments of operational Internet infrastructure, revealing the need for a community focus on reproducing and repeating previous methods.

6 INFERRING SPOOFED TRAFFIC AT IXPS

In the two previous chapters, we provided details of methodological challenges, including comparing our *Prefix-Level Customer Cone (PLCC)* with the *Full Cone (FC)* (LICHTBLAU et al., 2017). Moreover, we presented our Spoofer-IX methodology design to accurately detect the transmission of spoofed traffic by AS members of IXPs.

This chapter focuses exclusively on the results obtained from the Spoofer-IX methodology applied to traffic and topology data from the third largest IXP in Brazil, with more than 200 member ASes connected at the IXP switching fabric. We review the traffic classification results, including a temporal analysis of traffic flow snapshots two years apart.

The remainder of this chapter is structured as follows. In line with the datasets presented in Section 6.1, from Section 6.2 to 6.8, we show the results of a series of extensive analyses we did using Spoofer-IX methodology, including a comparison against the state-of-the-art. Last, in Section 6.9, we discuss our validation efforts regarding the inferences made and the results obtained.

6.1 Datasets

We now introduce our datasets and their collection methodologies. Table 6.1 summarizes the datasets we use. It contains three sets of datasets, *i.* Vantage Points, *ii.* Base Filtering, and *iii.* Correlation datasets. They are grouped by their source/application in the proposed methodology itself.

i. Vantage Points Datasets.

IXP-BR: traffic. We used SFlow (P. Phaal, S. Panchen, and N. McKee, 2001) traffic data from an IXP that belongs to the Brazilian IXP Ecosystem (IX.br) (IX.br, 2020). This IXP transports up to 200 Gbps of traffic among 200+ members. The IXP operators configured a sample rate of 1:4096 packets. We used two datasets, one from April 1 to June 5 2017 (10 weeks), and the other from May 1 to June 5 2019 (5 weeks), to evaluate our method.

Topology data over connectivity fabric. To identify the pair of adjacent ASes sending and receiving each flow across the IXP fabric, we used layer-2 information (i.e.,

Table 6.1: Datasets summary.

	Dataset	Source type	Data format
Vantage Points Datasets	IXP-BR	Passive Traffic Flow	Sflow (RFC3176, 2001), sampling 1:4096
	MAC-to-ASN	Device configuration files	Raw dump snapshots, monthly
	Looking Glass Servers	BGP announcements	Raw dump snapshots, daily
	Route Server	BGP announcements	Raw dump snapshots, monthly
Base Filtering Datasets	Bogons	Prefixes list	Prefixes ranges, stable set (Team CYMRU, 2018b)
	Unassigned	Prefixes feed	Prefixes ranges, updated every 4h (Team CYMRU, 2018a)
	Routers IPs	ITDK data	IP ranges (CAIDA, 2017)
	Public BGP data	BGP announcements	RIBs and Updates files (RIPE, 2018; ROUTEVIEWS, 2018)
	AS Siblings (AS-to-Org)	Organizations and ASes	Hash of organization & ASes
	IXP ASes / LAN prefixes	PeeringDB data	ASNs, prefixes ranges (PeeringDB, 2019)
Correlation Datasets	Available Blocks	IANA/RIRs	IP Ranges (IANA, 2018b)
	NetAcuity Edge	IP Geolocation	Historical IP Ranges (NETACUITY, 2019)
	Prefix-to-ASN	BGP announcements	Prefix to ASN (LUCKIE et al., 2013)

Source: by author (2019).

MAC addresses) since the source and destination IP addresses in the IP headers of the observed packets contain the communication endpoints. To map MAC addresses to sending and receiving ASes of each flow (the MAC-to-ASN mapping), we relied on information from the forwarding database of each switch that is part of the IXP switching fabric (recall Chapter 5 - §5.3).

BGP routing data. We enriched our BGP datasets with vantage point-specific BGP routing data from both the Looking Glass servers (GIOTSAS; DHAMDHERE; CLAFFY, 2016) and snapshots from the Route Server (RICHTER et al., 2014) for the same time window of our IXP traffic data collection.

ii. Base Filtering Datasets.

Bogons and Unassigned addresses. We used Team Cymru's Fullbogons feed (Team CYMRU, 2018a; Team CYMRU, 2018b) to filter out traffic with source IP addresses that are bogons (e.g., private, special use, reserved) (MOSKOWITZ et al., 1996; WEIL et al., 2013; COTTON et al., 2013) or unassigned. Unassigned prefixes are allocated by IANA to an RIR (IANA, 2018b; IANA, 2018c), but not currently assigned by the RIR to an end-user (e.g., an ISP) (NRO, 2020). We used the lists compiled by Team Cymru (which uses lists of prefixes maintained by each RIR (NRO, 2020) to update their feed) in each 4h interval per day for the same time windows as our IXP traffic data collection.

Router IP addresses. For comparability with previous work (LICHTBLAU et al., 2017), we used CAIDA's Internet Topology Data Kit (ITDK) to identify router interface IP addresses. We used the ITDK snapshot closest in time to the IXP traffic capture window (CAIDA, 2017; CAIDA, 2019b). We consider traffic from ITDK-inferred router interfaces to be *unverifiable* (recall Chapter 5 - §5.2) because the source IP address could

be from any of the interfaces of the router, which might be assigned by an AS whose address space is not in the customer cone of the router's owner.

Public BGP Data. Our traffic filters relied on customer cones inferred from public BGP routing table snapshots collected by Route Views (RV) and RIPE's Routing Information Service (RIS) (RIPE, 2018; ROUTEVIEWS, 2018). In Spoofer-IX, we downloaded one BGP RIB table per day from all available (18 and 16 in 2017, 19, and 18 in 2019 from RIS and RV, respectively) collectors for the same time windows as our traffic data. We extracted all AS paths in these tables that announced reachability to IPv4 prefixes, repeating this process for each week.

AS Siblings. We used CAIDA's AS to Organization classification of ASes into sets that likely belong to the same organizations (HUFFAKER et al., 2019). CAIDA's method parses the Regional Internet Registries' WHOIS dumps and delegation files to create a unified mapping between ASes and organization names, then uses hints in the name strings, delegation files, identifiers, and email addresses to infer AS sets with common ownership. For each measurement period, we used the AS-to-Organization mapping that CAIDA constructed closest to the traffic capture window.

IXP ASNs (BGP Route Servers) and LAN Prefixes. We collected a list of AS Numbers used by IXP Route Servers, querying PeeringDB (PeeringDB, 2019) for networks of type "Route Server" and extracted the ASN and also the IPv4 and IPv6 LAN prefixes used by the members to establish their inter-domain BGP sessions.

iii. Correlation Datasets.

Available Blocks. In order to identify the valid prefixes and its allocation status, e.g., if they are assigned to organizations and RIRs, we used IANA's and RIR's lists of blocks assignments (IANA, 2018a). These lists indicate assigned prefixes, as well as prefixes not yet assigned and the reserved ones by IETF.

Geolocation. We draw on two geolocation providers: NetAcuity Edge (NETACUITY, 2019) (henceforth referred to as NetAcuity) for all results presented in this thesis and MaxMind GeoLite2 (MAXMIND, 2019) (GeoLite) given it is a popular free offline geodatabase. NetAcuity's commercial database has an alleged accuracy of 99.9% on a country level, and 97% on a city level. GeoLite, in contrast, is free – it is a less accurate version of the commercial GeoIP2 database that is maintained by the same company.

Prefix-to-ASN mappings data (prefix2AS). We use IP prefix to AS mappings, i.e., denoted as IP prefix $\longrightarrow$ AS to associate IP addresses to ASes. These mapping files

are produced in the *Stage 1 Build the Customer Cone* (Chapter 5 - §5.1) based on public BGP data from RIPE RIS (RIPE, 2018) and RouteViews (ROUTEVIEWS, 2018).

6.2 Longitudinal Traffic Classification Based on Spoofer-IX

In order to apply the traffic classification processing to flow data, we first prepared all the necessary datasets (§6.1). In particular, regarding the first stage (Chapter 5 - §5.1), i.e., build the cones, Table 6.2 summarizes the key parameter values used for the rest of the analyses. For the Prefix-Level Customer Cone (PLCC), we used seven days of public BGP data as input, corresponding to each week of traffic data we classify following our methodology. Later, in §6.6, we present the parameters for Full Cone (FC) when we compare the results of both methods.

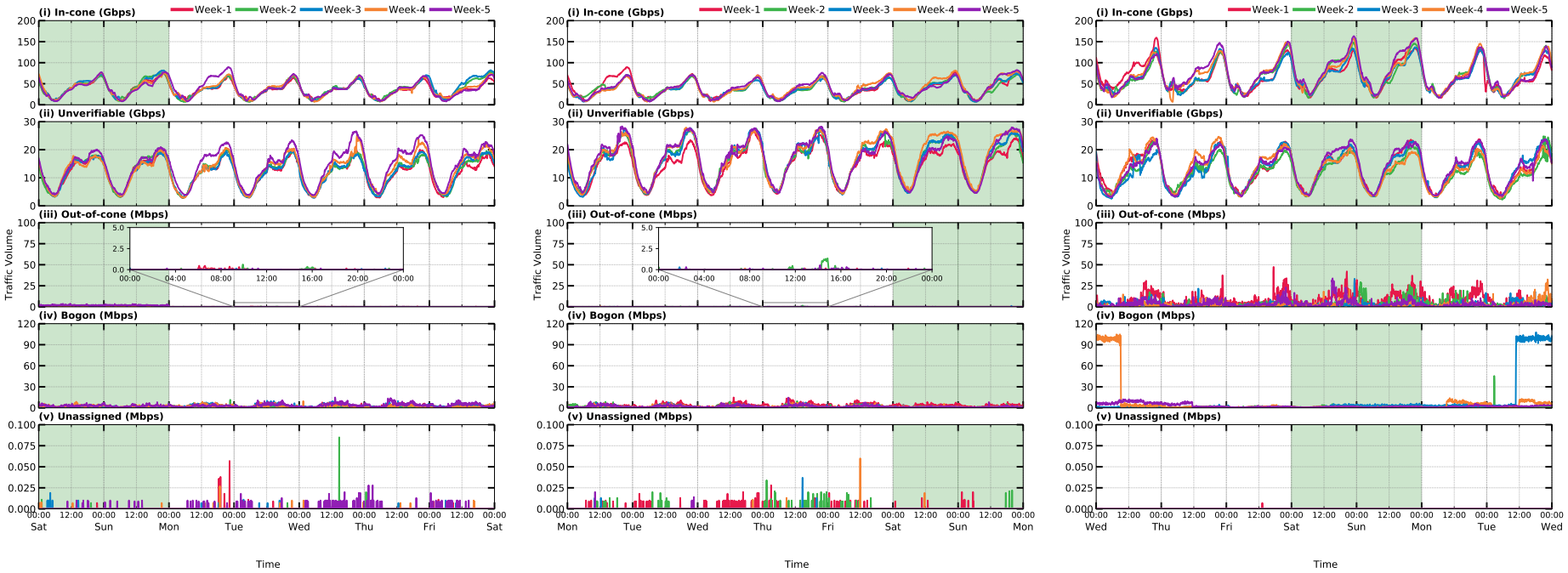
Table 6.2: Parameters for Spoofer-IX cone inference algorithm.

Parameters	**Prefix-Level Customer Cone (PLCC)**
(1) BGP input time window	7-days
(2) Number of monitors	RIPE Routing Information Service (RIPE RIS) & RouteViews project (RV) (18 and 16 in 2017, 19 and 18 in 2019 RIPE RIS and RV, respectively)
(3) Files per monitor	one file per day
(4) Use of RIBs / Updates files	RIBs only

Source: by author (2019).

Figure 6.1 shows the volumes of traffic we classified into multiple subplots with custom scales, one for each of the five categories defined in our methodology (i.e., In-cone, Out-of-cone, Unverifiable, Bogon, Unassigned) organized in descending order of traffic volume for two different years in 2017 and 2019. There are five curves in each category, each curve representing one week of the traffic classified. Additionally, the green-highlighted area in the figures set the weekend boundaries. We present these three distinct five-week periods to show our results are consistent, at least for these periods.

Figure 6.1: Longitudinal analyses, two-years – 2017 (April to Jun, ten weeks) and 2019 (May to Jun, five weeks of traffic) classified with Spoofer-IX. For all fifteen weeks, we inferred almost no Out-of-Cone traffic – in 2019 a maximum of 40Mbps for an IXP with a peak of 200Gbps.



(a) April 1 - May 5, 2017

(b) May 1 - Jun 5, 2017

(c) May 1 - Jun 5, 2019

Source: by author (2019).

Although not directly shown in the individual plots in Figure 6.1, the peak combined rate across the core switch during the period was 120Gbps in 2017, and 200Gbps in 2019. Our first observation regards the volume of traffic classified as In-cone and Out-of-cone, i.e., first (i) and third (iii) subplots in each plot of Figure 6.1. As expected the majority of the traffic across the exchange is classified as In-cone – average of 70.04% in 2017 and 84.66% in 2019 over the total volume of traffic, because it comes predominantly from large content providers (Carisimo et al., 2018).

During 2017, the peak Out-of-Cone traffic we inferred was less than 5Mbps (see inset zoom-in in Figure 6.1(b)), and in 2019, 40Mbps (Figure 6.1(c)), that is in average less than 0.01% of the total volume of traffic from each of the periods analyzed. We believe these values are upper-bounds of Out-of-Cone traffic at the IXP core switch, and we reached these volumes after investigating the underlying properties of traffic between pairs of members, in rank order of contribution to the Out-of-cone traffic volume at the IXP. We manually investigated the relationship between ASes exchanging packets unlikely to be spoofed, such as TCP packets carrying data or directed towards a known transport provider. In addition, we found 27 sibling ASes in 12 distinct organizations that were exchanging traffic across the IXP, but missing from CAIDA's public AS-to-Org dataset (see §6.1). To determine which ASes were siblings, we consulted the official website of those ASes to find information on their ownership, contacted the ASes directly to inquire, or contacted the IXP operators to understand the relationship between two ASes at the IXP. Further, through the IXP operators, using different communication strategies (email, phone, personally), we approached 36 members of the IXP and obtained clarifications from 34 of them.

Although the number of members was similar between 2017 and 2019 (208 and 203, respectively), 28 new members were present in the 2019 analysis. We found that the increase in Out-of-cone traffic between 2017 and 2019 was due to additional complex relationships and traffic transport agreements between members in the 2019 data that are not visible to the IP layer or in the BGP protocol (more details in §6.7). Table 6.3 summarizes the number of unique AS pairs we observed to exchange traffic for the five-week periods beginning 1 April 2017 and 1 May 2019. While we inferred more than 98% of the AS pairs had a p2p relationship, approximately 1.4% of AS pairs had a different class of relationship that impacts our ability to infer SAV policy of the transmitting AS.

Table 6.3: Unique AS pairs observed exchanging traffic at the IXP in each 5-week period. Approximately 1.4% of AS pairs had a non-p2p relationship. (This IXP was rearchitected in 2019, which may explain the drop in observed peers.)

Relationship	April 2017		May 2019	
p2p	19,161	(98.7%)	12,057	(98.4%)
p2c	222	(1.1%)	183	(1.5%)
s2s	21	(0.1%)	10	(0.1%)
total	19,404		12,250	

Source: by author (2019).

Next, we investigate Bogon and Unassigned volume of traffic, which also should not be routed on the Internet (as discussed previously in Chapter 2 - §2.3). Even though it represents 0.00543% of all traffic exchanged in 2019, we examined its properties to understand the motivation of network operators behind the usage of these prefixes. For example, in Figure 6.1 the peak volume of traffic with Bogon source addresses was approximately 100Mbps across the exchange for Wednesday at the end of Week-3 (Figure 6.1(c)). We found these networks make deliberate use of private addresses (defined by RFC1918 (MOSKOWITZ et al., 1996)) as sources when using tunneling protocols, e.g., Generic Routing Encapsulation (GRE) and IP over IP encapsulation (IPIP). It consisted in 61.14% of the traffic at that moment. According to the members involved, these communications were associated with IP Transport service over bilateral agreements through the shared switching fabric infrastructure of the IXP. This type of traffic was introduced in place of stacked VLANs IEEE 802.1ad (JEFFREE T., 2019), informally known as QinQ. This type of traffic is not allowed in many IXPs; according to (IX.br, 2019), it is enabled only in very special cases. Network operators therefore rely on alternative tunneling solutions, as our results show. In Section 6.4, we analyze more details regarding the protocols being used in the distinct traffic categories.

The Unassigned volume of traffic exhibits interesting behavior. In our longitudinal analysis, we captured a change of behavior, i.e., the disappearance of this type of traffic at the IXP. The peak rate was less than 0.1Mbps (or 100Kbps) on Week-2 in Figure 6.1(a) in 2017, and after a year, in 2019, it was not present. This might due to many factors, such as the result of many efforts regarding Internet security worldwide, which greatly enhanced and gained more traction from 2017 onwards (ISOC, 2018; Tech Accord, 2018; NIC.br,

2019). Besides that, the IPv4 address space depletion probably had an impact, making rare the unassigned space, year after year (NRO, 2020; NRO, 2019).

Finally, we analyze the Unverifiable traffic. The results are nearly the same when comparing 2017 and 2019, i.e., we observe that the classification shows a consistent time-of-day and day-of-week effects across all the fifteen weeks analyzed, which is also consonant with the In-cone traffic behavior. So, even though we cannot safely attest to the traffic nature, due to the restrictions already explained in Chapter 2 - §2.4, the fact that it matches the overall behavior of the In-cone traffic, suggests that we have more valid traffic than invalid. Following, we further analyze the composition of the Unverifiable traffic in detail.

Table 6.4: Unverifiable traffic sub-categories definitions, and the average traffic fraction breakdown. It shows the average traffic fraction of each sub-category over the Unverifiable traffic across the exchange in the first week of May 2019.

Sub-category	Traffic Fraction	Meaning
P2C-in-cone	32.44%	We isolate traffic sent from a Provider to a Customer across the exchange. Because a provider can transit packets from any source address in the Internet (§2.4), there are no invalid addresses that would allow the detection of spoofed packets. We apply the Customer Cone approach (In-cone / Out-of-cone) to the P2C traffic only as a matter of extra analysis.
P2C out-of-cone	29.42%	
Unknown Ingress MAC	0.94%	If the MAC-to-ASN mapping for either the source (ingress) or destination (egress) MAC addresses is missing, we can not proceed with the validation of the packet. This situation happens simply because the IXP lacks complete historical data for this mapping (§6.1).
Unknown Egress MAC	20.47%	
Remote Peering	14.10%	If we can not determine the destination AS because the destination (egress) MAC address and packet VLAN tag indicated the traffic was from a Remote Peering Provider.
Transport Provider	1.28%	When the participant is physically distant from the IXP or wants to interconnect in more than one IXP location, they often need to rely on a contract with a Transport Provider. The *Transport Provider* acts in the very same way as the Transit Provider, i.e., it can transport packets from any source address in the Internet, there are no invalid addresses that would allow detection of spoofed packets. This is a two-step validation process. First, we try to validate the traffic. Second, if it fails to be validated then, we check if the member is connected to more than one IXP and check if the ASN involved in the exchange is a known Transport Provider in the market.
Sibling-to-Sibling	0.06%	Traffic sent from a Sibling to a Sibling organization across the exchange. These organizations can transit packets from any source address they want (§2.4) between them. There are no invalid addresses that would allow the detection of spoofed packets.
Bogon in VLAN	0.18%	IXP participants can request VLANs from the IXP to isolate their communications (§2.5). When the prefixes of both categories (Bogon and Unassigned) exchange traffic in their isolated VLAN, there are no invalid addresses that would allow the detection of spoofed packets.
Unassigned in VLAN	0.0%	
Stray	1.11%	Traffic whose source IP is attributed to a router interface. It is not clear whether the router is spoofing this address, or has transmitted the packet on an outbound interface different from the one it used as the source address.

Source: by author (2019).

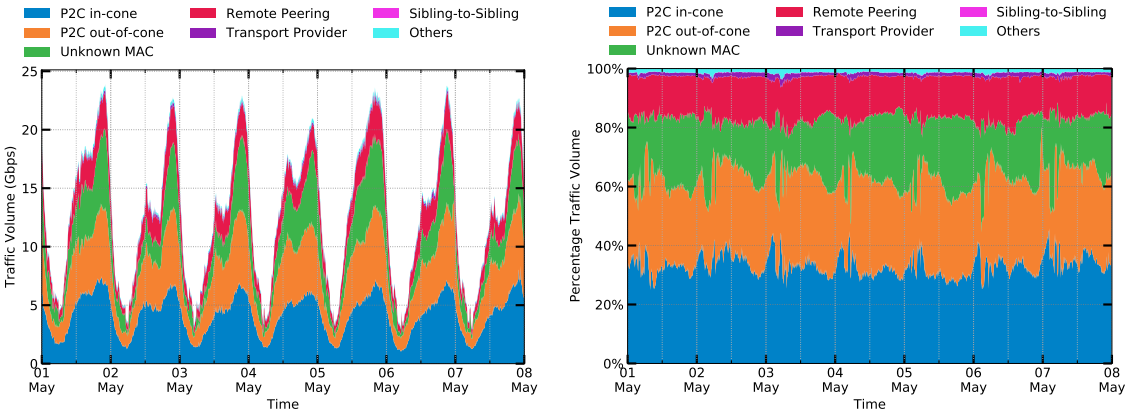
6.3 Unverifiable Traffic Breakdown

The definition of Unverifiable traffic sub-categories is not straightforward. We defined a taxonomy based on multiple factors: (i) investigations on the relationships between specific parties; (ii) discussions with the IXP operators on operational complexities; (iii) how distinct peering agreements impact traffic flows; and (iv) analyses performed over packets exchanged. As a result of this work, the Unverifiable traffic is composed of (i) flows, which we can not validate due to lack of information to correlate and (ii) flows with

properties that would not allow detection of spoofed packets by definition. In Table 6.4, we present the complete list of Unverifiable traffic sub-categories, their meanings, as well as the average fraction of traffic of each individual sub-category across the first week of May 2019.

For both the 2017 and 2019 observation periods, there was a peak of ≈25Gbps of Unverifiable traffic across the exchange, which represents ≈15.30% of the overall traffic passing at the IXP at that time (Figures 6.1(a), 6.1(b)and, 6.1(c) – ii subplot). Figure 6.2 provides a classification of the traffic involved for the first week of May 2019, in absolute (6.2(a)) and relative (6.2(b)) traffic volume values. 61.9% of the Unverifiable traffic was sent from a provider to a customer across the exchange, where no cone of valid addresses applies (Chapter 2 - §2.4).

Figure 6.2: Classification of Unverifiable traffic. 61.8% of the Unverifiable traffic was sent by a provider to a customer across the exchange. Because a provider can transit packets from any source address in the Internet, there are no invalid addresses which would allow detection of spoofed packets. For completeness, we further classify traffic from each provider as being in or out of their Customer Cone.



(a) Absolute Unverifiable Traffic. (b) Relative Unverifiable Traffic.

Source: by author (2019).

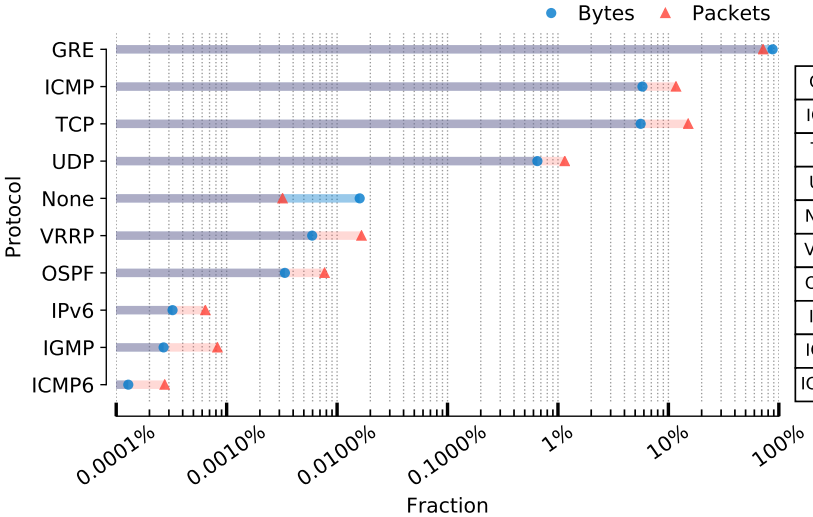
If we had applied the Customer Cone approach to this p2c traffic, we would have inferred 52% of it was from within the provider's customer cone, with the remaining 48% of traffic being from outside of the provider's customer cone. Because *a provider can transit packets from any source address in the Internet* (Chapter 2 - §2.4), there are no invalid addresses that would allow detection of spoofed packets. This potential for erroneous inference is why we must classify all packets from a Transit Provider to a Customer as Unverifiable. Another 21.41% of the Unverifiable traffic was because we did not have an ASN mapping for either the source or destination MAC addresses (the IXP

lacked historical data for this mapping), and for 14.10% of traffic we could not determine the origin AS because the source MAC address and VLAN tag indicated the traffic was from a remote peering provider. Finally, all of the other categories (i.e., Transport Provider, Sibling-to-Sibling, Bogon in VLAN, Unassigned in VLAN, and Stray) summed to only 2.62% of the traffic. Although these last categories do not represent much relatively, they prove to have significant impact on false positives/negatives on spoofing identification at IXPs.

6.4 Distilling Protocol Diversity from Distinct Traffic Categories

In addition to analyses on the traffic perspective, we studied the quantitative and qualitative characteristics of our distinct traffic categories. First, we explore the situation of transport protocols being employed in the exchange of traffic with Bogon prefixes, as previously discussed (§6.2). Following, we contrast characteristics of the traffic exchanged at our IXP leveraging a set of known application protocols identified as potential attack vectors in each of the distinct categories defined (US-CERT, 2019). For the application traffic analysis, we are aware that, unfortunately, using traffic flows and port numbers alone provide a severely limited mechanism for classifying applications (LABOVITZ et al., 2010; ROUGHAN et al., 2004).

Figure 6.3: Transport protocols mix seen in the Bogon traffic at the IXP (Week-1, May 2019), bytes and packets. List of protocols ordered by bytes.



	%bytes	%packets
GRE	87.8955%	72.0084%
ICMP	5.8189%	11.7116%
TCP	5.6082%	15.0982%
UDP	0.6513%	1.1526%
None	0.0160%	0.0032%
VRRP	0.0060%	0.0166%
OSPF	0.0034%	0.0077%
IPv6	0.0003%	0.0006%
IGMP	0.0003%	0.0008%
ICMP6	0.0001%	0.0003%

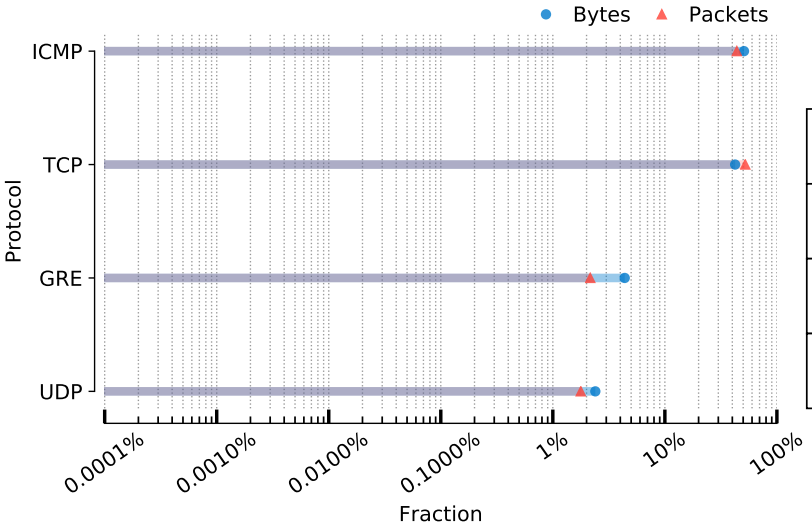
Source: by author (2019).

Transport protocols used to exchange Bogon traffic. In Figure 6.3, we can see the complete transport protocols list we identified when looking to the entire Bogon

traffic in the first week of May 2019. We observe that GRE clearly dominates the mix with a share of more than 87% bytes and 72% packets, representing 1.56 TB of absolute traffic contribution in that week alone. The other significant share of traffic is composed of the ICMP and TCP protocols, accounting for some 11.43% of the exchanged bytes. Other protocols such as TCP, VRRP, OSPF, etc. account for roughly 0.68% of the bytes exchanged at the IXP. The "None" entry is an indication of mal-formed packets, i.e., packets that did not contain valid data in the packet header. These usually happen due to some network equipment error during packet handling (e.g., processing overload, firmware bug).

We also looked if the same traffic proprieties hold if we isolate the multilateral Bogon traffic exchanged between members from the bilateral traffic. To enable this analysis, we re-run the classification tagging the Bogon traffic exchanged in bilateral sessions to our Unverifiable category. Figure 6.4 shows the results of the analysis. While we can see that there are Bogon packets exchanged in the multilateral peering, we observe less diversity on the transport protocols employed, being more pronounced ICMP and TCP suggesting the lack of adoption of the bogon static filters from the Best Current Practice (BCP) by a total of six members. Although GRE still appears, it is not the largest contributor to the category, confirming what the network operators explained on the reasons to use private addresses (see 6th paragraph on §6.2).

Figure 6.4: Transport protocols mix seen at the IXP (Week-1, May 2019) when applied filter to see only traffic exchange in multilateral agreements, bytes and packets. List of protocols ordered by bytes.



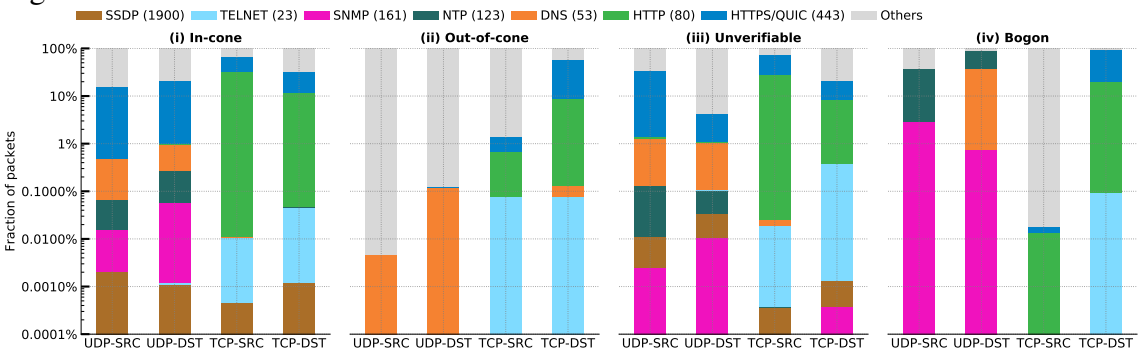
	%bytes	%packets
ICMP	50.8481%	43.8829%
TCP	42.3738%	52.1835%
GRE	4.3826%	2.1567%
UDP	2.3955%	1.7769%

Source: by author (2019).

Application protocols mix. Figure 6.5 shows a port-based application classification of packets considering four traffic categories, i.e., the ones which had traffic during

the analysis at the IXP. The data processed to generate this graph was May 1st 2019, but the overall behavior holds in a very similar way for other days. The overall traffic behavior, without considering the protocol level information in our categories, is best seen in Figure 6.1(c).

Figure 6.5: Traffic mix of application protocols seen at the IXP (May 1st 2019), fraction of packets for (i) In-cone, (ii) Out-of-cone, (iii) Unverifiable and (iv) Bogon traffic. We do not exhibit the Unassigned category because it has no traffic (as seen in Figure 6.1(c)). List of protocols is ordered from less to a more expressive presence across categories. Note the log scale in the shared Y axis.



Source: by author (2019).

We split our port-based classification according to (i) direction, i.e., SRC and DST port numbers, and (ii) the respective transport protocol, i.e., we focus on TCP vs. UDP. To achieve that, first we break the traffic into all protocol numbers located by analyzing the flow data in 5-min bins. Second, aiming to see if we locate attack traffic signatures into the distinct traffic categories, we pre-selected a set of application protocols identified as potential attack vectors based on the US-CERT alert list (US-CERT, 2019; ROSSOW, 2014). Then, we aggregate the traffic flow data of the whole day into each distinct application protocol located, and order by the amount of packets exchanged, select the seven most popular protocols and aggregate all the remaining traffic/port numbers into “others”. The list of protocols is ordered from less to a more expressive presence across categories. We note that port numbers in others are mostly randomly distributed, suggesting ephemeral port numbers. Note the log scale in this graph.

Before we discuss Figure 6.5 details, it is worth to mention that no network attack signs were identified, neither any misbehaved application protocol. Now, to understand how to read the figure, first concentrate on TCP traffic. Take the case of In-cone web traffic HTTP (80) and HTTPS/QUIC (443), which we expect to see both directions. Packets from clients to servers carry 80/443 in their packet DST field, and reply packets from servers carry 80/443 in their packet SRC field and an ephemeral port number in the DST field. Analyzing Figure 6.5 (i) In-cone TCP-SRC (HTTP+HTTPS correspond to 65.36% of the

traffic) and TCP-DST (HTTP+HTTPS correspond to 32.9% of the traffic) bars we can see that this interaction is well-reflected. The behavior just described needs to change if this were spoofed traffic, then it is expected to see traffic flowing in higher rates in only one direction (it can vary accordingly with the attack strategy and attack type employed, e.g., flooding or amplification). Looking to (ii) Out-of-cone and (iv) Bogon categories on TCP-DST packets with HTTP and HTTPS in the header we still can see both traffic directions, i.e., client/server communications happening (besides, remember that the traffic spike in that day was less than 30Mbps, as you can see in Week-1 Figure 6.1(c)). Otherwise, if we instead had seen only TCP-DST traffic direction and expressive traffic rates it could be a hint towards a flooding attack destined to HTTP/HTTPS servers, requiring further analyses.

For the UDP traffic, we can observe the same expected behavior, i.e., we find traffic flowing in both directions, requests to, and responses from key Internet service applications (e.g., DNS, NTP, SNMP). Recall that the set of protocols we analyzed are the most popular (US-CERT, 2019; ROSSOW, 2014) to be employed in DDoS attacks (Chapter 2 - §2.1). Interestingly, Figure 6.5 shows that the behavior of such ports for the In-cone, Out-of-cone, and Unverifiable, both SRC and DST, follow a very similar pattern, with very low fraction of packets in all them, with the only difference between these categories being which protocols appear in each one. Table 6.5 shows the percentage in the format of SRC/DST for traffic found in these categories, for each protocol.

Table 6.5: Percentage of UDP traffic mix of application protocols, per analysis of Figure 6.5. The percentages are shown in the format of SRC/DST traffic.

	DNS	**NTP**	**SNMP**	**SSDP**
(i) In-cone	0.42 / 0.69%	0.05 / 0.21%	0.01 / 0.05%	0.002 / 0.001%
(ii) Out-of-cone	0.004 / 0.12%	–	–	–
(iii) Unverifiable	1.12 / 0.95%	0.12 / 0.07%	0.002 / 0.01%	0.008 / 0.02%

Source: by author (2019).

The more intriguing case goes back to Bogon traffic. 36.65% of all Bogon UDP packets carry port number 53 as DST and, hence, are destined to DNS servers. However, as shown in Figure 6.1(c) we found very low traffic volume and packet rates (precisely 144 packets/s on DNS DST packets) which does not show signs of an attack, although it suggests bad configuration practices (DNS open resolvers (KüHRER et al., 2015)).

Next, we study how the results of our methodology, Spoofer-IX, cross-check with the CAIDA Spoofer Project (CAIDA, 2018c). The Spoofer Project has been collecting data on the deployment and character of IP source address validation on the Internet since

2005 (BEVERLY; BAUER, 2005; BEVERLY et al., 2009). These measurements are publicly available (CAIDA, 2018c; CAIDA, 2019a), allowing us to cross-check active measurement inferences of spoofability with our findings.

6.5 Lack of SAV Compliance Cross-Check

In 2005, Beverly et al. (BEVERLY; BAUER, 2005) developed a client-server technique to allow users to test networks to which they are currently attached and operationalized a platform to track trends from February 2005 to April 2009 (BEVERLY et al., 2009). Since 2015, when UCSD/CAIDA took over development and support of the spoofer infrastructure, the collected data accounts for 6845 autonomous systems (10% of the total routed ASes) in 207 countries (CAIDA, 2018c; LUCKIE et al., 2019) as of August 2019. This system required a user to download and execute the client software once per measurement, limiting coverage. Data from the project comes from participants who install the active probing client. The client automatically runs tests both periodically and when it detects a new network attachment point. We analyze the rich dataset of Spoofer tests for the same time windows as our traffic data.

Table 6.6: Congruity between CAIDA's public Spoofer dataset and inferences using the IXP. Of the 35 ASes that overlapped, CAIDA's Spoofer Project dataset inferred 54% of them had not deployed SAV, because CAIDA received a packet with a spoofed source address. Of the overlap, only 4 of the 35 (11%) were observed to forward an Out-of-cone packet into the IXP, and only 2 of these were also in CAIDA's Spoofer Project dataset as also not deploying SAV.

Spoofer-CAIDA	Spoofer-IX		Sum
	In-cone	Out-of-cone	
Spoof-received	17	2	19 (54.3%)
Spoof-blocked	14	2	16 (45.7%)
Sum	31 (88.6%)	4 (11.4%)	35

Source: by author (2019).

There were 203 members in the IXP we analyzed with Spoofer-IX in May 2019. We inferred spoofed traffic for 38 members (18.7%). A fraction of the 203 members, 17.2%, or 35 members, were also in CAIDA's public Spoofer dataset (CAIDA, 2018c; CAIDA, 2019a), which requires a volunteer to have been present in the network to run an active measurement test (Chapter 3 - §3.1). Table 6.6 summarizes the (in)congruity between the two datasets. We can see the results from the Spoofer project with these 35 ASes as ground truth. The Spoofer dataset indicated that 54% of the 35 members had

not deployed SAV, while 46% did. Considering the same set, Spoofer-IX detected only 4 members (11%) as sending Out-of-cone spoofed packets into the IXP. The result indicates that this specific IXP may not provide effective visibility into SAV deployment *because participants were not forwarding spoofed packets*, at least during our five-week observation window. It is not related to the accuracy of Spoofer-IX.

The quantitative differences in the measurements reflect both the different vantage points and the fundamental difference between the *ability to spoof and actual spoofing*, as carried out and visible in passive traffic flow traces. More generally, these results show the importance of our measurements. We imagine its utility as part of an expert system suite of cybersecurity services or compliance practices of modern IXPs.

6.6 Spoofer-IX vs State-of-the-art

This section compares the Spoofer-IX against the State-of-the-art in identifying spoofed traffic at IXPs, i.e., Full Cone (LICHTBLAU et al., 2017). To compare our approach with FC, we reproduce FC's methodology and analyze how reliably the method leads to the conclusions. We show that FC lacks the appropriate treatment of the fundamental challenges (Chapter 4), which can lead to incorrect identification of spoofed traffic.

The analyses and results presented here aim to: *i.* show the differences in the precision of the traffic classification results and *ii.* dive into the reasons for the differences that arise between the two methods. First, we describe the comparison procedure along with the summary of the algorithm's parameters (§6.6.1). Following, we analyze the traffic classification results, comparing both methods (§6.6.2). Last, we discuss the causes of discrepancies seen in the results (§6.6.3).

6.6.1 State-of-the-art Comparison Procedure

Table 6.7 presents side-by-side the parameters used in both algorithms to build the cones of each method before the traffic classification stage. For the Full Cone (FC) we used the same values as the authors defined – nine days (LICHTBLAU et al., 2017) of public BGP data as input; and for the Prefix-Level Customer Cone (PLCC) we used seven days, corresponding to each specific week of traffic data we classify, following our

methodology (Chapter 5). The analyses performed take as input the same traffic flow data for both methods, as described in §6.1 (Vantage Points Datasets, IXP-BR: traffic).

Table 6.7: Parameters that define the input data for both cones inference algorithms – Prefix-Level Customer Cone (PLCC) and Full Cone (FC).

Parameters	**Prefix-Level Customer Cone (PLCC)**	**Full Cone (FC)**
(1) BGP input time window	7-days	9-days
(2) Number of monitors	RIPE Routing Information Service (RIPE RIS) & RouteViews project (RV) (18 and 16 in 2017, 19 and 18 in 2019 RIPE RIS and RV, respectively)	
(3) Files per monitor	one file per day	all available files
(4) Use of RIBs / Updates files	RIBs only	RIBs and Updates

Source: by author (2019).

Even though Lichtblau et al. (2017) did not originally release the source code of the method, needed to reproduce it, we were able to obtain some portion of the code with the authors. Additionally, we did interactions by email and a remote meeting with two of the authors for clarifications. The portion of the code we obtained allowed us to reproduce their Full Cone methodology (LICHTBLAU et al., 2017). We had to develop our own version of their traffic classification scheme, and for that we followed the information available in the publication [1]. We did not obtain their datasets, which were protected by NDA agreements, like ours, so we applied both methods, ours and theirs, to our dataset.

Metrics for atypical network events analysis. We study the traffic classification results looking for atypical network events, which could hint to network attacks. We investigate two common attack strategies, namely flooding and amplification/reflection (recall Chapter 2 - §2.1). Recall that flooding attacks are often carried out using a wide range of source IPs, while amplification attacks require selective spoofing of source IPs of victims. To accomplish our goal, we use two distinct metrics to assess the IPv4 address activity and behavior over time (RICHTER et al., 2016; BENSON et al., 2015).

1. *Activity and Churn in active IPv4 addresses* allow to measure the volatility of address activity over time. To capture changes in the population of active addresses, we define a *gained event* if an address is not seen in a given window of time, e.g., 5-min or a day, but then is seen in the subsequent window. Conversely, a *lost event* occurs if an address is seen in a given window of time, but not seen in a subsequent window.

[1] The authors said that they "have hardcoded information of the IXP itself, making it hard to share" their traffic classification code.

The entry called *same event* measures how many addresses are active in a given window of time, and still active in a subsequent window.

2. *Spatio-Temporal Properties in active IPv4 addresses* captures aggregated properties of active IPs seen within each time-window, e.g., each 5-min, 1-hour, a day. We analyze the components of a given traffic category, through which we can glean considerable information from flow-level data. We quantify the number of unique ASNs, BGP prefixes, and countries. The number and diversity of sources facilitate insight into the overall traffic behavior. These numbers help identify components that enable opportunistic network inferences, characterize the frequency and granularity of traffic sources.

Analysis execution steps. In order to perform the comparison, we proceed with distinct full-runs of Spoofer-IX and Lichtblau et al. (2017) method. In Chapter 5 - §5.3, Figure 5.4 depicts five steps that should be executed for both methods, i.e., Spoofer-IX and Full Cone, to compute the metrics. In particular, the analysis requires first the complete execution of each cone construction method. Following, we classify the traffic in accordance with each proposal. Lastly, we proceed with the metrics computation. In other words, we derive the traffic categories classification results of each method, which are then used as input to the metrics computation, revealing the corresponding traffic behaviors.

6.6.2 Traffic Classification Comparison

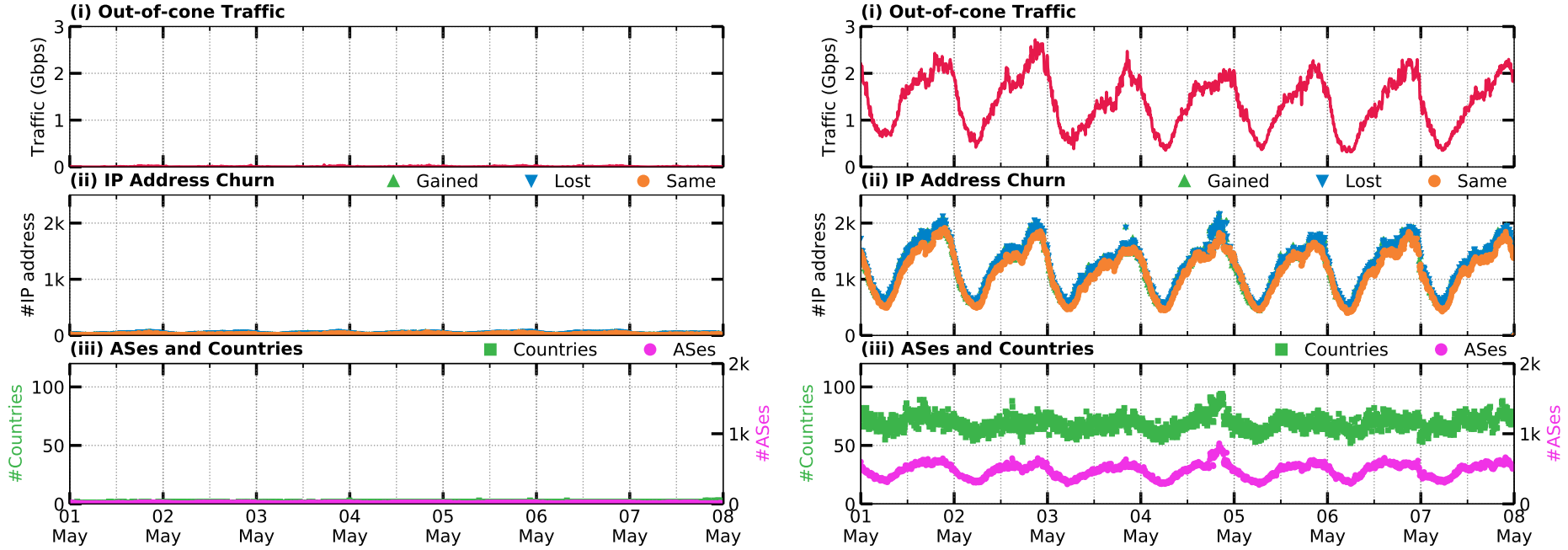
Figure 6.6 (in page 89 for better visualization) shows the volume of Out-of-cone traffic inferred by both the Spoofer-IX (6.6(a)) and Full Cone (6.6(b)) methods for traffic data captured during the first week of May 2019. There are two additional subplots with custom scales, one for each of the metrics defined in our analysis methodology (§6.6.1). We compute each metric per 5-minute window of traffic data and use the same range on Y axes between methods for ease of comparison.

The Spoofer-IX method infers a peak of (just) 40Mbps of Out-of-cone traffic (so small that it is visible only in §6.2 - Figure 6.1(c)), whereas the Full Cone method infers a peak of 2.5Gbps. The diurnal pattern of the inferred Out-of-cone traffic matches user-demand for content, with no observable peaks suggesting a volumetric spoofed-source attack launched from within member ASes of the IXP. The second row of Figure 6.6 shows churn in source IP addresses seen in each five-minute window, results obtained by the

Activity and Churn metric. For the Full Cone method, the absolute volume of source addresses observed follows the traffic volume profile (Figure 6.6(b)(i)) as a whole. Note that the gain and lost curves are superimpose, addresses are being replaced at the same rate, with a stable amount of addresses being used. In addition, looking to the third row we can see that the traffic is concentrated in 478 ASes and 69 countries per five minute window on average, as computed by the Spatio-Temporal Properties metric. This is not a typical pattern of attacks that utilize randomly-spoofed source addresses that would be spread throughout the address space. Actually it is the opposite: the observed behavior indicates normal traffic being delivered to members of the IXP by a variety of ASes spread in distinct countries. Following, we analyze what is behind the different results between the two methods.

(Note: intentionally left blank to provide in the next page the best visualization of Figure 6.6.)

Figure 6.6: Comparison of metrics for Out-of-cone traffic inferred by the Spoofer-IX and Full Cone for the first week of May 2019. We compute each metric per 5-minute window of traffic data, and use the same range on Y axes between methods to allow for comparison. For the IXP we studied, the Full Cone method inferred an average of 1.5Gbps of spoofed traffic, whereas our methodology inferred a maximum of 40Mbps (best seen in Figure 6.1(c))-iii.



(a) Spoofer-IX / Prefix-Level Customer Cone

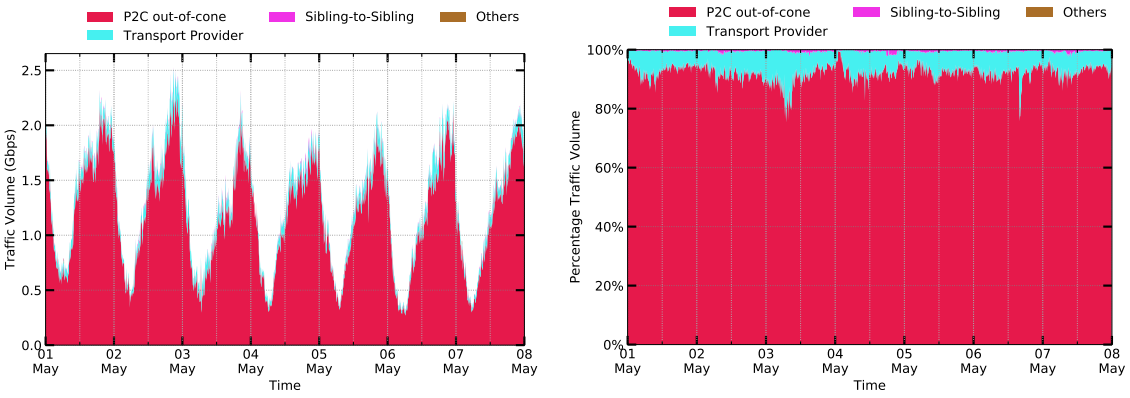
(b) Full Cone

Source: by author (2019).

6.6.3 Analysis on discrepancy of classification results

The discrepancy in the amounts of Out-of-cone traffic found between the Spoofer-IX and the state-of-the-art is because the latter classifies as Out-of-cone three types of traffic: Provider-to-Customer, Transport Provider to Customer, and Sibling-to-Sibling. In contrast, Spoofer-IX classifies these as Unverifiable. Figure 6.7 provides the classification breakdown of the Full Cone Out-of-cone traffic, as seen before in Figure 6.6(b) (i), by the lens of the Spoofer-IX method. The traffic breakdown is shown in absolute (6.7(a)) and relative (6.7(b)) volume values and refers to the first week of May 2019.

Figure 6.7: Classification of *Out-of-cone traffic for the Full Cone* through the lens of the Spoofer-IX. We infer that 92.6% of this traffic was from a provider to customer across the IXP. However, because a provider can transit traffic from any source IP address to their customer, it is incorrect to identify spoofed packets by their IP address crossing an IXP from a provider to a customer.



(a) Absolute Unverifiable Traffic. (b) Relative Unverifiable Traffic.

Source: by author (2019).

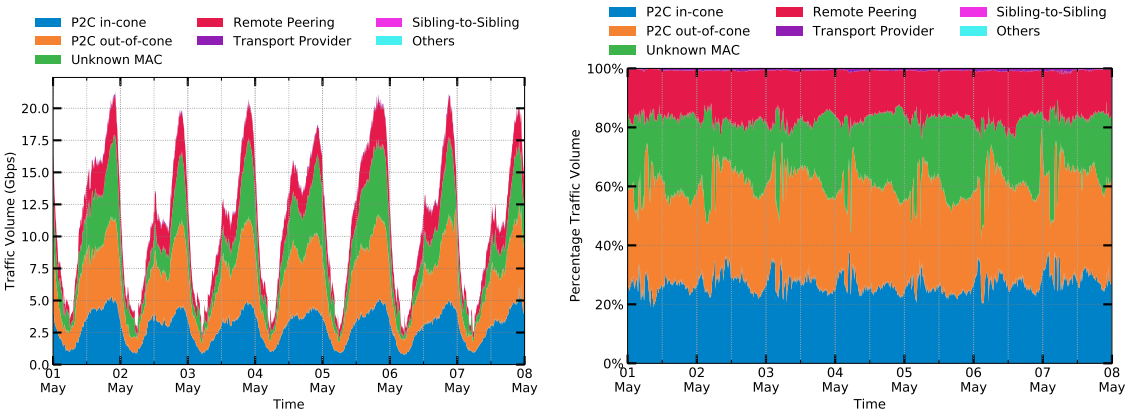
Figure 6.2 in §6.3 shows how Spoofer-IX classified Unverifiable traffic. That includes 1 — 5 Gbps of traffic from Providers to Customers (represented in orange). When we classified the Full Cone's Out-of-cone traffic using the Spoofer-IX method, 92.6% of the traffic was from a Provider to a Customer across the exchange, carrying 0.5 – 2 Gbps of traffic (Figure 6.7(a)).

Finally, the traffic volume classified as In-cone by the Full Cone method is larger than with Spoofer-IX. 85.5% of the traffic that the Full Cone method classified as In-cone was also classified as In-cone by the Spoofer-IX method, with the remaining 14.5% classified as Unverifiable by Spoofer-IX (which means that with the Full cone method there could be up to 14.5% false negatives, i.e. undetected spoofing). We analyzed the In-cone traffic results of Full Cone through the lens of the Spoofer-IX, as in Figure 6.8. Like Figure 6.2, the traffic breakdown is shown in absolute Figure 6.8(a) and relative

Figure 6.8(b) volume values referring to the first week of May 2019.

First, because the Full Cone includes all prefixes where an AS was observed in the AS path of a BGP route in an AS's Full Cone (recall Chapter 4 - §4.1.1), traffic from providers and peers of that AS was inferred as In-cone for that AS, corresponding to an average of 60% of their In-Cone Unverifiable traffic. Second, as they do not check for traffic direction and types of AS relationships, the fact of having Unknown MAC addresses (23.34% of traffic on average) does not interfere with their classification process since they only care to check if the packets SRC address matches or not the Full Cone. Third, they ignore the presence of Remote Peering agreements (16.02% of traffic on average) and classify the traffic regardless of not having the correct information. Putting together the results of both Figures 6.7 and 6.8 gives us the original values computed by the Spoofer-IX methodology, as seen in Figure 6.2 in §6.3.

Figure 6.8: Classification of *In-cone traffic for the Full Cone* through the lens of the Spoofer-IX. 60% of this traffic was from a provider to customer across the IXP. Because the Full Cone includes all prefixes where an AS was observed in the AS path of a BGP route in an AS's Full Cone, traffic from providers and peers of that AS was inferred as In-cone for that AS.



(a) Absolute Unverifiable Traffic. (b) Relative Unverifiable Traffic.

Source: by author (2019).

Both, Spoofer-IX methodology and Full Cone, are subject to False Positives (traffic labeled as spoofed when it is valid/In-cone) and False Negatives (traffic labeled as valid when it is spoofed/Out-of-cone). As the results above demonstrated, the work of Lichtblau et al. (2017) suffers from a considerable amount of False Positives and False Negatives, while Spoofer-IX provides accurate traffic classifications (i.e., minimize both False Positives and False Negatives). These results show that Spoofer-IX avoided wrong classifications through the definition of precise filters to isolate the portion of the traffic that is not feasible (e.g., Provider-to-Customer, Transport Provider-to-Customer, Sibling-to-Sibling) or uncertain (e.g., Stray, Remote Peering, Bogon in VLAN) to identify spoofed

packets, making them Unverifiable (check Table 6.4 in §6.3).

Even though we have inferred Out-of-cone traffic with Spoofer-IX at our IXP, there are still edge cases we have not yet discussed, as some of the traffic appears to have signatures of legitimate traffic (recall discussions in §6.4). Next, we look into the Out-of-cone traffic nature and discuss its properties.

6.7 Looking at the Out-of-cone Traffic Nature

In this section, we look specifically at the Out-of-cone traffic tagged by the classification process. As discussed in §6.2, we believe that the amount of Out-of-cone traffic crossing the IXP core switch is even lower. To confirm our intuition, we manually check if there are any signs of attack traffic behavior, such as flooding and amplification. We also investigate what are the potential factors/properties impacting on these results. We use space-filling Hilbert Curves (HEIDEMANN et al., 2019; WESSELS; CLAFFY, 2019) to generate a map of IPv4 address usage extracted from the Out-of-cone category and Sankey diagrams, exploring the properties of members exchanging this traffic.

Figure 6.9 shows Hilbert heatmaps, one per day for the entire Week-1 of May 2019. The heatmap presents the usage of the IPv4 address space in each day according to the source IP addresses of Out-of-cone packets resulting from our classification processing. The IPv4 address space is rendered in two dimensions using a space-filling continuous fractal Hilbert curve (HEIDEMANN et al., 2019; WESSELS; CLAFFY, 2019) of order 16. Each square in the figure represents a /8 IP prefix block; the numbers in each square indicate the number of the first IPv4 octet. Each colored dot represents how many IP sources generated traffic within a given /16 from each block, with blue and red meaning low (from 1) and high counts (above 255), respectively. The color black means no packets with a source address in the /16 block. The green rectangular shapes denote reserved address space blocks by IETF RFCs (IANA, 2018b).

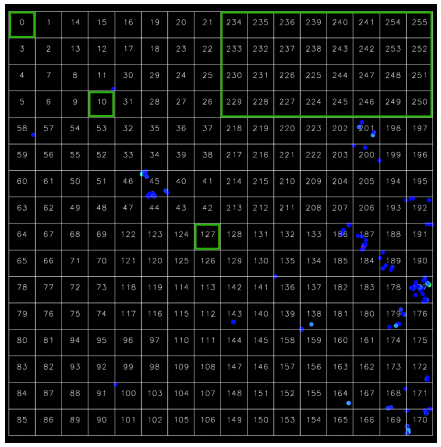
We observe a clear pattern in terms of IP address space usage across hours and days, suggesting that this Out-of-cone traffic is legitimate and not associated with attacks. Further, the plots showed no indication of random exploration of the IP space (e.g., multicast IP ranges, reserved blocks, and military prefixes), which might indicate an attack (MAJKOWSKI, 2018a).

Next, we examined the top five prefixes by usage of its IP space. We checked the AS owners of such prefixes, as well as the ASes' business type classification of the

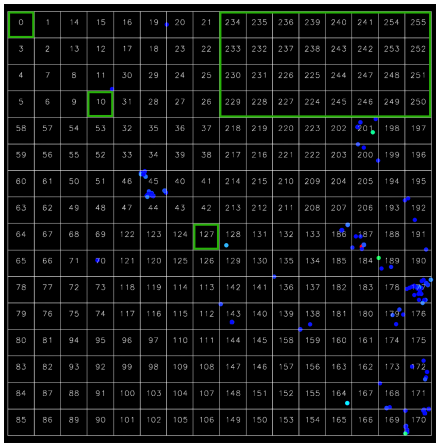
corresponding AS. For packets with source address within these top prefixes, we examined the respective ingress AS crossing the IXP infrastructure. In the list of ingress ASes, we found regional ISPs present in more than one IXP and being also part of complex relationships, i.e., group associations (variations also include partnerships/services exchange and franchising) leveraging transport providers to reach their intended destinations.

(Note: intentionally left blank to provide in the next page the best visualization of Figure 6.9, Hilbert heatmaps.)

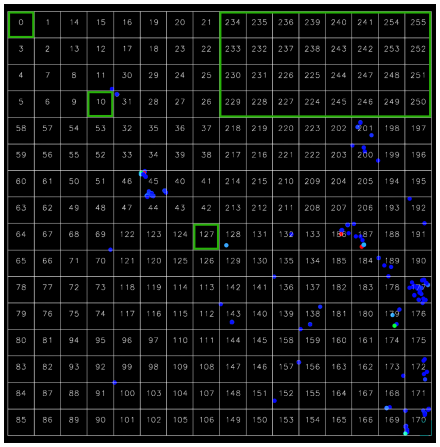
Figure 6.9: Hilbert heatmap visualization showing the utilization of the address space according to the Out-of-cone traffic that is left (Week-1, May 2019). The IPv4 address space is rendered in two dimensions using a space-filling continuous fractal Hilbert curve (HEIDEMANN et al., 2019; WESSELS; CLAFFY, 2019) of order 16. Each square in the figure represents a /8 IP prefix block; the numbers in each square signs the number of the first IPv4 octet. Each colored dot represents how many IP sources generated traffic within a given /16 from each block. The level of activity is indicated by colors, from blue (low) to red (high), with green, yellow and orange as moderate levels, and black meaning no packets with a source address in the /16 block. Green boxes denote reserved address space blocks by IETF RFCs.



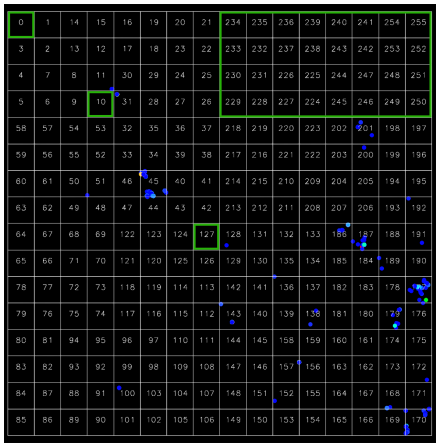
(a) May 01



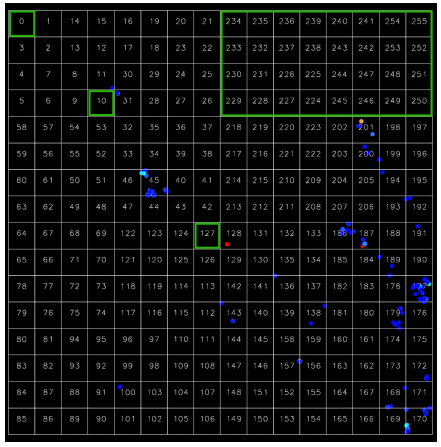
(b) May 02



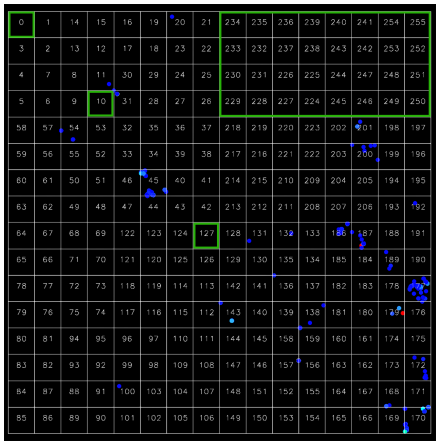
(c) May 03



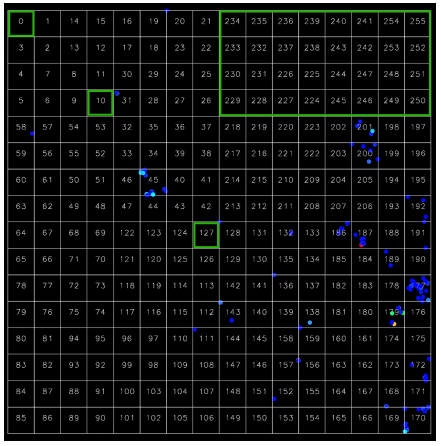
(d) May 04



(e) May 05



(f) May 06



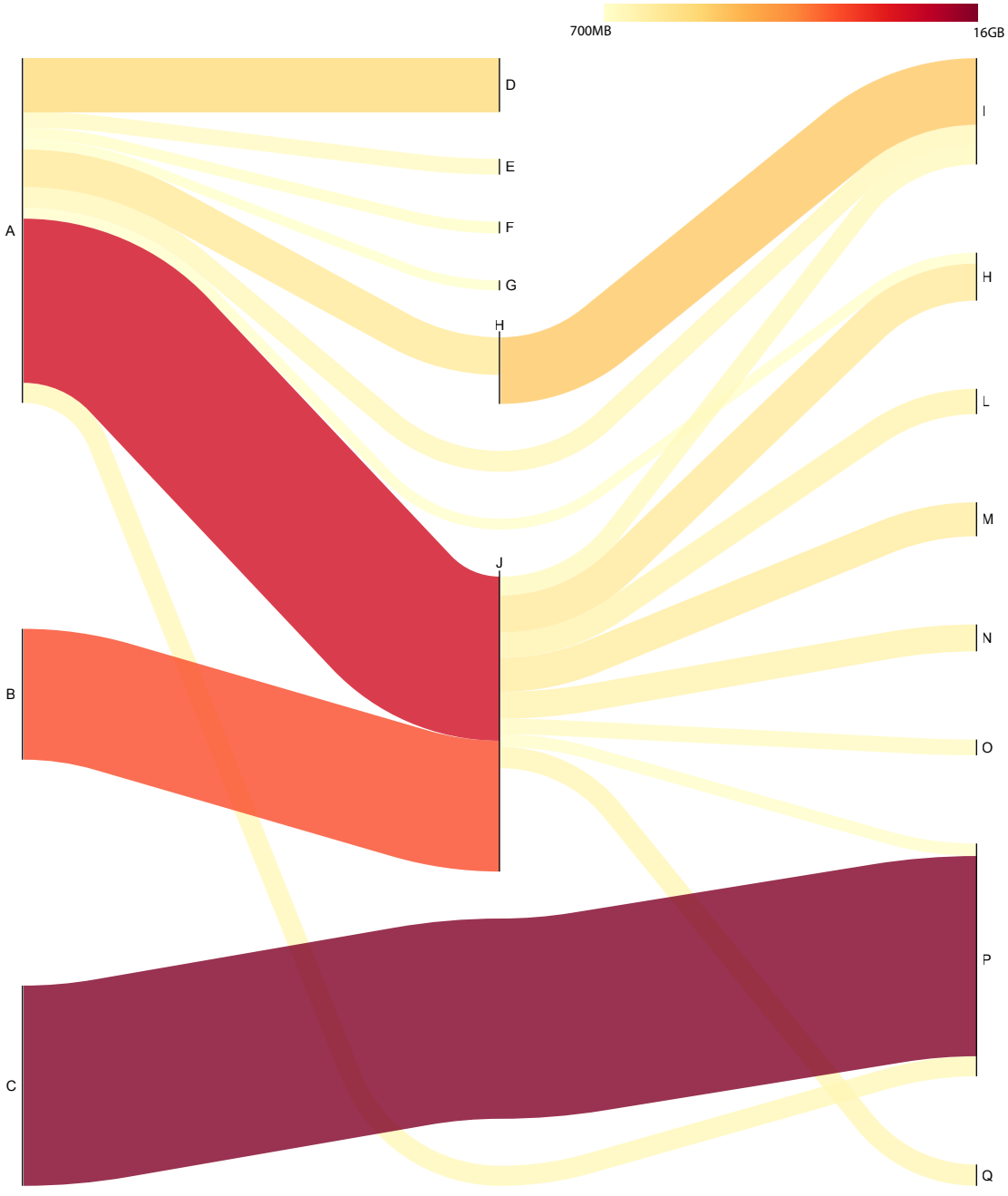
(g) May 07

Source: by author (2019).

Figure 6.10 shows a complementary view of the traffic properties, depicting the traffic exchanged between members through a Sankey diagram. It shows the top 20 heavy-hitter pairs of members by the total traffic volume exchanged in the Out-of-cone category on May 03, 2019 (where the peak of 40 Mbps is present). The width of the connections between AS pairs is proportional to the traffic volume exchanged. This view helps understand who are the members involved in the communications and how the traffic from Figure 6.9 reaches the IXP. In summary, by analyzing the data, we found that the regional ISPs mentioned before relied on Transport providers not previously known/mapped (A, B, C, and J in the diagram) to reach the peering fabric and exchange traffic. So, three traffic properties can contribute to false positives in the out-of-cone traffic, i.e., complex AS relationships, AS presence in multiple locations, and unknown transport providers carrying on the traffic via layer-2.

(Note: intentionally left blank to provide in the next page the best visualization of Figure 6.10, Sankey diagram.)

Figure 6.10: Sankey diagram with the top 20 pairs of members by the total Out-of-cone bytes exchanged in May 03, 2019. The width of the connections is proportional to the volume of bytes exchanged. Note that the ASNs were replaced by a letter due to a non-disclosure agreement.



Source: by author (2019).

6.8 Perspectives on Filtering Consistency by IXP Members

In the previous sections, we examined the traffic properties, made inferences about the lack of SAV and the incidence (or lack of) attacks exploiting spoofing in the periods considered. We now investigate the problem using a higher-level perspective to answer three questions. The first is how consistent is the overall application of filtering policies among all AS members of the IXP? This information offers opportunities for IXPs coordinators to help members improve compliance with missing practices or even to define better regulatory practices when new members connect to the infrastructure (§6.8.1).

The second question regards changes, hopefully, improvements in the adoption of SAV through time (§6.8.2). Even though our data is limited to two periods, 2017 and then 2019, it may be enough to determine a trend as long as a visible change in the adoption of SAV is present. Can we see positive changes in the behavior of members regarding SAV compliance? The results allow us to see members who (potentially) have never deployed or are not following the BCPs correctly.

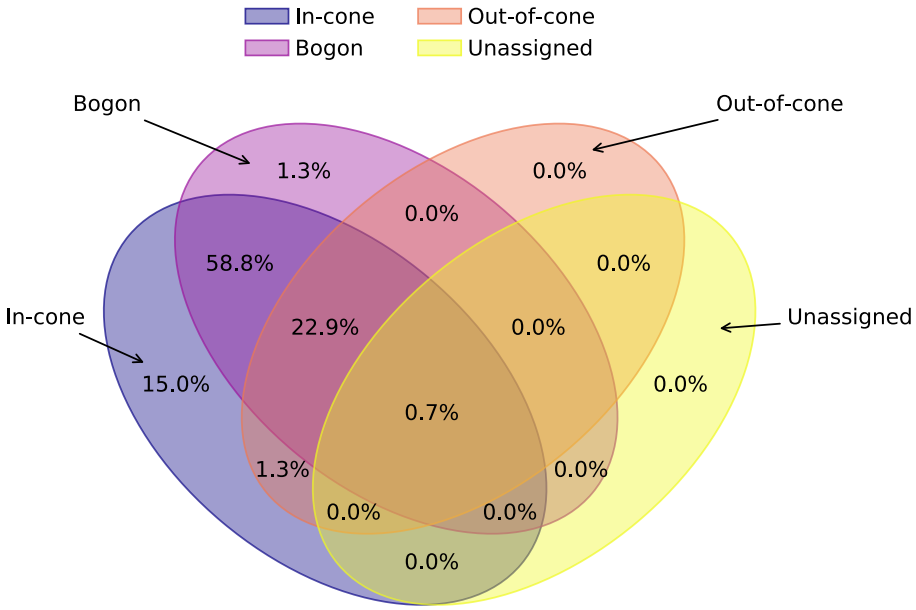
The third and last question relates to the switched fabric infrastructure itself, a profile analysis of the traffic behavior correlated with the physical Colocation Facilities, where members interconnect (§6.8.3). Is there any concentration of members failing to protect their networks among the CFs? We aimed to understand this because we believe that IXP coordinators should split with its CFs the burden of auditing traffic to confirm that each network performs SAV, therefore serving as regulatory agents (more on that on Chapter 7), as well as connectivity enablers.

6.8.1 Filtering Consistency Behavior

Figure 6.11 presents a Venn diagram with the percentage of members at the IXP contributing with traffic to the distinct categories, as well as intersections in contributions. The results in the plot refer to all packets exchanged in the traffic collected during the five-week period in 2019. For each packet, we inspect the associated category (as defined in Chapter 5, see classification diagram in Figure 5.3) and use MAC-to-ASN mapping (§6.1) to identify the member AS emitting these packets. As in previous work (LICHTBLAU et al., 2017), the percentages reflect lower bounds on which filtering strategy member ASes apply, as an AS may not send flows with spoofed source IP addresses across the IXP during our observation window. We argue that these lower bounds are usefully tight given

the length of the observation period (15 weeks, spanning two years).

Figure 6.11: Four-Set Venn Diagram with the percentage of members contributing traffic to the following categories: In-cone, Out-of-cone, Bogon and Unassigned. Analyses performed for the five week period of 2019, May 01 to June 05 2019. Note that the area sizes in the plot are not shown in proportion.



Source: by author (2019).

Interestingly, not all members appear as source of traffic. Out of 203 active members in 2019 during the five weeks, 154 (75.86%) members appeared as source of traffic at the IXP. From those 154 ASes, 15% did not send any traffic classified as either Out-of-Cone, Bogon, or Unassigned, i.e., their traffic was clean. On the other end of the spectrum, one AS (0.7%) contributed traffic to all four categories; it proves that there are indeed networks without any kind of filtering deployed. Surprisingly, around 1.3% of participants contribute with Bogon traffic only, the easiest to avoid due to its static nature. This indicates a gross misconfiguration and potential vulnerability; we responsibly notified these networks by means of their IXP. According to operators of the networks involved, the problem was caused by an updating procedure in routers accidentally deleting the filter for bogon ranges. A single AS member contributed packets in the Unassigned category. This same member also sent Out-of-cone, Bogon, and In-cone packets. Not surprisingly, no member contributed Out-of-cone traffic exclusively. Almost 23% (35 members) contributed with In-cone, Out-of-cone, and Bogon packets, while 58.8% (90 members) contributing with both In-cone and Bogon traffic. Lastly, 1.3% (2 members) contributed to In-cone and Out-of-cone.

Considering all 200+ members at the IXP, few contributed with potentially spoofed

traffic. Recall that for members to control the exchange of Bogon traffic requires relatively static filters at their network devices that do not need frequent updating as topology and customers change (Chapter 2 - §2.2). In contrast, SAV filtering (of Out-of-cone traffic) requires updating filters as business dynamics change. It thus surprised us to see more networks exchanging Bogon traffic than Out-of-cone traffic. We explained this mystery with our previous discovery that AS members occasionally use Bogon source IPs to exchange traffic via tunneling protocols (e.g., GRE, IP-in-IP) with another member at the same switching fabric. Nevertheless, the presence of Out-of-cone traffic suggests that those member ASes sending it do not strictly enforce SAV according to the BCPs (FERGUSON; SENIE, 2000; BAKER; SAVOLA, 2004).

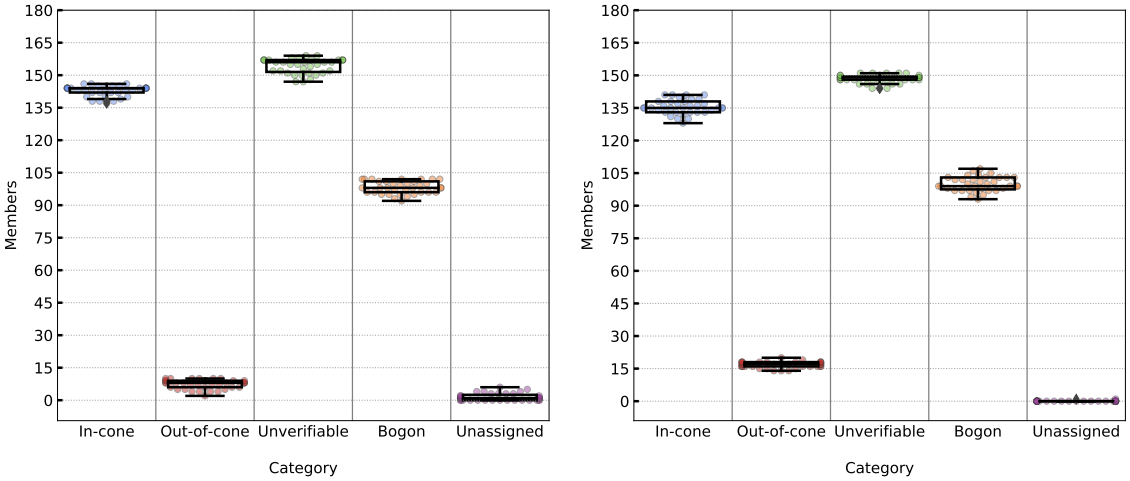
6.8.2 Trends of Filtering Over Time

This analysis investigates the trends of filtering configurations over time by the members at the IXP. We show the number of participants present in each category for the two years analyzed. Towards this end, Figure 6.12 shows a Swarm plot (i.e., a categorical scatterplot) overlapped with a Box plot considering the values of all categories (In-cone, Out-of-cone, Unverifiable, Bogon, Unassigned). Each circle in the swarm indicates the total number of members per day over five weeks of 2017 (Figure 6.12(a)) and another five weeks of 2019 (Figure 6.12(b)). Box plot values show the minimum, maximum, average (square), median (line inside square), lower (25th), and higher (75th) quartiles for the number of IXP members over time in each category.

As we can observe, even with a two-year difference between the traffic flow data analyzed, the overall behavior remained mostly equal for all categories. This result is not surprising, as typically operators minimize interventions to deployed operational devices to avoid service disruptions (LUCKIE et al., 2019). These results are consistent with those observed by active measurements (CAIDA, 2018c). In all of them, the box maximum and minimum values of each category have little or no difference. The highest standard deviation among categories is 3.38 members in Bogon category (Figure 6.12(b)), where we observed the deliberate use of these prefixes with tunneling protocols (discussed in §6.2 and §6.4). Moreover, there is no presence of outlier observations.

These results indicate that there has been no significant progress in terms of deployment or remediation of SAV in the networks evaluated in the period 2017-2019. We cannot state this behavior applies to the Internet as a whole or will persist, but based on the

Figure 6.12: Swarm Box plot reflecting configured filtering practices over time. It shows the scatterplot distribution per category of the total number of members per day over five weeks of 2017 (6.12(a)) and another five weeks of 2019 (6.12(b)). The points represent the results of each day being shifted horizontally (only along the categorical axis) to avoid overlap, while in the y-axis they show the total number of members. The overlapping Box plot presents the minimum, maximum, median (line inside square), lower (25th) and higher (75th) quartiles values for the number of IXP members over time in each category after the classification process.



(a) May 1 - Jun 5, 2017. (b) May 1 - Jun 5, 2019.

Source: by author (2019).

present evidence, we cannot be optimistic about networks increasing protection against IP spoofing attacks.

6.8.3 Filtering Behavior of Members by Colocation Facilities

In our previous analyses regarding the filtering behavior, we focused on how many members contribute to the distinct traffic categories independently of their points of connection to reach the IXP core switching fabric (definitions and related challenges are presented in Chapter 2 - §2.5 and Chapter 4 - §4.2). In this analysis, we separate the members per Colocation Facilities based on ground-truth data (MAC-to-AS mapping - §6.1) in order to observe the behaviors on filtering strategies. The leading question was: – is there any CF with substantially more members failing to protect their networks?

Table 6.8 shows the behavior seen along the five weeks in 2019 (May 01 - Jun 05 2019). As explained before (Chapter 2 - §2.5, Figure 2.9), the architecture of modern IXPs is composed of a switching fabric that interconnects other switches located in remote physical Colocation Facilities. As the table shows, the IXP we studied (§6.1) had at the period of this analysis ten distinct active CFs, with a varied amount of members, as seen

in the 2nd column [2]. Now, to answer our question, take into account the absolute number of members connected in each CF (2nd column) and look to our set of distinct categories (3rd to 7th columns). We can see that there is no concentration of cases among the CFs.

Table 6.8: Breakdown of IXP members presence per Colocation Facility and the traffic categories over five weeks in 2019 (May 01 - Jun 05 2019). The table presents the absolute total number of members which appeared in each category (nonexclusive, since a member can contribute to more than one category over time) in the five-weeks analysis along with the corresponding fraction over the total number of members connected to each CF. The last row shows a sum over the period analyzed. We anonymized the Colocation Facilities real names and sorted in descending order of members connected (2nd column).

Colocation Facility	Members Connected	In-cone	Out-of-cone	Bogon	Unassigned	Unverifiable
CF-1	70	47 (67.14%)	15 (21.42%)	38 (54.28%)	1 (1.42%)	46 (65.71%)
CF-2	34	25 (73.52%)	6 (17.64%)	19 (55.88%)	0 (0%)	24 (70.58%)
CF-3	27	20 (74.07%)	5 (18.51%)	18 (66.66%)	0 (0%)	20 (74.07%)
CF-4	26	23 (88.46%)	6 (23.07%)	21 (80.76%)	0 (0%)	26 (100%)
CF-5	16	13 (81.25%)	1 (6.25%)	11 (68.75%)	0 (0%)	13 (81.25%)
CF-6	14	10 (71.42%)	3 (21.42%)	9 (64.28%)	0 (0%)	10 (71.42%)
CF-7	10	8 (80%)	2 (20%)	8 (80%)	0 (0%)	9 (90%)
CF-8	3	3 (100%)	0 (0%)	1 (33.4%)	0 (0%)	3 (100%)
CF-9	2	1 (50%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)
CF-10	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Sum	203	151 (74.38%)	38 (18.71%)	128 (63.05%)	1 (0.49%)	154 (75.86%)

Source: by author (2019).

6.9 Discussion on Validation Efforts

The validation of results is a big challenge for the Spoofing problem, and the complexity only grows in the context of IXPs as vantage points. There are two key reasons. First, there is no global registry that contains ground truth on which addresses are valid source addresses for packets transited by an AS. Second, the IXPs are composed by dynamic infrastructures, interconnecting many networks which in turn are affected by economic factors that shape the networks market dynamics with buy/sell/merge/partnership operations, periodically changing the network organization landscape (Julio Wiziack, 2019; Bnamericas, 2019; David Shepardson, 2017; Steve Evans, 2014). Even so, we made distinct efforts to validate our results. Due to the complexity, we did the validation in a

[2]There are a lot of distinct factors that make one CF have more members connected than others. The top five factors, according with discussions with IXP operators and IXP members are: i. for how long the CF has been operational, ii. strategic physical location (in order to ease members' reach the facility), iii. additional services offered (e.g., strategic CDNs, transport providers, critical services co-location), iv. quality of services and technical support, and v. financial costs.

two independent phase process. First, the Spoofer-IX components based on inferences, and second the results obtained with the use of our method.

Spoofer-IX components. Recall Chapter 5 - Figure 5.1, where we illustrate an overview of the Spoofer-IX components. Note that we have two components that are built based on algorithm inferences; these are, in order of importance, the *Prefix-Level Customer Cones* and the *Sibling ASes*. To validate the results we obtained with our PLCC algorithm, we used Route Servers and Looking Glass Servers data (§6.1). These datasets helped us to check the resulting AS relationships and prefixes inferred aimed at the members of the IXP under analysis, i.e., allowed to see if we have missed something or incorrectly established an inference between ASes.

In practice, to check the cones results, we proceeded with three distinct validation operations in different moments and forms. First, we developed a code to check for missing prefix announcements in the resulting cones. We used the IXP BGP data (i.e., Route Server data) as a baseline. Second, we manually analyzed the top ten individual prefixes with the highest matching results after classifying traffic, double-checking if the prefix was valid (we checked its announcement). Last, we correlated and annotated our traffic classification using cones information, external mapping data (e.g., MAC-to-AS, prefix-to-ASN), and the IXP BGP data doing individual cases analysis when necessary.

Regarding the Sibling ASes, we found some missing from CAIDA's public AS-to-Org dataset (HUFFAKER et al., 2019) while we were digging into the properties of the Out-of-cone traffic category. So, we investigated which ASes were siblings in three complementary ways: by consulting the official website of those ASes to find information on their ownership, contacting the ASes directly to enquire, or the IXP operators to understand the relationship between two ASes at the IXP.

Spoofer-IX Results. After the components validation, we use Spoofer-IX to classify the traffic and assess its results. In this stage, we had limited ground truth to validate our inferences of spoofed traffic events. The IXP did not have sufficiently granular data on their traffic monitoring systems to use on detailed analysis for validation; neither they had a security system that cared about spoofability. So we cross-validated our traffic classification results, specifically the Out-of-cone traffic, with our two metrics *Activity and Churn in active IPv4 addresses* and *Spatio-Temporal Properties in active IPv4 addresses* (see details in §6.6). Moreover, we corroborate the observed behaviors with reports from the IXP coordinators and its members, which the IXP operators helped contact.

Besides these efforts, a diverse set of other investigations were carried in individual

members of the IXP. The members investigated were selected based on their potential atypical behavior, e.g., high traffic volume exchanged, number of packets, diversity of members that it communicates within a given time window. With the members selected, we performed their traffic analyses. The list of traffic properties considered in the analyses include the traffic exchange direction, protocols involved in the communications, both transport and application layers, as well as their communications patterns. In combination with those analyses, we also looked to the businesses conducted by the members, checking if their traffic behavior matches, their peering agreements established and the presence in multiple locations, and the potential usage of unconventional routing setups (e.g., use of tunneling protocols, software for routing automation (NOCTION, 2019)).

Alternative strategies of validation. Although we executed extensive validation efforts like the ones mentioned above to increase the confidence of the outcomes, three other strategies could have been considered. The first alternative would be the use of simulation. One could develop a simulator to generate synthetic traffic aiming to test the logic behind the proposed method. This would be useful to perform unit tests for our developed codebase. The second alternative would be to generate synthetic traffic and inject it into the real dataset. This way, one could create scenarios of potential attacks not captured in our traces and show that the method can accurately identify them. Finally, one could hire a real attack (DDOS-BLACK, 2019; DDOS-STRESS, 2019; SECURITY-PROF, 2019). We did not consider that approach because we were working with production networks, an IXP and its members on the Internet, and authorizations would be required from each player, as well as satisfying many Ethical requirements (Lisa Vaas, 2019; Thomas Brewster, 2018).

7 TOWARDS SCALABLE INTER-DOMAIN SPOOFING MONITORING

In the previous chapters, we discussed the challenges, existing methods, and the results of spoofed traffic inference at IXPs. For that we collected sFlow traffic data from the core switch of the third largest Brazilian IXP, with up to 200 Gbps of traffic among 200+ members. However, while we chose an IXP due to its locality and the amount of connected networks, our method is not limited to IXPs. In principle, every network on the inter-domain Internet can opt to apply our method to detect spoofing. In this chapter, we discuss the use of Colocation Facilities as alternative/complementary vantage points to scale the execution and help bootstrap Spoofer-IX in large-scale peering fabrics.

The remaining of this chapter is organized as follows. In Section 7.1 we present our study on the feasibility to use Spoofer-IX in larger infrastructures, and in Section 7.2 we discuss on our methodology generality and limitations. Finally, in Section 7.3 we close the chapter with some considerations.

7.1 Scaling Spoofer-IX to More Complex IXP Architectures

In this section, we explored practical application and generalizability of our Spoofer-IX method and implementation to larger and more complex IXP infrastructures. In this context we believe the critical question lies in the feasibility of splitting the flow data collection across switching peering fabrics. Our goal is to maximize the ability for any networks on the Internet to detect spoofed traffic, including IXPs with diverse interconnection practices and network topologies that hinder the deployment of IP-based measurement methodologies. Note however that is out of scope here to dissect the second IXP at the same level we did in Chapter 6 with the first mid-size IXP. First we explain the datasets used in this study (§7.1.1). Following, we present our procedure to achieve our goal (§7.1.2). Lastly, we show the results (§7.1.3).

7.1.1 Datasets

To explore a case study, we partnered with a much larger Brazilian IXP (IX.br, 2020) [1]. This second IXP, has over 1,600 members and transports up to 6 Tbps, allowing an evaluation at scale, focused on feasibility. We followed the same methodology to collect its traffic. We record traffic data using sFlow (P. Phaal, S. Panchen, and N. McKee, 2001) with a configured sample rate of 1:4096 packets. From this very large IXP, we examine traffic exchanged during one day (April 12, 2018) in three distinct Colocation Facilities part of its switching fabric infrastructure.

In line with Chapter 6 - §6.1, we collected datasets to the corresponding period of the traffic under analysis, and we built new specific cones, as well as the required MAC-to-ASN mappings (recall Chapter 5).

7.1.2 Analysis Procedure

At the time of the analysis, this large IXP had over 30 Colocation Facilities and 150 switches. For such larger network infrastructures, the diversity of interconnection practices and local network topology arrangements that can exist may hinder the deployment of any IP layer measurement method. To this end, we extended the Spoofer-IX implementation to be more flexible about input parameters, and to run using information (i.e., traffic flow data, topology information, and MAC-to-ASN mapping from members) from individual networks. As a proof-of-concept, we collaborated with three Colocation Facilities that are part of this second IXP. We collected traffic flow data from eight individual switches across these distinct facilities, as well as topology and MAC-to-ASN mapping information, and ran our method individually (i.e., per switch).

The traffic analysis per switch enables us to scale the execution to much larger peering fabrics and lower the barrier to deploy the method, providing localized SAV compliance enforcement (also discussed previously in Chapter 6 - §6.8.3). The steps which should be executed are exactly the same ones as we did for the first IXP where the traffic flow was collected at the core switch. The process of bootstrapping the execution of our methodology (step 1, Chapter 5 - §5.3, figure 5.4) is mostly shared between all runs, if the

[1] Many times people do not have any idea of how complex are and how long does it take these procedures and negotiations to obtain access do this kind of data. To give an idea, to have access to a small sample data of the second IXP took 2 years and a half.

Table 7.1: One day of traffic for individual switches of three distinct Colocation Facilities of a second IXP in Brazil classified with our method. We omit the Bogon, Unassigned and Out-of-cone classes since nothing was detected.

Facility	Switches	Max Traffic Rate	Average Traffic Rate	Average % In-Cone Traffic	Average % Unverifiable Traffic	Time to Execute
CF1	SW1	684Gbps	398Gbps	94.16%	5.84%	4066.28s
	SW2	99Gbps	32Gbps	68.18%	31.82%	2923.12s
CF2	SW1	7Gbps	5Gbps	88.36%	11.64%	865.28s
	SW2	10Gbps	7Gbps	90.2%	9.8%	537.65s
	SW3	43Gbps	28Gbps	73.88%	26.12%	777.54s
	SW4	33Gbps	20Gbps	88.14%	11.86%	1123.05s
CF3	SW1	341Gbps	192Gbps	86.53%	13.46%	3008.04s
	SW2	557Gbps	309Gbps	96.18%	3.81%	2967.50s

traffic flow period matches for all switches. The exceptions are local information regarding the network and the switch under analysis. The cones construction (step 2) is done one time and shared for all distinct executions, always matching the timeframe of traffic data and BGP data. Finally, with the datasets bootstrap finished, the next steps (3 to 5) are straightforward executions of our classification pipeline, followed by data transformations and metrics computation, leveraging the datasets prepared.

7.1.3 Results

Spoofer-IX can handle the analysis of much larger network infrastructures and not only IXP networks. Table 7.1 summarizes the classification results we observed during one day in April 2018 for each switch in the partnered CFs. To perform the traffic classification step we used the same server employed to the analysis of the first IXP (processor: 2x Intel Xeon E5-2640 v4 2.4GHz - 40 threads, RAM memory: 64GB RDIMM, storage: 1TB SSD SATA and 3x disks 1.2TB 10K RPM), which took on average 2400s (40 minutes) to classify one day of traffic flow data. The table shows the set of switches grouped by the Colocation Facility it belongs to, the max and average traffic rate in Gbps, the average percentage of traffic found in each category, and the time (in seconds) to execute the classification using the server mentioned. Through these three facilities we analyzed the traffic sent by 485 members in total. As expected, the majority of traffic was classified as valid (in-cone). Moreover, no traffic was classified as spoofed. Switches CF1-SW2 and CF2-SW3 had a higher average of Unverifiable traffic due increased Provider-to-Customer traffic going through them when compared to other switches. In contrast, CF1-SW1, CF3-SW1 and CF3-SW2 handled well the highest rates, being responsible for delivering the

traffic of big content providers to IXP members.

Through discussions with the IXP coordinators, we hypothesize that the strict set of policies adopted by this second IXP lead to a more secure infrastructure. Among their policies, they have a quarantine network for new members. It is an isolated network that every new member must first connect in order to perform a validation of the security properties and configurations, prior to join the shared switching fabric with all other members. Besides that they also implemented a policy to drop traffic matching bogon prefixes (IX.br, 2018).

7.2 Discussion on Methodology Generality and Limitations

We discuss three key points regarding our methodology. We start assessing the generality of the methodology, followed by limitations, and conclude with IXP emerging trends and impact on the detection of SAV.

Generality of the methodology. Assessing the generality of our approach requires applying our method to traffic collected from a large set of IXPs, which is challenging because it requires the assistance from other IXP operators. The system was designed and developed with generality in mind, by following the Best Current Operational Practices (BCOPs) defined by a group of IXPs (Euro-IX, 2019b; FREEDMAN et al., 2019; Internet Society, 2019) that describe how IXP operators should configure IXPs. Those documents describe how IXP operators should securely configure VLANs and route servers. Therefore, it might be straightforward to apply our methodology to other IXPs; more generally, any other method to infer spoofed traffic in IXP traffic data will have to address the same challenges we encountered.

Applying our methodology requires two data sets: the traffic data sets themselves, and the metadata that maps IXP infrastructure – VLAN tags on each packet, and MAC addresses to ASes. Our methodology is automated except for inference of the siblings (Chapter 6 - §6.2), which requires some manual effort. However, there are a wide variety of IXP architectures that affect traffic visibility (Chapter 4 - §4.2), and our traffic classification method may be impacted by new IXP architecture innovations to support advanced services. Moreover, our use of traffic characterization was limited to the packet headers available to us; full payload would enable improvements in traffic analysis, and additional cross-checks (as pointed in Chapter 6 - §6.4).

Methodology Limitations. Spoofer-IX relies on public BGP observations (RIPE,

2018; ROUTEVIEWS, 2018) to infer AS relationships, prefixes announcements, and the owner ASN in order to build the Prefix-Level Customer Cone. However, we may make mistakes in some of the inferences, as there are multiple possible explanations for the topological arrangements and inter-domain phenomenons observed by AS relationship and Customer Cone inference algorithms. Although we use state-of-the-art inferences techniques, we have identified two situations that no existing Internet cartography algorithm deals with today: dynamics of National Internet Registry (NIR) unassigned ASes, and hijacked prefixes. The former requires validation of the fine-grained level of assigned ASNs, other than only check at IANA level. That is due to regional ASNs, attributed from IANA to NIRs, which have been in turn unallocated locally at the NIR, but continue to be online in BGP announcements at public BGP collectors (stale information) but not anymore in the Route Servers/LG servers at the IXP. With respect to the latter, to avoid hijacked prefixes it is required the validation of every prefix identified in the public BGP observations through the lens of Validation of Route Origination using the Resource Certificate Public Key Infrastructure (RPKI) and Route Origin Authorizations (ROAs), which only recently had a strong push from the IETF community to all networks to implement by default (CHUNG et al., 2019; TESTART et al., 2019). As a result of the lack of prefix validation, some cones may include hijacked prefixes, affecting how the traffic will be classified.

Lastly, neither the Full Cone nor the Customer Cone handle the complexities that sibling ASes (ASes under the same administrative control) bring. In particular, because siblings may provide mutual transit to each other, the set of valid addresses that can transit between each AS is the entire routed address space. However, to observe this behavior in public BGP data, which both the FC and CC use, would require a view from each sibling AS. Current sibling relationship inference methods (CAI et al., 2010; HUFFAKER et al., 2019) use WHOIS data, which is not only inconsistently formatted across regions (AFRINIC, APNIC, ARIN, LACNIC, RIPE NCC), but also becomes stale if not updated as mergers occur, leading to false and missing inferences (HUFFAKER et al., 2019).

IXP emerging trends and impact on the detection of SAV. New IXP services allow networks to self-provision private, on-demand bandwidth in seconds between data center locations (a.k.a, colocation facilities) or cloud service providers, (MARCOS et al., 2018; EPSILON, 2019; MEGAPORT, 2019; Packet Fabric, 2019; CONSOLE, 2019). In 2019, AMS-IX, DE-CIX and LINX joined to develop an API to provision and configure interconnection services at multiple IXPs (BUCKINGHAM, 2019). The resulting IX-

API (AMS-IX; DE-CIX; LINX, 2019) will allow users to manage their interconnection services, from ordering new ports, to configuring, changing, and canceling services at multiple IXPs. These proposals share a common goal: enable a more dynamic interconnection environment, where networks and IXPs can adapt to changing conditions. They do not propose to change methods to implement the configurations tackled in this thesis, but rather create abstractions to facilitate configuration changes.

7.3 Considerations

Applying the Spoofer-IX method and system to both IXPs was a frustrating experience, requiring that we overcame many challenges, including: (1) policy enforcement, e.g. NDA agreements to obtain access to traffic and topology data; (2) evolving processes and architecture within the IXPs, e.g. obtaining up to date topology information; (3) interfacing with running systems and distinct device manufacturers; and (4) handling system failures and data problems. These challenges will characterize any modern interconnection environment, and navigating them is an integral aspect of successfully executing this sort of analysis.

However, we see great potential in enabling execution of our methodology across as broad a set of networks as possible, including IXPs distributed across many colocation facilities and switch fabrics. The modular decomposition of our approach, including bootstrapping and data preparation steps, promotes this generalizability and broad impact. It helps to reduce the time and complexity to bootstrap the deployment of our method. It also benefits the overall shared infrastructure within members by having multiple localized SAV compliance enforcement in the distinct network attachment points. This case study demonstrated that the Spoofer-IX methodology and system implementation can handle the analysis of much larger network infrastructures, even beyond IXPs.

8 FINAL CONSIDERATIONS

In this chapter we first summarize the thesis and present the conclusions (§8.1). Then, we discuss prospective directions for future research (§8.2).

8.1 Concluding Remarks

The use of IXPs as a focal point to help on SAV deployment has received recent attention by both the research (LICHTBLAU et al., 2017) and policy communities (ISOC, 2018; Tech Accord, 2018; NIC.br, 2019). However, inferring SAV deployment at an IXP is remarkably challenging, more so than has been captured in the literature, due to a combination of operational complexities that characterize today's interconnection ecosystem, and the inherently heuristic nature of topology and traffic inferences on persistently opaque network infrastructure. Many of our discoveries were eye-opening, although not cause for optimism for those interested in infrastructure protection.

In this thesis, we discovered several methodological challenges for inferring spoofed packets at IXPs. We recognized the importance of using the semantics of AS relationships, which is conceptually straightforward but even more painstakingly complicated in practice than we expected. We designed, implemented, and applied a new methodology, called Spoofer-IX, that accounts for both epistemological and operational challenges to accurately classify spoofed traffic in the inter-domain level by processing heavily aggregated Internet traffic data. In our research, we also showed how Spoofer-IX reveals inaccuracies in previous methods that are agnostic to AS relationship semantics. In addition, we described the potential pitfalls in using BGP-based filtering to infer the presence of spoofed packets in traffic crossing links at an IXP. The complexity underlying this particular BGP-based inference is subtle, and attempts to minimize false positives can easily come at the expense of significant false negatives. The five main contributions of this thesis and a brief summary follow.

First contribution. A detailed analysis of methodological challenges for inferring spoofed packets at IXPs.

Summary. This contribution forms the basis of our methodology. The extensive studies (MARCOS et al., 2018; MULLER et al., 2019a) performed on shared switching fabrics, such as IXPs and Colocation Facilities, enabled us to identify and analyze the methodological challenges and their implications for applying

BGP-based SAV inference methods to modern IXP connectivity fabrics. The findings from these studies, combined with a comprehensive analysis of previous work that attempted to tackle this inference problem, showed inaccuracies in previous methods.

Second contribution. Designed and developed a methodology to classify traffic flows for the purpose of accurately inferring spoofed traffic.

Summary. We design and implement Spoofer-IX (MULLER et al., 2019a), a novel methodology to detect the transmission of spoofed traffic (which implies lack of source address validation) by AS members of IXPs. Spoofer-IX addresses two fundamental issues overlooked in the existing literature (LICHTBLAU et al., 2017). First, Spoofer-IX considers the type of relationship between neighbors at an IXP when determining which source addresses are valid in IP packets crossing the IXP. Second, Spoofer-IX considers asymmetric routing and traffic engineering, by designing a novel Prefix-Level Customer Cone that includes addresses that may be valid source addresses for an AS to transit. The accuracy of this method depends on the quality of BGP data and AS relationship inferences, which we know to be imperfect (LUCKIE et al., 2013). However, our method is congruent with what network operators do when configuring static access control lists to deploy SAV (Internet Society, 2019; FREEDMAN et al., 2019; Job Snijders, 2016).

Third contribution. The application of our methodology to classify and analyze packets extensively at a medium-sized IXP, considering two periods, two years apart.

Summary. We applied our method to traffic and topology data from one of the largest IXPs in Brazil, with more than 200 member ASes using the IXP switching fabric. We reported insights from the extensive analyses over the traffic classifications conducted and our interactions with IXPs and network operators of their member ASes. We investigated the impact of different filtering choices on inferred valid address space, and the likelihood of false negatives when classifying traffic according to different filtering choices. We also compared our methodology with a recently proposed method (LICHTBLAU et al., 2017) that did not consider AS relationships in its inference of spoofed traffic, reporting that the majority of members at the IXP sent spoofed packets, and demonstrate

the inaccuracies of this approach (MULLER et al., 2019a).

Fourth contribution. Assess the deployment of Spoofer-IX to distinct networks.

Summary. We partnered with a second Brazilian IXP with over one thousand members to assess and explore practical application, generalizability, and scalability of our Spoofer-IX methodology and implementation to larger and more complex IXP infrastructures. We discuss how to scale the analysis by observing traffic per switch and how networks could independently adopt our methodology to detect and filter spoofed traffic.

Fifth contribution. Find evidence that epistemological and cross-validation challenges remain. Describe and publish our code to promote further work.

Summary. We also found epistemological challenges remain. While we inferred Out-of-cone traffic with our methodology at the mid-sized IXP, there are still more complex edge cases, as some of the traffic appears to have signatures of legitimate traffic. We publicly release our code (MULLER et al., 2019b) in hopes that other researchers, IXPs, and other networks will use it to further improve our collective ability to measure and expand deployment of SAV filtering. Finally, the overall work presented in this thesis illustrates the deep subtleties of scientific assessments of operational Internet infrastructure, which exemplifies the persistent tension between the need for reproducibility of methods and results (BAJPAI et al., 2019b; BAJPAI et al., 2019a), and the opacity of commercial infrastructure.

8.2 Future Research Directions

In spite of the progresses reported in the thesis, promising opportunities for future research remain. In the following, we discuss the most prominent ones.

- **Dormant address space analysis.** Spoofer-IX can be refined to handle the analysis of packets exploring what we call the dormant address space, i.e., BGP prefixes allocated to ASes but not announced over the globe as seen by the public BGP collectors from known projects (ROUTEVIEWS, 2018; RIPE, 2018; PCH, 2020).

To begin with, it would be necessary to create a method to infer such address space, as well as define the update frequency that should be considered to this resulting dataset.

- **IPv6 spoofing inference.** In this thesis, we focus on IPv4 traffic exclusively, as native IPv6 traffic still ranges below 5% at the vantage points of interest (IX.br, 2020; AMS-IX, 2020a; DE-CIX, 2020). However, with the increasing adoption of IPv6 (IANA, 2018c; NRO, 2019), we can expect more IPv6 traffic, as well as spoofing attacks. The study of IPv6 traffic requires an in-depth study, in the same way as we did for IPv4 traffic in this thesis. Additional challenges include how to deal with the changing nature of IPv6 ASes topology (GIOTSAS et al., 2015a) in order to best capture inferences, starting on AS Relationships and Customer Cones.
- **Refine methodology to academic networks / PoP level.** Our methodology can also be deployed at the Points of Presence (PoPs) of National Research and Education Networks (NRENs) (e.g., RNP, GEANT, Internet2) however, it needs to be refined to handle the related operational complexities. For example, many of these networks make use of virtual circuits, provided through systems such as OSCARS (ESnet, 2020) or Internet2 Advanced Layer 2 Services (AL2S) (Internet2, 2020), when linking campuses or research facilities together, and sometimes they may use reserved prefixes in the traffic exchanges. Virtual circuits and other research related services properties should be considered during the traffic classification processing pipeline.
- **Develop a real-time version of Spoofer-IX to perform active traffic analyses.** We developed Spoofer-IX to perform on-demand analyses of passive traffic flow data collected by network administrators. However, it will also be valuable for networks to have our methodology always running, evaluating, and giving feedback in real-time of their traffic. It would be necessary to cope with many challenges, such as how to deal with continually evolving network conditions, as well as obtaining fresh datasets for accurate results.

All the previous lines of work will expand the relevance of our work in the measurement research community, as well as the technical community, and further our insights into the spoofing problem in new scenarios.

REFERENCES

AGER, B. et al. Anatomy of a Large European IXP. In: **SIGCOMM**. New York, NY, USA: ACM, 2012. p. 163–174.

AMS-IX. **AMS-IX – Amsterdam Internet Exchange**. 2018. Available from Internet: <https://ams-ix.net/>.

AMS-IX. **AMS-IX Partner Program**. 2019. Available from Internet: <https://www.ams-ix.net/ams/partners>.

AMS-IX. **AMS-IX Private Interconnect Service**. 2019. Available from Internet: <https://www.ams-ix.net/ams/service/private-interconnect>.

AMS-IX. **AMS-IX sFlow statistics**. 2020. Available from Internet: <https://stats.ams-ix.net/sflow/index.html>.

AMS-IX. **Amsterdam Internet Exchange (AMS-IX)**. 2020. Available from Internet: <https://www.ams-ix.net/>.

AMS-IX; DE-CIX; LINX. **IX-API Simplify your IX services**. 2019. Available from Internet: <https://ix-api.net/>.

ANWAR, R. et al. Investigating interdomain routing policies in the wild. In: **Proceedings of the 2015 Internet Measurement Conference**. New York, NY, USA: [s.n.], 2015. (IMC '15), p. 71–77.

APACHE. **Apache Avro**. 2019. Available from Internet: <https://avro.apache.org>.

APNIC. **Weekly Routing Table Report**. 2020. Available from Internet: <http://thyme.apnic.net/current/data-summary>.

BAJPAI, V. et al. Encouraging Reproducibility in Scientific Research of the Internet. **Dagstuhl Reports**, v. 8, n. 10, p. 41–62, Jan 2019.

BAJPAI, V. et al. The dagstuhl beginners guide to reproducibility for experimental networking research. **SIGCOMM Comput. Commun. Rev.**, ACM, New York, NY, USA, v. 49, n. 1, p. 24–30, feb. 2019.

BAKER, F.; SAVOLA, P. **Ingress Filtering for Multihomed Networks**. 2004. RFC 3704 (BCP 84).

BATTISTA, G. D.; PATRIGNANI, M.; PIZZONIA, M. Computing the types of the relationships between autonomous systems. In: **IEEE INFOCOM**. Piscataway, NJ, USA: IEEE Press, 2003. v. 1, p. 156–165 vol.1.

BELLOVIN, S. M. Security Problems in the TCP/IP Protocol Suite. **ACM SIGCOMM CCR**, ACM, v. 19, n. 2, abr. 1989.

BELLOVIN, S. M. A look back at "security problems in the tcp/ip protocol suite". In: **Proceedings of the 20th Annual Computer Security Applications Conference**. USA: IEEE Computer Society, 2004. (ACSAC '04), p. 229–249. ISBN 0769522521. Available from Internet: <https://doi.org/10.1109/CSAC.2004.3>.

BENSON, K. et al. Leveraging internet background radiation for opportunistic network analysis. In: **Proceedings of the 2015 Internet Measurement Conference**. [S.l.]: ACM, 2015. (IMC '15).

BEVERLY, R.; BAUER, S. The Spoofer Project: Inferring the extent of source address filtering on the Internet. In: **SRUTI**. Berkeley, CA, USA: USENIX, 2005.

BEVERLY, R. et al. Understanding the Efficacy of Deployed Internet Source Address Validation Filtering. In: **IMC**. New York, NY, USA: ACM, 2009.

Bnamericas. **América Móvil Brasil folds NET into Claro**. 2019. Available from Internet: <https://www.bnamericas.com/en/news/america-movil-brasil-folds-net-into-claro>.

BRITO, S. H. B. et al. An analysis of the largest national ecosystem of public internet exchange points: The case of brazil. **Journal of Communication and Information Systems**, v. 31, n. 1, Oct. 2016. Available from Internet: <https://jcis.sbrt.org.br/jcis/article/view/371>.

BUCKINGHAM, L. **IX-API For the Good of the Internet**. 2019. Available from Internet: <https://www.linx.net/ix-api-for-the-good-of-the-internet/>.

BYERS, K. **Netmiko**. 2019. Available from Internet: <http://ktbyers.github.io/netmiko/>.

CAI, X. et al. Towards an as-to-organization map. In: **Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement**. New York, NY, USA: ACM, 2010. (IMC '10), p. 199–205.

CAIDA. **The CAIDA Internet Topology Data Kit (ITDK-2017-08)**. 2017. Available from Internet: <http://www.caida.org/data/internet-topology-data-kit>.

CAIDA. **CAIDA ASRank**. 2018. Available from Internet: <http://as-rank.caida.org>.

CAIDA. **CAIDA Inferred AS to Organization Mapping Dataset**. 2018. Available from Internet: <https://www.caida.org/data/as-organizations/>.

CAIDA. **CAIDA Spoofer Project**. 2018. Available from Internet: <https://www.caida.org/projects/spoofer/>.

CAIDA. **CAIDA Spoofer Project Public API**. 2019. Available from Internet: <http://tyr.caida.org:9000>.

CAIDA. **The CAIDA Internet Topology Data Kit (ITDK-2019-04)**. 2019. Available from Internet: <http://www.caida.org/data/internet-topology-data-kit>.

Carisimo, E. et al. Studying the evolution of content providers in the internet core. In: **2018 Network Traffic Measurement and Analysis Conference (TMA)**. [S.l.: s.n.], 2018.

CASTRO, I. et al. Remote peering: More peering without internet flattening. In: **Proceedings of the 10th ACM International on Conference on Emerging Networking Experiments and Technologies**. [S.l.]: ACM, 2014.

CHATZIS, N. et al. On the Benefits of Using a Large IXP As an Internet Vantage Point. In: **Proceedings of the 2013 Conference on Internet Measurement Conference**. New York, NY, USA: ACM, 2013.

CHATZIS, N.; SMARAGDAKIS, G.; FELDMANN, A. **On the Importance of Internet eXchange Points for Today's Internet Ecosystem**. 2013. Available from Internet: <https://arxiv.org/abs/1307.5264>.

CHUNG, T. et al. Rpki is coming of age: A longitudinal study of rpki deployment and invalid route origins. In: **Proceedings of the Internet Measurement Conference**. New York, NY, USA: ACM, 2019. (IMC '19), p. 406–419.

CISCO. **IPv6 Extension Headers Review and Considerations**. 2006. Available from Internet: <https://www.cisco.com/en/US/technologies/tk648/tk872/technologies_white_paper0900aecd8054d37d.pdf>.

CLAISE, B. **Cisco Systems NetFlow Services Export Version 9**. 2004. RFC 3954 (Informational).

COHEN, R.; RAZ, D. The Internet Dark Matter - on the Missing Links in the AS Connectivity Map. In: **IEEE INFOCOM 2006**. Piscataway, NJ, USA: IEEE Press, 2006. p. 1–12.

COMARELA, G.; GüRSUN, G.; CROVELLA, M. Studying Interdomain Routing over Long Timescales. In: **IMC**. New York, NY, USA: ACM, 2013.

CONSOLE. **Console - The Cloud Connection Company**. 2019. Available from Internet: <https://www.consoleconnect.com/>.

COTTON, M. et al. **Special-Purpose IP Address Registries**. 2013. RFC 6890 (BCP 153). Updated by RFC 8190.

DAINOTTI, A. et al. Estimating internet address space usage through passive measurements. **ACM SIGCOMM CCR**, ACM, v. 44, n. 1, dec. 2013.

DAINOTTI, A. et al. Lost in Space: Improving inference of IPv4 address space utilization. **IEEE JSAC**, v. 34, n. 6, June 2016.

David Shepardson. **U.S. regulator approves CenturyLink's $24 billion Level 3 acquisition**. 2017. Available from Internet: <https://www.reuters.com/article/us-level-3-communi-m-a-centurylink/u-s-regulator-approves-centurylinks-24-billion-level-3-acquisition-idUSKBN1CZ25V>.

DDOS-BLACK. **Order DDoS attack on the website - Service DDoS**. 2019. Available from Internet: <https://ddos-black.info>.

DDOS-STRESS. **DDoS Service**. 2019. Available from Internet: <https://ddos-stress.cc/>.

DE-CIX. **DECIX – Internet Exchange in Frankfurt**. 2018. Available from Internet: <https://www.de-cix.net/en/about-de-cix>.

DE-CIX. **DEC-IX MetroVLAN**. 2019. Available from Internet: <https://www.de-cix.net/en/de-cix-service-world/metrovlan>.

DE-CIX. **DE-CIX Internet Exchange**. 2020. Available from Internet: <https://www.de-cix.net/en/>.

DHAMDHERE, A.; DOVROLIS, C. The internet is flat: Modeling the transition from a transit hierarchy to a peering mesh. In: **Proceedings of the 6th International COnference**. New York, NY, USA: Association for Computing Machinery, 2010. (Co-NEXT '10). ISBN 9781450304481. Available from Internet: <https://doi.org/10.1145/1921168.1921196>.

DHAMDHERE, A.; DOVROLIS, C. Twelve years in the evolution of the internet ecosystem. **IEEE/ACM Trans. Netw.**, IEEE, Piscataway, NJ, USA, 2011.

DHAMDHERE, A.; DOVROLIS, C. Twelve years in the evolution of the internet ecosystem. **IEEE/ACM Transactions on Networking**, v. 19, n. 5, p. 1420–1433, Oct 2011. ISSN 1558-2566.

DIMITROPOULOS, X. et al. As relationships: Inference and validation. **SIGCOMM Comput. Commun. Rev.**, ACM, New York, NY, USA, v. 37, n. 1, p. 29–40, jan. 2007. ISSN 0146-4833.

DUAN, Z.; YUAN, X.; CHANDRASHEKAR, J. Constructing Inter-Domain Packet Filters to Control IP Spoofing Based on BGP Updates. In: **INFOCOM**. Piscataway, NJ, USA: IEEE Press, 2006.

EPSILON. **Epsilon Telecommunications Limited – Connectivity made simple**. 2019. Available from Internet: <www.epsilontel.com/>.

ESnet. **OSCARS - On-Demand Secure Circuits and Advance Reservation System**. 2020. Available from Internet: <http://es.net/engineering-services/oscars/>.

Euro-IX. **IXP BCOPs (Best Current Operational Practices)**. 2019. Available from Internet: <https://www.euro-ix.net/en/forixps/set-ixp/ixp-bcops/>.

Euro-IX. **IXP BCOPs (Best Current Operational Practices), Technical Recommendations**. 2019. Available from Internet: <https://www.euro-ix.net/en/forixps/set-ixp/ixp-bcops/technical-recommendations/>.

FAYAZ, S. K. et al. Efficient network reachability analysis using a succinct control plane representation. In: **Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation**. USA: USENIX Association, 2016. (OSDI'16), p. 217–232. ISBN 9781931971331.

FERGUSON, P.; SENIE, D. **Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing**. 2000. RFC 2827 (BCP 38). Updated by RFC 3704.

FLACH, T. et al. An internet-wide analysis of traffic policing. In: **Proceedings of the 2016 ACM SIGCOMM Conference**. New York, NY, USA: Association for Computing Machinery, 2016. (SIGCOMM '16), p. 468–482. ISBN 9781450341936. Available from Internet: <https://doi.org/10.1145/2934872.2934873>.

FREEDMAN, D. et al. **Mutually Agreed Norms for Routing Security (MANRS) Implementation Guide**. 2019. Available from Internet: <https://www.ripe.net/publications/docs/ripe-706>.

GAO, L. On inferring autonomous system relationships in the internet. **IEEE/ACM Trans. Netw.**, IEEE Press, v. 9, n. 6, p. 733–745, dec. 2001. ISSN 1063-6692.

GIOTSAS, V.; DHAMDHERE, A.; CLAFFY, k. Periscope: Unifying Looking Glass Querying. In: **Passive and Active Network Measurement Workshop (PAM)**. New York, NY, USA: Elsevier North-Holland, Inc., 2016.

GIOTSAS, V. et al. IPv6 AS Relationships, Clique, and Congruence. In: **PAM**. New York, NY, USA: Springer-Verlag, 2015.

GIOTSAS, V. et al. Mapping peering interconnections to a facility. In: **Proceedings of the 11th ACM Conference on Emerging Networking Experiments and Technologies**. New York, NY, USA: Association for Computing Machinery, 2015. (CoNEXT '15). ISBN 9781450334129. Available from Internet: <https://doi.org/10.1145/2716281.2836122>.

GOOGLE. **TextFSM**. 2019. Available from Internet: <https://github.com/google/textfsm>.

HAAG, P. **nfdump**. 2019. Available from Internet: <https://github.com/phaag/nfdump>.

HARDIN, G. The tragedy of the commons. **Science**, American Association for the Advancement of Science, 1968.

HAWKINSON, J.; BATES, T. **Guidelines for creation, selection, and registration of an Autonomous System (AS)**. 1996. RFC 1930 (BCP 6). Updated by RFC 6996, 7300.

HEIDEMANN, J. et al. **ANT Censuses of the Internet Address Space**. 2019. Available from Internet: <https://ant.isi.edu/address/index.html>.

Huffaker, B. **ASRank**. 2018. Slideset presented at Internet Initiative Japan (IIJ). Available from Internet: <http://www.caida.org/publications/presentations/2018/asrank_iij/asrank_iij.pdf>.

HUFFAKER, B. et al. **CAIDA inferred AS to organization mapping dataset.** 2019. Available from Internet: <https://www.caida.org/data/as-organizations/>.

IANA. **Autonomous System (AS) Numbers**. 2018. Available from Internet: <https://www.iana.org/assignments/as-numbers/as-numbers.xml>.

IANA. **IANA IPv4 Address Space Registry**. 2018. Available from Internet: <https://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xhtml>.

IANA. **Internet Protocol Version 6 Address Space**. 2018. Available from Internet: <https://www.iana.org/assignments/ipv6-address-space/ipv6-address-space.xhtml>.

IBGE. **IBGE lança mapa de Densidade Demográfica de 2010**. 2013. Available from Internet: <http://www.brasil.gov.br/governo/2013/10/ibge-lanca-mapa-de-densidade-demografica-de-2010/mapa-ibge.gif/view>.

Internet Society. **MANRS IXP Programme**. 2019. Available from Internet: <https://www.manrs.org/ixps/>.

Internet2. **AL2S - Internet2 Advanced Layer 2 Service**. 2020. Available from Internet: <https://www.internet2.edu/products-services/advanced-networking/layer-2-services/>.

ISOC. **MANRS IXP Programme**. 2018. Available from Internet: <https://www.manrs.org/participants/ixp/>.

ISOC. **MANRS Observatory**. 2019. Available from Internet: <https://observatory.manrs.org/>.

ISOC. **Anti-Spoofing MANRS Implementation Guide**. 2020. Available from Internet: <https://www.manrs.org/isps/guide/antispoofing/>.

ISOC. **Mutually Agreed Norms for Routing Security (MANRS)**. 2020. Available from Internet: <https://www.manrs.org/>.

ITU. **ITU Interactive Transmission Map**. 2019. Available from Internet: <http://www.itu.int/itu-d/tnd-map-public>.

IX Reach. **IX Reach Remove Peering Services**. 2019. Available from Internet: <https://www.ixreach.com/services/remote-peering/>.

IX.br. **Programa por uma Internet mais Segura – Ações no IX.br (PT-BR)**. 2018. Available from Internet: <http://ix.br/doc/acoes-seguranca-ix-br-20180927.pdf>.

IX.br. **CIX - Solução de entroncamento para participantes**. 2019. Available from Internet: <http://old.ix.br/doc/Solucao_de_entroncamento_para_participantes.pdf>.

IX.br. **IX.br – Internet Exchange Brazil**. 2020. Available from Internet: <http://ix.br>.

IX.br Forum 12. **Remote Peering Panel with DEC-IX, AMS-IX, LINX and IX.br**. 2019. Available from Internet: <https://www.youtube.com/watch?v=K283b3AKZ94>.

JASINSKA, E. et al. **Internet Exchange BGP Route Server**. 2016. RFC 7947.

JEFFREE T. **IEEE 802.1ad official page**. 2019. Available from Internet: <http://www.ieee802.org/1/pages/802.1ad.html>.

Job Snijders. **Practical everyday BGP filtering: Peer Locking (NANOG67)**. 2016. Available from Internet: <https://www.youtube.com/watch?v=CSLpWBrHy10>.

JONKER, M. et al. Millions of targets under attack: A macroscopic characterization of the dos ecosystem. In: **IMC**. New York, NY, USA: ACM, 2017. (IMC '17), p. 100–113.

Julio Wiziack. **Agravamento da crise da Oi chega ao Palácio do Planalto**. 2019. Available from Internet: <https://www1.folha.uol.com.br/mercado/2019/08/agravamento-da-crise-da-oi-chega-ao-palacio-do-planalto.shtml>.

KATZ-BASSETT, E. et al. Studying black holes in the internet with hubble. In: **Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation**. USA: USENIX Association, 2008. (NSDI'08), p. 247–262. ISBN 1119995555221.

KLABA, O. **1.3Tbps DDoS mitigated by our VAC**. 2018. Available from Internet: <https://twitter.com/olesovhcom/status/969328679410110466>.

KLöTI, R. et al. A comparative look into public ixp datasets. **SIGCOMM Comput. Commun. Rev.**, ACM, 2016.

KOTTLER, S. **February 28th DDoS Incident Report**. 2018. Available from Internet: <https://githubengineering.com/ddos-incident-report/>.

KüHRER, M. et al. Going Wild: Large-Scale Classification of Open DNS Resolvers. In: **Proceedings of the 2015 Internet Measurement Conference**. New York, NY, USA: ACM, 2015. (IMC '15).

KüHRER, M. et al. Exit from Hell? Reducing the Impact of Amplification DDoS Attacks. In: **Proceedings of the 23rd USENIX Conference on Security Symposium**. Berkeley, CA, USA: USENIX Association, 2014.

LABOVITZ, C. et al. Internet inter-domain traffic. In: **Proceedings of the ACM SIGCOMM 2010 Conference**. New York, NY, USA: ACM, 2010. (SIGCOMM '10), p. 75–86.

LICHTBLAU, F. et al. Detection, Classification, and Analysis of Inter-domain Traffic with Spoofed Source IP Addresses. In: **IMC**. New York, NY, USA: ACM, 2017.

LINX. **ConneXions at London Internet Exchange Point**. 2019. Available from Internet: <https://www.linx.net/join-linx/connexions/>.

LINX. **LINX Private VLAN**. 2019. Available from Internet: <https://www.linx.net/products-services/private-vlan/>.

LINX. **London Internet Exchange (LINX)**. 2020. Available from Internet: <https://www.linx.net/>.

Lisa Vaas. **FBI crackdown on DDoS-for-hire sites led to 85attack sizes**. 2019. Available from Internet: <https://nakedsecurity.sophos.com/2019/03/21/fbi-crackdown-on-ddos-for-hire-sites-led-to-85-slash-in-attack-sizes/>.

LIU, B.; BI, J.; VASILAKOS, A. V. Toward incentivizing anti-spoofing deployment. **IEEE ToIFS**, v. 9, n. 3, March 2014.

LIU, X. et al. Passport: Secure and Adoptable Source Authentication. In: **NSDI**. Berkeley, CA, USA: USENIX, 2008.

LODHI, A. et al. Complexities in internet peering: Understanding the "black"in the "black art". In: **2015 IEEE Conference on Computer Communications (INFOCOM)**. [S.l.: s.n.], 2015. p. 1778–1786. ISSN 0743-166X.

LODHI, A. et al. Using peeringdb to understand the peering ecosystem. **SIGCOMM Comput. Commun. Rev.**, ACM, 2014.

LONE, Q. et al. Using Loops Observed in Traceroute to Infer the Ability to Spoof. In: **PAM**. New York, NY, USA: Elsevier North-Holland, Inc., 2017.

Lone, Q. et al. Using crowdsourcing marketplaces for network measurements: The case of spoofer. In: **2018 Network Traffic Measurement and Analysis Conference (TMA)**. [S.l.: s.n.], 2018.

LUCKIE, M. Spurious routes in public bgp data. **SIGCOMM Comput. Commun. Rev.**, ACM, New York, NY, USA, v. 44, n. 3, p. 14–21, jul. 2014. ISSN 0146-4833. Available from Internet: <http://doi.acm.org/10.1145/2656877.2656880>.

LUCKIE, M. et al. Network hygiene, incentives, and regulation: Deployment of source address validation in the internet. In: **Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security**. New York, NY, USA: Association for Computing Machinery, 2019. (CCS '19), p. 465–480. ISBN 9781450367479. Available from Internet: <https://doi.org/10.1145/3319535.3354232>.

LUCKIE, M. et al. Challenges in inferring internet interdomain congestion. In: **Proceedings of the 2014 Conference on Internet Measurement Conference**. New York, NY, USA: [s.n.], 2014. (IMC '14), p. 15–22.

LUCKIE, M. et al. AS Relationships, Customer Cones, and Validation. In: **IMC**. New York, NY, USA: ACM, 2013.

MA, R. T. et al. Interconnecting eyeballs to content: A shapley value perspective on isp peering and settlement. In: **Proceedings of the 3rd International Workshop on Economics of Networked Systems**. New York, NY, USA: Association for Computing Machinery, 2008. (NetEcon '08), p. 61–66. ISBN 9781605581798. Available from Internet: <https://doi.org/10.1145/1403027.1403041>.

MAJKOWSKI, M. **The real cause of large DDoS - IP Spoofing**. 2018. Available from Internet: <https://blog.cloudflare.com/the-root-cause-of-large-ddos-ip-spoofing/>.

MAJKOWSKI, M. **The rise of multivector DDoS attacks**. 2018. Available from Internet: <https://blog.cloudflare.com/the-rise-of-multivector-amplifications/>.

MARCOS, P. et al. Dynam-IX: A Dynamic Interconnection eXchange. In: **Proceedings of the 14th International Conference on Emerging Networking EXperiments and Technologies**. New York, NY, USA: ACM, 2018. (CoNEXT '18), p. 228–240. ISBN 978-1-4503-6080-7.

MARDER, A. et al. Pushing the boundaries with bdrmapIT: Mapping router ownership at internet scale. In: **Proceedings of the Internet Measurement Conference 2018**. New York, NY, USA: Association for Computing Machinery, 2018. (IMC '18), p. 56–69. ISBN 9781450356190. Available from Internet: <https://doi.org/10.1145/3278532.3278538>.

MAUCH, J.; SNIJDERS, J.; HANKINS, G. **Default External BGP (EBGP) Route Propagation Behavior without Policies**. 2017. RFC 8212. Updates 4271.

MAXMIND. **GeoLite2 Free Downloadable Databases**. 2019. Available from Internet: <https://dev.maxmind.com/geoip/geoip2/geolite2/>.

MCCONACHIE, A. **Anti-Spoofing, BCP 38, and the Tragedy of the Commons**. 2014. Available from Internet: <https://www.internetsociety.org/blog/2014/07/anti-spoofing-bcp-38-and-the-tragedy-of-the-commons/>.

MEGAPORT. **Megaport - A Better way to connect**. 2019. Available from Internet: <https://www.megaport.com>.

MORALES, C. **NETSCOUT Arbor Confirms 1.7 Tbps DDoS Attack; The Terabit Attack Era Is Upon Us**. 2018. Available from Internet: <https://www.arbornetworks.com/blog/asert/netscout-arbor-confirms-1-7-tbps-ddos-attack-terabit-attack-era-upon-us/>.

MORRIS, R. A Weakness in the 4.2BSD Unix TCP/IP Software Technical Report 117, AT&T Bell Laboratories. 1985.

MOSKOWITZ, R. et al. **Address Allocation for Private Internets**. 1996. RFC 1918 (BCP 5).

MOTAMEDI, R. et al. On mapping the interconnections in today's internet. **IEEE/ACM Trans. Netw.**, IEEE Press, v. 27, n. 5, p. 2056–2070, oct. 2019. ISSN 1063-6692. Available from Internet: <https://doi.org/10.1109/TNET.2019.2940369>.

MULLER, L. et al. Challenges in Inferring Spoofed Traffic at IXPs. In: **Proceedings of the 15th International Conference on Emerging Networking Experiments And Technologies**. New York, NY, USA: Association for Computing Machinery, 2019. (CoNEXT '19), p. 96–109. ISBN 9781450369985. Available from Internet: <https://doi.org/10.1145/3359989.3365422>.

MULLER, L. et al. **Spoofer-IX sourcecode**. 2019. <https://github.com/spoofer-ix/spoofer-ix>.

MULLER, L. et al. **Routing Coverage Analysis At IXPs in Brazil**. 2016. In: Semana de Infraestrutura da Internet no Brasil (IX-Forum 10). Available from Internet: <https://www.youtube.com/watch?v=93ckgAFRMa8>.

NETACUITY. **NetAcuity Trusted Geolocation Data**. 2019. Available from Internet: <https://www.digitalelement.com/solutions/netacuity-edge-premium/>.

NETSCOUT. **Insight into the Global Threat Landscape - NETSCOUT Arbor's Annual Worldwide Infrastructure Security Report**. 2019. Available from Internet: <https://www.netscout.com/report/>.

NIC.br. **Programa por uma Internet mais segura**. 2019. Available from Internet: <https://bcp.nic.br/i+seg/>.

NLNETLABS. **RPKI Tools**. 2018. Available from Internet: <https://www.nlnetlabs.nl/projects/rpki/funding/>.

NOCTION. **Noction Network Intelligence**. 2019. Available from Internet: <https://www.noction.com/>.

NOMIKOS, G. et al. O peer, where art thou?: Uncovering remote peering interconnections at ixps. In: **Proceedings of the Internet Measurement Conference 2018**. New York, NY, USA: ACM, 2018. (IMC '18), p. 265–278.

NRO. **Internet Number Resource Status Report – Q3 September 2019**. 2019. Available from Internet: <https://www.nro.net/wp-content/uploads/NRO-Statistics-2019-Q3.pdf>.

NRO. **The Number Resource Organization Extended Allocation and Assignment Reports**. 2020. Available from Internet: <https://www.nro.net/about/rirs/statistics/>.

OLIVEIRA, R. et al. The (in)Completeness of the Observed Internet AS-level Structure. **IEEE/ACM Trans. Netw.**, IEEE Press, Piscataway, NJ, USA, v. 18, n. 1, p. 109–122, feb. 2010.

ORSINI, C. et al. Bgpstream: A software framework for live and historical bgp data analysis. In: **Proceedings of the 2016 Internet Measurement Conference**. New York, NY, USA: ACM, 2016. p. 429–444.

P. Phaal, S. Panchen, and N. McKee. **InMon Corporation's sFlow: A Method for Monitoring Traffic in Switched and Routed Networks**. 2001. RFC 3176 (Informational).

Packet Fabric. **Packet Fabric**. 2019. Available from Internet: <https://www.packetfabric.com/>.

PCH. **Packet Clearing House (PCH) Raw Routing Data**. 2020. Available from Internet: <https://www.pch.net/resources/Raw_Routing_Data/>.

PeeringDB. **PeeringDB**. 2019. Available from Internet: <https://www.peeringdb.com>.

POSTEL, J. **Internet Protocol**. 1981. RFC 791.

Registro.br. **Registro.br**. 2020. Available from Internet: <https://registro.br/>.

REKHTER, Y.; LI, T. **A Border Gateway Protocol 4 (BGP-4)**. 1994. RFC 1654.

RICHTER, P. et al. Peering at Peerings: On the Role of IXP Route Servers. In: **Proceedings of the 2014 Conference on Internet Measurement Conference**. New York, NY, USA: ACM, 2014.

RICHTER, P. et al. Beyond counting: New perspectives on the active ipv4 address space. In: **Proceedings of the 2016 Internet Measurement Conference**. [S.l.]: ACM, 2016. (IMC '16), p. 135–149.

RIPE. **RIPE RIS**. 2018. Available from Internet: <http://www.ripe.net/ris/>.

RIPE NCC. **BGPdump**. 2019. Available from Internet: <https://bitbucket.org/ripencc/bgpdump-hg/wiki/Home>.

ROSSOW, C. Amplification Hell: Revisiting Network Protocols for DDoS Abuse. In: **Network and Distributed System Security (NDSS**. San Diego, CA, USA: Internet Society, 2014.

ROUGHAN, M. et al. Class-of-service mapping for qos: A statistical signature-based approach to ip traffic classification. In: **Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement**. [S.l.]: ACM, 2004. (IMC '04).

ROUGHAN, M. et al. 10 lessons from 10 years of measuring and modeling the Internet's Autonomous Systems. **IEEE JSAC**, v. 29, n. 9, October 2011.

ROUTEVIEWS. **U. Oregon Route Views Project**. 2018. Available from Internet: <http://www.routeviews.org/>.

RYBA, F. J. et al. **Amplification and DRDoS Attack Defense – A Survey and New Perspectives**. Open Archive: arXiv.org, 2015. Available at <http://arxiv.org/abs/1505.07892>.

SCHEID, T. **Defending the Olympics from DDoS**. 2016. Available from Internet: <https://blog.apnic.net/2016/10/17/defending-olympics-ddos/>.

SCRAPINGHUB. **Scrapy**. 2019. Available from Internet: <https://scrapy.org>.

SECURITY-PROF. **Professional DDoS Service**. 2019. Available from Internet: <http://security-prof.info>.

SHANI, T. **Botnet-led DDoS Attacks Are Hitting Record Intensities**. 2019. Available from Internet: <https://www.imperva.com/blog/botnet-led-ddos-attacks-are-hitting-record-intensities-imperva-is-mitigating-all-of-them/>.

Sirota, J. **Report Anual (2018) Equipe de Engenharia – IX.br**. 2018. Available from Internet: <http://forum.ix.br/files/apresentacao/arquivo/395/2.1%20-%20IX.br_atualiza%C3%A7%C3%A3o_2018.pdf>.

Sirota, J. **Report Anual (2019) Equipe de Engenharia – IX.br**. 2019. Available from Internet: <https://forum.ix.br/files/apresentacao/arquivo/747/09%2020%20%20Julio.pdf>.

Steve Evans. **Telefonica expands Brazil presence with 7.2bn Euros GVT acquisition**. 2014. Available from Internet: <https://www.zdnet.com/article/telefonica-expands-brazil-presence-with-eur7-2bn-gvt-acquisition/>.

SUBRAMANIAN, L. et al. Characterizing the Internet hierarchy from multiple vantage points. In: **INFOCOM**. Piscataway, NJ, USA: IEEE Press, 2002. v. 2, p. 618–627 vol.2.

Team CYMRU. **IPv4 Fullbogons**. 2018. Available from Internet: <https://www.team-cymru.org/Services/Bogons/fullbogons-ipv4.txt>.

Team CYMRU. **The Bogon Reference**. 2018. Available from Internet: <http://www.team-cymru.com/bogon-reference.html>.

Tech Accord. **Cybersecurity Tech Accord**. 2018. Available from Internet: <https://cybertechaccord.org/>.

TELEGEOGRAPHY. **Submarine Cable Map**. 2019. Available from Internet: <http://www.submarinecablemap.com>.

TESTART, C. et al. Profiling bgp serial hijackers: Capturing persistent misbehavior in the global routing table. In: **Proceedings of the Internet Measurement Conference**. New York, NY, USA: Association for Computing Machinery, 2019. (IMC '19), p. 420–434. ISBN 9781450369480. Available from Internet: <https://doi.org/10.1145/3355369.3355581>.

Thomas Brewster. **Cops Take Down World's Biggest 'DDoS-For-Hire' Site They Claim Launched 6 Million Attacks**. 2018. Available from Internet: <https://www.forbes.com/sites/thomasbrewster/2018/04/25/massive-ddos-attack-service-webstresser-org-taken-down/>.

TIMBERG, C. **A Flaw In The Design**. 2015. Available from Internet: <https://www.washingtonpost.com/sf/business/2015/05/30/net-of-insecurity-part-1>.

US-CERT. **Protecting Against Malicious Code - Security Tip (ST18-004)**. 2018. Available from Internet: <https://www.us-cert.gov/ncas/tips/ST18-271>.

US-CERT. **UDP-Based Amplification Attacks**. 2019. Available from Internet: <https://www.us-cert.gov/ncas/alerts/TA14-017A>.

VERISIGN. **Q2 2018 DDoS Trends Report: 52 Percent of Attacks Employed Multiple Attack Types**. 2018. Available from Internet: <http://www.circleid.com/posts/20180927_q2_2018_ddos_trends_report_52_percent_of_attacks_multiple_types/>.

WEIL, J. et al. **IANA-Reserved IPv4 Prefix for Shared Address Space**. 2013. RFC 6598 (BCP 153).

WESSELS, D.; CLAFFY kc. **Mapping the IPv4 Address Space**. 2019. Available from Internet: <https://www.caida.org/research/id-consumption/census-map/>.

WOODCOCK, B.; FRIGINO, M. 2016 survey of internet carrier interconnection agreements. **Packet Clearing House, November**, 2016.

XIA, J.; GAO, L. On the evaluation of as relationship inferences. In: **GLOBECOM '04.** Piscataway, NJ, USA: IEEE Press, 2004. v. 3, p. 1373–1377 Vol.3.

YAAR, A.; PERRIG, A.; SONG, D. StackPi: New packet marking and filtering mechanisms for DDoS and IP spoofing defense. **IEEE JSAC**, IEEE, v. 24, n. 10, Oct 2006.

ZARGAR, S. T.; JOSHI, J.; TIPPER, D. A survey of defense mechanisms against distributed denial of service (ddos) flooding attacks. **IEEE Communications Surveys Tutorials**, IEEE Press, Piscataway, NJ, USA, v. 15, n. 4, p. 2046–2069, 2013.

ZHANG, B. et al. Collecting the Internet AS-level topology. **ACM SIGCOMM CCR**, ACM, New York, NY, USA, v. 35, n. 1, jan. 2005.

ZHU, L. et al. Connection-Oriented DNS to Improve Privacy and Security. In: **Proceedings of the 2015 IEEE Symposium on Security and Privacy**. Washington, DC, USA: IEEE Computer Society, 2015.

A ACHIEVEMENTS

A.1 Peer-reviewed publications

Directly related to thesis

The development of this thesis has led to the publication of the following peer-reviewed/journal papers:

- **Journal**: Elsevier Computer Networks (COMNET).
  - **Title**: Spoofed Traffic Inference at IXPs: Challenges, Methods and Analysis
  - **Authors**: **MÜLLER L.**, LUCKIE M., HUFFAKER B., CLAFFY K., BARCELLOS M.
  - **Qualis**: A1
  - **ISSN**: 1389-1286
  - **Status**: Submitted
- **Conference**: 15th International Conference on emerging Networking EXperiments and Technologies (ACM CoNEXT 2019)
  - **Title**: Challenges in Inferring Spoofed Traffic at IXPs
  - **Type**: Main track (full-paper)
  - **Authors**: **MÜLLER L.**, LUCKIE M., HUFFAKER B., CLAFFY K., BARCELLOS M.
  - **Qualis**: A1
  - **Date**: December 9-12, 2019
  - **Location**: Orlando, Florida, U.S.
  - **Digital Object Identifier (DOI)**: <https://doi.org/10.1145/3359989.3365422>

Collaboration projects

In addition to the aforementioned main outcomes of this thesis, we further authored/coauthored some others studies on correlated large network measurement studies, Internet eXchange Points opportunities, and routing problems. These publications are listed next.

- **Conference**: Passive and Active Measurement: 20th International Conference (PAM-2019).

  - **Authors**: MAZOLLA F., **MÜLLER L.**, OLIVEIRA R. and BARCELLOS M.
  - **Title**: A Decade of Backbone Evolution of the Brazilian Academic Network: observations from the perspective of the routers
  - **Type**: PAM 2019 Ph.D School
  - **Qualis**: A1
  - **Date**: March 27-29, 2019
  - **Location**: Puerto Varas, Chile

- **Conference**: 14th International Conference on emerging Networking EXperiments and Technologies (ACM CoNEXT-2018).
  - **Authors**: MARCOS P., CHIESA M., **MÜLLER L.**, KATHIRAVELU P., DIETZEL C., CANINI M. and BARCELLOS M.
  - **Title**: Dynam-IX: a Dynamic Interconnection eXchange
  - **Type**: Main track (full-paper)
  - **Qualis**: A1
  - **Date**: December 4-7 2018
  - **Location**: Heraklion/Crete, Greece
  - **Digital Object Identifier (DOI)**: <https://doi.org/10.1145/3281411.3281419>

- **Conference**: ACM Special Interest Group on Data Communication (ACM SIGCOMM-2018).
  - **Authors**: MARCOS P., CHIESA M., **MÜLLER L.**, KATHIRAVELU P., DIETZEL C., CANINI M. and BARCELLOS M.
  - **Title**: Dynam-IX: a Dynamic Interconnection eXchange
  - **Type**: ACM SIGCOMM Posters and Demos
  - **Qualis**: A1
  - **Date**: August 21-23, 2018
  - **Location**: Budapest, Hungary
  - **Digital Object Identifier (DOI)**: <https://doi.org/10.1145/3234200.3234218>

A.2 Invited Talks

Directly related to thesis

- **MÜLLER, L.**, HUFFAKER B., LUCKIE M., CLAFFY K. AND BARCELLOS M. – *Challenges in Inferring Spoofed Traffic at IXPs*. In: 2nd International Workshop on Darkspace and UnSolicited Traffic Analysis (DUST-2019), San Diego, California, September 2019.
- **MÜLLER, L.**, HUFFAKER B., LUCKIE M., CLAFFY K. AND BARCELLOS M. – *Using IXPs to Measure Improvements of Source Address Validation Filtering in Inter-Domain Traffic*. In: CAIDA Workshop on Active Internet Measurements (AIMS-2018), San Diego, California, March 2018.
- **MÜLLER, L.** – *Internet Infrastructure Measurement: Trends and Challenges*. In: Seminario Internacional en Ciencias de la Computación (SiCC-2017), Medellín, Colombia, October 2017.
- **MÜLLER, L.** and MARCOS, P. B. and OLIVEIRA, R. R. and BERTHOLDO, L. M. and BARCELLOS, M. P. – *Análise de Abrangência dos IXs no Brasil (Routing Coverage Analysis At IXPs in Brazil)*. In: Semana de Infraestrutura da Internet no Brasil (IX-Forum), São Paulo, December 2016.

Related to collaboration projects

- MARCOS P., CHIESA M., **MÜLLER L.**, KATHIRAVELU P., DIETZEL C., CANINI M. and BARCELLOS M. – *Dynam-IX: a Dynamic Agreement Marketplace on Internet eXchange Points*. In: 13th European Peering Forum, Athens, Greece, September 2018.
- MARCOS P., CHIESA M., **MÜLLER L.**, KATHIRAVELU P., DIETZEL C., CANINI M. and BARCELLOS M. – *Dynam-IX: a Dynamic Agreement Marketplace on Internet eXchange Points*. In: ACM, IRTF & ISOC Applied Networking Research Workshop 2018, Montreal, Canada, July 2018.
- MARCOS P., CHIESA M., **MÜLLER L.**, KATHIRAVELU P., DIETZEL C., CANINI M. and BARCELLOS M. – *Dynam-IX: a Dynamic Agreement Marketplace on Internet eXchange Points*. In: RIPE 76, Marseille, France, May 2018.

A.3 Research Projects

- *GT-IPÊ Analytics: Transforming Raw Monitoring Data to Generate Valuable Information for Network Management*. Funding source: Rede Nacional de Ensino e Pesquisa (RNP).
  - 2017-2018 (Phase 1)
  - 2018-2019 (Phase 2)
  - 2019-2020 (Phase 3)
- *Uncovering the Hidden Dynamics of Changing Internet Interconnections*. Funding source: Microsoft Azure, Data Science Grant 2017-2018.

A.4 Research Grants

- *Mapping Interconnection in the Internet: Colocation, Connectivity and Congestion*. Funding source: NSF CNS-1414177 (CAIDA, 2018).

A.5 Supervised Final BSc Paper

- Alessandra Helena Jandrey. *Estudo Comparativo de Bases de Geolocalização IP: Mapeando os Sistemas Autônomos Brasileiros*. 2017. Trabalho de Conclusão de Curso. (Graduação em Engenharia da Computação) - Universidade de Santa Cruz do Sul. Orientador: Lucas Müller.
- Tiago Silva Leal. *Detecção e Análise Proativa de Anomalias no Tráfego de Rede*. 2016. Trabalho de Conclusão de Curso. (Graduação em Ciência da Computação) - Universidade de Santa Cruz do Sul. Orientador: Lucas Müller.
- Cassiano de Mello Padilha. *Avaliação Experimental de Tecnologias SDN para Implantação em Redes de Produção*. 2016. Trabalho de Conclusão de Curso. (Graduação em Engenharia de Computação) - Universidade de Santa Cruz do Sul. Orientador: Lucas Müller.
- Vinícius Martins de Souza. *Um Estudo Experimental do Plano de Controle em Redes Definidas por Software*. 2016. Trabalho de Conclusão de Curso. (Graduação em Ciência da Computação) - Universidade de Santa Cruz do Sul. Orientador: Lucas

Müller.

A.6 Co-Supervised Final BSc Paper

- Rodrigo Favalessa Peruch. *Medindo a incidência de spoofing no contexto de tráfego IPv6 inter-domínio em um IXP*. 2019. Trabalho de Conclusão de Curso. (Graduação em Ciência da Computação) - Universidade Federal do Rio Grande do Sul. Orientador: Marinho Barcellos. Co-orientador: Lucas Müller
- Fabricio Martins Mazzola. *Um Estudo Sobre os Efeitos de Mudanças de Configuração de Roteadores ao Longo do Tempo*. 2018. Trabalho de Conclusão de Curso. (Graduação em Ciência da Computação) - Universidade Federal do Rio Grande do Sul. Orientador: Marinho Barcellos. Co-orientador: Lucas Müller
- Rafael Rafael da Fonte Lopes da Silva. *A Study on SDN Control Plane Distribution*. 2015. Trabalho de Conclusão de Curso. (Graduação em Ciência da Computação) - Universidade Federal do Rio Grande do Sul. Orientador: Marinho Barcellos. Co-orientador: Lucas Müller
- Erick Malmann Wagner. *Avaliação de Desempenho de Controladores OpenFlow*. 2015. Trabalho de Conclusão de Curso. (Graduação em Engenharia de Computação) - Universidade Federal do Rio Grande do Sul. Orientador: Marinho Barcellos. Co-orientador: Lucas Müller
- Gustavo Miotto. *AppFlow: Suporte a Regras da Camada de Aplicação em Arquiteturas SDN OpenFlow*. 2015. Trabalho de Conclusão de Curso. (Graduação em Engenharia de Computação) - Universidade Federal do Rio Grande do Sul. Orientador: Marinho Barcellos. Co-orientador: Lucas Müller

B PAPER PUBLISHED AT ACM CONEXT 2019

- **Title**: Challenges in Inferring Spoofed Traffic at IXPs
- **Type**: Main track (full-paper)
- **Authors**: Lucas Muller (UFRGS), Matthew Luckie (U. Waikato), Bradley Huffaker (CAIDA/UCSD), Kc Claffy (CAIDA/UCSD), Marinho Barcellos (UFRGS)
- **Qualis**: A1
- **Date**: December 9-12, 2019
- **Location**: Orlando, Florida, U.S.
- **Digital Object Identifier (DOI)**: https://doi.org/10.1145/3359989.3365422

Challenges in Inferring Spoofed Traffic at IXPs

Lucas Müller
UFRGS / CAIDA
lfmuller@inf.ufrgs.br

Matthew Luckie
University of Waikato
mjl@wand.net.nz

Bradley Huffaker
CAIDA / UC San Diego
bradley@caida.org

kc claffy
CAIDA / UC San Diego
kc@caida.org

Marinho Barcellos
UFRGS and University of Waikato
marinho@inf.ufrgs.br

ABSTRACT

Ascertaining that a network will forward spoofed traffic usually requires an active probing vantage point in that network, effectively preventing a comprehensive view of this global Internet vulnerability. Recently, researchers have proposed using Internet Exchange Points (IXPs) as observatories to detect spoofed packets, by leveraging Autonomous System (AS) topology knowledge extracted from Border Gateway Protocol (BGP) data to infer which source addresses should legitimately appear across parts of the IXP switch fabric. We demonstrate that the existing literature does not capture several fundamental challenges to this approach, including noise in BGP data sources, heuristic AS relationship inference, and idiosyncrasies in IXP interconnectivity fabrics. We propose a novel method to navigate these challenges, leveraging *customer cone* semantics of AS relationships to guide precise classification of inter-domain traffic as in-cone, out-of-cone (*spoofed*), unverifiable, bogon, and unassigned. We apply our method to a mid-size IXP with approximately 200 members, and find an upper bound volume of out-of-cone traffic to be more than an order of magnitude less than the previous method inferred on the same data. Our work illustrates the subtleties of scientific assessments of operational Internet infrastructure, and the need for a community focus on reproducing and repeating previous methods.

CCS CONCEPTS

• **Networks** → **Network measurement**; **Network security**.

KEYWORDS

IP spoofing, Internet eXchange Point, Denial-of-service, Network filtering

ACM Reference Format:
Lucas Müller, Matthew Luckie, Bradley Huffaker, kc claffy, and Marinho Barcellos. 2019. Challenges in Inferring Spoofed Traffic at IXPs. In *The 15th International Conference on emerging Networking EXperiments and Technologies (CoNEXT '19), December 9–12, 2019, Orlando, FL, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3359989.3365422

CoNEXT '19, December 9–12, 2019, Orlando, FL, USA

ACM ISBN 978-1-4503-6998-5/19/12…$15.00
https://doi.org/10.1145/3359989.3365422

1 INTRODUCTION

Networks that allow spoofed source Internet Protocol (IP) addresses in packets are a cybersecurity risk on the global Internet, because they enable attacks such as spoofed denial-of-service (DoS) attacks that are operationally infeasible to trace back to the actual source. Recognizing that lack of *source address validation* (SAV) is fundamentally an architectural limitation [10, 60], the Internet Engineering Task Force (IETF) introduced best current practices recommending that networks block packets with spoofed source addresses [9, 29]. Compliance with these filtering practices has misaligned incentives i.e., it protects the *rest* of the Internet from attacks being sourced from the network that must pay a non-trivial cost for deploying and accurately maintaining the filters. Thus, despite many attempts to improve SAV deployment and mitigate the impact of DoS attacks, some of the most damaging DoS attacks in the Internet still leverage IP spoofing as a vector, setting new records each year for the volume of traffic launched at even highly provisioned networks, disrupting access to those networks [43, 44, 59, 71].

Identifying networks that do not filter spoofed packets is critical to global network infrastructure protection, because it provides a focus for remediation and policy interventions [53]. However, identification of these networks is challenging at Internet scale. The definitive method requires an active probing vantage point in each network being tested, to see if a spoofed packet successfully traverses the network [13, 15]. Since there are approximately 65K independently routed networks on the Internet in 2019 [6, 75], this method has limited feasibility for a comprehensive assessment of Internet spoofing.

Broader visibility into the spoofing problem may lie in the capability to infer lack of SAV compliance from large, heavily aggregated Internet traffic data, such as traffic observable at Internet Exchange Points (IXPs). Most Autonomous Systems (ASes) connect to an IXP to exchange traffic between their customers, i.e., via peering relationships where neither AS pays the other for transit. For these ASes, legitimate source addresses in packets will belong to direct or indirect customers of the AS sending the packets across the IXP fabric to their peers.

However, inferring SAV deployment at an IXP is remarkably challenging, more so than has been captured in the literature, due to a combination of operational complexities that characterize today's interconnection ecosystem. First, determining which source addresses are valid in packets arriving at a given port of an IXP switch fabric is challenging, because there is no registry of which addresses networks should forward; in practice, we must infer valid

source addresses. Second, while the original role of IXPs was to promote peering between ASes, networks now also use IXPs to obtain IP transit services from a provider [1], and we have found evidence of organizations joining their sibling network ASes across an IXP. For ASes offering transit across the IXP, and for sibling networks, it is infeasible to infer invalid source addresses from IXP traffic data – the set of valid addresses is potentially the entire address space. Third, while IXPs may be thought of as a single switching fabric, in practice IXPs and resellers offer complex services, including remote peering, layer-2 transport, and virtualized segmenting of traffic into multiple Virtual Local Area Networks (VLANs). These interconnection practices occur below and are thus not visible to the IP layer or in the Border Gateway Protocol (BGP).

Accurately inferring SAV deployment at an IXP requires navigating all of these aspects. In this paper, we describe a methodology that does so. One of our discoveries does not bode well for the ability to automate this method: identifying the myriad cases that explain patterns in traffic at a given IXP is largely manual in nature, and must be repeated at each IXP to accommodate IXP-specific architectural engineering and business decisions. However, we imagine its utility as part of an expert system suite of cybersecurity services or compliance practices of modern IXPs.

This paper makes the following contributions:

(1) We provide a detailed analysis of methodological challenges for inferring spoofed packets at IXPs. Based on IP routing, addressing, and IXP concepts, we analyze methodological challenges and their implications for building IP spoofing detection capabilities at IXPs (§2). We include a comprehensive analysis of previous work which also inferred spoofing at IXPs. We also analyze challenges specific to applying BGP-based SAV inference methods to modern IXP connectivity fabrics (§3).

(2) We develop a methodology to classify traffic flows for the purposes of accurately inferring spoofed traffic. We design and implement Spoofer-IX, a novel methodology to detect the transmission of spoofed traffic (which implies lack of source address validation) by AS members of IXPs (§4). Spoofer-IX addresses two fundamental issues not addressed in the existing literature [45]. First, Spoofer-IX considers the type of relationship between neighbors at an IXP when determining which source addresses are valid in IP packets crossing the IXP. Second, Spoofer-IX considers asymmetric routing and traffic engineering, by designing a prefix-level customer cone that includes addresses that may be valid source addresses for an AS to transit. The accuracy of this method depends on the quality of BGP data and AS relationship inferences, which we know to be imperfect [54]. However, our method is congruent with what network operators do when configuring static access control lists to deploy SAV [30, 37, 42].

(3) We use our methodology to classify packets at a IXP in Brazil with approximately 200 members. We apply our method to traffic and topology data (described in §5) from one of the largest IXPs in Brazil, with more than 200 member ASes using the IXP switching fabric. We report our analysis findings, and results of our interactions with IXP and network operators to validate the findings (§6). We investigate the impact of different filtering choices on inferred valid address space, and the likelihood of false negatives when classifying traffic according to different filtering choices. We also compare our method with a recently proposed method [45] that did not consider AS relationships in its inference of spoofed traffic, reporting that the majority of members at the IXP sent spoofed packets, and demonstrate pitfalls of this approach. Indeed, at the medium-sized IXP we studied, with approximately 200 members, this previous method inferred spoofed traffic coming from 62.3% of addresses over a one-week period in May 2019, but our AS-relationship-aware method inferred spoofed traffic coming from less than 1 in 5 (18.7%) member ASes during our five-week observation period in May 2019.

(4) We find evidence that epistemological and cross-validation challenges remain, and we publish our code to promote further work. When we compared our results with CAIDA's crowdsourced measurements, we found that CAIDA received positive spoofing tests (lack of SAV) in 54% of the member ASes at this IXP. This is not necessarily inconsistent, since even at a heavily aggregating exchange point, one cannot detect lack of SAV without actually observing spoofed packets, which CAIDA's crowdsourced approach explicitly injects. We conclude our paper with a discussion of lessons learned (§8), including that we believe further work is required to understand the degree to which IXPs can be used as a lens into SAV deployment, and why we think such work is important to future cybersecurity efforts. Our conclusions highlight the persistent tension between the need for reproducibility of methods and results [7, 8], and the opacity characteristic of commercial infrastructure. We publicly release our code [62] in hopes that other researchers and IXPs will use it to further improve our collective ability to measure and expand deployment of SAV filtering.

2 BACKGROUND AND RELATED WORK

2.1 Source Address Validation

The Internet architecture provides no explicit mechanism to prevent packets with forged headers from traversing the network. This vulnerability allows IP spoofing attacks, i.e., when hosts send IP packets using fake source addresses that cannot feasibly be traced back. To reduce the incidence of this type of attack, network operators can configure their routers to identify and block spoofed packets before these packets leave their networks. Such filtering is well-specified and a standardized IETF best current practice [29], frequently referred to as Source Address Validation (SAV) [38]. Network operators often implement SAV by using *ingress filters* in routers, which drop packets with source addresses outside the locally valid address space before they enter the global Internet.

2.2 Address Space Fundamentals

For the purposes of this study, we distinguish three main categories of IP address space: Bogon, Unassigned, and Routed. *Bogon* addresses are reserved by the IETF [22, 61] for specific uses such as private networks and loopback interfaces; they do not uniquely identify any host, and should not be routed on the Internet. *Unassigned* addresses [34, 35] have not been assigned by an Internet registry to an AS and should not be used or routed by anyone. *Routed* addresses have been assigned to some AS, and are thus potentially valid source addresses in inter-domain traffic.

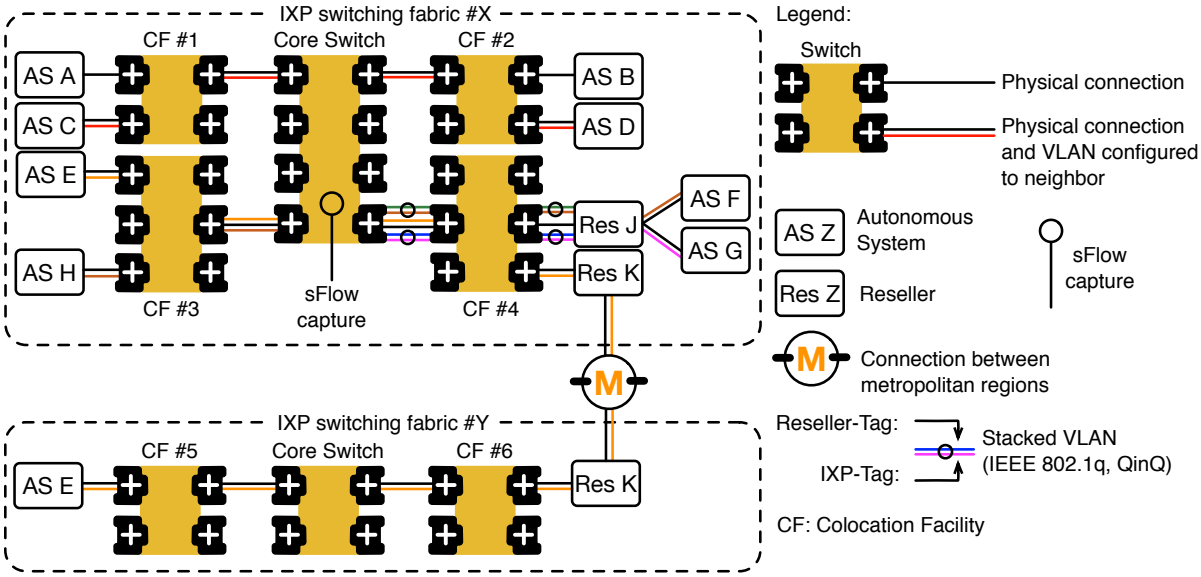


Figure 1: Illustration of the architecture of modern IXPs. Modern IXPs typically construct a switching fabric using a core switch that interconnects other switches located in remote colocation facilities. ASes typically connect to a switch located in a colocation facility, and can form bilateral peering relationships with neighbors. These ASes may request a VLAN to isolate their traffic from other members at the IXP. Resellers can provide services such as remote peering and layer-2 transport.

2.3 IXPs as Observatories

IXPs are attractive vantage points to observe signals of SAV deployment, as hundreds of ASes may be present at a single logical location. The IXP operator assigns each member a unique IP address from a prefix controlled by the operator, which the member assigns to their router interface connected to the IXP, and uses to establish BGP routing with other members. When a member AS's router transmits a packet across the Ethernet switching fabric, the source and destination media access control (MAC) addresses in the Ethernet frame uniquely identify the AS pair exchanging the packet, and its direction.

Figure 1 illustrates the architecture of many modern IXPs [4, 23, 28, 30, 40, 41, 48]. The figure contains two separate IXPs and their switching fabrics #X and #Y, with a core switch for each IXP. While some IXPs may consist of a single core switch where participants interconnect, operators achieve the scale of modern large IXPs by placing switches at distinct physical colocation facilities, any of which can serve as an IXP attachment point. The figure shows that the switches are adjacent, but in practice colocation facilities are usually in different buildings. IXP operators often use sFlow [66] or NetFlow [19] to collect traffic flow statistics. A comprehensive view of all traffic from all services at the IXP would require flow data captured from all switches in the switching fabric, as traffic between participants at a single colocation facility will not travel to the core switch.

Participants can exchange traffic directly across the switching fabric in a bilateral session. In figure 1, ASes A and B exchange traffic directly. However, modern IXPs often use VLANs to provide logical isolation between different types of interconnection [18, 27]. For example, an IXP may provide a route server, but only offer that route server on a specific VLAN. Similarly, traffic between two participants may be sufficiently sensitive or high volume that members request a VLAN from the IXP to isolate their communications [3, 24, 47]. In figure 1, ASes C and D exchange traffic in their own isolated VLAN.

To foster IXP growth and enable more networks to interconnect, IXPs have supported resellers, which provide value-added services at an IXP, such as remote peering and layer-2 transport [17, 39, 58, 64]. A reseller provides remote peering services so that an AS that is not physically present at a colocation facility can still reach other members at the IXP, without the AS incurring colocation facility fees or port charges from the IXP operator. These resellers require some cooperation with the IXP, e.g., [2, 46]. The IXP assigns the remote peers any VLAN tags they require to participate at the exchange as local members do.

An IXP may use different technical approaches to support remote peering providers [17, 41, 64]. A reseller can bridge Ethernet networks so that the MAC address of the customer router's interface will uniquely identify the origin of traffic in the peering fabric. A second approach is for a reseller to push a tag (reseller-tag) to uniquely identify their specific customer AS to the IXP, so that the MAC address of the Ethernet frame corresponds to the reseller's router. Figure 1 illustrates this second approach, where reseller J allows customer ASes F and G to reach other members. When the reseller transmits these packets into the IXP, the reseller also pushes a tag (reseller-tag) to uniquely identify their specific customer AS to the IXP. The IXP bridges traffic into the IXP switching fabric by removing the outer-most reseller-tag while keeping the IXP-tag. In figure 1, the sFlow tap sees the IXP-tag and the MAC address of the reseller, which uniquely identifies the AS that sent the packet.

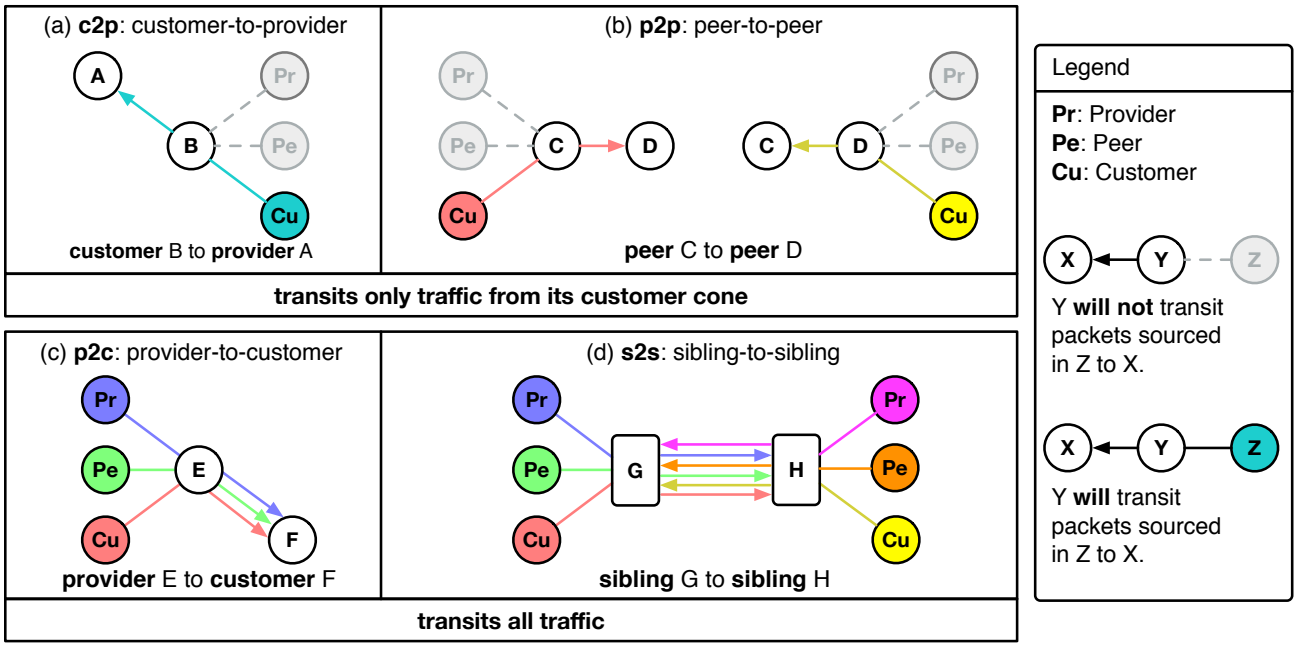


Figure 2: The customer cone constrains the set of source addresses expected in valid inter-domain traffic transiting an AS behaving rationally in a c2p or p2p relationship. In the c2p relationship shown in (a), B transits traffic from its customers to A, but not its peers and providers. Similarly, in the p2p relationship shown in (b), C only transits traffic from its customers to D (likewise, from D to C). However, as shown in (c), the p2c relationship does not constrain the source addresses transited by E to F, and neither does the s2s relationship between G and H in (d).

A reseller can also provide remote peering to members colocated at one IXP that want to reach members in a different IXP. Figure 1 shows a more complicated example, where AS E bridges their network between metropolitan regions using the services of a reseller (K) present at both IXPs.

2.4 AS Relationships and Customer Cones

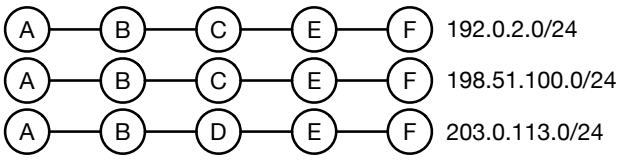
The three primary classes of AS relationships are customer-provider (c2p, p2c), peering (p2p) and sibling (s2s). In a c2p relationship (also known as transit), a customer buys access to achieve global reachability to all routed Internet address space. In a p2p relationship, two ASes agree to exchange traffic destined to prefixes they or their customers own, typically without either AS paying the other [31]. In a s2s relationship, a single organization operates both ASes, and may transit packets received from any source.

An AS's *customer cone* includes all ASes reachable through its customer ASes, i.e., direct and indirect customer ASes (in other words, ASes reachable only through p2c links) [54]. The customer cone constrains which source IP addresses one should see in valid inter-domain traffic transiting from a customer to its provider, or between peers. Figure 2 illustrates the subtleties: an AS in a c2p or p2p relationship with another AS should only send packets with a source address from within its customer cone – respectively, (a) and (b) in figure 2. In contrast, a link between a provider to its customer or between two siblings may forward packets with *any* routed source address – (c) and (d) in figure 2.

2.5 Measuring Deployment of SAV

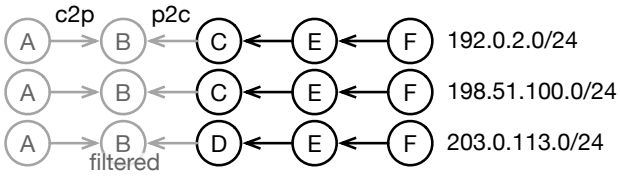
Many academic research efforts have described techniques to promote deployment of SAV [25, 49, 50, 77]. Fewer efforts have tried to empirically measure SAV compliance for networks attached to the global Internet. In 2005, Beverly, *et al.* developed a client-server technique to allow users to test networks to which they are currently attached [12], and operationalized a platform to track trends over time [13, 15]. The platform allows for inference of deployed SAV policy, but has limited coverage, because it relies on users downloading and running measurement software. To overcome this limitation, researchers have recently investigated techniques to infer lack of SAV using macroscopic Internet data sets. In 2017, Lone *et al.* reported a technique to infer spoofed traffic in massive traceroute archives, based on the assumption that an edge network should never appear to be providing transit in a traceroute path [51]. This method is limited by whatever appears in the traceroute archives, and can be hampered by traceroute artifacts caused by inconsistent Internet Control Message Protocol (ICMP) implementations in routers [57].

Most closely related to our study, in 2017 Lichtblau *et al.* used a large IXP as a vantage point for inferring which networks at the IXP had not deployed SAV [45]. For each member at the IXP, their method infers a set of IP prefixes containing addresses that may legitimately appear in the source field of IP packets crossing an IXP. They infer that a member AS that sends a packet into the IXP switching fabric with a source address outside of those prefixes has not deployed SAV. They argued against using AS relationships and



ASes	Full Cones:
A, B, E, F	192.0.2.0/24, 198.51.100.0/24, 203.0.113.0/24
C	192.0.2.0/24, 198.51.100.0/24
D	203.0.113.0/24

Figure 3: Example full cones (§3.1.1) for six ASes given these BGP paths. The full cone for an AS includes every prefix that contains that AS in the path for all routes observed by public route collectors, regardless of the underlying relationships.



ASes	AS-Level Customer Cones:
A, B, F	(contain only themselves with these paths)
C	C, E, F
D	D, E, F
E	E, F

ASes	Prefix-Level Customer Cones:
A, B	(empty with these paths)
C, D, E, F	192.0.2.0/24, 198.51.100.0/24, 203.0.113.0/24

Figure 4: Example customer cones (§3.1.2) for six ASes using the same BGP paths from figure 3. In customer cone construction, we annotate each AS link with a c2p, p2c, or p2p relationship before inferring the prefix-level customer cone.

AS customer cones which they claimed did not address asymmetric routing. However, their method did not consider ASes forming customer-provider or sibling relationships at the IXP, where all routed addresses may be legitimate source addresses in IP packets crossing an IXP – (c) and (d) in figure 2. In these cases, there is no way to infer SAV deployment across these links at the IXP.

3 TACKLING METHODOLOGICAL CHALLENGES

We describe the core of our methodology in the context of two complex challenges to inferring spoofed traffic in IXP traffic data. The first challenge (§3.1) is determining which addresses are valid source addresses in traffic transiting a given neighbor AS, i.e., packets with a source address that is *in-cone* for that AS. An incomplete set of valid addresses could yield false inferences of failure to deploy SAV, should a valid address appear in the observed packets but not be in the in-cone set, i.e., be *out-of-cone* for that AS. The second challenge (§3.2) is navigating the analytical implications of modern IXP interconnection practices that can impede the visibility of both topology and traffic. These practices complicate the analysis of which ASes exchanged traffic and their routing relationship. Once we address these challenges, the remainder of our method is IXP-specific but straightforward, and we describe it in §4.

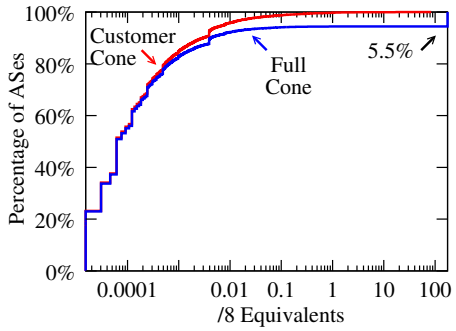
3.1 Subtleties in Cone Construction

Inferring the set of valid source addresses for packets traveling from a specific AS to a specific adjacent AS at an IXP requires navigating a multidimensional parameter space. Precision in this process is crucial. Mistakenly excluding valid addresses could result in a misclassification of an AS as not performing source address validation (false positive). Similarly, including invalid source addresses could result in spoofed packets going undetected (false negatives). As mentioned in §1, there is no global registry that contains ground truth on which addresses are valid source addresses for packets transited by an AS; instead, we must infer them from BGP routing data sources [65, 68, 70], even though these sources may contain spurious announcements [52].

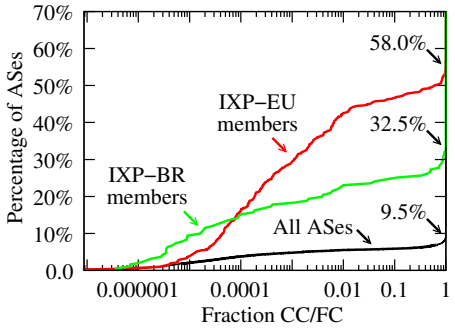
3.1.1 Full Cone. The full cone (used in [45]) is the more permissive of the two construction methods. Aiming to minimize false positives, Lichtblau *et al.* chose to *"not distinguish between peering/sibling, customer-provider and provider-customer links. Rather, whenever [the algorithm sees] two neighboring ASes on an AS path, [the algorithm] presumes a directed link between the two, where the left AS is considered upstream of the right AS."* The resulting cone for an AS, which they call its *full cone (FC)*, includes every prefix that contains that AS in the BGP route's AS path [45], for all routes observed by public route collectors in Routing Information Base (RIB) snapshots and updates during the measurement period.

They acknowledged that this method intentionally sacrifices specificity, i.e., inflating the address space considered legitimate for each AS pair, in the interest of avoiding false positives, i.e., avoiding mistakenly attributing a failure to deploy SAV. Using this method, a stub AS that provides a public BGP view containing all prefixes it received from its peers and providers will have *all* of these prefixes included in its full cone, i.e, the entire routed address space will be deemed valid. Figure 3 illustrates the full cones for six ASes; if A were a stub AS and a customer of B, all three prefixes would be included in A's full cone even though no system in A should originate packets with those source addresses.

3.1.2 Customer Cone. The customer cone is the more restrictive of the two construction methods; it takes into account the semantics of AS relationships. As described in §2, the AS-level customer cone defines the set of ASes reachable using customer links from the AS, including the AS itself [54]. We use the *provider/peer-observed customer cone* (PPCC) algorithm defined in [54] to build an AS-level customer cone. Using the paths in figure 4, the PPCC method constructs the cone of AS C using routes observed from its providers and peers. The PPCC method accommodates hybrid relationships,



(a) Absolute Cone Sizes.



(b) Relative Cone Sizes.

Figure 5: The cone construction approach significantly impacts the source addresses each method will consider valid. In (a) we show that 5.5% of all ASes had the equivalent of all routed addresses (175 /8 equivalents) in their full cone in April 2017. In (b) we show that while 90.5% of ASes had (full and customer) cones covering the same set of addresses, 58% of the IXP-EU members would have covered more addresses, with 42% of ASes having a full cone 100 times larger than their customer cone. Note, per discussion in §3.2, an AS announcing 0.01% /8 equivalents is announcing less than 0.006% of the routed address space.

where an AS may not propagate all of its customer routes to all of its peers and providers. Customer cone inference critically relies on accurate routing relationship inferences; a customer link incorrectly inferred to be a peer link will result in address space that the provider AS transits being incorrectly excluded from its customer cone. Figure 4 illustrates the AS-level customer cones for the same ASes and paths as figure 3, with link annotations to identify the inferred routing relationships between ASes. However, an AS-level customer cone does not define the set of valid source addresses in traffic transiting a given neighbor AS.

Once we have the AS-level customer cone for C, we transform it into its corresponding prefix-level cone by including all prefixes originated by ASes in the AS-level customer cone for C during the same observation window. This novel prefix-level cone construction accommodates traffic engineering practices, where an AS may announce different prefixes through different providers, but forward traffic from within these prefixes according to the best route to the destination. To illustrate, in figure 4, we include 203.0.113.0/24 in C's prefix-level customer cone, even though that prefix is not observed in any BGP paths involving C, because F is in C's customer cone. Importantly, we do not include these three prefixes in A's customer cone, because A has no customers. We also combine the prefix-level customer cones of siblings, because a sibling C may transit packets from the customer cone of any of C's siblings to C's peers or providers.

3.1.3 Impact of the Cone Construction Method. Figure 5 shows how the choice of cone construction method impacts inference of valid address space for all ASes (figure 5a) and for the ASes at the IXP-EU used in [45] and the IXP-BR in our study (figure 5b), in both cases using traffic and BGP data from April 2017 (see §5 for further detail on the datasets we used). In particular, 5.5% of all ASes in the Internet had a full cone that contained all routed address space. For 90.5% of ASes, the full cone and customer cone were congruent (included the same addresses), but 58% of IXP-EU member ASes had full cones covering more addresses than the customer cone,

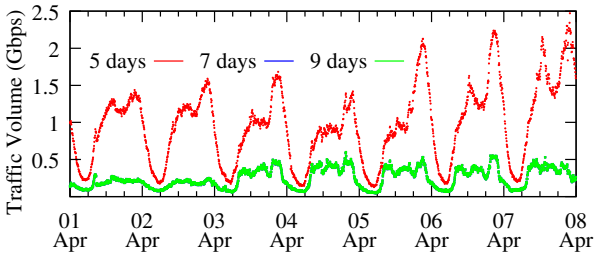


Figure 6: The inferred out-of-cone traffic volume for the full cone is sensitive to changing BGP observation window sizes in the construction of the cone. While the 7 and 9 day lines are almost identical, the 5-day line contains an order of magnitude more traffic because the set of valid addresses for each AS is smaller.

and 42% of ASes had an FC 100 times larger than their CC. This disparity of cone sizes for all ASes compared to those at the IXP is because while over 80% of the Internet's ASes are stubs, i.e., do not provide transit, these are less likely to peer at an IXP. Further, IXPs are popular places to operate public route collectors because the collector can obtain BGP routing views from multiple ASes at a single place. Therefore, those ASes at an IXP that provide a routing view will have all of the prefixes they announce in routes to the collector, including those from their peers and providers, in their full cone. Figure 6 shows how the choice of BGP observation window impacts [20] the inference of out-of-cone traffic at our IXP in Brazil in April 2017 using the full cone. This effect is because of the FC's permissive nature, which exposes the cone inference to announcements across the whole Internet.

Neither the full cone nor the customer cone handle the complexities that sibling ASes (ASes under the same ownership) bring. Because siblings may provide mutual transit to each other, the set of valid addresses that can transit between each AS is the entire

routed address space. To observe this behavior in public BGP data, which both the FC and CC use, would require a view from each sibling AS. Current sibling relationship inference methods [14, 32] use WHOIS data, which is not only inconsistently formatted across regions, but also becomes stale if not updated as mergers occur, leading to false and missing inferences [32].

3.2 Topology and Traffic Visibility

While the original role of IXPs was to promote peering between ASes physically present and connected to a switching fabric, in practice IXP services have become more complicated. For example, many networks now obtain transit services from a provider at the IXP [1]. Or, an organization can connect its sibling networks using the IXP switching fabric. IXPs may also offer services such as remote peering and layer-2 transport, as well as virtualized segmenting of traffic into multiple VLANs. These services present three challenges to accurate inference of SAV deployment.

First, the BGP routing relationship between two IXP members impacts whether the customer cone can constrain inference of valid source address space. As discussed in §2.4, a provider AS may forward packets with a source address from any routed prefix in the Internet to their customer, and a sibling may forward packets from the provider of one sibling to the customer of another sibling. In these cases, we cannot apply a cone of valid addresses to infer the SAV policy of the transmitting member. We can only make this inference when that member has a peering or transit relationship with another member. In contrast to prior work [45], we consider the routing relationship between the two IXP member ASes exchanging traffic when evaluating the source address of a packet crossing the IXP.

Second, there are traffic visibility impediments. As discussed in §2.3, traffic between members connected to the same switch will stay within the switch. In a distributed switching fabric, observing all member traffic requires traffic capture from all switches. Similarly, ASes may establish private interconnections with other ASes at the same colocation facility; their traffic exchange does not use the core IXP switching fabric. Further, to infer SAV policy of an IXP member, we require hosts in the cone of the IXP member to attempt to send spoofed packets to hosts they would reach across the IXP. Because most ASes peer at an IXP, only destinations in the customer cone of the receiving AS would receive that packet, i.e., the victim or the amplifier must be reached via the IXP. Because most customer cones are small (figure 5a, where only 5% of ASes have more than 0.006% of the routed address space in their customer cone) the chance of a victim or amplifier also being reached via a peering relationship at the IXP is small; a victim or amplifier is more likely to be reached via a transit relationship at the IXP.

Third, shared use of IXP ports creates attribution challenges. While the IXP can supply the AS number of record for a given port, with the associated Ethernet MAC address, that port does not necessarily uniquely identify the sending AS when a reseller uses the port to provide layer-2 transport, in cases of remote peering and port resale (§2.3), or when the port connects to another exchange. Prior work has illustrated measurement challenges of inferring remote peering [17, 64]. In this work, the IXP provided us the reseller and IXP tags they used to bridge remote peers. This IXP-specific knowledge exemplifies why we believe a customer-cone-based approach to SAV inference will ultimately be integrated into expert system capabilities rather than be amenable to complete layer-3 automation.

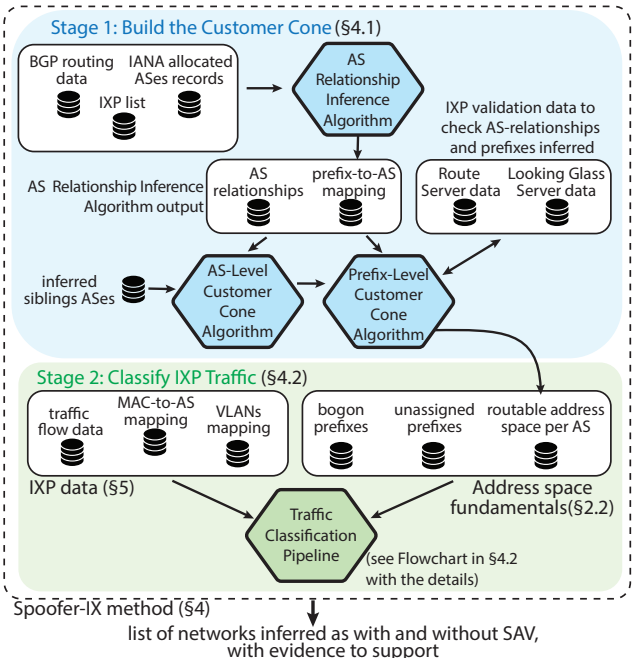


Figure 7: Spoofer-IX Inference Method Overview.

4 IMPLEMENTING CLASSIFICATION PIPELINE

The customer cone construction method described in §3 underpins our traffic classification method - how we infer invalid source addresses (presumably spoofed) in packets crossing an IXP, and the ASes responsible for transmitting them. We describe how these pieces fit together in our system implementation, which relies on IXP traffic measurements and topological information, i.e., BGP data and IXP switching fabric forwarding databases. The implementation, illustrated in figure 7, has two stages: (1) build an accurate *prefix-level customer cone* from BGP data, and (2) verify that the customer cone can serve to constrain our inference, and if so classify traffic as *in* or *out* of the transmitting AS's customer cone.

4.1 Stage 1: Build the Customer Cone

The first stage has three phases, as follows.

Phase 1: Filter and Sanitize AS Paths. To avoid incorrectly identifying non-existent links between ASes, we use the method from [54] to discard paths with artifacts, such as loops, non-adjacent Tier-1 ASes, and reserved/unassigned ASes [33]. We also discard paths to prefixes longer than /24 or shorter than /8.

Phase 2: Infer AS Relationships. We use the sanitized AS Paths from phase 1 to derive AS relationships on a weekly basis, also according to the algorithm in [54]. This algorithm applies heuristics

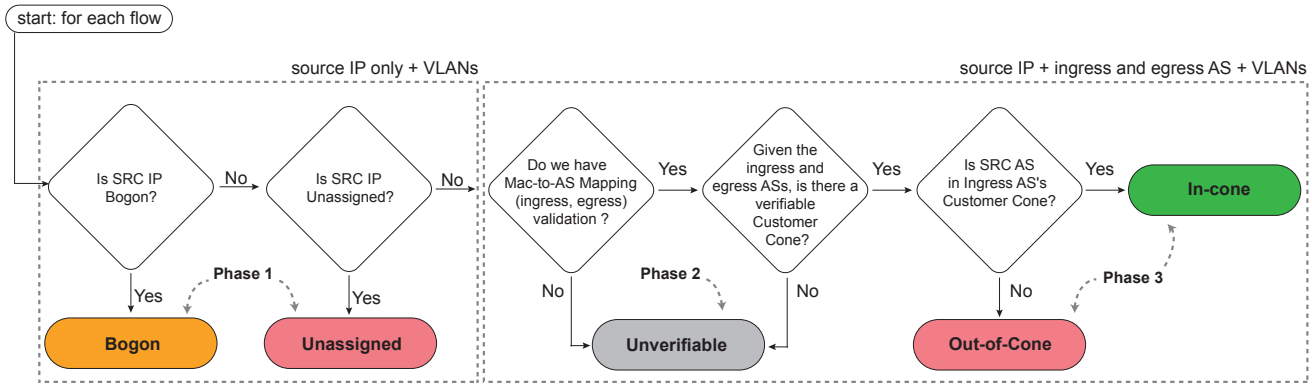


Figure 8: Flowchart showing our traffic classification pipeline (stage 2 of methodology, described in §4.2).

to annotate each AS link with either a transit (*C2P, P2C*) or peering (*P2P*) relationship.

Phase 3: Construct the Prefix-Level Customer Cone. An AS's *prefix-level customer cone* is the set of prefixes covering source addresses from the AS and its customers, for which the AS will transit traffic. Conceptually, constructing this cone is the most complicated part of our method, and where mistakes can impact its accuracy. We construct a prefix-level customer cone using the method we described in §3.1.2.

4.2 Stage 2: Classify IXP Traffic

The second stage has three phases, illustrated in figure 8.

Phase 1: Filter Bogon and Unassigned Addresses. We first classify traffic with *bogon* and *unassigned* source IP addresses, according to Team Cymru [73], as described in §5. Networks sending packets with unassigned source IP addresses are unlikely to have implemented SAV correctly, since the most obvious implementation blocks traffic from such addresses because they are not routed, therefore have no feasible return path. This phase is independent of any routing semantics, unlike the subsequent two phases, which consider the sending and receiving ASes for the monitored link, the routing relationship between them, and the prefix-level customer cone of the sending AS.

Phase 2: Filter Unverifiable Packets. This phase classifies traffic flows as suitable to inference of spoofing using the customer cone, marking unsuitable traffic as *Unverifiable*. Verifiable traffic must satisfy all of the following:

(1) It must have a valid MAC-to-AS mapping for both the sending and receiving MAC addresses.
(2) It must not have a known router IP address in the source IP address of the packet. Such a source IP address could be from any interface on the router, which might be assigned by an AS whose address space is not in the customer cone of the router's owner.
(3) It must not have a known IP address of the IXP LAN prefix. These prefixes are assigned to the IXP operator and should not be publicly announced, but sometimes member ASes mistakenly announce them.
(4) It must not have a source MAC address from a remote peer or layer-2 transport provider.
(5) It must not have a source MAC address from a known provider or sibling of the receiving AS.

Phase 3: Classify Packets with Customer Cone. The remaining traffic has a MAC-to-AS mapping, and is either transmitted by a customer of a transit provider at the IXP, or by a peer of another AS at the IXP. If a relationship was not visible in BGP, then we assume the traffic between these members was p2p and use the cones to classify the traffic exchanged. For these transmitting ASes, we classify traffic as *in-cone* or *out-of-cone* using the prefix-level customer cone (henceforth *customer cone* or *CC*) created in the previous stage. We classify a packet whose source IP belongs to the sending AS's customer cone address space as *in-cone*. Otherwise, we classify the packet as *out-of-cone*.

5 DATASETS

IXP-BR: traffic and routing data. We used sFlow [66] traffic data from a Brazilian IXP [40]. This IXP transports up to 200 Gbps of traffic among 200+ members. The IXP operators configured a sample rate of 1:4096 packets, and we used two datasets from 1 April to 6 May 2017, and 1 May to 5 June 2019, to evaluate our method.

Topology data over connectivity fabric. To identify the pair of adjacent ASes sending and receiving each flow across the IXP fabric, we used layer-2 information (i.e., MAC addresses) since the source and destination IP addresses in the IP headers of the observed packets contain the communication endpoints. To map MAC addresses to sending and receiving ASes of each flow (the MAC-to-AS mapping), we relied on information from the forwarding database of each switch that is part of the IXP switching fabric.

Router IP addresses. For comparability with previous work [45], we used CAIDA's Internet Topology Data Kit (ITDK) [16] to identify router interface IP addresses. We used the ITDK snapshot closest in time to the IXP traffic capture window. We consider traffic from ITDK-inferred router interfaces to be *unverifiable* (§4.2) because the source IP address could be from any of the interfaces of the router, which might be assigned by an AS whose address space is not in the Customer Cone of the router's owner (§4.2).

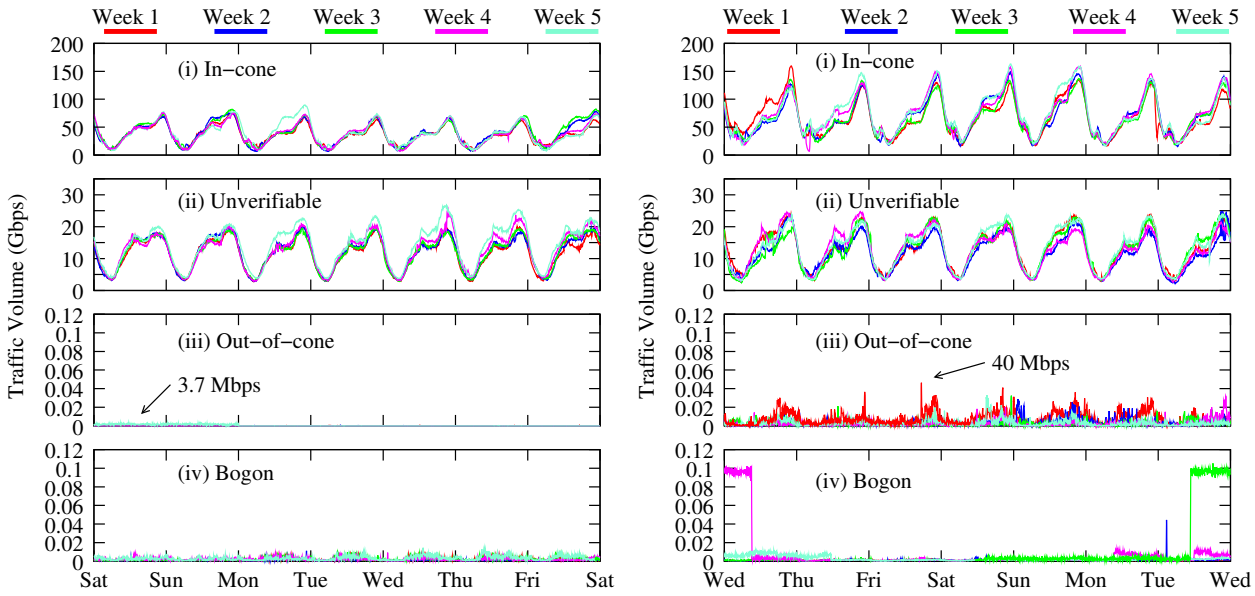


Figure 9: Five weeks of traffic for 2017 and 2019 classified with our method. We omit the unassigned class, which is negligible. For all ten weeks, we inferred almost no out-of-cone traffic – a maximum of 40 Mbps for an IXP with a peak of 200 Gbps.

Bogons and unassigned addresses. We used Team Cymru's Full-bogons feed [72, 73] to filter out traffic with source IP addresses that are bogons (e.g., private, special use, reserved) [22, 61, 76] or unassigned. Unassigned prefixes are allocated by IANA to an RIR [34, 35], but not subsequently assigned by the RIR to an end-user (e.g., an ISP) [75]. We used the lists compiled by Team Cymru in each 4h interval per day for the same time windows as our IXP traffic data collection.

Public BGP Data. Our traffic filters rely on Customer Cones inferred from public BGP routing table snapshots collected by Route Views (RV) and RIPE's Routing Information Service (RIS) [65, 70]. We downloaded one BGP RIB table per day from all available (18 and 16 in 2017, 19 and 18 in 2019 from RIS and RV, respectively) collectors for the same time windows as our traffic data. We extracted all AS paths in these tables that announced reachability to IPv4 prefixes, repeating this process for each week.

AS Siblings. We used CAIDA's AS to Organization classification of ASes into sets that likely belong to the same organizations [32]. CAIDA's method parses the Regional Internet Registries' WHOIS dumps and delegation files to create a unified mapping between ASes and organization names, then uses hints in the name strings, delegation files, identifiers, and email addresses to infer AS sets with common ownership. For each measurement period, we used the AS-to-Organization mapping that CAIDA constructed using WHOIS data collected closest to the traffic capture window.

6 RESULTS

Figure 9 shows the volumes of traffic we classified into each category for two different five-week periods in 2017 and 2019. We present these two five-week periods to show our results are consistent at least for these time periods. In 2017, the peak rate across the core switch during the period was 120 Gbps; in 2019 the peak had grown to 200 Gbps, and as expected the majority of the traffic across the exchange is classified as in-cone.

In 2017, the peak out-of-cone traffic we inferred was 3.7 Mbps, and in 2019, 40 Mbps. We believe these values are upper-bounds for out-of-cone traffic at the IXP core switch, and we derived these volumes after investigating the underlying properties of traffic between pairs of members, in rank order of contribution to the out-of-cone traffic volume at the IXP. For packets that had a signal they were not spoofed – e.g., a Transmission Control Protocol (TCP) packet with payload, or packets towards a known transport provider, we manually investigated the relationships between the parties. We found 27 sibling ASes in 11 distinct organizations that were exchanging traffic across the IXP, but missing from CAIDA's public AS-to-Org dataset (§5). To determine which ASes were siblings, we consulted the official website of those ASes to find information on their ownership, contacted the ASes directly to enquire, or contacted the IXP operators to understand the relationship between two ASes at the IXP. Further, through the IXP operators, we approached 36 members of the IXP, and 34 of those members responded with explanations of the behavior we saw.

Although the number of members was similar between 2017 and 2019 (208 and 203, respectively), 28 new members were present in the 2019 analysis. Because we focused our manual investigations on the 2017 data, we believe that there are additional sibling relationships and routing behaviors in the 2019 data that we have not discovered yet. We hypothesize that these missing sibling inferences are the likely cause of the increase in out-of-cone traffic between 2017 and 2019. Table 1 summarizes the number of unique AS pairs we observed to exchange traffic for the five week periods beginning 1 April 2017 and 1 May 2019. While we inferred more

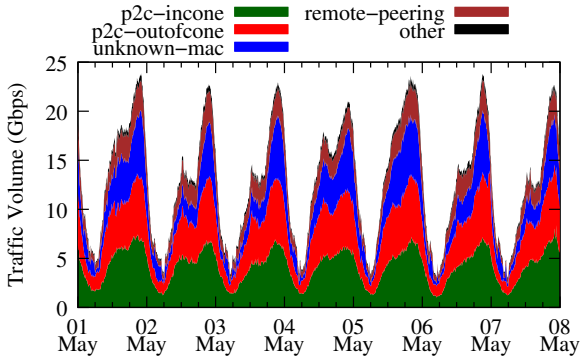


Figure 10: Classification of unverifiable traffic. 61.8% of the unverifiable traffic was sent by a provider to a customer across the exchange. Because a provider can transit packets from any source address in the Internet, there are no invalid addresses which would allow detection of spoofed packets. For completeness, we further classify traffic from each provider as being in or out of their customer cone.

Relationship	April 2017		May 2019	
p2p	19,161	(98.7%)	12,057	(98.4%)
p2c	222	(1.1%)	183	(1.5%)
s2s	21	(0.1%)	10	(0.1%)
total	19,404		12,250	

Table 1: Unique AS pairs observed exchanging traffic at the IXP in each 5-week period. Approximately 1.4% of AS pairs had a non-p2p relationship. (This IXP was rearchitected in 2019, which may explain the drop in observed peers.)

than 98% of the AS pairs had a p2p relationship, approximately 1.4% of AS pairs had a different class of relationship that impacts our ability to infer SAV policy of the transmitting AS.

Figure 9 also shows the volume of traffic with bogon source addresses, with a peak of approximately 100 Mbps across the exchange for the Wednesday at the end of week 3 (9b-iv). We found these networks deliberately used RFC1918 private addresses as source addresses of packets used to tunnel traffic between members – Generic Routing Encapsulation (GRE) and IP-in-IP represented 61.1% of the traffic, while the other 38.9% were ICMP, TCP, and User Datagram Protocol (UDP).

For both the 2017 and 2019 observation periods, there was a peak of approximately 25 Gbps of unverifiable traffic across the exchange, representing 15.3% of total traffic exchanged at the IXP (figures 9a-ii and 9b-ii). Figure 10 provides a classification of the traffic involved for the first week of May 2019. 61.9% of the unverifiable traffic was sent from a provider to a customer across the exchange, where no cone of valid addresses applies (§2.4). If we had applied the customer cone approach to this p2c traffic, we would have inferred 52% of it was from within the provider's customer cone, with the remaining 48% of traffic being from outside of the provider's customer cone. Because a provider can transit packets from any source address in the Internet (§2.4), there are no invalid addresses that would allow detection of spoofed packets. This potential for erroneous inference

Spoofer-CAIDA	Spoofer-IX		Sum
	In-cone	Out-of-cone	
Spoof-received	17	2	19 (54.3%)
Spoof-blocked	14	2	16 (45.7%)
Sum	31 (88.6%)	4 (11.4%)	35

Table 2: Congruity between CAIDA's public spoofer dataset and inferences using the IXP. Of the 35 overlapping ASes, CAIDA's spoofer dataset inferred 54% of them had not deployed SAV, because CAIDA received a packet with a spoofed source address. Only 4 of these 35 (11%) were observed to forward an out-of-cone packet into the IXP; 2 of these 4 were in CAIDA's spoofer dataset as not deploying SAV.

is why we must classify all packets from a transit provider to a customer as unverifiable. Another 21.4% of the unverifiable traffic was because we did not have an AS mapping for either the source or destination MAC addresses (the IXP lacked historical data for this mapping), and for 14.1% of traffic we could not determine the origin AS because the source MAC address and VLAN tag indicated the traffic was from a remote peering provider. Finally, all of the other categories summed to only 2.6% of the traffic, so we do not discuss these categories further.

We inferred out-of-cone traffic for 38 of the 203 members (18.7%) at the IXP between 1 May and 5 July 2019. Of the 203 members, 35 (17.2%) were also in CAIDA's public spoofer dataset [15], which requires a volunteer to have been present in the network to run an active measurement test that explicitly sends packets with spoofed source addresses to CAIDA's servers to test SAV deployment of the volunteer's network (§2.5). Table 2 summarizes the (in)congruity between the two datasets. Of the 35 ASes that overlapped, CAIDA's spoofer dataset indicated 54% of them had *not* deployed SAV. Only 4 of these 35 ASes (11%) were inferred by Spoofer-IX to forward an out-of-cone packet into the IXP, implying that this IXP may not provide effective visibility into SAV deployment, because participants were not forwarding spoofed packets, at least during our five-week observation window.

Figure 11 shows the volume of out-of-cone traffic inferred by both the Spoofer-IX and full cone methods for traffic data captured during the first week of May 2019. The Spoofer-IX method inferred a peak of 40 Mbps of out-of-cone traffic (best seen in figure 9b), whereas the full cone method inferred a peak of 2.5 Gbps. The diurnal pattern of the inferred out-of-cone traffic matches user-demand for content, with no observable peaks suggesting a volumetric spoofed-source attack launched from within member ASes of the IXP. The second row of figure 11 shows churn in source IP addresses [11, 69] seen in each five minute window. For the full cone method, the absolute volume of source addresses observed follows the traffic volume profile as a whole, and is concentrated in 20-40 ASes per five minute window, which is not a typical pattern of attacks that utilize randomly-spoofed source addresses.

The discrepancy between the size of the traffic classified as out-of-cone by the full cone and Spoofer-IX methods is because the full cone classified some provider-to-customer traffic as being out-of-cone (§2.5), whereas Spoofer-IX classified provider-to-customer traffic as unverifiable. Figure 10 shows Spoofer-IX classified 1 – 5

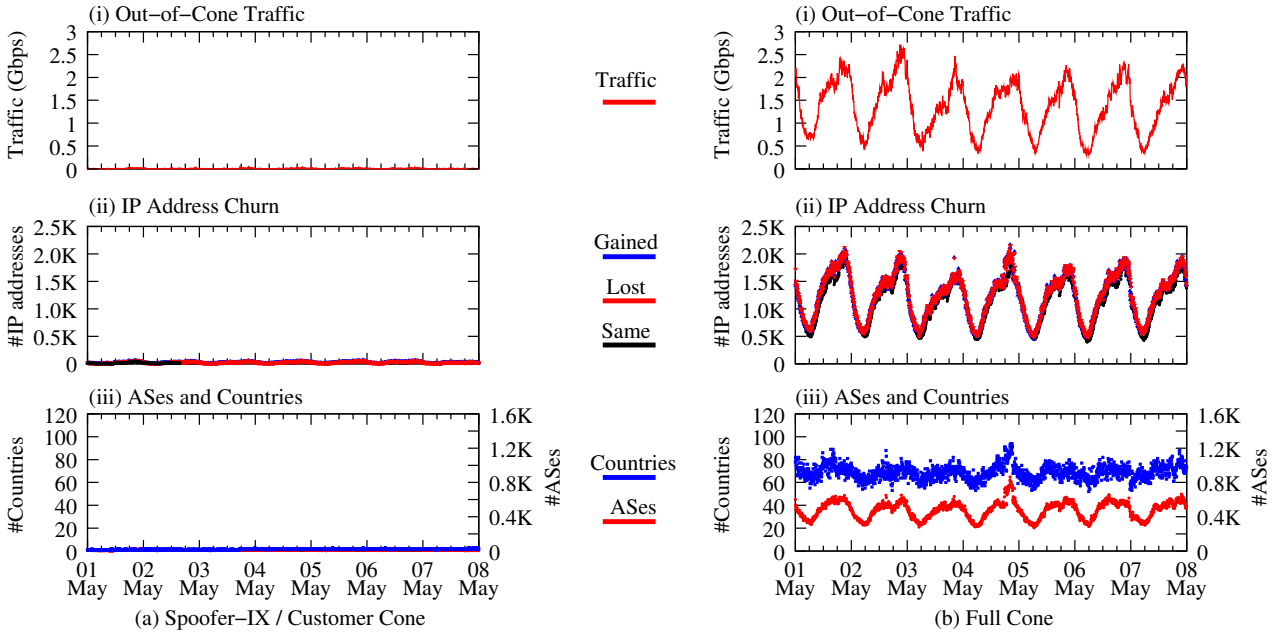


Figure 11: Comparison of metrics for out-of-cone traffic inferred by the Spoofer-IX and full cone methods for the first week of May 2019. We compute each metric per 5-minute window of traffic data, and use the same range on Y axes between methods to allow for comparison. For the IXP we studied, the full cone method inferred an average of 1.5 Gbps of out-of-cone traffic, whereas our method inferred a maximum of 40 Mbps (best seen in figure 9b-iii).

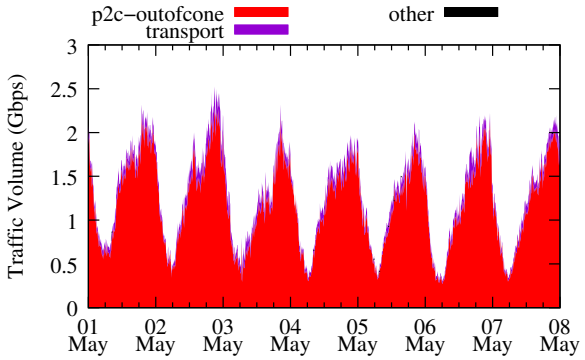


Figure 12: Spoofer-IX classification of traffic classified as out-of-cone by the full cone method. Spoofer-IX infers that 92.6% of this out-of-cone traffic was from a provider to customer across the IXP, and therefore unverifiable, because a provider can transit traffic from any source IP address to their customer, and it is therefore not feasible to identify spoofed packets by their IP address alone.

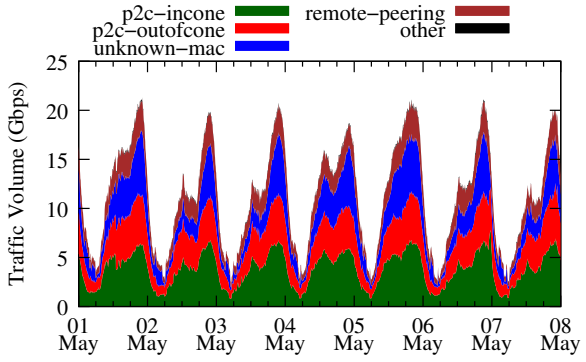


Figure 13: Classification of in-cone traffic for the full cone that Spoofer-IX classified as unverifiable. The traffic profile is similar to that in figure 10, with some unverifiable provider-to-customer traffic classified as out-of-cone by the full cone method (figure 12).

Gbps of out-of-cone traffic from providers to customers as part of the unverifiable traffic that Spoofer-IX classified. When we classified the full cone's out-of-cone traffic using the Spoofer-IX method, 92.6% of the traffic was from a provider to a customer across the exchange, carrying 0.5 – 2 Gbps of traffic (figure 12).

Finally, the traffic volume classified as in-cone by the full cone method is larger than that by the Spoofer-IX method. 85.5% of the traffic that the full cone method classified as in-cone was also classified as in-cone by the Spoofer-IX method, with the remaining 14.5% classified as unverifiable by Spoofer-IX. Figure 13 shows how the Spoofer-IX method classified 59.9% of this unverifiable traffic as from a provider to a customer across the IXP, and 26.4% of the unverifiable traffic as out-of-cone for the provider. We hypothesize

that this traffic is classified as in-cone for the full-cone method because some provider ASes (or their customers) provided a BGP view, so the full cone included these addresses as in-cone for these provider ASes (§3.1.3). Note that the traffic profiles in figure 10 and figure 13 are similar: the discrepancy is mostly due to the full cone method classifying some of Spoofer-IX's unverifiable provider-to-customer traffic as out-of-cone (figure 12). However, all routed addresses may be legitimate source addresses in IP packets crossing an IXP from a provider to customer, and no cone of valid addresses can infer the SAV policy of the provider for these packets.

7 DISCUSSION AND INSIGHTS

Challenges of Validation. We could not acquire ground truth data to validate our results, in part due to the negligible amount of out-of-cone traffic we observed, and the challenge of asking any network to validate a small volume of packets. Due to lack of accessible ground truth, we instead verified that our prefix-level customer cone inferences (§3.1.2) were consistent with BGP data extracted from the IXP's route servers. The only inconsistencies we found were due to ASes that had been returned to their RIR and still appeared in public BGP announcements, but did not appear in routes from the IXP route servers.

Generality of the methodology. Assessing the generality of our approach requires applying our method to traffic from other IXPs, which is challenging because it requires the cooperation of other IXP operators. However, we believe our method is generalizable, as we designed and developed Spoofer-IX to accommodate the Best Current Operational Practices (BCOPs) defined by a group of IXPs [28, 37] that describe how IXP operators should configure IXPs. These documents describe how IXP operators should securely configure VLANs and route servers. As such we believe our methodology can be applied to other IXPs; more generally, any other method to infer spoofed traffic in IXP traffic data will have to address the same challenges we encountered.

Applying our method requires two data sets: the traffic data sets themselves, and the metadata that maps IXP infrastructure – VLAN tags on each packet, and MAC addresses to ASes. Our method is automated except for inference of the siblings (§6), which requires some manual effort. However, there are a wide variety of IXP architectures that affect traffic visibility (§3.2), and new IXP architecture innovations to support advanced services will require careful consideration of their impact on our method. Our use of traffic characterization was limited to the packet headers available to us; full payload would enable improvements in traffic analysis, and additional cross-checks.

Emerging IXP trends and their impact on the inference of SAV policy. New IXP services allow networks to self-provision private, on-demand bandwidth in seconds between data center locations (a.k.a, colocation facilities) or cloud service providers, [21, 26, 56, 58, 67]. In 2019, AMS-IX, DE-CIX and LINX joined to develop an API to provision and configure interconnection services at multiple IXPs [55]. The resulting IX-API [5] will allow users to manage their interconnection services, from ordering new ports, to configuring, changing, and canceling services at multiple IXPs. These proposals share a common goal: enable a more dynamic interconnection environment, where networks and IXPs can adapt to changing conditions. They do not propose to change methods to implement the configurations tackled in this paper, but rather create abstractions to facilitate configuration changes.

8 LESSONS LEARNED

The use of IXPs as a focal point for SAV deployment has received recent attention by both the research [45] and policy communities [36, 63, 74]. However, inferring SAV deployment at an IXP is remarkably challenging, more so than has been captured in the literature, due to a combination of operational complexities that characterize today's interconnection ecosystem, and the inherently heuristic nature of topology and traffic inferences on persistently opaque network infrastructure. Many of our discoveries were eye-opening, although not cause for optimism for those interested in infrastructure protection.

First, although we approached this project aware of several methodological challenges for inferring spoofed packets at IXPs, the reality was even more daunting. We recognized the importance of using the semantics of AS relationships, which is conceptually straightforward but even more painstakingly complicated in practice than we expected. We designed, implemented, and applied a method that accounts for both epistemological and operational challenges, and showed how this method reveals inaccuracies in methods that are agnostic to AS relationship semantics.

But we also found epistemological challenges remain. While we infer out-of-cone traffic with our method at our IXP, there are still edge cases we have not yet explained, as some of the traffic appears to have signatures of legitimate traffic. More importantly, we believe further effort is required to understand the degree to which any IXP could be used as a SAV deployment lens. We publicly release our code [62] in hopes that other researchers and IXPs will use it to further improve our collective ability to measure and expand deployment of SAV filtering. Finally, this work illustrates the deep subtleties of scientific assessments of operational Internet infrastructure, which exemplifies the persistent tension between the need for reproducibility of methods and results [7, 8], and the opacity of commercial infrastructure.

ACKNOWLEDGMENTS

We thank the anonymous reviewers and our shepherd, Sergey Gorinsky, for their valuable feedback on our paper. We are also thankful to Leandro Bertholdo, Bruno Lorensi, Cesar Loureiro, Julio Sirota, Milton Kashiwakura, Demi Getschko – all from IX.br – for their support, feedback, and discussions that allowed this work to be possible. We are also thankful to Anja Feldmann and Franziska Lichtblau who helped to improve our work. This material is based in part on research sponsored by the Department of Homeland Security (DHS) Science and Technology Directorate, Homeland Security Advanced Research Projects Agency, Cyber Security Division via contracts D15PC00188 and 140D7018C0010, the National Science Foundation (NSF) via awards OAC-1724853 and OIA-1937165, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant 310408/2017-2, and by CAPES/Brazil via Finance Code 001. The published material represents the position of the authors and not necessarily that of the sponsors.

REFERENCES

[1] Bernhard Ager, Nikolaos Chatzis, Anja Feldmann, Nadi Sarrar, Steve Uhlig, and Walter Willinger. 2012. Anatomy of a Large European IXP. In *ACM SIGCOMM*. 163–174.

[2] AMS-IX. 2019. AMS-IX Partner Program. https://www.ams-ix.net/ams/partners.

[3] AMS-IX. 2019. AMS-IX Private Interconnect Service. https://www.ams-ix.net/ams/service/private-interconnect.

[4] AMS-IX. 2019. Amsterdam Internet Exchange (AMS-IX). https://www.ams-ix.net/.

[5] AMS-IX and DE-CIX and LINX. 2019. IX-API Simplify your IX services. https://ix-api.net/.

[6] APNIC. 2019. Weekly Routing Table Report. http://thyme.apnic.net/current/data-summary

[7] V. Bajpai, O. Bonaventure, k. claffy, and D. Karrenberg. 2019. Encouraging Reproducibility in Scientific Research of the Internet. *Dagstuhl Reports* 8, 10 (Jan 2019), 41–62.

[8] Vaibhav Bajpai, Anna Brunstrom, Anja Feldmann, Wolfgang Kellerer, Aiko Pras, Henning Schulzrinne, Georgios Smaragdakis, Matthias Wählisch, and Klaus Wehrle. 2019. The Dagstuhl Beginners Guide to Reproducibility for Experimental Networking Research. *ACM SIGCOMM Computer Communication Review (CCR)* 49, 1 (Feb. 2019), 24–30.

[9] F. Baker and P. Savola. 2004. Ingress Filtering for Multihomed Networks. RFC 3704 (BCP 84).

[10] S. M. Bellovin. 1989. Security Problems in the TCP/IP Protocol Suite. *ACM SIGCOMM Computer Communication Review (CCR)* 19, 2 (April 1989), 32–48.

[11] Karyn Benson, Alberto Dainotti, kc claffy, Alex C. Snoeren, and Michael Kallitsis. 2015. Leveraging Internet Background Radiation for Opportunistic Network Analysis. In *ACM Internet Measurement Conference (IMC)*. 423–436.

[12] R. Beverly and S. Bauer. 2005. The Spoofer Project: Inferring the Extent of Source Address Filtering on the Internet. In *USENIX Steps to Reducing Unwanted Traffic on the Internet (SRUTI)*.

[13] Robert Beverly, Arthur Berger, Young Hyun, and k claffy. 2009. Understanding the Efficacy of Deployed Internet Source Address Validation Filtering. In *ACM Internet Measurement Conference (IMC)*. 356–369.

[14] Xue Cai, John Heidemann, Balachander Krishnamurthy, and Walter Willinger. 2010. Towards an AS-to-organization Map. In *ACM Internet Measurement Conference (IMC)*. 199–205.

[15] CAIDA. 2019. CAIDA Spoofer Project. https://www.caida.org/projects/spoofer/.

[16] CAIDA. 2019. The CAIDA Internet Topology Data Kit. http://www.caida.org/data/internet-topology-data-kit.

[17] Ignacio Castro, Juan Camilo Cardona, Sergey Gorinsky, and Pierre Francois. 2014. Remote Peering: More Peering Without Internet Flattening. In *ACM Conference on emerging Networking EXperiments and Technologies (CoNEXT)*. 185–198.

[18] Nikolaos Chatzis, Georgios Smaragdakis, Anja Feldmann, and Walter Willinger. 2013. There is More to IXPs than Meets the Eye. *ACM SIGCOMM Computer Communication Review (CCR)* 43, 5 (Oct. 2013), 19–28.

[19] B. Claise. 2004. Cisco Systems NetFlow Services Export Version 9. RFC 3954.

[20] Giovanni Comarela, Gonca Gürsun, and Mark Crovella. 2013. Studying Interdomain Routing over Long Timescales. In *ACM Internet Measurement Conference (IMC)*. 227–234.

[21] Console. 2019. Console - The Cloud Connection Company. https://www.consoleconnect.com/.

[22] M. Cotton, L. Vegoda, Ed. R. Bonica, and B. Haberman. 2013. Special-Purpose IP Address Registries. RFC 6890 (BCP 153). Updated by RFC 8190.

[23] DE-CIX. 2019. DE-CIX Internet Exchange. https://www.de-cix.net/en/.

[24] DE-CIX. 2019. DEC-IX MetroVLAN. https://www.de-cix.net/en/de-cix-service-world/metrovlan.

[25] Z. Duan, X. Yuan, and J. Chandrashekar. 2006. Constructing Inter-Domain Packet Filters to Control IP Spoofing Based on BGP Updates. In *IEEE INFOCOM*. 1–12.

[26] Epsilon. 2019. Epsilon Telecommunications Limited – Connectivity made simple. www.epsilontel.com/.

[27] Euro-IX. 2019. IXP BCOPs (Best Current Operational Practices). https://www.euro-ix.net/en/forixps/set-ixp/ixp-bcops/.

[28] Euro-IX. 2019. IXP BCOPs (Best Current Operational Practices), Technical Recommendations. https://www.euro-ix.net/en/forixps/set-ixp/ixp-bcops/technical-recommendations/.

[29] P. Ferguson and D. Senie. 2000. Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing. RFC 2827 (BCP 38). Updated by RFC 3704.

[30] David Freedman, Brian Foust, Barry Greene, Ben Maddison, Andrei Robachevsky, Job Snijders, and Sander Steffann. 2019. Mutually Agreed Norms for Routing Security (MANRS) Implementation Guide. https://www.ripe.net/publications/docs/ripe-706.

[31] Lixin Gao. 2001. On Inferring Autonomous System Relationships in the Internet. *IEEE/ACM Transactions on Networking* 9, 6 (Dec. 2001).

[32] B. Huffaker, K. Keys, R. Koga, M. Luckie, and kc claffy. 2019. CAIDA inferred AS to organization mapping dataset. https://www.caida.org/data/as-organizations/.

[33] IANA. 2019. Autonomous System (AS) Numbers. https://www.iana.org/assignments/as-numbers/as-numbers.xml.

[34] IANA. 2019. IANA IPv4 Address Space Registry. https://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xhtml.

[35] IANA. 2019. Internet Protocol Version 6 Address Space. https://www.iana.org/assignments/ipv6-address-space/ipv6-address-space.xhtml.

[36] Internet Society. 2019. IXP Participants. https://www.manrs.org/participants/ixp/.

[37] Internet Society. 2019. MANRS IXP Programme. https://www.manrs.org/ixps/.

[38] Internet Society. 2019. Mutually Agreed Norms for Routing Security (MANRS). http://www.manrs.org/manrs.

[39] IX Reach. 2019. IX Reach Remove Peering Services. https://www.ixreach.com/services/remote-peering/.

[40] IX.br. 2019. IX.br – Internet Exchange Brazil. http://ix.br.

[41] IX.br Forum 12. 2019. Remote Peering Panel with DEC-IX, AMS-IX, LINX and IX.br. https://www.youtube.com/watch?v=K283b3AKZ94.

[42] Job Snijders. 2016. Practical everyday BGP filtering: Peer Locking (NANOG67). https://www.youtube.com/watch?v=CSLpWBrHy10.

[43] O. Klaba. 2019. 1.3Tbps DDoS mitigated by our VAC. https://twitter.com/olesovhcom/status/969328679410110466.

[44] S. Kottler. 2019. February 28th DDoS Incident Report. https://githubengineering.com/ddos-incident-report/.

[45] Franziska Lichtblau, Florian Streibelt, Thorben Krüger, Philipp Richter, and Anja Feldmann. 2017. Detection, Classification, and Analysis of Inter-domain Traffic with Spoofed Source IP Addresses. In *ACM Internet Measurement Conference (IMC)*. 86–99.

[46] LINX. 2019. ConneXions at London Internet Exchange Point. https://www.linx.net/join-linx/connexions/.

[47] LINX. 2019. LINX Private VLAN. https://www.linx.net/products-services/private-vlan/.

[48] LINX. 2019. London Internet Exchange (LINX). https://www.linx.net/.

[49] B. Liu, J. Bi, and A. V. Vasilakos. 2014. Toward Incentivizing Anti-Spoofing Deployment. *IEEE ToIFS* (2014), 436–450.

[50] Xin Liu, Ang Li, Xiaowei Yang, and David Wetherall. 2008. Passport: Secure and Adoptable Source Authentication. In *USENIX Symposium on Networked Systems Design and Implementation (NSDI)*. 365–378.

[51] Qasim Lone, Matthew Luckie, Maciej Korczyński, and Michel van Eeten. 2017. Using Loops Observed in Traceroute to Infer the Ability to Spoof. In *Passive and Active Measurement (PAM)*. 229–241.

[52] Matthew Luckie. 2014. Spurious Routes in Public BGP Data. *ACM SIGCOMM Computer Communication Review (CCR)* 44, 3 (July 2014), 14–21.

[53] Matthew Luckie, Robert Beverly, Ryan Koga, Ken Keys, Joshua A. Kroll, and k claffy. 2019. Network Hygiene, Incentives, and Regulation: Deployment of Source Address Validation in the Internet. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*.

[54] Matthew Luckie, Bradley Huffaker, Amogh Dhamdhere, Vasileios Giotsas, and kc claffy. 2013. AS Relationships, Customer Cones, and Validation. In *ACM Internet Measurement Conference (IMC)*. 243–256.

[55] Lynsey Buckingham. 2019. IX-API For the Good of the Internet. https://www.linx.net/ix-api-for-the-good-of-the-internet/.

[56] Pedro Marcos, Marco Chiesa, Lucas Muller, Pradeeban Kathiravelu, Christoph Dietzel, Marco Canini, and Marinho Barcellos. 2018. Dynam-IX: a Dynamic Interconnection eXchange. In *ACM Conference on emerging Networking EXperiments and Technologies (CoNEXT)*.

[57] Alex Marder, Matthew Luckie, Amogh Dhamdhere, Bradley Huffaker, Jonathan Smith, and k claffy. 2018. Pushing the Boundaries with bdrmapIT: Mapping Router Ownership at Internet Scale. In *ACM Internet Measurement Conference (IMC)*.

[58] Megaport. 2019. Megaport - A Better way to connect. https://www.megaport.com.

[59] C. Morales. 2019. NETSCOUT Arbor Confirms 1.7 Tbps DDoS Attack; The Terabit Attack Era Is Upon Us. https://www.arbornetworks.com/blog/asert/netscout-arbor-confirms-1-7-tbps-ddos-attack-terabit-attack-era-upon-us/.

[60] R. Morris. 1985. A Weakness in the 4.2BSD Unix TCP/IP Software Technical Report 117, AT&T Bell Laboratories. (1985).

[61] R. Moskowitz, D. Karrenberg, Y. Rekhter, E. Lear, and G. de Groot. 1996. Address Allocation for Private Internets. RFC 1918 (BCP 5).

[62] L. Muller, M. Luckie, B. Huffaker, kc claffy, and M. Barcellos. 2019. Spoofer-IX sourcecode. https://github.com/spoofer-ix/spoofer-ix.

[63] NIC.br. 2019. Programa por uma Internet mais segura. https://bcp.nic.br/i+seg/.

[64] George Nomikos, Vasileios Kotronis, Pavlos Sermpezis, Petros Gigis, Lefteris Manassakis, Christoph Dietzel, Stavros Konstantaras, Xenofontas Dimitropoulos, and Vasileios Giotsas. 2018. O Peer, Where Art Thou?: Uncovering Remote Peering Interconnections at IXPs. In *ACM Internet Measurement Conference (IMC)*. 265–278.

[65] University of Oregon. 2019. Route Views Project. http://www.routeviews.org/.

[66] P. Phaal and S. Panchen, and N. McKee. 2001. InMon Corporation's sFlow: A Method for Monitoring Traffic in Switched and Routed Networks. RFC 3176.

[67] PacketFabric. 2019. PacketFabric. https://www.packetfabric.com/.

[68] PCH. 2019. Raw Routing Data. https://www.pch.net/resources/Raw_Routing_Data/.
[69] Philipp Richter, Georgios Smaragdakis, David Plonka, and Arthur Berger. 2016. Beyond Counting: New Perspectives on the Active IPv4 Address Space. In *ACM Internet Measurement Conference (IMC)*. 135–149.
[70] RIPE. 2019. Routing Information Service (RIS). http://www.ripe.net/ris/.
[71] T. Scheid. 2016. Defending the Olympics from DDoS. https://blog.apnic.net/2016/10/17/defending-olympics-ddos/.
[72] Team CYMRU. 2019. IPv4 Fullbogons. https://www.team-cymru.org/Services/Bogons/fullbogons-ipv4.txt.
[73] Team CYMRU. 2019. The Bogon Reference. http://www.team-cymru.com/bogon-reference.html.
[74] Tech Accord. 2019. Cybersecurity Tech Accord. https://cybertechaccord.org/.
[75] The Number Resource Organization. 2019. NRO Extended Allocation and Assignment Reports. https://www.nro.net/statistics/.
[76] J. Weil, V. Kuarsingh, C. Donley, C. Liljenstolpe, and M. Azinger. 2013. IANA-Reserved IPv4 Prefix for Shared Address Space. RFC 6598 (BCP 153).
[77] A. Yaar, A. Perrig, and D. Song. 2006. StackPi: New Packet Marking and Filtering Mechanisms for DDoS and IP Spoofing Defense. *IEEE Journal on Selected Areas in Communications (JSAC)* (2006), 1853–1863.