ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2014-ENST-0090

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Informatique »

présentée et soutenue publiquement par

Pierre-Antoine VERVIER

le 19 décembre 2014

Attaques par détournement BGP malveillant : détection, analyse et contre-mesures

Directeur de thèse : **Marc DACIER**

Jury

Michael BAILEY, Professeur associé, University of Illinois at Urbana-Champaign	Rapporteur
Vern PAXSON, Professeur, UC Berkeley	Rapporteur
Michael BEHRINGER, Ingénieur éminent, Cisco Systems Inc.	Examinateur
Ernst BIERSACK, Professeur adjoint, EURECOM	Examinateur
Olivier BONAVENTURE, Professeur, Université catholique de Louvain	Examinateur
Hervé DEBAR, Professeur, Télécom SudParis	Examinateur
Marc DACIER, Chercheur principal, Qatar Computing Research Institute	Directeur de thèse

TELECOM ParisTech
École de l'Institut Télécom - membre de ParisTech

THÈSE

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2014-ENST-0090

EDITE - ED 130

PhD Thesis

submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy of

TELECOM ParisTech

Specialty: Computer Science

Pierre-Antoine VERVIER

Defense date: 19 Decembre 2014

Detection, analysis and mitigation of malicious BGP hijack attacks

Advisor: **Marc DACIER**

Committee in charge:

Michael BAILEY, Associate professor, University of Illinois at Urbana-Champaign	Reviewer
Vern PAXSON, Professor, UC Berkeley	Reviewer
Michael BEHRINGER, Distinguished engineer, Cisco Systems Inc.	Examiner
Ernst BIERSACK, Adjunct professeur, EURECOM	Examiner
Olivier BONAVENTURE, Professor, Université catholique de Louvain	Examiner
Hervé DEBAR, Professor, Télécom SudParis	Examiner
Marc DACIER, Principal scientist, Qatar Computing Research Institute	Advisor

TELECOM ParisTech
École de l'Institut Télécom - member of ParisTech

THÈSE

Abstract

The vulnerability of the Internet inter-domain routing infrastructure against BGP hijacking has gained a lot of attention in the last few years due to several hijack incidents being reported. In 2006, Ramachandran et al. presented evidence of blocks of IP addresses being stolen by BGP hijackers to launch spam campaigns. They coined the expression "BGP spectrum agility" to refer to this threat. Since then, only a very few anecdotal cases have been reported. However, it is a common belief among network operators and ISPs that these attacks could be taking place but, so far, no one has produced evidence to back up that claim. The main goal of this thesis is to determine whether BGP spectrum agility is still, as of today, a problem worth of consideration. If yes, we further aim at rigorously assessing the frequency and the prevalence of these attacks and characterise the attackers' modus operandi.

A wide range of tools have been proposed to help network operators defend against accidental BGP hijacks but they either suffer from high deployment cost or are cluttered with high false-positives rate, which limits their usage to network operators who have the ground truth about their network. The contribution of this thesis is *threefold*.

First, motivated by the lack of tool readily available to study at large scale the phenomenon of malicious BGP hijacks, we propose our own data collection and analysis framework, called SpamTracer. The data collection part consists in collecting a combination of control plane (BGP) data and data plane (traceroute) measurements related networks generating malicious network traffic, such as spam emails. These network traces are further enriched with registration information from IRRs to provide a comprehensive set of features to characterise the routing behavior of the offending networks. For the data analysis part we propose a novel approach to identify and validate possible cases of malicious BGP hijacks from the SpamTracer dataset. The methodology consists of a multi-stage scoring and filtering process, whose results are enriched by means of external data sources and feedback from network operators to validate candidate hijack cases.

Secondly, we applied our methodology on data collected over a period of almost two years and reveal what we believe to be more than 2,000 malicious hijacks, which have taken place on a regular basis over the whole period of the experiment. Some of them were confirmed by the victim network owners and an ISP who was unwittingly involved in several hijack cases.

Thirdly, we unveil a sophisticated *modus operandi* used by cybercriminals to stealthily hijack blocks of IP addresses. Our results show that the identified attacks were rather successful at circumventing BGP hijack and spam mitigating techniques. In the light of these findings, we propose some directions to defend more effectively against this emerging threat and take a final step towards helping mitigate such attacks by leveraging characteristics of the identified hijacks into a real-time blacklist of hijacked IP address blocks.

Finally, this thesis aims at being an eye opener for the community to the fact that frequent, persistent and stealthy BGP hijack attacks have taken place in the Internet for months or even years. We also hope that it will spur new research to understand why these hijacks are taking place and how they can be mitigated.

Résumé

Ces dernières années, la vulnérabilité aux « attaques par détournement BGP » (BGP hijacking) de l'infrastructure de routage dans l'Internet a suscité davantage l'attention, en raison notamment de plusieurs incidents ayant été rapportés. En 2006, Ramachandran et al. ont apporté la preuve que des pirates utilisent ce type d'attaques pour voler des blocs d'adresses IP pour ensuite les utiliser afin d'envoyer du spam. Ils ont appelé ce phénomène « BGP spectrum agility ». Depuis lors, seuls quelques cas anecdotiques de ce phénomène ont été observés. Néanmoins, il est communément admis au sein de la communauté des opérateurs réseau et Fournisseurs d'Accès à Internet (FAIs) que des attaques de ce type ont lieu régulièrement dans l'Internet. Personne ne peut cependant en fournir la preuve. L'objectif principal de cette thèse est de déterminer si, à l'heure actuelle, ce phénomène de « BGP spectrum agility » est réel et constitue une menace pour la sécurité de l'Internet. Si cela se confirme, nous souhaitons également determiner la fréquence ainsi que le prévalence de ce type d'attaques.

Un large éventail d'outils existent afin permettre aux opérateurs réseau de se protéger contre ces attaques mais ils souffrent soit d'un coût de déploiement prohibitif soit d'un taux trop élevé de fausses alertes. La contribution de cette thèse est *triple*.

Premièrement, afin de combler le manque d'outil disponible pour une étude à large échelle des attaques par détournement BGP malveillant, nous proposons un nouveau système de collecte et d'analyse de données, appelé SpamTracer. La collecte d'informations de routage issus du plan de contrôle (BGP) ainsi que de mesures réseaux actives (traceroute) en rapport avec des blocs IP ayant émis du trafic réseau malveillant, par exemple du spam, consistitue la partie *collecte de données* de SpamTracer. La partie *analyse de données* de SpamTracer s'appuie sur un procédé novateur de filtrage et d'évaluation d'anomalies de routage extraites des données collectées afin d'identifier et de valider des cas suspects d'attaques par détournement BGP malveillant.

Deuxièmement, nous avons analysé presque deux ans de données et dévoilons plus de 2.000 attaques par détournement BGP malveillant qui ont eu lieu de façon régulière pendant toute la période de l'expérimentation. Un grand nombre de ces attaques ont été confirmées par des victimes ou par des FAIs impliqués involontairement.

Troisièmement, nous révélons un modus operandi sophistiqué utilisé par les cybercriminels afin de subrepticement prendre le contrôle de blocs d'adresses IP sans l'autorisation de leur propriétaire. Nos résultats montrent que les attaques identifiées ont réussi à mettre en échec des mesures de prévention contre le spam et les attaques par détournement BGP. À la lumière de ces résultats, nous proposons des pistes afin de mieux se protéger contre cette menace émergente. Nous tirons également parti des caractéristiques des attaques observées pour concevoir un système de détection en temps réel de détournements de blocs IP.

Enfin, cette thèse invite la communauté réseau à prendre conscience que des attaques par détournement BGP malveillant ont eu lieu dans l'Internet, et ce de façon récurante et persistante, pendant des mois, voire des années. Nous espérons que ce travail inspirera de nouvelles recherches afin de comprendre mieux encore la motivation des pirates derrières ces attaques ainsi que le moyen de s'en prémunir.

Acknowledgments

A PhD thesis is, unlike what it might seem, not just the result of a one-man job over a few years. In fact, it is be more of a long and arduous journey paved with milestones that would be impossible to complete without the help and support from many people.

First of all, I would like to thank my advisor, Marc Dacier, for having given me the chance to work on such an incredibly interesting topic. His strong expertise in so many areas and his guidance were extremely helpful and contributed to a large extent at making the completion of this thesis possible. I really enjoyed all these long hours spent brainstorming new ideas and discussing technical matters. For having given me his passion to do research I am very grateful to him.

I would also like to thank Olivier Thonnard who provided me with so much technical advice throughout this work. His deep knowledge in security and data analytics areas was key to achieve part of this work. Besides that it was also a pleasure to work as colleagues during all these years.

I also wish to thank Michael Bailey (University of Illinois at Urbana-Champaign) and Vern Paxson (University of California, Berkeley) for having accepted to review this work and for providing me with their extremely insightful comments and feedback. In addition to the reviewers, I would also like to thank the examiners Michael Behringer (Cisco Systems Inc.), Ernst Biersack (Eurecom), Olivier Bonaventure (Université catholique de Louvain) and Hervé Debar (Télécom SudParis) for having accepted to sit in my thesis committee and for having shared their expertise with me via their comments and thoughts on this work.

I also wish to thank Sanjay Sawhney and Matthew Elder from Symantec Research Labs for having sponsored and supported this research work until its very end.

A big thank also goes to all the people I have worked with, during my stay at Symantec.cloud (United Kingdom) and afterwards at Eurecom (France), in the various academic and industrial collaborative partnerships, and in the VIS-SENSE project, who gave me feedback on my ongoing research and allowed for some interesting joint works. In particular, thank you to my colleagues at Symantec and my fellow PhD students at Eurecom.

Of course I could not forget to mention how important friends are when, in good times but mostly in hard times, they are there to relax from work and have fun.

Last but foremost, I would like to thank and express my deepest gratitude to my family and, in particular, my parents who provided me with the moral support and the encouragement needed to achieve this PhD journey. They have always been so patient and understanding. All this would have never been possible without them.

To them and to all the people who played a role in this achievement I dedicate this thesis.

A toutes et à tous, un tout grand merci!

Contents

1 Introduction **1**
- 1.1 Problem statement . . . 2
  - 1.1.1 (P1) Correlation of security-related and routing data . . . 2
  - 1.1.2 (P2) Assessment of the existence of malicious BGP hijacks . . . 2
  - 1.1.3 (P3) If existent, assessment of the prevalence of malicious BGP hijacks and characterisation of the attackers' behavior . . . 3
- 1.2 Research objectives . . . 3
- 1.3 Structure of the thesis . . . 6

2 Background and related work **7**
- 2.1 The facts about malicious BGP hijacks . . . 8
  - 2.1.1 Previous studies . . . 8
  - 2.1.2 Case studies . . . 9
- 2.2 Tracing back malicious activities on the Internet . . . 10
  - 2.2.1 Spam emails . . . 11
  - 2.2.2 Other malicious activities . . . 12
- 2.3 The Internet inter-domain routing infrastructure . . . 12
  - 2.3.1 IP prefixes and AS numbers . . . 12
  - 2.3.2 The Border Gateway Protocol . . . 13
  - 2.3.3 BGP updates and routing policies . . . 14
  - 2.3.4 Security issues . . . 15
  - 2.3.5 Securing BGP . . . 22
  - 2.3.6 Detecting and mitigating IP prefix hijacking . . . 26
  - 2.3.7 BGP monitoring and analysis . . . 31
- 2.4 Conclusion . . . 33

3 SpamTracer **35**
- 3.1 On the identification of malicious BGP hijacks . . . 36
  - 3.1.1 Dataset 1: control plane data . . . 36
  - 3.1.2 Dataset 2: data plane measurements . . . 38
  - 3.1.3 Dataset 3: malicious activities logs . . . 40
  - 3.1.4 Correlation of datasets . . . 40
- 3.2 SpamTracer . . . 40
  - 3.2.1 Routing data collection . . . 47
  - 3.2.2 Multi-stage scoring and data filtering . . . 50
  - 3.2.3 Validation of candidate hijacks . . . 58
  - 3.2.4 Root cause analysis . . . 63
- 3.3 Conclusion . . . 65

4 The malicious BGP hijacks phenomenon **67**
4.1 Routing data collection results . . . 68
4.2 Multi-stage scoring and data filtering results . . . 69
4.3 Validation of candidate hijack results . . . 71
4.3.1 Long-lived hijacks . . . 84
4.3.2 Short-lived hijacks . . . 86
4.4 Root cause analysis results . . . 94
4.5 Summary of findings . . . 96
4.6 Effectiveness of current countermeasures . . . 98
4.6.1 BGP hijack detection . . . 98
4.6.2 BGP hijack prevention . . . 99
4.7 Operationalizing SpamTracer . . . 100
4.7.1 BGP hijack attacks characteristics . . . 101
4.7.2 Real-time blacklist of hijacked networks . . . 102
4.7.3 Real-world deployment . . . 104
4.8 Conclusion . . . 106

5 Conclusion and future challenges **109**
5.1 Research contributions . . . 109
5.2 Future research and perspectives . . . 112
5.2.1 Expanding the scope of the security-related data . . . 112
5.2.2 On the routing data collection . . . 113
5.2.3 On the multi-stage scoring and data filtering . . . 113
5.2.4 On the validation of candidate BGP hijacks . . . 114
5.2.5 Exploitation of results . . . 114
5.2.6 Monitor the IPv6 Internet . . . 114

6 Résumé en français **117**
6.1 Introduction . . . 118
6.1.1 Exposé du problème . . . 119
6.1.2 Objectifs de recherche . . . 120
6.1.3 Structure de la thèse . . . 122
6.2 Positionnement par rapport à l'état de l'art . . . 122
6.2.1 Les attaques par détournement BGP malveillant . . . 122
6.2.2 Le filtrage anti-spam . . . 123
6.2.3 La sécurité du routage inter-domaine dans l'Internet . . . 123
6.3 SpamTracer . . . 124
6.3.1 La collecte de données de routage . . . 125
6.3.2 Processus d'extraction et d'évaluation des anomalies de routage 126
6.3.3 Validation des cas suspects de détournements BGP . . . 128
6.3.4 Analyse de l'origine des détournements BGP . . . 130
6.4 Résultats . . . 131
6.4.1 Résultats : la collecte des données de routage . . . 131
6.4.2 Résultats : processus d'extraction et d'évaluation des anomalies de routage . . . 131
6.4.3 Résultats : validation des cas suspects de détournements BGP 132

6.4.4 Résultats : analyse de l'origine des détournements BGP . . . 139
6.4.5 Efficacité des contres-mesures . . . 140
6.4.6 Opérationnalisation de SpamTracer . . . 143
6.4.7 Enseignements . . . 145
6.5 Conclusion et perspectives futures de recherches . . . 146
6.5.1 Contributions scientifiques . . . 146
6.5.2 Perspectives futures de recherche . . . 148

Bibliography **151**

1

Introduction

The current Internet routing infrastructure is known to be vulnerable to BGP hijacking, which consists in taking control of blocks of IP addresses without any consent of the legitimate owners. This is due to the fact that BGP [150], the de facto inter-domain routing protocol, relies on the concept of trust among interconnected autonomous systems (ASes). Accidental, not necessarily malicious, BGP hijack incidents are known to occur on the Internet. They are generally attributed to misconfigurations. A few cases have received public disclosure on network operational mailing lists, such as NANOG, or blog posts [40, 44, 46, 91]. Techniques to detect these BGP hijacks have been proposed to help network operators monitor their own prefixes to react quickly to such possible outages. These approaches suffer from a very high false-positive rate [96, 116, 157, 194], which is still acceptable to these users since they are only interested in alerts related to the networks they own. Other proposals aim at preventing BGP hijacks [101, 102, 119] but their large-scale adoption and deployment are hindered by their implementation cost.

In 2006, Ramachandran et al. [148] introduced a new phenomenon called "BGP spectrum agility", which consists of spammers advertising for a short period of time (*i.e.,* less than one day) BGP routes to large (*i.e.,* /8) previously unannounced blocks of IP addresses and, subsequently, using the available IP addresses for spamming. Later, some other authors also identified the emission of spam emails coming from hijacked prefixes [70, 96]. Furthermore, complementing the work done in [154], we have described in [47, 178] a special case of hijack in which a couple of IP address blocks were stolen and used to send spam. Most recently, we have also shown in [177], thanks to another real-world case, that correlating routing anomalies with malicious traffic, such as spam, is not sufficient to decisively prove the existence of a malicious BGP hijack.

Besides these sparse cases and despite the apparent desire of some owners to detect whether their own IP address block could ever be stolen, to the best of our knowledge there is no documented evidence that BGP attacks are a threat worth being investigated, since no one has shown that hackers have the possibility to routinely use that modus operandi to commit nefarious activities. If they were capable of it, this would constitute a very serious threat to the Internet since this would enable them not only to send spam emails while defeating the classical IP blacklists

but, more importantly, to run large scale DDoS at almost no cost or run man-in-the-middle attacks against almost any target of their choosing. Therefore, we feel that there is a need to rigorously assess the existence and prevalence of this potential threat.

This raises several unanswered questions that this thesis aims at addressing. Namely, as of 2014, do cybercriminals use BGP hijacking to perform other malicious activities from the stolen IP address blocks? If no, under what assumptions do we conclude that this phenomenon does not exist? If yes, how prevalent are these attacks on the Internet and what is the attackers' modus operandi?

1.1 Problem statement

To answer the previous questions, the three following problems need to be tackled: (i) how to correlate malicious activities with routing information, (ii) how to demonstrate the existence or the absence of malicious BGP hijacks and, (iii) if the phenomenon exists, how to assess its prevalence and characterise the attackers' behavior.

1.1.1 (P1) Correlation of security-related and routing data

In order to be able to correlate malicious activities and routing anomalies on the Internet routing infrastructure, there is a need to collect data about security events, *e.g.,* spam messages, and the state of the routing infrastructure related to the networks emitting the malicious network traffic at the same time. Moreover, both the security-related and routing datasets should be as *representative* as possible to maximise the likelihood of observing malicious activities performed from a hijacked network if they exist. It should also be as *comprehensive* as possible to precisely characterise the malicious activities performed as well as the routing behaviors observed. As already pointed out in several works on BGP hijack detection [59, 165, 194], *distributed vantage points* should be used to gather an as complete as possible view of the Internet topology. The quality of the collected data is thus of critical importance to demonstrate the existence or the absence of malicious BGP hijacks.

1.1.2 (P2) Assessment of the existence of malicious BGP hijacks

In order to demonstrate the existence or the absence of malicious BGP hijacks using the collected data, we first need to extract routing anomalies resulting from known BGP hijack scenarios. However, many routing anomalies due to BGP hijacks can also result from abnormal, though legitimate, BGP practices or from misconfigurations and operational errors [126, 144]. It is thus necessary to *use and extend current BGP hijack detection techniques* [59, 96, 116, 144, 167, 189, 194] to take advantage of all the features of the collected data to identify possible malicious routing behaviors. A solution also has to be found to the problem of the *validation of suspicious hijack cases*, which is a challenge because of the lack of ground-truth data usually only

available from the owner of a victim network. This is mainly due to the fact that inter-AS routing on the Internet is usually governed by routing policies set up and kept private by network owners because they reflect business agreements between them. We finally need to rigorously and scientifically evaluate the assumptions made and the ability of our methodology under these assumptions to detect such attacks from the data collected and the routing anomalies extracted using it.

1.1.3 (P3) If existent, assessment of the prevalence of malicious BGP hijacks and characterisation of the attackers' behavior

If the conjecture turns out to be true, there is a need to assess the prevalence of these attacks to determine the potential effect on current security systems relying on IP reputation and in the Internet routing infrastructure. Upon detection of instances of BGP hijacks performed to support other malicious activities, the attack scenarios should also be leveraged to *refine the data collection process* by adding new discriminative features of the attackers' behavior. It should also be leveraged to *improve detection and mitigation techniques.*

1.2 Research objectives

This work aims at addressing the previously introduced challenges by building an environment for collecting and analysing data for the study of malicious BGP hijack attacks.

> **Claim:** *Despite the fact that several techniques and deployed environments exist for monitoring the inter-domain routing infrastructure and detecting BGP hijack attacks, there is no environment readily usable for the study of BGP hijacks specifically performed to support other malicious activities such as spamming, hosting of phishing or malware distributing websites, or launching DDoS attacks.*

This claim, addressed thoroughly in Chapter 2, stems from three limitations of current approaches to detect and mitigate BGP hijacking attacks on the Internet, which limit their use for solving the problems described here above. This claim can be articulated in the following three points.

1. Most current approaches to monitor the inter-domain routing infrastructure for BGP hijack attacks are solely based on monitoring the control plane [59, 116, 144, 189] and thus suffer from the high similarity between routing anomalies resulting from BGP hijacks and those resulting from benign BGP practices or misconfigurations. These BGP hijack *detection* techniques are thus primarily useful to network operators who want to monitor their own network since they have the ground-truth information about it. Other methods [96, 157, 194] leverage data plane traces to complement BGP-based routing anomaly detection by assessing the impact of BGP routing changes on the forwarding

plane in an effort to reduce the number of false-positive alarms. However, the usage of additional data plane traces does not necessarily result in a reduced, manageable number of false-positive alerts. Moreover, these techniques are not able to deal with the wide range of BGP hijack types that need to be looked at.

2. Alternatively, BGP hijack *prevention* techniques, such as the RPKI framework [120], are developed to bring security into the trust-based BGP ecosystem, usually by means of cryptography-based protocol extensions, but they currently fail to be widely adopted and deployed due to required router software changes and more expensive hardware requirements.

3. In 2006, Ramachandran et al. [148] reported on the observation of "BGP spectrum agility", a phenomenon where spammers hijack, for a very short period of time (*i.e.,* less than a day) large blocks (*i.e.,* /8's) of *unannounced* IP addresses. However both BGP hijack detection and prevention techniques prove to be ineffective against this type of hijack.

In order to address the challenges described here above, this work makes the following assumptions.

> **Assumption 1:** *The first observations of "BGP spectrum agility" described spammers carrying out BGP hijacking attacks to steal blocks of IP addresses and use them to launch spam campaigns in an effort to hinder their traceability and circumvent IP-based blacklisting. In an effort to assess the existence, as of 2014, of "BGP spectrum agility" in the wild Internet, we envision spammers use BGP hijacking.*
> **Assumption 2:** *Hijacking a block of IP addresses requires an attacker to modify the routes taken by data packets so that they reach its physical network instead of the victim's one. Consequently, monitoring both the control and data planes from sufficiently distributed viewpoints should expose such hijack attacks.*

This thesis builds upon these two assumptions to solve the research problems in the following ways.

To address the first **problem (P1)** and based on the first observations of the "BGP spectrum agility" phenomenon, our approach consists in correlating malicious activities and abnormal routing behaviors. Our objective is to design and implement a comprehensive data collection system that will enable us to study the routing behavior of networks performing malicious activities and demonstrate rigorously and scientifically the existence or the absence of malicious BGP hijacks. The system relies on the collection of (i) data plane network traces towards offending networks, (ii) control plane (BGP) data, and (iii) registration information related to these networks extracted from Internet Routing Registry (IRR) databases.

To address the second **problem (P2)**, we anticipate that existing BGP hijack detection techniques can be extended and improved with new algorithms to detect

potential malicious BGP hijack cases. Validating a hijack case based solely on routing changes observed from outside a monitored network can be very challenging due to the lack of ground-truth data. To achieve this goal, we are considering the use of external data sources, such as Internet Routing Registries (IRRs) providing registration information related to IP and AS resources on the Internet as well as feedback from network owners.

To address the third **problem (P3)**, that is if malicious BGP hijacks really occur on the Internet, we want to assess their prevalence both in terms of the amount of malicious activities performed from stolen networks and in terms of the number of networks involved in those attacks. This will determine the impact of these attacks on the Internet routing infrastructure and on the current security systems and, in particular, the ones relying on IP reputation, such as spam sender IP-based blacklists. We also want to take advantage of the heterogenous collected data, *i.e.,* data plane network traces, BGP and registration data, to be able to gain more insights into the attackers' behavior and characterise the modus operandi used to hijack blocks of IP addresses and use the available IP addresses to launch other types of attacks from them.

Thesis: this thesis will demonstrate the following statements.

- Correlating security-related data, in particular spam emails, with routing data related to networks emitting this malicious traffic enables us to uncover candidate hijacked networks. The collected data enriched with registration information from IRRs and direct feedback from network operators enable us to validate the malicious nature of these attacks.
- As of 2014, *BGP spectrum agility* can be observed in the real-world in the form of stealthy and persistent campaigns of malicious BGP hijacks.
- There exists a specific class of agile spammers that are able to hijack routinely, persistently and -quite likely- automatically a large number of blocks of IP addresses to send spam from and host scam websites on the stolen IP address space.
- Malicious BGP hijacks are carried out using a stealthy modus operandi consisting in hijacking previously unadvertised blocks of IP addresses and announcing these blocks in BGP using AS numbers stolen from their legitimate owner to hinder traceability. Such hijacks proves to be rather successful at circumventing traditional BGP hijack and spam protection techniques.

- Some characteristics of uncovered malicious BGP hijack campaigns, such as the AS numbers used by attackers, can successfully be leveraged to detect other attack instances and mitigate, for example by means of a blacklist of hijacked IP address blocks, their effects on the inter-domain routing infrastructure and the security of networks and end hosts on the Internet.

This work has led to the development of SpamTracer, a framework for the study of malicious BGP hijacks on the Internet via the collection and analysis of BGP data and traceroute measurements related to networks sourcing malicious network traffic, in particular spam emails.

1.3 Structure of the thesis

The rest of this thesis is organised as follows. In Chapter 2, we review the **relevant literature** related to past reported cases of malicious BGP hijacks and current technologies leveraging IP reputation for tracing back malicious activities on the Internet. We also provide a comprehensive survey of the numerous techniques developed to detect, mitigate and prevent BGP hijacking and motivate our approach by demonstrating their limitations to address the challenges introduced by this thesis.

Chapter 3 attempts to tackle the research problems **P1** and **P2** by providing the **methodological** contribution of this thesis. We present our environmental setup dubbed SpamTracer that we have built to study malicious BGP hijacks, namely the data collection and analysis processes. First we present our approach to correlating routing- and security-related data (**P1**) by collecting BGP and traceroute measurements related to networks originating malicious network traffic. Second we propose a novel technique to identify, from the collected data, possible cases of malicious BGP hijack attacks (**P2**).

Chapter 4 addresses the research problems **P2** and **P3** by going into the details of the experimental **results** obtained when using the SpamTracer environment for almost two years. First we expose a large corpus of validated real-world malicious BGP hijack attacks, which provides an experimental validation of our approach (**P2**). Second, we offer some insights from the in-depth analysis of the attacks we have found, in particular we reveal the existence of stealthy and persistent campaigns of malicious BGP hijacks. We elaborate on the modus operandi used by the attackers and on the effectiveness of current counter-measures to defeat the attacks we have observed. To close this chapter we explore possible ways to use the uncovered attacks characteristics to help mitigate future instances of such devious attacks (**P3**).

Finally, we **conclude** this thesis in Chapter 5 by presenting our solution to the three research problems and providing an answer to the research questions posed at the beginning of this work. We also present in this final chapter new research challenges and future work perspectives to improve our approach and our understanding of the malicious BGP hijacks phenomenon.

2

Background and related work

Contents

2.1 The facts about malicious BGP hijacks 8
2.1.1 Previous studies 8
2.1.2 Case studies 9
2.2 Tracing back malicious activities on the Internet 10
2.2.1 Spam emails 11
2.2.2 Other malicious activities 12
2.3 The Internet inter-domain routing infrastructure 12
2.3.1 IP prefixes and AS numbers 12
2.3.2 The Border Gateway Protocol 13
2.3.3 BGP updates and routing policies 14
2.3.4 Security issues 15
2.3.5 Securing BGP 22
2.3.6 Detecting and mitigating IP prefix hijacking 26
2.3.7 BGP monitoring and analysis 31
2.4 Conclusion 33

The main contribution of this thesis consists in validating (or invalidating) the conjecture that cybercriminals carry out BGP hijacking attacks at large scale to steal networks from their legitimate owners and use the related network IP addresses to launch other malicious activities from them. In this chapter we provide background information and review previous work related to the three facets of this work, namely (i) previous conjectural and anecdotal reports of BGP hijacking attacks performed by cybercriminals to hinder their traceability and remain stealthy, (ii) the prevalent use of IP addresses to identify and track malicious activities on the Internet, and (iii) the vulnerability of the current Internet inter-domain routing infrastructure to IP prefix hijacking attacks.

2.1 The facts about malicious BGP hijacks

In this section we review previous studies and case studies that have reported (possible) BGP hijack attacks carried out to perform other malicious activities.

2.1.1 Previous studies

Understanding the Network-Level Behavior of Spammers. In 2006, Ramachandran et al. [148] introduced a new phenomenon called "BGP spectrum agility", which consists of spammers advertising for a very short period of time (*i.e.,* less than one day) BGP routes to large (*i.e.,* /8) previously unannounced *hijacked* blocks of IP addresses and using the available IP addresses for spamming.

These observations were made by correlating BGP announcements with spam at a university campus network. During a period of four months, they identified a small but persistent set of spammers *"(1) advertising (in fact, hijacking) large blocks (i.e., /8's) of IP addresses, (2) sending spam from the IP addresses widely distributed throughout the blocks, and (3) withdrawing the route to the blocks shortly after spam is sent"*. Moreover, they observed that the hijacked IP address blocks were allocated though unannounced and unused by the time they were hijacked. They also witnessed unallocated AS numbers (ASN's) in the AS paths for the hijacked IP address blocks suggesting further attempt from the spammers to hinder traceability. Finally, authors claim that, during periods when such spammers were observed, spam from short-lived prefixes accounted for 1% to a maximum of 10% of all spam received at their sinkhole.

Discussion: This work provides the first evidence of the existence of the "BGP spectrum agility" phenomenon and, as such, unveiled a new spammers behavior and brought the security threat it poses into the spotlight. It is also the first work that documents such a phenomenon describing the type of hijack performed and the impact on a popular spam mitigation technique, *i.e.,* spam sender blacklists. It also reports on the prevalence, at that time, of this phenomenon. Unfortunately, authors do not offer details when explaining the reasons why they consider the IP address blocks to be hijacked. It is important to determine whether networks were indeed hijacked because hijacking address blocks hinder traceability of attackers and can lead to misattributing attacks. In fact, receiving spam from short-lived networks does not necessarily imply the networks were hijacked. If IP address blocks were not hijacked, the short-lived announcements might, for example, just be networks experiencing unstable reachability. Unfortunately, without access to the dataset used for the experiments it is difficult to re-examine these findings. Finally, authors acknowledge that "some short-lived announcements may be misconfigurations [126]".

Accurate Real-Time Identification of IP Prefix Hijacking. Hu et al., in [96], in the context of the validation of their prefix hijack detection system, claimed to have correlated identified suspicious hijacks with spam sources from the spam dataset used by Ramachandran et al. in [148].

Discussion: Although this work seems to confirm the first observations of the "BGP

spectrum agility" phenomenon, authors actually do not really provide any confirmation of the hijacks. First, while Ramachandran et al. described hijacks involving previously unannounced IP address space, Hu et al. claims to have identified hijacks (of type 1-2, 4-5 in Section 2.3.4) involving IP address space already announced at the time of the hijack. Thus, their findings do not seem to be consistent with the previous ones. Finally, they do not offer any validation of the hijacks other than reporting suspicious prefixes. Such a validation could, for instance, have been done by checking prefixes and ASes involved in hijacks against Internet Routing Registries (IRRs) or by getting direct feedback from the victim prefixes owner.

An Empirical Study of Behavioral Characteristics of Spammers: Findings and Implications. Recently, Duan et al. performed in [70] a study of the network-level behavior of spammers in an effort to better characterise and identify them. Among other things, they looked at the reachability properties of spam only, versus non-spam only, networks and discovered that a small portion of the spam only networks were only reachable during a short period of time which coincided with the arrival of spam from these networks. In particular, they found in their two months spam dataset that 4% and 10% of spam only networks were only reachable for at most one day and one week, respectively. However, authors also show that non-spam only and mixed, spam and ham networks also, for some of them, exhibit such short-lived characteristic.

Discussion: While these observations might correspond to the "BGP spectrum agility" phenomenon described by Ramachandran et al., there is no conclusive evidence to confirm it or not.

Spamhaus [34] maintains a blacklist called Don't Route or Peer (DROP) which is a subset of one of its other lists, Sender BlockList (SBL), and which consists of "IP address blocks that are hijacked or leased by professional spam or cyber-criminal operation". This list is meant to be used by network operators to either drop all traffic related to these networks or remove them from routing tables. Unfortunately, regardless of the succinct list description, little is known about the exact listing policy used by Spamhaus. Nevertheless, each IP address block listed in the DROP list is associated with a SBL record which can sometimes help uncover the reasons behind the listing of the block.

2.1.2 Case studies

Link Telecom hijack

In [154], Schlamp et al. performed a forensic analysis of a validated case of BGP hijack where a couple of IP address blocks belonging to the Russian telecommunication company LinkTelecom were hijacked for five months and the stolen IP addresses used to perform malicious activities, such as sending spam, hosting web services, scanning remote hosts for vulnerabilities and generating IRC traffic. Using archived BGP routing data, IRR dumps and NetFlow data, they were able to reconstruct the whole story of the hijack and uncover the modus operandi of the attacker.

In April 2011, when the hijack started, Link Telecom (AS31733) had suspended its activity, leaving its network prefixes unannounced. The attacker took advantage of that to hijack and advertise the prefixes via the US ISP Internap (AS12812) and remain unnoticed. In August 2011, network operators at Link Telecom realised their network had been hijacked and started to complain, first to Internap and then on the North American Network Operators' Group (NANOG) mailing list [46]. Shortly after, Link Telecom regained control over their prefixes and the hijack stopped.

Discussion: This validated malicious BGP hijack case study, though isolated, shows that malicious actors have the motive, the means and the opportunity to exploit the Internet routing infrastructure to cover their tracks while performing other malicious activities. However, the Link Telecom hijack is quite different from those described by Ramachandran et al. in [148], namely (i) it lasted five months instead of one day or less and (ii) it involved much smaller IP address blocks, *i.e.,* /16's to /21's instead of /8's. Finally, such an isolated case is not enough to confirm or not the existence of the "BGP spectrum agility" phenomenon as the systematic use of BGP hijacking used by cybercriminals to hinder their traceability and remain stealthy.

Reports on network operational mailing lists

In order to favour and facilitate the communication among network operators on the Internet, the operational community uses public mailing lists, such as the North American Network Operators' Group (NANOG) mailing list [22] or the RIPE Working Groups mailing lists [27]. Such mailing lists are sometimes used by network operators to report BGP hijack incidents, as it was observed in the case of the Link Telecom hijack described here above.

Over the last few years, a few cases of suspect malicious BGP hijacks were reported on the NANOG and RIPE mailing lists [3, 46, 48, 49, 51]. Such reports sometimes provide evidence to conclusively confirm the maliciousness of the attacks, *e.g.,* details about the prefixes and AS numbers involved, the time period of the attacks, logs describing the malicious activities performed from the hijacked IP address blocks. While some of these cases indeed turn out to be suspicious cases of malicious BGP hijacks, only few of them can be validated due to the lack of ground-truth information or appropriate (*e.g.,* historical) data required to perform a forensics investigation. Nevertheless, validated cases can be leveraged to study the modus operandi of attackers, especially how they hijack IP address blocks, and also provide insights into the type of malicious activities performed from the hijacked networks.

2.2 Tracing back malicious activities on the Internet

Attributing, in the sense of *tracking back* or *source tracking*, cyber-attacks in the Internet consists in "determining the identity or location of an attacker or an attacker's intermediary" [182]. A resulting identity may be a person's name, an account, an alias, or similar information associated with a person. A location may include physical (geographic) location, or a virtual location such as an IP address or Ethernet

address. Such action is critical to be able to mitigate or prevent future attacks from the same attacker but also to take legal actions against him. Many security systems on the Internet nowadays leverage *IP addresses* to traceback cyberattacks. The efficiency of these systems thus relies on the correctness of the association between an IP address and the attacker using it, *i.e.,* the IP identity of the attacker. In this context, IP prefix hijacking is a serious threat to IP traceback since it results in an attacker stealing the IP identity of his victim.

In this thesis, our main goal is to study the "BGP spectrum agility" phenomenon. The first observations of the phenomenon by Ramachandran et al. [148] reported suspicious BGP hijacks were used by spammers to launch spam campaigns from the stolen IP space and remain stealthy. Other reports corroborated these observations about spammers behaving this way [96, 142, 143, 154]. Schlamp et al. in [154] showed however that hijacked networks can be used for other malicious purposes than sending spam. Here below, we look at the impact that cybercriminals using "BGP spectrum agility" would have on the ability of current techniques to trace back malicious activities on the Internet. We first consider the impact of "BGP spectrum agility" on the identification and mitigation of spam emails and then expand our scope to other malicious activities.

2.2.1 Spam emails

Spam mitigation techniques can be categorised into two main categories: pre-acceptance and post-acceptance [141]. Pre-acceptance solutions, such as email sender IP-based reputation systems, take advantage of low-level network features to identify spam traffic before it reaches the mail server. Those techniques are usually lightweight so they serve as a first layer of defense in spam filters.

The idea behind IP-based reputation systems is simply to build a database of IP addresses associated with spamming activities. Upon reception of an email, this knowledge base can be queried to help determine if this is spam or not. Relying on an IP reputation database is relatively effective since one just needs to accept, delay or reject emails from already known bad hosts.

Spam sender IP blacklisting [11, 21, 33, 34, 42, 87, 149, 161] is probably the most popular spam mitigation technique. Despite the cost of maintaining those blacklists, their incompleteness [148, 149] and possible inaccuracies [160, 161], these are still heavily used in commercial spam filters [10, 35, 39]. On the other hand post-acceptance techniques leverage features extracted from the content of email messages, such as text patterns or embedded URLs. Those techniques are usually efficient at identifying spam but require heavier processing. In a typical spam filter deployment, lightweight, pre-acceptance, techniques will be used to filter out obvious spam while more expansive, post-acceptance, solutions will be used to discard the remainder [35, 39].

2.2.2 Other malicious activities

Similar to spam sender IP blacklists, other malicious activities can be traced back using the IP address of malicious hosts involved. Malicious domain names are often mitigated using blacklists, some of which are provided as a list of IP addresses to which the blacklisted domains resolve to [127]. Additional blacklists contains IP address of hosts involved in various cyber-attacks: botnet C&C servers [1, 32], firewall alerts [13] or networks under the control of cybercriminals [41]. Finally, some of these blacklists are provided so that they can be directly used in security systems such as firewalls or IDS's [14].

2.3 The Internet inter-domain routing infrastructure

2.3.1 IP prefixes and AS numbers

The content of this section is inspired from [85, 100]. IP addresses are either 32-bit (IPv4) or 128-bit (IPv6) words that represent unique identifiers for hosts to communicate on the Internet. IPv4 addresses are usually represented using the dotted-decimal notation where the decimal value of each of the four bytes is separated by a dot, *e.g.,* 74.125.239.33. IPv6 addresses are longer, hence they are commonly represented as eight groups of four hexadecimal digits separated by colons, *e.g.,* 2607:f8b0:4010:0801:0000:0000:0000:1008. IP addresses are allocated or assigned to organisations in blocks of contiguous addresses commonly referred to as *IP prefixes* and represented using the Classless Inter-Domain Routing (CIDR) notation [74]. For example, the IPv4 prefix 74.125.239.0/24 belongs to Google Inc. and designates a block of 256 IPv4 addresses starting at 74.125.239.0 and ending at 74.125.239.255. Similarly, the IPv6 prefix 2607:f8b0::/32 (groups of 0's and leading 0's can be omitted) belongs to Google Inc. and designates a block of 2^{32} IPv6 addresses starting at 2607:f8b0:: and ending at 2607:f8b0:ffff:ffff:ffff:ffff:ffff:ffff. Considering IP addresses in blocks mainly aims at preventing the inflation of the size of routing tables by reducing the number of routes to individual destination IP addresses [74, 99]. CIDR IP prefixes also allows routing tables to be further shrunk by aggregating IP prefixes when possible, *e.g.,* an ISP being allocated a /16 IPv4 prefix can assign or sub-allocate several sub-prefixes to its customers and only announce the aggregated /16 prefix in BGP.

Each Autonomous System (AS) in the Internet is assigned an AS number (ASN) for use as a unique identifier in the inter-domain routing infrastructure. Initially coded as 16-bit integers, AS numbers are 32-bit integers since 2007 [181]. Among the 2^{32} possible ASNs, AS0, AS65535 and AS4294967295 are reserved for special purposes by the IANA. Moreover, AS64512-65534 and AS4200000000-4294967294 are reserved for private ASes [131] that are willing to participate in inter-domain routing but do not require public visibility of their AS in the global Internet routing infrastructure, usually because they are only connected to the Internet via one or more Internet Service Providers (ISPs) acting as proxies for the routing process.

Finally, AS23456 ("AS_TRANS") is reserved for interoperability between routers supporting only 16-bit ASNs and those supporting both 16- and 32-bit ASNs.

Originally, IP address blocks used to be allocated or assigned to organisations directly by the IANA. In 1997, the Internet Assigned Numbers Authority (IANA) delegated the responsibility of registration and administration of IP addresses and AS numbers to five Regional Internet Registries (RIRs). The *Réseaux IP européens* (RIPE) is responsible for the allocations and assignments in Europe and Middle-East. The *American Registry for Internet Numbers* (ARIN) handles IP and AS resource registrations in North America. The *Asia Pacific Network Information Center* (APNIC) manages IP and AS resources in the Asian and Pacific regions. The *Latin America and Caribbean Network Information Center* (LACNIC) and the *African Network Information Center* (AfriNIC) are responsible for the Latin and South American region, and Africa, respectively. Each RIR is allocated IP address space and AS number ranges that can be further allocated to Local Internet Registries (LIRs) or directly assigned to end-user organisations. LIRs are usually ISPs or academic institutions who receive IP and AS allocations from RIRs for further assignment to customers. Overall, the allocation and assignment of IP addresses and AS numbers is organised in a hierarchical fashion with the IANA acting as the root.

Finally, *bogon* or *martian* IP prefixes are special purpose IP address blocks reserved by the IANA and are thus not supposed to be routed in the global Internet [104, 103]. For instance, the IPv4 prefix 192.168.0.0/16 or the IPv6 prefix fc00::/7 are meant to be exclusively used internally in networks [90, 151]. Additional prefixes can be assimilated to bogons [38] and correspond to IP address space not allocated by the IANA to any Regional Internet Registry (RIR) or not allocated/assigned by any RIR to an LIR or end-user organisation. Like bogons these addresses should not appear in the global Internet routing infrastructure.

2.3.2 The Border Gateway Protocol

The content of this section is inspired from [100, 174]. The Internet is comprised of tens of thousands of smaller interconnected networks called Autonomous Systems (ASes), each of them belonging to a single administrative entity (*e.g.,* an Internet Service Provider, a company, a university). An AS is basically a network of end hosts, such as servers and computers, and routers responsible for forwarding network traffic to the appropriate destination inside or outside the AS. Consequently, each AS holds and manages a set of IP addresses for use in communications internal to the AS as well as IP addresses to communicate with the outside world. IP-routing is the process in networks that consists for a network entity (*e.g.,* AS, router, end host) in advertising to others its IP address(es) and receive the routes to reach the other entities in the network. In the Internet, routing is organised in a two-layer hierarchy: intra-AS, or intra-domain routing is achieved using an Interior Gateway Protocol (IGP) such as RIP, OSPF, IS-IS, IGRP or EIGRP and allows end hosts within an AS to communicate with each other. Inter-AS, or inter-domain routing uses an Exterior Gateway Protocol (EGP) such as the Border Gateway Protocol (BGP) to allow ASes to advertise to others the IP addresses of their network and

receive the routes to reach the other ASes. The main reason why intra- and inter-AS routing are dissociated and achieved using different protocols is to (i) help routing protocols scale with the size and complexity of the networks they are applied on by using ASes as an abstraction layer and (ii) to allow network owners to control the information they disclose to the outside world on their relationships with other ASes as well as the internal network configuration within their AS.

The security of intra-AS routing represents a very interesting area of research as some of the protocols in use are impacted by vulnerabilities that allow attackers to tamper with the routing process and, for instance, influence the routes taken by data packets within an AS [106, 132]. However, to hijack a block of IP addresses and use them to perform other malicious activities, an attacker needs to manipulate how data packets flow in the global Internet infrastructure, *i.e.,* between ASes, hence we focus on the study of inter-AS routing security.

The Border Gateway Protocol (BGP) is the de-facto inter-domain routing protocol on the Internet. The version of BGP currently in use is version 4 and is specified in RFC 4271 [150]. BGP is a decentralised distance vector routing protocol which allows ASes to exchange network IP addresses reachability information with each other.

2.3.3 BGP updates and routing policies

The content of this section is inspired from [100, 174]. Inter-domain routing is primarily driven by *routing policies* defined at ASes allowing them to translate inter-AS business relationships and contractual agreements into routing decisions, such as accepting or not routes, advertising or not routes, and even modifying route attributes. Routing policies eventually aim at influencing how data packets flow between ASes. For example, a multi-homed customer will generally not advertise prefixes of one of his ISP to the other ISP to avoid acting as a transit AS between the two ISPs, wasting bandwidth for traffic unrelated to his network.

A BGP-speaking router establishes BGP sessions with *peer* BGP speakers in other ASes. BGP speakers within the same AS can also peer with each other to exchange routing information learned from other ASes. BGP speakers build their routing table incrementally by advertising to and receiving from their peers network reachability information related to their own network as well as other networks on the Internet. During a BGP session routers essentially exchange BGP update messages allowing them to *advertise* or *withdraw* routes to destination IP prefixes.

A BGP route advertisement carries (i) *Network Level Reachability Information (NLRI)*, namely the destination network IP prefixes advertised by this update message and, (ii) a set of BGP attributes used by routers to select the *best* route among all routes received towards a given destination network. BGP attributes mainly include (ii.a) a *local preference* set only for BGP speakers within the same AS to indicate a preferred point of exit from the AS, (ii.b) the *AS path* corresponding to the sequence of ASes already traversed by the BGP advertisement message used to prevent routing loops, (ii.c) the *origin* indicating where the route was learned from,

e.g., BGP or an IGP, (ii.d) a *multi-exit discriminator (MED)* set for BGP speakers outside the AS to indicate a preferred point of entry in the AS and, (ii.e) *communities* which provide a way to group together different routes based on the routing policies that apply to them, *i.e.,* the routing action that should be taken at the receiving router. In theory, a router should select the route with the lowest local preference, the smallest AS path, the lowest origin and the highest MED in that order. Each time a router propagates a BGP advertisement, it *appends* its AS number to the AS path. In practice, routing policies may influence the final decision. The first AS in the AS path is commonly called the *origin AS* and subsequent ASes are referred to as the *upstream ASes*. Because IP prefixes can be of different length, it is common to observe some prefixes overlapping with others in BGP advertisements. In this case, a router will select the route to the longest (more specific) prefix over the routes to the shorter (less specific) ones.

A BGP route withdrawal contains only the destination network IP prefixes (NLRI) whose routes should not exist anymore and should be removed from the routing table.

2.3.4 Security issues

BGP faces two classes of security issues [65] that stem from (i) its reliance on other protocols, such as TCP, over which BGP sessions are run and (ii) the trust-based exchange of reachability information between interconnected ASes making the protocol vulnerable to IP prefix hijacking attacks. IETF's Operational Security Capabilities for IP Network Infrastructure (OPSEC) working group, in [60], presents general threats on routing protocols and proposes a collection of Best Current Practices (BCP's) for BGP operations and security in [72].

Attacks on BGP sessions

Attacks on BGP sessions consist for the attacker in tampering with the exchange of BGP messages between two BGP-speaking routers out of the attacker's control. To exchange routing information BGP-speaking routers establish BGP sessions over TCP on port 179 with peer routers [150]. While TCP provides BGP with a reliable data channel it also introduces security issues which can impact BGP operations in a similar way it can impact any application-layer protocol relying on TCP.

(i) TCP data is transferred in clear text which means that anyone sitting in between two communicating BGP speakers can eavesdrop on the messages exchanged and gain information about routing policies between ASes. (ii) An adversary can also inject new packets in, modify or remove packets from, an existing TCP connection thus affecting any BGP session running on top of it and eventually the routing state and stability of routers and ASes. Such attacks on the TCP connection can result in a denial of service (DoS) of routers failing to communicate with each other anymore [65]. It can also lead to routers selecting specific routes chosen by attackers [65]. For instance, one can drop all packets from a given router in a BGP session leaving one to believe the other router went down. The former will remove all routing table entries received from the latter. When repeated, this attack can trigger the *route*

flap damping [180] BGP mechanism which removes all flapping routes from a peer in an effort to preserve routing stability.

Solutions proposed to mitigate attacks on BGP sessions aim at either providing BGP with an underlying secure data channel, for example by relying on IPsec [111] or extending the protocol with, usually lightweight, security measures [58, 81, 89, 173].

IPsec [111] is a suite of protocols providing security at the network layer. If used between two routers, it can protect all BGP sessions against attacks on the confidentially, authenticity and integrity of BGP messages. However the use of IPsec requires major changes to routers software and hardware which has currently hindered its deployment.

The *Generalized TTL security mechanism (GTSM)* (or "BGP security hack") [81] relies on the premise that most BGP sessions occur between IP-level adjacent routers to prevent a remote attacker to set up or tamper with BGP sessions by automatically dropping packets having the IPv4 Time To Live (TLL) or IPv6 Hop Limit lower than a given threshold. For instance, given the maximum IP TTL value is 255, dropping packets with a TTL lower than 254 enforces every peer to be directly connected. Unfortunately, this mechanism becomes weaker when multi-hop BGP sessions are used. GTSM is actually deployed and used by network operators [102]. It has also been proposed to compute a Message Authentication Code (MAC) of the TCP header and BGP packet and distribute it via the *TCP MD5 signature option* [89] or the more recent and generalized *TCP authentication option (AO)* [173] to bring authentication and integrity in BGP sessions. Finally, a mechanism like *unicast Reverse Path Forwarding* [58] which consists in discarding packets for which no route exists to the source IP address can also be used to mitigate (D)DoS attacks using packets with spoofed source IP addresses. Though it is hard to assess the prevalence of these mechanisms, they are available for network operators to use in their routers.

IP prefix hijacking

In BGP, each AS implicitly trusts the peer ASes it exchanges routing information with. *IP prefix hijacking* (or BGP hijacking) is an attack against the Border Gateway Protocol that consists in injecting, thanks to a participating genuine BGP router, erroneous reachability information into the routing infrastructure in order to take control of blocks of IP addresses owned by a given organisation without its authorisation. This attack is possible because there is no widely deployed security mechanism in BGP that allows a router to verify (i) that an AS originating a prefix is entitled to (Origin Validation) and (ii) that a router belongs to the AS it claims to (Path Validation). This enables the attacker to perform other malicious activities, such as spamming, phishing, malware hosting, using hijacked IP addresses belonging to somebody else. It also allows an attacker to perform a denial-of-service attack on the victim network or intercept traffic destined to it. To perform IP prefix hijacking an attacker needs to control a router, either by hacking into an existing one [142] or by setting up a BGP peering relationship with an ISP or another network already connected to the Internet. The impacts of erroneous BGP announcements on the routing

infrastructure varies (i) with the content of the announcements [65, 96, 134], including the IP prefix advertised and the AS path, and (ii) with the location of the misbehaving router and of the victim network in the Internet topology [117, 118, 190, 191]. Next we propose a new classification of the different BGP hijack types inspired from previous work [65, 96, 134] and based on the following attack properties:

- **Method.** The different IP prefix hijacking methods are characterised by the IP prefix that is advertised by the attacker and the AS path [96].
- **Objective.** The objective behind prefix hijacking depends on the attacker's goal (*e.g.,* DoS a network, impersonate a bank's website, send spam using stolen IP addresses, etc). Three objectives are usually considered [134], which are (i) blackholing, (ii) impersonating the victim network, and (iii) intercepting traffic related to the hijacked prefix.
  Blackholing consists in attracting and then dropping all traffic related to a network making it unreachable. This is equivalent to a denial of service (DoS) attack.
  Impersonating by hijacking consists for an attacker in taking control over a network prefix either to replace services hosted on the victim network by others under the attacker's control or to use the available IP addresses to perform, possibly malicious, network activities on behalf of the victim network owner. This is equivalent to an IP identity theft.
  Intercepting by hijacking consists in attracting traffic related to a network in order to eavesdrop or tamper with network packets, and then forwarding it back to the owner of the victim network. This is equivalent to a Man-In-The-Middle (MITM) attack [139].
- **Effectiveness.** We define the effectiveness of an IP prefix hijack as the likelihood of the attack to pollute the entire Internet, *i.e.,* to deceive the maximum number of ASes into selecting the bogus route to the prefix instead of the genuine one.
- **Disruptions.** A BGP hijack attack on IP address space that is already *advertised* and *used* can affect the communications between the victim network and other networks on the Internet (*e.g.,* cause a denial of service (DoS) of services hosted on the victim network, etc). We consider here the potential maximum disruptions each BGP hijack type can cause.

Table 2.1 and Figure 2.1 respectively summarises and illustrates the different IP prefix hijack types based on the method employed, the attacker's objective, the hijacking attack effectiveness and the potential disruptions caused. Here below we provide a more detailed description of the different IP prefix hijack types along with the method employed, the objective of the attacker and, the effectiveness of and potential disruptions caused by the attack.

Type 1: Same prefix, new origin AS: In this scenario, an attacker *Mallory* originates an IP prefix p (66.102.0.0/20 in Figure 2.1) using its own AS ({AS_Mallory}

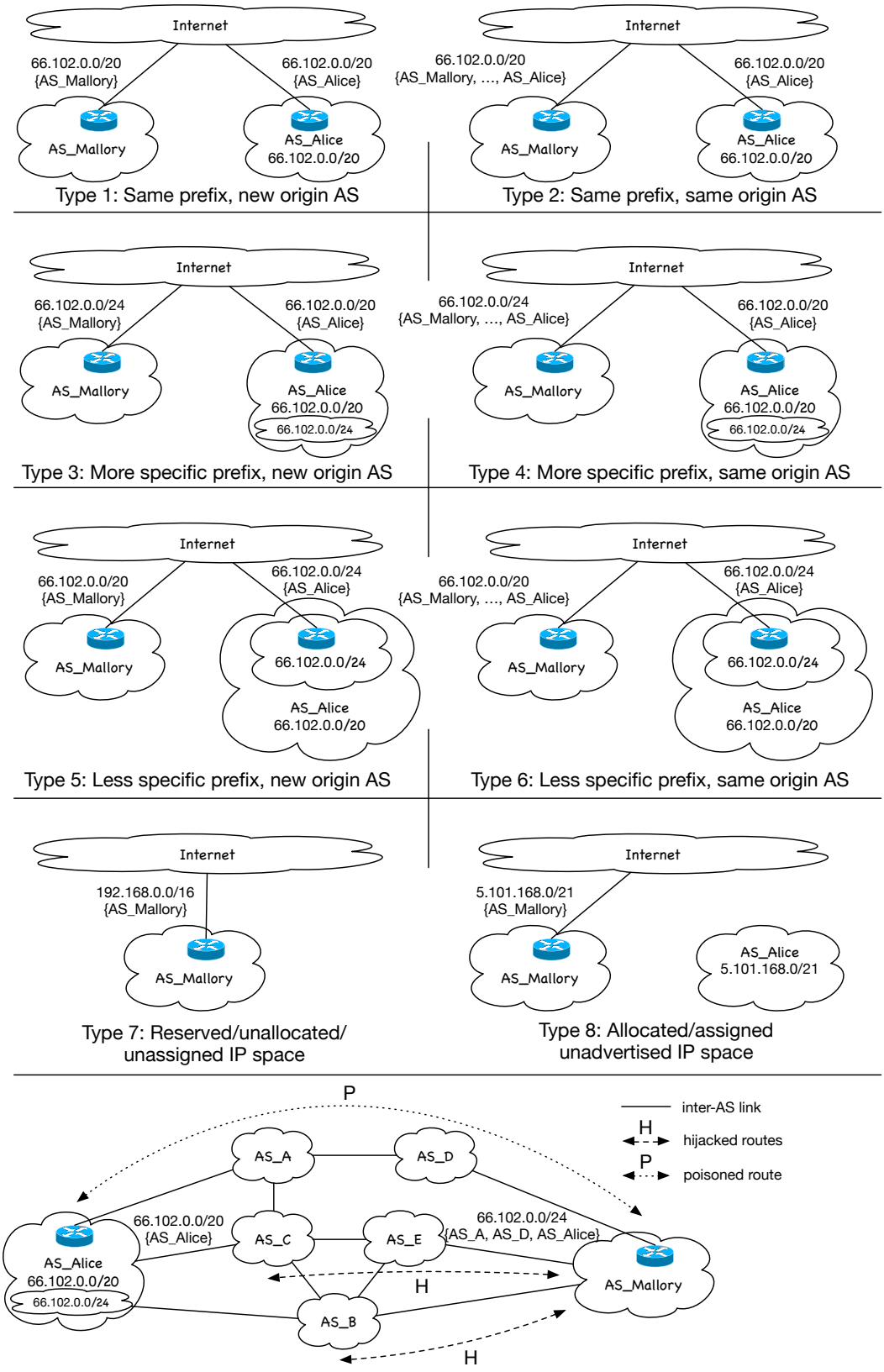


Figure 2.1 – IP prefix hijack types

Method		Objective			Effectiveness	Disruptions
		Black.	Impers.	Interc.		
T1:	Same prefix New origin AS	✓	✓		Low	High
T2:	Same prefix Same origin AS	✓	✓		Low	High
T3:	More specific prefix New origin AS	✓	✓		High	High
T4:	More specific prefix Same origin AS	✓	✓		High	High
T5:	Less specific prefix New origin AS	✓	✓		Low	Low
T6:	Less specific prefix Same origin AS	✓	✓		Low	Low
T7:	Reserved/unallocated/ unassigned IP space		✓		Low	Low
T8:	Allocated/assigned Unadvertised IP space		✓		High	Low
T9:	Man-In-The-Middle			✓	High	Low

Table 2.1 – IP prefix hijack types: the attack method, objective, effectiveness and potential disruptions caused.

in Figure 2.1) while p is already advertised by its legitimate owner Alice with its AS ({AS_Alice} in Figure 2.1). The IP prefix p is then originated by multiple ASes, creating a so-called Multiple-Origin AS (MOAS) conflict [192]. The attacker reveals its own AS making it easier to trace it back. By advertising the same prefix, the attacker's hijack will likely not pollute the whole Internet, since some ASes will choose the correct route while others will be deceived into selecting the erroneous one. If the victim network is used, it is likely to observe large disruptions due to all traffic from polluted ASes being diverted to the attacker's own network infrastructure, and being dropped in case of blackholing. The attacker might impersonate some of the victim network hosted services reducing the disruptions caused by the attack.

Type 2: Same prefix, same origin AS: This hijack type differs from *type 1* by the fact that *Mallory* does not originate p (66.102.0.0/20 in Figure 2.1) using its own AS but instead *forges* part of the AS path by prepending ASN's, eventually including *her*'s and *Alice*'s to the AS path ({AS_Mallory,...,AS_Alice} in Figure 2.1). We refer to this practice as *AS path forgery*. This deceives other ASes receiving such a BGP announcement into believing the announcement has been propagated by all the ASes in the AS path, especially *Alice* who is the expected origin for the BGP announcements. Such a trick is possible because there is no way for other ASes to verify that each AS in a BGP announcement AS path indeed propagated the announcement. While hindering attacker's traceability, this hijacking method is similar to *type 1* with respect to its effectiveness and the potential disruptions caused.

Type 3: More specific prefix, new origin AS: This case is similar to *type 1* except that *Mallory* advertises an IP prefix p_m (66.102.0.0/24 in Figure 2.1) which is more specific than p (66.102.0.0/20 in Figure 2.1), still using its own AS ({AS_Mallory} in Figure 2.1) as the origin. This method is much more effective that *type 1* because BGP routers will usually select the route to the most specific prefix. Hence, such hijack will cause all traffic from a large portion of (if not the entire) Internet to go towards the attacker's network. If that subnet is used by the victim then the hijack will create large disruptions, but if it is unused the attack will not cause any disruption.

Type 4: More specific prefix, same origin AS: This method combines *type 3* and AS path forgery. While hindering attacker's traceability, its effectiveness and the potential disruptions caused are similar to the hijack *type 3*.

Type 5: Less specific prefix, new origin AS: In this method, *Mallory* advertises a prefix p_l (66.102.0.0/20 in Figure 2.1) which is less specific than p (66.102.0.0/24 in Figure 2.1) using its own AS ({AS_Mallory} in Figure 2.1). In the example in Figure 2.1 *Alice* owns p_l which includes p but she only advertises p. However, p_l and p are not required to have the same owner and p_l can include multiple subnets having different owners. This method builds upon the premise that advertising the less specific prefix p_l does not affect the routing state of p due to the longest prefix matching rule. Instead, it will route all IP space in p_l not covered by p to *Mallory*. While such a hijack is not guaranteed to equally pollute the whole Internet, due for instance to aggregation of prefixes at some ASes, it is unlikely to produce disruptions to the victim network since the hijacked IP address space was not advertised.

Type 6: Less specific prefix, same origin AS: This method combines *type 5* and AS path forgery. While hindering attacker's traceability, its effectiveness and the potential disruptions caused are similar to the hijack *type 5*.

Type 7: Reserved/unallocated/unassigned IP space: In this scenario, *Mallory* advertises a route in BGP towards IP address space (192.168.0.0/16 in Figure 2.1) that is reserved or that has not been allocated or assigned to any ISP or end-user organisation. Such a hijack can turn out to be very ineffective when BGP advertisements for such IP address space are filtered, for instance when using bogon filters [38]. It is also unlikely to create any disruption since the hijacked IP address space is not supposed to be advertised and used as publicly routable IP address space. Nevertheless, if such IP address space is used internally by an AS that does not filter out incoming announcements related to this space then a external hijack will affect internal routes, likely causing disruptions.

Type 8: Allocated/assigned but unadvertised IP space: This method consists in advertising a route in BGP towards a block of IP addresses (5.101.168.0/21 in Figure 2.1) that has been allocated or assigned but which is not currently publicly advertised by its legitimate owner. This hijacking method is by far the stealthiest and most effective way for an attacker to take full control over a

block of IP addresses it does not own. This type of hijack can be referred to as *IP squatting*. Similarly to *type 7*, if such IP address space is used internally by an AS that does not filter out incoming announcements related to this space then a external hijack will affect internal routes, likely causing disruptions.

Type 9: Man-In-The-Middle: The BGP MITM attack [59, 139] is a very stealthy attack allowing to intercept traffic destined to a network IP prefix, for the purpose of eavesdropping or tampering with the communications related to this prefix. In the scenario described in [139], it is carried out by an attacker *Mallory* by (i) advertising a prefix p_m (66.102.0.0/24 in Figure 2.1) which is more specific than the victim *Alice*'s prefix p (66.102.0.0/20 in Figure 2.1), (ii) prepending the AS path with the ASes along a path from AS_Mallory to AS_Alice ({AS_A,AS_D,AS_Alice} in Figure 2.1). Step (i) allows *Mallory* to hijack and intercept traffic for p_m ($\xrightarrow{H}$ in Figure 2.1) and step (ii) allows her to forward this traffic back to *Alice* by securing a path to her using *AS path poisoning* ($\xrightarrow{P}$ in Figure 2.1). AS path poisoning builds upon the premise that routers observing their AS number in the AS path of a BGP announcement will discard it to prevent routing loops. It allows *Mallory* to pollute all ASes but those along the path from her to *Alice*. Performed as is, this hijack is very effective and will unlikely cause any disruptions.

The main goal of this thesis is to study the "BGP spectrum agility" phenomenon. In this context, it is noteworthy that suspect hijacks reported in the first observations of the phenomenon by Ramachandran et al. [148] corresponded to the hijack type 8 where allocated/assigned though unadvertised IP address space was hijacked. This hijack method was also described as a method of choice for cybercriminals to use BGP hijacking as a way to perform other malicious activities from stolen IP space [93, 133, 142]. Moreover, from our classification, hijack type 8 appears to be the most effective and least disruptive method for cybercriminals take control over IP address space for further use in other malicious activities. It it thus particularly important to consider this hijack type in our methodology to study the "BGP spectrum agility" phenomenon. Nevertheless, since the phenomenon remains hardly documented, it is also important to consider other types of hijack (1-7 and 9).

BGP hijack incidents are known to occur on the Internet and while a few get public disclosure on mailing lists, such as NANOG [44], or blog posts [40] these are generally attributed to misconfigurations.

IP prefix hijacking defense proposals are twofold: (i) some techniques aim at *securing BGP* by providing the protocol with new mechanisms to prevent hijacking and (ii) other *hijack mitigation* techniques aim at monitoring the Internet routing infrastructure and trigger alarms upon abnormal routing changes. Unlike hijack prevention techniques, mitigation methods usually require no changes to router software which makes them readily and easily deployable.

2.3.5 Securing BGP

Some techniques have been proposed to bring security into BGP [102] and prevent IP prefix hijacking attacks, most of them using cryptography to sign some elements of BGP updates and ensure routing information authenticity and integrity. Here below we review some techniques proposed by actors in the operational and research routing communities.

Secure-BGP (S-BGP) [110] is an approach which aims at securing *route origination* and *route propagation*. It relies on a Public Key Infrastructure (PKI) which bears the hierarchical structure of the IP address and AS number delegation system and which provides certificates binding IP prefixes, AS numbers and routers to the public key of their respective owner. It introduces two types of so-called attestations, namely *address attestations* and *route attestations*, which correspond to signed datum that binds respectively IP prefixes to the authorised origin AS(es) and the source AS to destination ASes of a BGP update. Address attestations implement secured route origination. Secured route propagation is achieved by nesting route attestations each time a BGP update is propagated. A router can verify the validity of a BGP update by checking all attestations against the PKI. Address attestations are not stored in the PKI but instead are distributed to routers out-of-band and route attestations are embedded in BGP updates. Its deployment was substantially hindered by the required update of routers hardware to support S-BGP functionalities [102]. S-BGP is currently not being deployed.

Secure Origin BGP (soBGP) [183] similarly aims at providing secured *route origination* and *route propagation*. It uses three types of certificates: (i) an *Entity Certificate* binds an AS to its owner's public key, (ii) an *Authorization Certificate* binds an IP prefix to an authorized origin AS, and (iii) a *Policy Certificate* ties a (set of) prefix(es) or AS(es) to routing policies its owner wants to enforce. Policy certificates are exchanged between ASes and allows each of them to actually build a topology of its view of the network. Route origination is secured through authorization and prefix policy certificates and route propagation is weakly secured via AS policy certificates. A limitation of soBGP is that it does not protect against forged AS paths compliant with routing policies. Unlike S-BGP where the certification structure follows the IP address and AS number delegation hierarchy, soBGP relies on a web-of-trust bootstrapped using a few root-certificates from trusted entities to build the PKI. soBGP certificates are also conveyed using a new "BGP security message". soBGP is currently not being deployed.

Pretty Secure BGP (ps-BGP) [137] combines a classic PKI with an AS-based reputation system to provide secured *route origination* and *route propagation*. In ps-BGP, the PKI is used to map ASes to their owner's public key and its structure follows the AS number delegation chain. Each AS statically associates a degree of confidence in the routing information trustworthiness of all other ASes. Moreover, it builds and distribute via BGP a Prefix Assertion List (PAL) containing its and possibly some of its peers' asserted prefix ownership, subsequently used by other ASes

to build a prefix to origin AS mapping and to infer an inter-AS topology. Route origination is secured by checking the consistency of BGP announcements with the PALs of the origin AS's peers. The AS-based reputation weakens the security of ps-BGP when facing colluding adversary ASes though [65]. Secured route propagation is implemented in the same way as S-BGP by having routers signing propagated BGP updates and using the PKI to validate the sequence of nested signatures. The AS-based reputation allows partial verification of AS paths and route origins which help reduce the computational cost of the protocol. ps-BGP is currently not being deployed.

Interdomain Route Validation (IRV) [83] builds upon the premise that each AS on the Internet is authoritative to validate routing information pertaining to its network, such as the IP prefixes it is allowed to originate or its BGP neighbours. IRV proposes to set up *IRV servers* in each AS. *Route origination* and *route propagation* can be secured by having routers, upon reception of a BGP update, contact the IRV server at each AS along the path to validate the origin AS of the advertised prefixes and the AS adjacencies. The main advantage of IRV is that it does not require major changes to router software and hardware due to the validation process being carried out by IRV servers. The limitations of the protocol include (i) its dependency on a secured {AS,IRV server} mapping service and (ii) its dependency on existing routing to IRV servers. IRV is currently not being deployed.

Secure Path Vector (SPV) [97] is another proposal to secure BGP *route origination* and *route propagation*. It proposes to secure route origination by leveraging a PKI bearing the hierarchical structure of the IP address delegation system. The PKI allows to produce certificates binding IP prefixes to their owner's public key. SPV does not specify any certificate distribution mechanism though. SPV secures route propagation by having a router in the AS originating a prefix to generate one-time signatures which are propagated with the BGP updates to allow routers along the path to validate and sign the AS path. It aims at preventing adding, removal and modification of an individual AS in an AS path. The upside of SPV's approach is that it uses symmetric key cryptography for one-time signatures which is less computationally expensive than public-key cryptography usually used in other approaches. However one major downside of the approach is that one-time signatures are only valid for a limited period of time introducing a strong timing constraint on the advertisement of BGP updates. SPV is not currently being deployed.

BGP Route Origin Verification via DNS (ROVER) [77, 78] aims at providing secured *route origination* by taking advantage of the DNS infrastructure coupled with DNSSEC to allow secured publishing and querying of route origin data. To implement it, Gersch et al. propose to store IP address blocks and their authorised BGP origin ASes in the reverse-DNS and secure the zones using DNSSEC. To do that they introduce a new naming convention to specify CIDR IP prefixes in the reverse-DNS [79] and two new DNS response types, namely Route lock (RLOCK) and Secure Route Origin (SRO), indicating, for a queried IP prefix, that it is secured

using ROVER and its associated authorised BGP origin ASes, respectively. Routers can validate the origin of BGP routes by checking BGP announcements against the published authorised data in the reverse-DNS. ROVER is currently not being deployed but can be tested by network operators through a testbed via shadow copies of the reverse-DNS zone [30].

Listen and Whisper [166] is a scheme aiming at providing only secured *route propagation*. Their motivation lies in avoiding the use of a complex PKI to support cryptographic verifications of routing information. Instead, they introduce two techniques to prevent AS path tampering attacks: Whisper and Listen. Whisper monitors the control plane by verifying the consistency of one-way hash signatures (*e.g.,* RSA or SHA) computed on the AS paths towards a given destination IP prefix. The origin AS of a prefix generates a first signature and embeds it in the BGP updates for that prefix. Each AS along the AS path updates the signature, which can be verified against adding, removal and modification of ASes. Whisper only checks the consistency between pairs of routes and may not trigger an alarm in case of a pair of bogus routes. Listen monitors the data plane by looking for incomplete TCP connections indicating reachability problems with a given destination IP prefix. It is used to reduce the number of false-positives given by Whisper. It thus combines cryptography and BGP anomaly detection to secure path propagation. Listen and Whisper is effective at identifying isolated adversaries, but is of limited use when facing multiple or colluding adversaries. It is not currently being deployed.

Secure Traceroute [138] is, unlike previous proposals, not intended to secure BGP by validating routing updates but instead to detect data plane routing problems caused by a faulty or malicious node in between two communicating hosts. To achieve this they propose a new traceroute method which builds upon the classic traceroute and introduces new features: (i) nodes along the path provide the IP address of the next-hop in their reply, (ii) the source and each node along the path exchange (supported by a PKI) an encryption key and some shared secret to securely identify each other's traceroute packets, and (iii) traceroute replies are authenticated using a Message Authentication Code (MAC). Unfortunately, secure traceroute introduces too much overhead for the security gain preventing it from being deployed and used.

RPKI. Unfortunately, the techniques described here above have never been widely deployed due to an inadequate balance between the security gain and the deployment overhead. However, in the last few years network operators have started to adopt and deploy a BGP hijack prevention framework commonly referred to as the RPKI system. Supported and developed by actors in the operational community, this framework has been gaining more momentum than others in the last few years and relies on a Resource Public Key Infrastructure (RPKI), codified in RFC 6480 [120], to prevent the injection of bogus routing announcements [101, 119]. The RPKI used in this scheme consists of a distributed database of certificates of four types: (i) a type A called Route Origin Authorisation (ROA) binds an IP address block to its authorised

BGP origin AS(es), and (ii) a type B that binds a router IP address to the AS number it belongs to, and (iii-iv) certificates C and D that binds respectively IP addresses and AS numbers to the public key of their respective owner. The certification chain in the RPKI follows the IP address and AS number delegation chain, with the IANA acting at the root certificate authority for RIRs' certificates, a RIR is then acting as the certificate authority for ISPs' or end-user organisations' certificates, etc. Each certificate is signed with the private key of its holder and also embeds its public key. The framework proposes two separate techniques to secure BGP: (i) secured *route origination* codified in RFC 6483 [101] and (ii) secured *route propagation* [119]. (i) Secured route origination uses ROAs (type A certificates) to verify that a given IP address block is originated by the authorised AS(es). A router is then able to verify the validity of a received BGP update for a given IP address block and BGP origin AS by (i-a) querying the RPKI for a ROA related to the IP address block and verifying its cryptographic validity, and, (i-b) if the ROA is cryptographically valid, verifying that the origin AS observed in the BGP update matches the authorised one(s) found in the ROA. This prevents an attacker from announcing a block he does not own. (ii) Secured route propagation aims at preventing AS path forgery by ensuring that each AS in the AS path was not impersonated. This is done by having each router signing a BGP update it propagates so that subsequent routers can verify, using type B certificates from the RPKI, that all routers which have signed the update indeed knows the secret key of the AS (they speak for) in the path.

Assuming ROAs would bind IP address blocks with their legitimate BGP origin AS, hijacks can still be successful if attackers forge the BGP AS path and prepend the legitimate BGP origin AS to it (*i.e.,* hijack types 2, 4 and 6).The only solution to prevent BGP AS path forgery is thus to have secured routed propagation deployed. Unfortunately, this solution is much more invasive and cannot be deployed without substantial software and hardware updates on all BGP routers.

Secured route propagation is currently still at an early development stage, not yet standardized and not yet being deployed. Secure route origination however is progressively being deployed. Major router vendors, such as Cisco [50] and Juniper [107], now support BGP route origin validation using the RPKI and ROAs. Additionally, RIRs and ISPs also progressively support and deploy BGP route origin validation [31]. According to the RIPE NCC [152] there is currently 4.1% of the IPv4 address space covered by ROAs.

Summary

While the RPKI framework coupled with ROAs and BGPsec appears to currently be the most promising solution to bring security into BGP, its deployment is still at an early stage. Despite the effort of the operational and research routing communities to develop and deploy a security extension to BGP, the current Internet routing infrastructure remains vulnerable to IP prefix hijacking attacks.

2.3.6 Detecting and mitigating IP prefix hijacking

While working on solutions to secure BGP, the research routing community has explored an alternate way of mitigating IP prefix hijacking attacks by means of BGP monitoring and anomaly detection. Hereafter we describe some IP prefix hijacking detection and mitigation techniques, highlighting their strengths and weaknesses to identify the different types of hijacks considered in Section 2.3.4.

Control plane-based approaches

Some existing proposals [113, 115, 116, 144, 159] aim at detecting IP prefix hijacking by *passively* monitoring the routing infrastructure. However due to the strong similarity between IP prefix hijacking and some legitimate routing changes those methods suffer from many false-positives. Such legitimate routing changes include BGP engineering practices, *e.g.,* MOAS conflicts created by some multi-homing configurations or by IP anycast [96, 192]. These prefix hijack detection techniques are then mostly useful for network owners who have the ground-truth information about changes related to their network required to discard false alarms. Unfortunately, due to their high false-alarm rate they cannot be used to systematically study the prevalence of the "BGP spectrum agility" phenomenon.

PHAS: Prefix Hijack Alert System. In 2006, Lad et al. proposed PHAS [116], a system designed to detect prefix hijacks where an AS originates a prefix it does not own (*i.e.,* hijack type 1 in Section 2.3.4). The idea behind PHAS is to (i) build an initial mapping at time t_0 of every IP prefix P with its set of originating AS(es) $O(P, t_0)$ and then (ii) check this mapping against real-time BGP updates to detect prefix origin changes and notify P's owner when $O(P, t-1) \neq O(P, t)$. To deal with prefixes exhibiting origin oscillations, *e.g.,* due to unstable Internet connectivity, and the resulting large number of notifications, PHAS introduces the notion of windowed origin set for a prefix $O(P, t, k)$: the announcement of a new origin for a prefix is immediately reported since it may indicate a hijack. However, in the case of the withdrawal of an origin for a prefix, the removal of the origin from $O(P, t, k)$ is delayed by k units of time with $k = 1h$ by default and increasing with the number oscillations observed for the prefix P. As an additional service, PHAS offers network owners to register ground-truth origin AS(es) for their prefixes.

Limitations: First, PHAS only deals with hijacks involving origin AS conflicts (*i.e.,* hijack type 1) and is thus not able to detect hijacks involving a sub- or super-prefix or AS path forgery (hijack types 2-6). Second, it is blind to hijacks involving previously unannounced IP address space (hijack type 7-8) since no state in the model exists for such networks. Finally, PHAS is basically a routing change reporting tool and, as such, it does not try to infer the legitimacy of these changes usually resulting in a lot of alerts which can only be ruled out by the owner of a network.

Topology-Based Detection of Anomalous BGP Messages. In [115], Kruegel et al. describes a method to detect prefix hijacks involving either prefix ownership

violation or invalid AS paths. The detection of prefix violation hijacks simply consists in triggering an alarm upon origin AS changes for a prefix, discarding cases of aggregation (marked in the BGP updates or via communities) or involving prefixes smaller then /8's. To detect invalid AS paths, they first model the Internet AS-level topology (inferred from BGP data) as a graph and classify each AS as core or periphery AS based on its number of peers. BGP updates are checked against the model and an AS path is considered valid only if (i) it contains no more than one subsequence of core ASes (*i.e.,* periphery ASes do not transit traffic between core ASes, that is the path is topologically valley-free) and (ii) all pairs of previously unconnected periphery ASes are within a given geographical distance (arbitrarily set to 300 km).

Limitations: Like other systems relying on an initial AS-level topology model to detect anomalies, it does not consider the update of the model which is problematic in a dynamic routing system. Moreover, it is blind to hijacks involving previously unannounced IP address space (hijack type 7-8) since no state in the model exists for such networks. The prefix ownership violation detection merely reports origin changes and does not attempt any origin validation. Finally, the use of geolocation can be too coarse-grained and introduce inaccuracies in the invalid AS paths detection.

Detecting Bogus BGP Route Information: Going Beyond Prefix Hijacking. The system proposed by Qiu et al. in [144] aims at detecting prefix ownership and invalid AS path hijacks. They propose a system based on the assumption that "although routing is dynamic, the (prefix, owner) associations and inter-AS relationships are stable over time". They use only AS-level topology information to evaluate relationships between ASes. Their system works as follows: (i) it builds the prefix-origin AS associations and the AS-level topology (with directed AS-links) from BGP updates and (ii) considers an AS path valid if it is valley-free (*i.e.,* if the path contains, in that order, zero or more customer-to-provider links, followed by zero or more peer-to-peer links, followed by zero or more provider-to-customer links) and contains no inter-AS link not in the model or in the wrong direction. They further refine their learning/detection algorithm with a few heuristics based on the attackers' behavior and common legitimate BGP engineering practices.

Limitations: Like other systems relying on an initial AS-level topology model to detect anomalies, it does not consider the update of the model which is problematic in a dynamic routing system. Moreover, it is blind to hijacks involving previously unannounced IP address space (hijack type 7-8) since no state in the model exists for such networks.

Neighborhood Watch for Internet Routing: Can We Improve the Robustness of Internet Routing Today? Siganos et al. developed in [159] an approach to validate the origin AS of IP prefixes based on Internet Routing Registries (IRRs). IRRs are databases provided and maintained by Regional Internet Registries (RIRs) and some Internet Service Providers (ISPs) to allow network owners to share with others registration as well as routing policy information related to their network.

They leverage different information fields related to ASes and IP address blocks in IRR's to score each (prefix, origin AS) pair. From experiments run on real-world BGP data, they conclude that (i) although routing registries are useful to validate route origination (ii) (even minor) improvements to IRRs by RIRs, ISPs and other network owners would (greatly) improve their reliability. Finally (iii) ASes seem unprepared to handle routing misbehaviour; reacting too late to large scale events and not at all to small scale events.

Limitations: While providing interesting insights into route origin validation using IRRs, in the context of prefix hijack detection it can only detect those involving prefix ownership violation. Moreover, IRRs are known to be incomplete and to contain inaccurate and stale data [45, 86, 112, 158, 163, 164], which hinder their reliability for detecting and mitigating BGP hijacks. However, unlike previous approaches to detect hijacks based solely on BGP data, it does not leverage an initial model of the Internet AS-level topology and can potentially identify hijacks of previously unannounced IP address space (hijack type 7-8).

Pretty Good BGP: Improving BGP by cautiously adopting routes. Pretty Good BGP (pgBGP) is a system proposed by Karlin et al. in [108] to prevent prefix ownership violation. The key idea behind this approach is to detect and penalise suspicious new routes to a prefix for a given time period (24 hours by default) to let enough time for the prefix owner to validate it or not. The system works as follows: (i) it monitors BGP update messages, (ii) accepts updates with origin ASes for a prefix seen within the past few days, (iii) marks as suspicious and quarantines for 24 hours updates with new origin ASes, as they can be due to a hijack, and (iv) marks as suspicious and quarantines for 24 hours updates with new sub-prefixes, as they can be due to a hijack. Routes with quarantined ASes are given low local preference in BGP and routes with quarantined sub-prefixes are ignored. Because pgBGP doesn't provide any mechanism for identifying valid and invalid routes, authors have set up the Internet Alert Registry [17] to which network operators can subscribe to find if they are involved in a hijack.

Limitations: Because pgBGP is basically a routing change reporting tool, it can produce false positives resulting from legitimate situations, *e.g.,* a provider change, a previously unseen auxiliary provider. Bogus routes can also be accepted after the delay period. Finally, it does not deal with hijacks involving AS path subversion and previously announced IP address space (hijack type 2, 4, 6-9).

Data plane-based approaches

Other proposals [96, 108, 157, 167, 189, 194] leverage *active probing* of networks together with passive monitoring to improve the detection by assessing the impact of BGP routing changes on the data plane.

Accurate Real-Time Identification of IP Prefix Hijacking. In [96], Hu at al. propose a system which aims at detecting in real time any type of hijack that can potentially occur on the Internet. Their system (i) detects abnormal routing

changes using BGP-based anomalies (*e.g.,* MOAS conflicts, new AS links, etc) and (ii) checks data plane fingerprints to assess the impact of abnormal routing changes on the data plane and reduce false positives. It builds upon the assumption that besides all symptoms of IP prefix hijacking, if a network is hijacked, host and network fingerprints from distant sources (affected by the hijack) and from sources close to or within (not affected by the hijack) the hijacked network would likely be different. Upon detection of origin changes or new AS-level links, they perform various checks including fingerprinting (*e.g.,* Nmap) of hosts within the network, idle scanning [95], evaluation of geographic distance and inter-AS business relationship between two newly connected ASes.

Limitations: While providing a comprehensive BGP hijacking detection system, Hu et al.'s approach is blind to hijacks of previously unannounced IP address space (hijack type 7-8). Moreover, the real-time system deployment requirements are rather high: monitoring hosts distributed across the Internet requires to have access to local BGP routing information and to be able to run fingerprinting measurements. Moreover, the accuracy of the system if highly dependent on the accuracy of fingerprinting measurements possibly affected by, for instance, firewalls, NATs, etc. Finally, some fingerprinting measurements, such as Nmap, are particularly aggressive towards the scanned networks.

Detecting prefix hijackings in the internet with Argus. Argus [157], developed by Shi et al., is a system for detecting any type of prefix hijack in real time. Its detection technique relies on the assumption that the hijack of a prefix will modify the reachability of hosts within the hijacked network differently from other legitimate routing changes or operational faults. Their system builds an initial model of the Internet AS-level topology and then monitors routing updates for changes in prefix origin ASes or abnormal inter-AS links. It launches *ping* measurements from a large number of distributed probing hosts across the Internet to assess the reachability impact of abnormal routing changes on the victim network and triggers an alarm when the result of probing hosts affected by the routing change sufficiently differs from the result of the other probing hosts. Argus is currently the only system providing alerts publicly[1]. Network operators can also subscribe to alerts related to their network.

Limitations: Argus alerts are cluttered with many false positives, mainly due to ping measurements inaccuracies and the choice of inappropriate target IP addresses to ping within a network. Moreover, Argus relies on an initial model to detect abnormal routing updates making it blind to hijacks of previously unannounced IP address space (hijack type 7-8). Finally, by relying on a reachability anomaly to detect hijacks, it is blind to interception attacks (hijack type 9) and some impersonation attacks.

iSPY: Detecting IP Prefix Hijacking on My Own. iSPY, proposed by Zhang et al. in [189], is an approach to detect prefix hijacking from inside an AS. It is thus only meant for network operators to monitor their own network. The technique

[1] http://argus.csnet1.cs.tsinghua.edu.cn/alarms/

consists of a cyclic probing of transit ASes, which interconnect the main parts of the Internet, from within the monitored AS. The system aims to successfully probe at least one live IP address per AS. This is done with a combination of *traceroute* and *ping*. A database is maintained by the system that collects and updates live IP addresses and their related ASes. The idea behind it is that the traffic redirection resulting from a hijack would redirect the answers of the probing to the new bogus AS where the attacker is located. Thus the answer would never reach the correct origin, where the IP prefix belongs and the probes were sent. To avoid false alarms the received pattern of one cycle is compared with the results of the previous cycle and above a given number of unreachable, or cut off, ASes, a hijack alarm is triggered.

Limitations: First, iSPY is only intended to be used by network operators to monitor their own network. Second, its detection accuracy is limited by the accuracy of ping and traceroute measurements as well as by the choice of live target IP addresses in the probed ASes. Finally, it is blind to interception attacks (hijack type 9).

A light-weight distributed scheme for detecting IP prefix hijacks in real-time. Zheng et al. proposed in [194] a prefix hijack detection system relying solely on data plane measurements. It builds upon the assumption that the location of an IP prefix in the Internet AS-level topology changes rarely. As a result, their detection technique assumes (i) that the hop count from a vantage point to a target network should be stable and (ii) that the AS-level traceroute from a vantage point to a "reference point" topologically close to the target network should be a super path of the AS-level traceroute to the target network. They assume that legitimate routing changes will modify topological location of the reference point and the target IP address whereas a hijack will only change the target IP address topological location.

Limitations: Unlike previous studies combining control plane and data plane information, data plane measurements are not triggered by BGP-based routing anomalies introducing a potentially high overhead of performing the measurements, depending on how often they are performed. It is thus intended to monitor a limited set of chosen networks. Data plane measurements can also suffer form inaccuracies, which can consequently affect the accuracy of the technique. Finally, the sole use of data plane measurements can make the system blind to hijacks involving a sub- or super-prefix of a monitored prefix when the monitored target IP address is not in the hijacked part of the prefix.

A study of prefix hijacking and interception in the Internet. In [59], Ballani et al. study prefix hijacking and Internet traffic interception attacks (hijack type 9). They first examine different hijacking and interception scenarios and evaluate the probability of success of such attacks based on the topological location of the attacker and the prefix he advertises. In particular, they show that, although small ASes can hijack and intercept traffic to a prefix, the higher in the hierarchy (*e.g.,* tier-1 is higher than tier-2, etc) an AS sits the easier. Then they actually perform a real-world interception of traffic related to their prefixes therefore demonstrating the possibility and ease of carrying out such attacks. They develop a signature for traffic interception: they assume that an attacker will advertise a route in BGP with its AS

as next-hop to the victim origin AS to attract traffic. However, to forward the traffic back the victim, the attacker needs a valid route with a correct next-hop AS. They then look at AS-level traceroutes to the intercepted prefix to detect the invalid next-hop AS. Finally, they apply their signature on real-world routing data and uncover no interception case.

Limitations: The accuracy of their traffic interception detection is affected by different types of routing anomalies that can appear when comparing BGP and traceroute paths [128] and that can be mistaken for interception attacks, *e.g.,* Internet eXchange Points (IXPs), sibling ASes, customers using provider IP address space, multi-homing, etc. Finally, the traffic interception attack scenario is not complete and differs from the one presented by Pilosov et al. at Defcon 2008 [139], which is more likely to be successful.

Summary

Several other techniques [92, 94, 95, 113, 129, 145, 158, 163, 167, 188, 193], similar to those we described above with respect to their methodology and limitations, have also been proposed to tackle the detection and mitigation of BGP hijack attacks. Unfortunately, each approach tackles one or a few types of hijack and thus does not provide a complete BGP hijack detection scheme. As pointed out in Section 2.3.4, when studying the "BGP spectrum agility" phenomenon it is important to be able to identify hijacks of allocated/assigned but unadvertised IP address space. However, only the technique proposed by Siganos et al. [159] leveraging IRRs to validate BGP BGP route origination is potentially able to identify such hijacks.

Besides techniques to secure BGP or detect and mitigate BGP hijacks, a few other works [76, 80, 82, 118, 146, 147] have focused on assessing the effectiveness, *i.e.,* the success likelihood, of BGP hijack attacks based (i) on the type of hijack performed and (ii) on the location in the Internet AS-level topology of the victim's AS with respect to that of the attacker's. Such findings are then leveraged to devise optimal deployment strategies for BGP hijack defense systems.

2.3.7 BGP monitoring and analysis

Apart from techniques to secure BGP and mitigate IP prefix hijacking attacks, a collection of tools and datasets are available to monitor and analyse the state of the Internet routing infrastructure. Essentially developed by the operational routing community to monitor the state of the Internet routing infrastructure and help troubleshoot routing issues, these tools can also be leveraged to monitor and investigate IP prefix hijacking attacks.

A BGP *looking glass (LG)* is an application that is installed on a network and provides the view of the Internet routing infrastructure from this network. A LG usually allows to retrieve BGP routes and perform network measurements like ping or traceroute to any given network on the Internet. It is often provided by ISPs for network diagnostic and troubleshooting purposes. A collection of publicly available LGs accessible via a web interface or telnet is available at [6].

As one of the five Regional Internet Registries, RIPE offers services to help monitor the routing infrastructure. The RIPE Routing Internet Service (RIS) is an infrastructure that collects BGP routing data from routers in different locations in the world, *i.e.,* not only in Europe. The dataset contains BGP updates and Routing Information Base (RIB) (or routing table) dumps. The data is archived and publicly available [26]. Such data is intensively used in research studies as well as in the operational community for network diagnostic and troubleshooting because it is publicly available and provides good visibility into small scale routing events thanks to its widely deployed vantage points. The routing data collected by the RIS can also be accessed via RIPEstat [28], a web interface allowing to retrieve and visualise routing information related to IP and AS resources. RIPEstat also provides a REST API which can be leveraged to script routing analyses.

The University of Oregon RouteViews project [43] is a similar infrastructure to the RIPE RIS which collects BGP updates and RIB dumps from routers from different locations in the world and makes the archived dataset publicly available. Similarly to RIPE RIS data, RouteViews data is intensively used by researchers and the operational community to study the Internet routing infrastructure.

The BGP Monitoring System (BGPmon) [5, 185] is a BGP routing data collection infrastructure initiated by the Collorado State University and designed to provide public access to live and archived BGP updates and RIBs (or routing tables) from many routers distributed around the world. BGPmon allows new participants willing to contribute with BGP data from their network to join and increase the diversity of BGP routes. Recently, RouteViews and BGPmon joined forces to provide an even more diverse (in terms of vantage points) and comprehensive BGP routing data feed.

Robtex [29] is a general-purpose Internet resources analysis tool which provides information related to domain names, IP addresses and AS numbers. It can be particularly useful in the analysis of routing problems but the absence of archived data, the undocumented data collection process and the absence of a query API limits its use in a systematic Internet routing analysis.

Huricane Electric BGP Toolkit [16] is another web tool similar to Robtex by the use cases and limitations.

Cyclops [12, 67] is a tool developed by the University of California, Los Angeles that monitors the routing infrastructure to detect anomalous inter-AS connectivity changes and prefix originations, possibly due to misconfigurations or prefix hijacks. It is particularly useful for network operators as they can provide ground-truth information about the AS-level connectivity of their network and be notified upon detection of a violating routing change. Although archived data is available, Cyclops only provides to the public a web interface for browsing and visualising routing information.

PeeringDB [23] is an online database where network operators can openly share with others information regarding their BGP peering policies, their presence at certain Internet eXchange Points (IXPs) or other peering facilities and properties of their network, *e.g.,* bandwidth and data link type. It is primarily intended for network operators to determine where and with whom to peer as well as to check the

status of their current peers. Although it is not complete and does not provide archived information, PeeringDB can be used as an accurate [162] source of peering information to analyse inter-AS connectivity.

BGPinspect [8] by Merit Network Inc. is an online web interface to the Route-Views routing data. It allows to view the routing tables from different collectors in the world and verify the global visibility of an IP prefix or AS.

Internet Routing Registries (IRRs) [20, 56, 130] are databases provided and maintained by Regional Internet Registries (RIRs) and some Internet Service Providers (ISPs) to allow network owners to share with others registration as well as routing policy information related to their network. Initially created as an IP address and AS number allocation/assignment registry, IRRs are now used by network operators to exchange routing policy information and use it in the configuration of their BGP routers, *e.g.,* to build filters. Internet routing policies are modelled using different objects (*e.g.,* `aut-num` describing an AS, `inetnum` describing an IPv4 address block, `route` binding an `aut-num` and an `inetnum` to describe a BGP route origination, route `import` and `export` rules between ASes, etc) and expressed using the Routing Policy Specification Language (RPSL) codified in RFC 2622 [53], which is used by everyone except ARIN who uses the Shared WHOIS Project (SWIP) [19]. IRRs provided by the different RIRs and ISPs [20] can usually be queried using the `whois` command or downloaded as a one-file text-format dump. While they constitute the only source of registration and routing policy information useful to detect and investigate anomalous routing events such as BGP hijacks, IRRs are also known to be incomplete and to contain inaccurate and stale data [45, 86, 112, 164], which hinder their reliability for detecting and mitigating BGP hijacks.

Renesys [25], now part of Dyn, is a company that provides, among other things, a routing alarms service to which network operators can subscribe to receive notifications upon detection of suspicious routing changes related to their network. Renesys is also highly active in the monitoring of the global Internet routing infrastructure, especially in security issues such as BGP hijacks. They regularly publish blog articles and contribute to different technical operational meetings (*e.g.,* NANOG, RIPE, etc).

BGPmon.net [7] is another company that provides a routing alarms service to subscribed customer network operators. They also provide, via their website blog, analysis and intelligence reports on Internet-scale routing issues and BGP hijack incidents.

2.4 Conclusion

In this chapter we have discussed previous work related to the *three* facets of this thesis.

(i) Introduced by Ramachandran et al. and then corroborated by other reports, the "BGP spectrum agility" phenomenon suggests that cybercriminals are able to run BGP hijacking attacks to use the stolen IP space to perform other malicious

activities. However, based on the evidence gathered thus far, the existence of such a phenomenon *as a new way* for cybercriminals to perform malicious activities while remaining stealthy remains to be demonstrated. This is first goal of this thesis. Evidence gathered thus far about the "BGP spectrum agility" phenomenon consists mostly of isolated hijack cases and does not enable us to assess the prevalence of the phenomenon. This is the second goal of this thesis. Finally, we really need a more thorough documentation of the global modus operandi of cybercriminals using "BGP spectrum agility", if they do so, as this is a key to design effective countermeasures. This is the third goal of this thesis.

(ii) We then looked at the potential impact of cybercriminals using "BGP spectrum agility" on current countermeasures for malicious activities on the Internet. It appears that if malicious actors were capable of running routinely such attacks, this would constitute a very serious threat to the Internet since this would enable them not only to send spam while defeating the classical IP blacklists but, more importantly, to run large scale DDoS at almost no cost or run man-in-the-middle attacks against almost any target of their choosing.

(iii) Finally, we reviewed the current state of the security of the Internet routing infrastructure as its vulnerability to BGP hijacking attacks is the base for the "BGP spectrum agility". First we have seen that despite the long effort to provide BGP with security extensions to defend against BGP hijacking and the current deployment of the RPKI with secured route origination, the routing infrastructure is still a long way from being secured. Second, we have reviewed techniques and systems that have been developed to fill the need to mitigate BGP hijacks while we are waiting for a global secured routing infrastructure. In particular, we have seen that many systems have been designed to mitigate hijacks using various data sources, usually BGP data and data plane measurements, but they generally suffer from high false positives and do not cover all hijack scenarios. In particular, most prefix hijack detection schemes only deal with hijacks of announced IP address space. However, the first reported malicious BGP hijacks all involved IP address space that was unannounced when it was hijacked. Overall, there is currently no existing technique or system ready to be used for a large-scale study of the "BGP spectrum agility" on the Internet. There is thus a need to a develop a system specifically for the study of BGP hijacks related to networks involved in malicious activities. This is the topic of the next chapter of this thesis.

3

SpamTracer

Contents

3.1 On the identification of malicious BGP hijacks 36
3.1.1 Dataset 1: control plane data 36
3.1.2 Dataset 2: data plane measurements 38
3.1.3 Dataset 3: malicious activities logs 40
3.1.4 Correlation of datasets 40
3.2 SpamTracer 40
3.2.1 Routing data collection 47
3.2.2 Multi-stage scoring and data filtering 50
3.2.3 Validation of candidate hijacks 58
3.2.4 Root cause analysis 63
3.3 Conclusion 65

Recall from Chapter 1 that the first objective of this thesis is to answer the question: *as of 2014, are there BGP hijacking attacks carried out by cybercriminals on the Internet for the purpose of using the stolen IP address space to perform malicious activities, such as sending spam, distributing malware, launching DDoS attacks, etc?* In Chapter 2 we surveyed evidence of previously reported malicious BGP hijacks and showed that there is a need to rigorously assess the existence and prevalence of this potential threat. We also assessed the lack of data or techniques for collecting it readily available to answer our question. We thus decided to build our own experimental infrastructure called SpamTracer [178]. In this chapter, we start by reviewing the types of data that are required to identify possible malicious BGP hijacks and subsequently derive the data collection requirements for our experimental infrastructure. We then move on to the description of the SpamTracer data collection infrastructure. Finally, we present a methodology we devised specifically for the analysis of data collected by SpamTracer to identify and study possible instances of malicious BGP hijacks.

3.1 On the identification of malicious BGP hijacks

The task of identifying malicious BGP hijacks is twofold: (i) determining whether a network is hijacked and (ii) finding whether a network is involved in malicious activities. As explained in Chapter 2, two types of data are commonly leveraged to study the Internet's inter-domain topology and, in particular, BGP hijacking [153]: *control plane* (BGP) data and *data plane* (*e.g.,* ping, traceroute) measurements.

3.1.1 Dataset 1: control plane data

A first approach to detect BGP hijacks is to directly leverage BGP routing (control plane) data. This data can be either (i) *BGP update messages* exchanged between BGP-speaking routers or (ii) snapshots (*i.e.,* instant pictures) of routers *BGP routing table* (or BGP Routing Information Base (RIB)). The former (i) correspond to raw reachability information (network IP prefix advertisement and withdrawal) messages exchanged as part of the protocol operation. The latter (ii) correspond to BGP routes that have been selected by a BGP speaker's decision process (*i.e.,* Local Routing Information Base or Loc-RIB per the conceptual BGP model described in [150]), which include BGP routes received from BGP peers or locally sourced as well as internal and static routes. Control plane routing data is collected at BGP-speaking routers. It is commonly exported in the Multi-threaded Routing Toolkit (MRT) routing information export format [63], like in the RIPE RIS [26] and RouteViews [43] archives. Although MRT is a standardized format that facilitates the analysis of routing data, for example via developed parsing tools (bgpdump) and libraries (libbgpdump) available for different programming languages, it is not compulsory and control plane routing data can be exported in any other format, for example XML in BGPmon [5]. As presented in Section 2.3.7 of Chapter 2, several real-time BGP routing datasets are made available to the network operational and research communities. The RIPE RIS, RouteViews and BGPmon routing datasets provide routing data collected at a large number of BGP-speaking routers distributed worldwide and have already been used in a large number of Internet routing studies, for instance [116, 123, 129, 157, 192, 194]. These BGP data collection projects use a special purpose BGP collector referred to as *route server*. In that context, a route server is a router that peers with routers in other ASes, usually distributed worldwide, and whose only purpose is to collect BGP routes from BGP peers without advertising anything to these peers. Besides providing archived BGP updates messages and BGP RIBs, RouteViews also provides direct access, via `telnet`, to its BGP collectors to obtain their view of the Internet routing infrastructure in real time. Finally, besides the popular publicly available BGP datasets provided by RIPE, RouteViews and BGPmon, one can collect BGP routing data at any BGP-speaking router under one's control.

When collecting and using BGP routing data to study the inter-domain routing infrastructure it is of tremendous importance to consider both the quality of the data with respects to the characteristics of the phenomenon we want to study and its intrinsic limitations. In the context of this thesis, we are interested in detecting BGP

hijack cases. Based on previous studies [96, 148, 157] and past hijack incidents [48, 121], we identified three key metrics to assess the quality of BGP routing data for use to study BGP hijacking: (i) the *visibility* of the BGP collector(s), (ii) the *time scale* of the data, and (iii) the *availability latency*.

(i) Recall from Section 2.3.3 in Chapter 2 that inter-domain routing decisions are primarily driven by routing policies, usually kept private, established between ASes, which affect the propagation of reachability information into the global routing infrastructure, for example some BGP routes can be visible to only one AS or a group of ASes. Besides that, normal BGP operational practices, such as IP prefix aggregation or IP prefix-based BGP route filtering, influence the propagation of reachability information from ASes to ASes. This results in a situation where all ASes do *not* have the same view of the inter-domain routing infrastructure. In this work, we define the visibility of a BGP collector as its ability to see BGP routes advertised on the Internet. As showed by Zhang et al. in [187], the more BGP collectors are used the higher the likelihood is to see a BGP hijack that is not globally visible.

(ii) Besides being of variable visibility, BGP routes also have variable durations [105, 148], ranging from a few seconds to several years. The duration of a BGP route to a given network IP prefix refers to the time period the route is advertised without interruption. In order to detect short-lived, possibly hijacking, BGP routes it is necessary to collect and use BGP routing data with a small enough time scale. The data time scale refers to the time granularity of the data.

(iii) Finally, the availability latency of BGP data is the time elapsed between two batches of data. For real-time BGP data processing the availability latency should be as small as possible.

Evaluating the quality of BGP routing data with respect to the three key requirements for studying BGP hijacks, namely the *visibility* of the BGP collector(s), the *time scale* and the *availability latency* of the data, is a challenging task due to the lack of ground-truth data. Table 3.1 lists the quality metrics for the three main publicly available BGP datasets. The most appropriate dataset should be chosen based on the requirements of the routing study, *e.g.,* a system aiming at detecting routing anomalies from BGP update messages in real time would probably choose the live BGP feed from BGPmon [157]. Some studies have attempted to assess the quality of different BGP datasets, including RIPE RIS, RouteViews and other private ones, with respect to some of their use cases, such as the inference of Internet's AS-level connectivity maps [84, 136, 186, 187] and the study of routing anomalies, including BGP hijacking [187].

Dataset	**Visibility**		**Time scale**			**Availability latency**		
	# collectors	# distinct peers (ASes)	RIB	`telnet`	Updates	RIB	`telnet`	Updates
RIPE RIS	17	515	8h	NA	0s	8h	NA	5m
RouteViews	19	242	2h	0s	0s	2h	0s	15m
BGPmon	1	95	0s	NA	0s	0s	NA	0s

Table 3.1 – Quality metrics for the three main public BGP datasets.

In Section 2.3.6 of Chapter 2 we presented different BGP hijacking detection

techniques taking solely advantage of control plane (BGP) data. Unfortunately, even last-generation tools yield output cluttered with alerts corresponding to benign network events, *e.g.,* a multi-homed network switching from one provider to another, a leak of (part of) a customer's routing table to its provider(s), etc. For the targeted audience of such tools (ISPs, prefix owners, etc), this is not a problem since they know the expected behaviour of their prefix. With this ground truth, they can make an informed decision on the value of the alert, and take appropriate action, if needed. Moreover, since they only monitor their own prefixes, the number of alerts they receive is manageable. Recently, Roughan et al. [153] also discussed the intrinsic limitations of BGP data for studying the AS-level Internet, which stems from the fact that BGP was designed to allow network operators to implement routing policies by exchanging reachability information with other ASes while hiding the AS-internal configurations. BGP is thus designed to hide some routing information rather than to expose all of them. Moreover, in the current Internet architecture the routing process takes place at the control plane while the forwarding process takes place at the data plane. In theory, the latter should be based upon the former with forwarding tables derived from routing tables, in other words control plane paths should match data plane paths. For example, one could expect that the AS-level path taken by an IP packet between two hosts in two different ASes would match the BGP AS path between these ASes. As shown in [55], this is not necessarily the case. One could also expect that any IP address block not present in the BGP routing tables should not be reachable, *i.e.,* IP packets sent to an IP address within that prefix should never reach any network. As shown in [64, 123], this is not necessarily the case either. These incongruities stem from the dissociation of the control and data planes.

It thus appears that while BGP data provides an intuitive means of monitoring the inter-domain routing infrastructure for BGP hijack attacks, it fails to provide enough information to be able to differentiate benign from malicious routing events.

3.1.2 Dataset 2: data plane measurements

In theory, a BGP hijack should always be observed in BGP since this is where the inter-domain routing process takes place. However, due to the lack of ground-truth data, relying solely on control plane information to detect and analyse routing changes is challenging. However, data plane measurements can be leveraged to determine the impact of a routing change on the forwarding paths to a network. Data plane measurements used in previous works on the detection of BGP hijacking attacks include the popular traceroute [189, 194] and `ping` [157, 189] but also various network and host fingerprints, such as *port scanning* using *idle scan* [95, 96], and *device fingerprinting* [96]. Unlike BGP routing data, to the best of our knowledge, there is no project providing publicly available, archived, data plane measurements datasets readily available for use in the study of the malicious BGP hijacks. A first major obstacle to collecting such a dataset is that there exists many different types of data plane measurements related to networks and hosts and performing them at the scale of the Internet and the IPv4 and IPv6 address spaces would be prohibitive. As of August 2014, there are about 513,000 IPv4 prefixes (equivalent to $2,710,135,328$

individual IPv4 addresses) and about 19,000 IPv6 prefixes (equivalent to 2.64×10^{33} individual IPv6 addresses) in BGP routing tables, *i.e.,* advertised on the Internet [99]. More importantly, while BGP is an event-driven protocol, data plane measurements capture only the instant state of a network or host and, depending on the application, may need to be performed repeatedly, which would also introduce scalability issues. A few projects including DIMES [155], CAIDA Archipelago (Ark, former Skitter) [66] and iPlane [124] aim at collecting traceroute and ping measurements for the purpose of mapping the Internet's AS-level topology. These projects provide access to the data plane measurements datasets either publicly or, at least, to researchers. Each dataset is built using a different methodology. Hence, the value and relevance of these datasets highly depend on the application and its requirements with respect to the data plane measurements. For instance, if an application requires some traceroute measurements to a given target to be performed several times a day, or be event-triggered then the previous datasets would not be appropriate. In 2013, Durumeric et al. [73] introduced an open source network scanning tool called *Zmap*[1]. The tool is able to achieve a scan of the entire IPv4 address space, with one port per IP address, in about 45 minutes. Several datasets built using the tool are made publicly available on the Internet-Wide Scan Data Repository website[2]. However the value of the tool really resides in the fact that it can be integrated into a custom data collection process to collect the most appropriate data for a given application. Furthermore, there is no standard format, like MRT for BGP data, to store and archive data plane measurements, which makes it harder to switch from one of the above datasets to another or from one of these datasets to the collection of one's own dataset. Due to the obstacles discussed here above, we believe that existing data plane measurements datasets are hardly usable in a study of malicious BGP hijacks.

Besides the valuable routing-related information data plane measurements can bring, this type of data may also suffer from inaccuracies, which stem from the limitations of the measurement techniques, *e.g.,* traceroute and `ping`. A lot of research has thus been done on designing alternate measurement techniques to eliminate the limitations of the previous ones [57, 138] and on identifying and mitigating inaccuracies introduced by the measurement techniques [54, 55, 122, 128]. In [57] Augustin et al. designed a new traceroute measurement technique called *Paris traceroute* aiming at eliminating inaccuracies introduced by load-balancing routers along the paths. Padmanabhan et al. proposed in [138] *Secure traceroute*, a traceroute technique relying on cryptography to ensure the correctness of IP-level hop addresses reported in the paths. In [54, 55, 122, 128] authors study the different incongruities that can appear in traceroute paths, such as unresponsive IP-level hops, IP addresses belonging to other networks than those traversed (third-parties), etc. Finally, like BGP routing data, traceroute measurements gain visibility into the diversity of the possible BGP routes to network IP address blocks when they are performed from several, topologically distributed vantage points [156].

In the end, although data plane measurements provide valuable complementary

[1] `https://zmap.io/`
[2] `https://scans.io/`

routing data to a BGP dataset, they appear to be hardly usable as the sole dataset of routing information for the study of malicious BGP hijacks.

3.1.3 Dataset 3: malicious activities logs

As mentioned in the introduction to this Chapter, our first goal is to determine whether BPG hijacking attacks are performed in preparation of other malicious activities, such as sending spam, hosting scam and malware distributing websites, performing DDoS attacks, etc. In Sections 2.2.1 and 2.2.2 of Chapter 2 we reviewed different types of malicious activities that can be traced back to IP addresses enabling further examination to determine whether the hosting network was hijacked at the time the nefarious activities were perpetrated.

Unfortunately, the size of the malicious activities datasets, *e.g.,* 24,000 distinct spam emitting network IP address blocks observed per day in the spam dataset collected at Symantec.cloud [35] spamtraps, makes it hardly feasible to identify malicious networks involved in BGP hijacks without considering any routing-related information to reduce the number of suspicious cases.

3.1.4 Correlation of datasets

We have seen in the previous sections that the challenge of studying and identifying malicious BGP hijacks can theoretically be tackled by leveraging routing- or security-related datasets separately. However, after reviewing some of these datasets it appeared that considering them separately would likely result in a high number of false-positive alarms due to intrinsic limitations in the data or the collection process. Instead, we advocate that *correlating* the different datasets with each other helps balance the weaknesses of some datasets with the strengths of the others, enabling us to produce meaningful candidate hijack cases and to eventually determine whether these attacks are taking place in the wild Internet today.

3.2 SpamTracer

As explained in Section 2.3.6 of Chapter 2 on page 26, despite the fact that several techniques and deployed environments exist to monitor the routing infrastructure and detect BGP hijack attacks, existing techniques and systems currently have several drawbacks, which seriously limit their use for studying malicious BGP hijacks. Furthermore, Roughan et al. recently showed in [153] the importance, in Internet routing research, of collecting data specific to the problem we study instead of using existing datasets, which may have been built for a different application.

We have thus set up a comprehensive experimental environment to study malicious BGP hijacks. We call it SpamTracer. Its complete workflow is depicted in Figure 6.1. Our goal here is to ① collect routing data related to networks originating malicious network traffic, ② extract from this data IP address blocks exhibiting an abnormal routing behavior and retain the ones most likely indicating they might

result from a BGP hijack, ③ manually (in)validate each candidate hijack by taking advantage of external data sources, and finally ④ investigate the root cause behind some validated malicious BGP hijacks to obtain new insights into the hijackers' behavior.

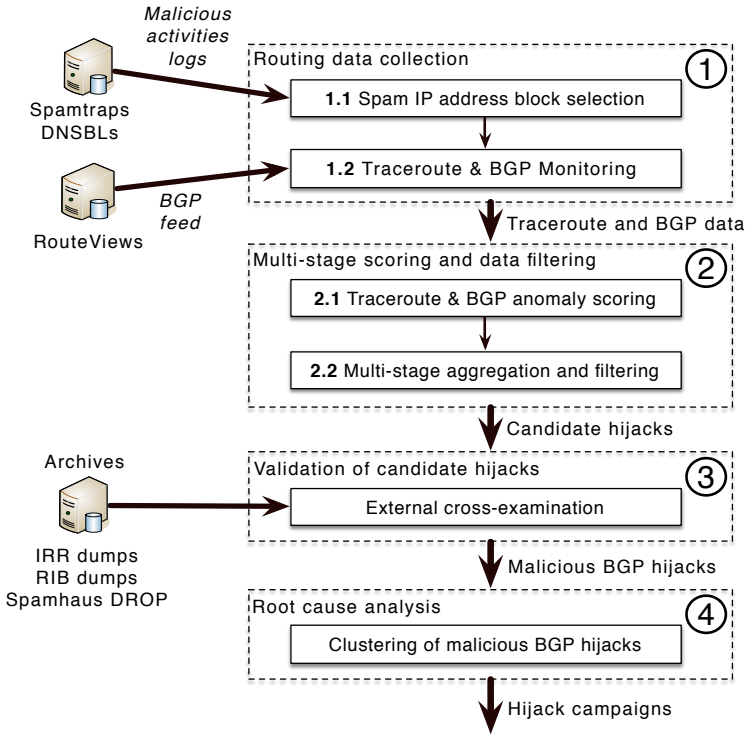


Figure 3.1 – SpamTracer: workflow.

The global distributed SpamTracer system architecture is depicted in Figure 3.2. It comprises *two* types of hosts:

- 1 **backend.** This server provides, on a hourly basis, each SpamTracer instance with a list of network IP prefixes to monitor (input) and fetches the traceroute and BGP data collected by each instance.

- N **instances.** SpamTracer instances are measurement nodes that are meant to be distributed in various ASes and countries (or regions of the world) on the Internet. They collect network traces (traceroute and BGP data) related to the monitored networks. Each instance also keeps a backup of the network traces collected at that instance.

The first prototype of SpamTracer used to comprise a single instance running together with the backend on a single host. This setup had three limitations that have prevented us from performing, initially, a large scale study of malicious BGP hijacking attacks:

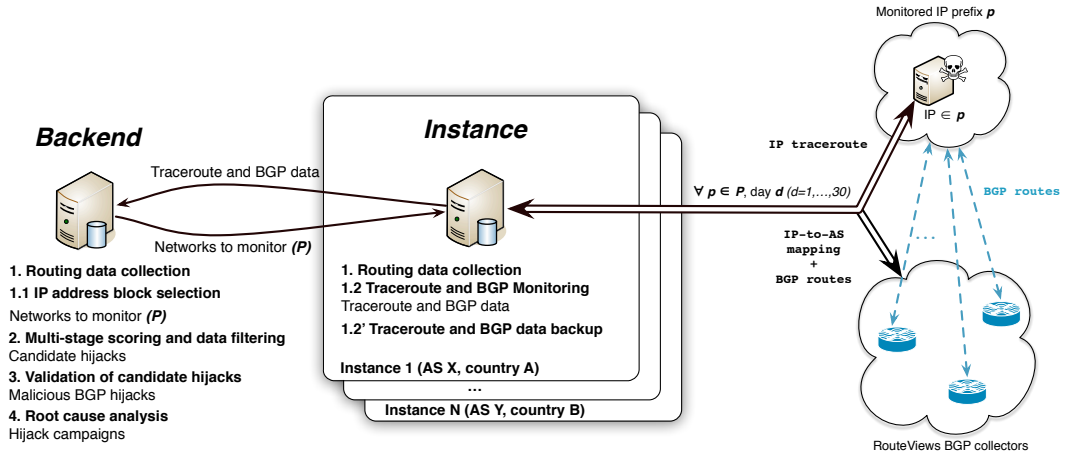


Figure 3.2 – SpamTracer: system architecture.

Number of viewpoints V: Network traces collected from a *single viewpoint* ($V = 1$) can prevent us from observing changes in the routing state of the monitored networks that are only visible from certain viewpoints in the Internet.

Monitoring capacity C: Only a *small fraction* of all networks observed everyday in a spam feed provided by Symantec.cloud could be monitored. In fact from the 60,000 networks that should be monitored everyday (each network was monitored for one week), a single SpamTracer instance is able to monitor only about 8,000 ($C = 8,000$) of them (*i.e.,* 13.33%) due to limited available hardware resources and network connection latencies.

Monitoring period P: Each network is monitored for a period of only seven days ($P = 7$), which prevents us from detecting a possible hijack if the revenant routing change occurs after this period.

We thus decided to update the SpamTracer system architecture to remedy these limitations. Here below we discuss (i) the sizing of the large-scale deployment, *i.e.,* the determination of the number N of instances to enable us to study at large scale the routing-level behavior of networks originating malicious network traffic, and (ii) the concrete implementation of the SpamTracer distributed architecture.

Sizing of the deployment: determining N.
To determine the number N of SpamTracer instances required to monitor the networks extracted from our IP address feeds, we simulated the *IP address block selection* algorithm of SpamTracer on several months of spam data (*i.e.,* spam networks). We used only spam data in our simulations since it represents the primary input feed (80%) to SpamTracer.

The objective of the SpamTracer IP address block selection algorithm is to monitor networks during a period of time so as to maximise the likelihood of observing a relevant routing change possibly due to a hijack and minimise the amount of

networks monitored, in other words *maximize* the relevance of the collected traces for our study of malicious BGP hijacks. Finding this best monitoring period for a given IP network is not easy but some features, such as BGP-related features, can be of help to properly adjust the monitoring period. We thus designed two versions of the IP address block selection algorithm: (i) a first version (v1) not taking advantage of BGP-related features and (ii) a second version (v2) relying on BGP-related features to optimize the monitoring period of IP address blocks.

The IP address block selection takes as input a set ι of IP addresses from the various IP address feeds used in SpamTracer. It gives as output a set θ of IP address blocks to monitor. The parameter C (resp. P) refers to the monitoring capacity of the instance (resp. the monitoring period). In our distributed deployment, we aim at monitoring each network from several viewpoints (or vantage points) so $\frac{N}{V}$ gives the number of instances to deploy at each vantage point. The number of instances N should also satisfy

$$N = \frac{C \times V}{8,000}. \tag{3.1}$$

The first version of the IP network selection algorithm, detailed in Algorithm 1 relies only on features from the IP address feeds. For each IP address ip, we fetch the longest matching IP prefix p currently in the routing tables. If p is already monitored, we restart monitoring it for P days (the default length of the monitoring period). Otherwise, we add it to the set θ of IP prefixes to monitor.

Algorithm 1 IP address block selection (version 1)

Require: Set ι of input IP addresses
Ensure: Set θ of output IP address blocks to monitor
procedure IPAddressBlockSelectionV1(C, P)
  $\theta = \{\}$
  for all $ip \in \iota$ **do**
    while $|\theta| < C$ **do**
      $p =$ getPrefixFromIP(ip)
      if $p \notin \theta$ **then**
        $\theta = \theta \cup \{p\}$
      else if $p \in \theta$ **then**
        resetTLL(p, P)
      end if
    end while
  end for
end procedure

The second version of the IP selection algorithm is described in Algorithm 2 and leverages BGP-related[3] features to adjust the monitoring period of an IP address block in an effort to maximize the relevance of the collected traces. It is similar to the first version but includes the following optimizations:

[3]BGP data consists of archived BGP routing table (RIB) dumps from RIPE RIS and RouteViews.

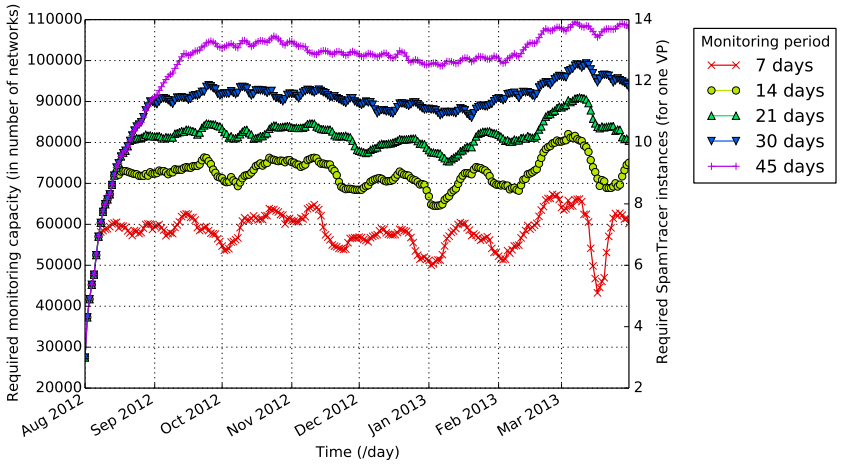


Figure 3.3 – IP address block selection algorithm **version 1**: required daily monitoring capacity C and SpamTracer instances N (for $V = 1$) for a varying length for the monitoring period $P = 7, 14, 21, 30, 45$ days.

- If an origin AS change is observed while an IP address block is monitored then its TTL (*i.e.,* its monitoring period) is reset. A change of origin AS for an IP prefix can be considered suspicious so we restart monitoring the prefix, for P days, when such an event occurs.
- If an IP address block disappears from the routing tables while it is monitored then its TTL (*i.e.,* its monitoring period) is set to expire the next day as it is unnecessary to continue monitoring an unadvertised network.

We thus ran the two versions of the IP address block selection algorithm on spam data between August 2012 and March 2013. The objective of our simulations is to determine the number of instances N we would need to monitor *all* networks seen in out dataset. We thus consider, in the algorithm, a theoretical monitoring capacity $C = \infty$ and a varying length for the monitoring period $P = 7, 14, 21, 30, 45$ days. The result of these simulations for the algorithm version 1 (resp. version 2) is depicted in Figure 3.3 (resp. Figure 3.4) The left y-axis depicts the number of networks that SpamTracer has to monitor, *i.e.,* the required monitoring capacity, on a *daily* basis. The right y-axis depicts the corresponding number of instances required to monitor these networks. The number of SpamTracer instances is computed based on a daily monitoring capacity of 8,000 networks for a single instance. Recall that we are considering the deployment of instances at different locations, *i.e.,* different viewpoints, hence to obtain the total number of instances N the required number of instances on each plot must be multiplied by the number of viewpoints V.

Counter intuitively, there appears to be almost no difference in the *required monitoring capacity* between the two versions of the algorithm. The required monitoring capacity is only slightly lower for the second version of the algorithm due to the fact that networks stop being monitored if they disappear from the routing tables. Nevertheless, these cases do not occur frequently and thus does not help reduce the

Algorithm 2 IP address block selection (version 2)

Require: Set ι of input IP addresses
Ensure: Set θ of output IP address blocks to monitor
  procedure IPAddressBlockSelectionV2(C, P)
    $\theta = \{\}$
    for all $ip \in \iota$ **do**
      while $|\theta| < C$ **do**
        $p =$ getPrefixFromIP(ip)
        if $p \notin \theta$ **then**
          $\theta = \theta \cup \{p\}$
        else if $p \in \theta$ **then**
          resetTTL(p, P)
        end if
      end while
    end for
    #Beginning of optimizations
    for all $p \in \theta$ **do**
      if getOriginAS($p, today$) $= \emptyset$ **then**
        #If p is not routed anymore then **stop** monitoring it after one day
        setTTL($p, 1$)
      end if
      if getOriginAS($p, yesterday$) $\neq$ getOriginAS($p, today$) **then**
        #If p changes origin AS then continue monitoring it and **reset** its TTL
        resetTTL(p, P)
      end if
    end for
    #End of optimizations
  end procedure

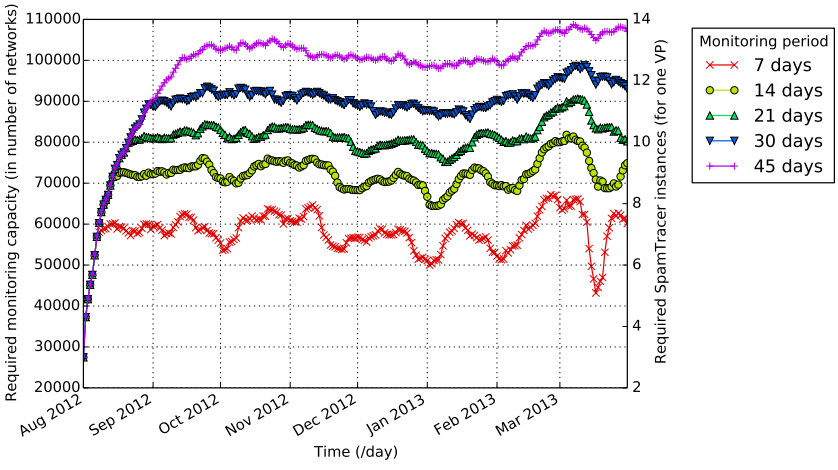


Figure 3.4 – IP address block selection algorithm **version 2**: required daily monitoring capacity C and SpamTracer instances N (for $V = 1$) for a varying length for the monitoring period $P = 7, 14, 21, 30, 45$ days.

amount of networks monitored. We thus use the first version of the IP address block selection algorithm in the current SpamTracer deployment.

Overall, for the two versions of the algorithm, we can see that at least 8 instances per viewpoint are required to monitor all networks for 7 days. This corresponds to the minimal setup. Considering a monitoring period of 45 days (*i.e.,* one and a half months), as many as 14 instances per viewpoint would be required.

Implementation of the distributed architecture.
The proposed solution to address the monitoring capacity limitation of a SpamTracer instance is thus to run multiple instances of it. The proposed solution to address the viewpoints limitation of a SpamTracer instance is to distribute several of them across different, topologically distributed locations. The proposed combined solution to address the two limitations is then to distribute *multiple instances* of SpamTracer across *different locations* by leveraging computing and storage resources in the public Internet cloud, such as Amazon Elastic Compute Cloud (Amazon EC2) [2]. We decided to leverage Internet cloud resources, instead of infrastructures like PlanetLab [24], to benefit from stable and constant hardware and networking resources required by SpamTracer.

From the previous discussion on the sizing of the deployment, we need a address a trade-off between the *monitoring capacity* C, the *monitoring period* P, the *number of viewpoints* V and the financial cost of the Internet cloud resources required to run the SpamTracer instances. Ideally *all networks* should be monitored from at least *3 viewpoints* (*e.g.,* U.S., Europe and Asia) so we consider $V = 3$. Then, we also want to increase the monitoring period to $P = 30$ days. In order to monitor all networks in our dataset, we derive from our simulations and the Equation 3.1 $N = \frac{100{,}000 \times 3}{8{,}000} = 37.5$ instances. Unfortunately, the cost of deploying so many instances was prohibitive. Instead of reducing the number of viewpoints V or the monitoring period P, we chose a daily monitoring capacity $C = 40,000$ instead of $100,000$ which is required to monitor all networks. The number of instances required thus becomes $N = \frac{40{,}000 \times 3}{8{,}000} = 15$, which consequently gives 5 instances per vantage point.

Resource	**Requirement (for ∼8K networks/day)**	**Amazon EC2 VPS config.**[4] **(for ∼40K networks/day)**
CPU	Single-core 1.5GHz	Equivalent to two cores of an Intel Xeon E5-2670 2.6 GHz
RAM	1GB	15 GB
Data storage	65MB/day or 25GB/year	1.68TB
Network I/O	3.3GB/day or 1.2TB/year	10.8TB/day
Operating system	Linux	Linux
User account type	Administrator (root)	Administrator (root)

Table 3.2 – SpamTracer system requirements vs. Amazon EC2 VPS configuration.

Based on the SpamTracer system requirements, detailed in Table 3.2, for the monitoring of 8,000 networks per day, with one IP address monitored per network,

we decided to deploy **three** SpamTracer instances on **three** Amazon EC2 virtual private servers (VPSs) located in the US-East, Europe-Ireland and Asia-Singapore regions. The system configuration of each Amazon EC2 VPS, also detailed in Table 3.2, enables to execute the equivalent of **five** original SpamTracer instances and monitor a total of 40,000 networks per day, with one IP address monitored per network. Thus the large scale deployment of SpamTracer via the distribution of traceroute measurement nodes (instances) in the public Internet cloud enabled us to increase the monitoring period P to 30 days, the number of traceroute viewpoints (or collectors) V to 3 and the monitoring capacity C to 40,000 networks per day and per viewpoint.

3.2.1 Routing data collection

The SpamTracer approach is based upon the first observation of the "BGP spectrum agility" phenomenon reported by Ramachandran et al. in [148] where spammers hijacked previously unannounced IP address blocks for a short period of time (*i.e.,* less than a day) to launch spam campaigns. The assumption behind this approach is thus that when an IP address block is hijacked for the purpose of performing other malicious activities from it for a short period of time then a routing change will be observed when the block is released by the hijacker to remain stealthy. Since we start monitoring a network when we observe malicious activities from it, we look for a routing change from the hijacked state of the network to the normal state of the network. The goal here is not to build a stand-alone BGP hijack detection system but instead to collect, in real time, routing data associated with malicious networks in order to identify cybercriminals operating from temporarily (*i.e.,* less than one day) hijacked IP address blocks. In the remainder of this section we describe the different parts of our experimental environment in more details.

The *data collection* module of SpamTracer (① in Figure 6.1) is based on a linear data flow where a feed of IP addresses to monitor is given as input and a series of enriched traceroutes and BGP routes are produced as output from which routing anomalies can be uncovered. The feed mainly consists of IP addresses which were used to send spam in the last hour to Symantec.cloud spamtraps [35]. Because the spam feed consists of around 3,500,000 spam emails per day, a sampling is performed and around 40,000 IP addresses are tracerouted every day. Bogon prefixes (unallocated or reserved IP blocks) seen originating spam are automatically selected for monitoring as they represent unused IP space that spammers may have hijacked. Building the AS-level routes enables to look at network routes from the same perspective as BGP, which matters when studying IP prefix hijacking. The IP-to-AS mapping is performed using live BGP feeds from six RouteViews [43] servers which are distributed worldwide. The view of the routing in the Internet can differ from one location to another so geographic distribution of BGP as well as traceroute collectors is important. The BGP AS paths from the BGP collectors to the monitored

[4]The system configuration of the Amazon EC2 VPS (instance type m1.xlarge) was inferred from system configuration details provided by Amazon (http://aws.amazon.com/ec2/previous-generation/ and http://aws.amazon.com/ec2/instance-types/) and from real-world performance tests (http://www.pythian.com/blog/virtual-cpus-with-amazon-web-services/).

networks are also collected. Finally, further information is collected on the monitored networks and the different IP hops and ASes traversed (*e.g.,* geolocation [15], `whois` [37], allocation status [38]).

IP address block selection

The main input of the SpamTracer framework are lists of IP addresses that should be monitored. For each feed the duration of the monitoring period in number of days can be set depending on the feed profile, *e.g.,* networks likely hijacked in the future would be assigned long monitoring periods. The IP address feeds currently used in SpamTracer along with their contribution to the daily number of monitored networks are detailed in Table 3.3.

Feed	Description	Contribution (%)
symantec.cloud	Hosts sending spam to Symantec.cloud spamtraps	80.0
shadowserver	C&C servers (source: Shadowserver [32])	3.0
spamhaus drop	Networks allegedly hijacked by cybercriminals (source: Spamhaus [41])	8.0
dshield	Malicious hosts (source: DShield [13])	3.0
russian business network	Hosts identified as belonging to the RBN cybercriminal organisation (source: emergingthreats.net [14])	3.0
malware domain list	Malicious hosts (source: Malware Domain List [127])	3.0

Table 3.3 – Feeds of IP addresses of networks originating malicious network traffic used as input to SpamTracer. The contribution refers to the proportion of each feed in the ∼40,000 daily monitored networks.

Our primary dataset is a live feed of spam emails collected at spamtraps. Every day we receive about 3,500,000 spam emails from about 24,000 distinct IP address blocks. The spam feed is updated on an hourly basis. The other feeds are updated on a hourly or daily basis based on the frequency at which the feed providers update them, *e.g.,* Spamhaus publishes a new version of the DROP list every day. Due to the overhead imposed by traceroute measurements and by querying the BGP collectors, our system can currently monitor about 40,000 IP address blocks on a daily basis. A random sample of IP address blocks is extracted from each feed every hour. Prior to extracting new IP address blocks to monitor, we map individual IP addresses in each feed to their IP address blocks currently announced in BGP using archived routing information bases (RIB's) from RouteViews and RIPE RIS. Because we monitor each block for 30 days, ∼1,300 ($= \frac{40{,}000}{30}$) new IP address blocks are added to the system everyday. When selecting blocks to monitor we prioritize the *recently announced* ones as they are good candidates for short-lived hijacks as suggested in [148]. We consider to be *recently announced* any IP address block in our spam dataset that became routed within the last 24 hours, based on archived routing information bases (RIB's) from RouteViews and RIPE RIS. It is noteworthy that only a handful of IP address blocks are identified as recently announced on daily basis, out of the ∼1,300 newly monitored IP address blocks. This, thus, only marginally biases the random sampling.

Traceroute and BGP monitoring

IP-level traceroute. A customized version of the classic traceroute function is used and is implemented in Python using the packet manipulation library Scapy[5]. For each destination host, 30 probe packets with incremented TTLs starting at 1 up to 30 are sent. Probe packets are sent in parallel to speed up the process. The base probe packet type is ICMP but when no reply is received for a given TTL, a second round is performed using UDP (port 33435) probe packets. For TTLs from which still no reply was received at the second round using UDP, TCP (port 80) probe packets are used for a third round. Using different types of probe packets aims at increasing the likelihood of at least one probe to reach the destination host. Moreover, according to Bush et al. [64], ICMP probes are the most likely to reach their destination host, followed by UDP and TCP probes, in that order. The port numbers for UDP and TCP probe packets were chosen based on the study performed by Luckie et al. in [122] where the authors observed that packets destined to these ports were more likely to reach their destination host. For each round for a given TTL, three probe packets are sent before trying with the following probe type or giving up with the TTL.

Paths uncovered using traceroute may have holes (*i.e.,* unresponsive hosts along the path) where no ICMP reply packet was received for some TTLs. Also, when no reply is received from a destination host, several IP addresses in the destination IP prefix are "pinged" to find a reachable host in the same network. Such technique allows to record a traceroute path that is more complete than the previous one and that still reaches the network IP prefix and AS of the original destination host.

AS-level traceroute (IP-to-AS mapping). Due to the many artefacts that can be found in IP-level routes uncovered using traceroute, studying anomalies in the Internet routing infrastructure using only such routes is a complicated task. Looking at the AS-level (i) enables to look at network routes from the same perspective as BGP which matters when studying IP prefix hijacking and (ii) hides some artefacts of IP-level routes by looking at the network from a higher-level view, *e.g.,* load-balancing inside ASes.

The IP-to-AS mapping is performed using live BGP data queried from six Route-Views [43] route servers distributed worldwide (on five continents). Moreover, these six route servers aggregate the majority of all RouteViews peering ASes. The route servers' BGP routes to the monitored network blocks are queried from their BGP route table (RIB) via `telnet`. Because traceroute is a live measurement and to enable the AS-level path to be as accurate as possible, it is important that each IP host is mapped to the AS announcing its IP prefix at that moment. Also the view of the routing in the Internet can differ from one location to another so geographic distribution of BGP collectors is important. Each IP-level hop is mapped to its IP prefix and the AS originating this prefix, as seen by the different BGP collectors. The BGP AS path as well as other BGP-related information related to the monitored

[5] http://www.secdev.org/projects/scapy/

network is also collected.

IP/AS-level traceroute enrichment. Further registration information extracted from IRRs [20] and geolocation information obtained from Maxmind [15] is collected for the monitored network and the different IP- and AS-level hops traversed:

IP-level hops information: Information collected about the IP-level hops traversed by traceroute paths includes the *domain name* and the *IP-level geolocation*.

AS-level hops information: Information collected about the ASes traversed by traceroute paths includes the *ASN*, the *AS-level geolocation*, the *AS allocation date* and *registry (RIR)*, and the *AS owner*.

Monitored IP address block information: Information collected about the monitored network includes the *IP address block allocation date* and *registry (RIR)*, the *IP address block owner*, and the presence of the IP address block in the *Team Cymru Bogon list (reserved or unallocated IP blocks)* [38] and the *Spamhaus DROP list* [41].

3.2.2 Multi-stage scoring and data filtering

As the amount of IP address blocks to monitor increased significantly over time, we needed a mechanism to automatically investigate them. This is why we have designed a multi-stage scoring and filtering system that analyses the raw data, identifies abnormal routing events, assigns individual scores based on a consistent set of criteria, and then aggregates all scores to eventually highlight IP blocks most likely indicating possible BGP hijacks. We describe here the main components of this multi-stage scoring and filtering system. We introduce a set of novel heuristics to identify abnormal routing behaviors from the routing data collected about the monitored networks to find cases likely resulting from a malicious BGP hijack. For this task, the following data features are available for each monitored network for a period of $n = 30$ days following the observation of malicious activities from the network:

- The sequence $T = \{t_{c,d} | c = 1, \ldots, 3; d = 1, \ldots, n\}$ of daily IP/AS traceroutes from our three traceroute vantage points to the network over a period of n consecutive days;
- The sequence $B = \{b_{c,d} | c = 1, \ldots, 6; d = 1, \ldots, n\}$ of daily BGP AS paths from six RouteViews servers to the network over a period of n consecutive days;
- The registration and geolocation information of IP- and AS-level hops in traceroutes.

The *data analysis* module of SpamTracer (② in Figure 6.1) is depicted in Figure 3.5. It takes as input the collected data about a network and gives as output

the degree of suspiciousness of the routing behavior of the analysed network. Our routing anomaly detection and analysis technique is based on two assumptions: (i) the routing anomalies must be observed by at least one of the six RouteViews servers and (ii) provided that they are impacted by the routing anomalies, the traceroutes provide the required input data to assess the suspiciousness of the routing anomalies detected in BGP.

In the remainder of this section we describe the Multi-stage scoring and data filtering module (② in Figure 6.1) including the Traceroute and BGP anomaly scoring (2.1 in Figure 6.1) and MCDA-based Multi-stage aggregation and filtering (2.2 in Figure 6.1) sub-modules. The Multi-stage scoring and data filtering module is depicted in Figure 3.5. This module aims at identifying candidate BGP hijacks based on routing anomalies uncovered from the collected traceroute and BGP data.

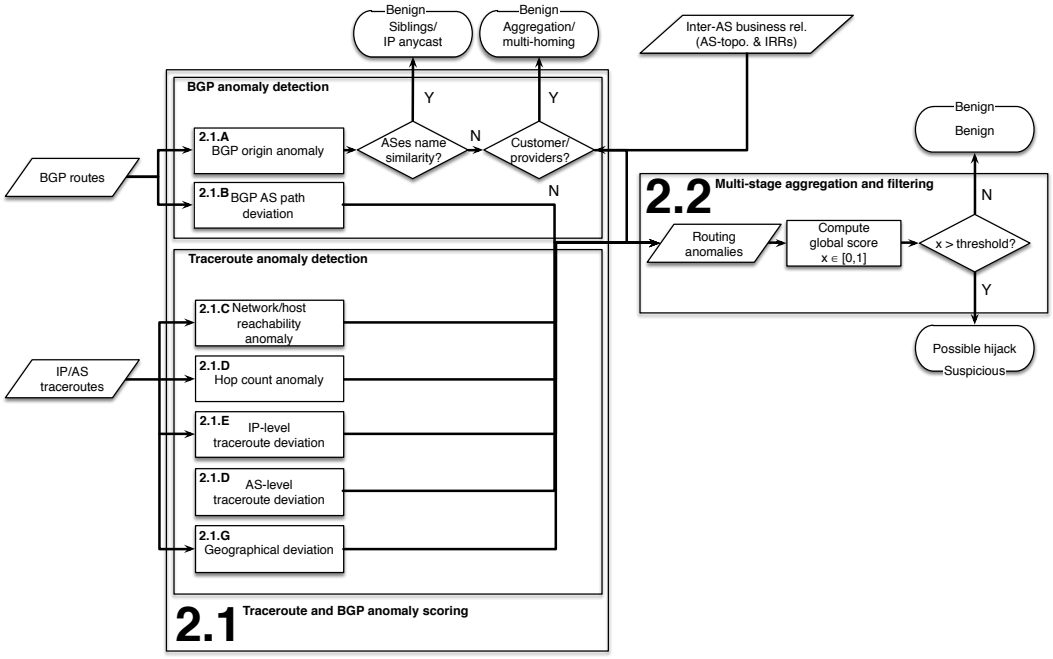


Figure 3.5 – SpamTracer: Multi-stage scoring and data filtering.

Traceroute and BGP anomaly scoring

BGP anomaly detection.
BGP anomalies provide a view from the control plane on the routing behavior of monitored networks and are extracted from the set of daily BGP AS paths.

(2.1.A) BGP origin anomaly: refers to an IP address block being announced by more than one AS. This anomaly is also commonly referred to as a Multiple Origin AS (MOAS) conflict. However, BGP engineering practices like aggregation and multi-homing may introduce a legitimate MOAS situations [96]. The definition of a MOAS states that a *single* IP prefix is originated by multiple ASes [192]. We identify the following three BGP engineering practices introducing MOAS conflicts:

(i) IP space advertised by a (single or multi-homed) customer and by one of its providers (aggregated or not). This corresponds to the "peering MOAS" introduced in [105].

(ii) IP space advertised by multiple ASes owned by a single organisation. Such ASes are commonly referred to as "sibling ASes", per the classification of the inter-AS business relationships model introduced by Gao in [75];

(iii) IP anycast addressing (with multiple origin ASes) [192];

As described in [96], other BGP engineering practices exist, including, for example, multi-homed networks using a static link with one provider or using a private AS number. In the first BGP practice (i) described here above, the conflicting ASes have a customer-provider relationship so they are direct neighbours in the AS path. Recently, Jacquemart et al. showed in [105] that such "peering MOASes" amount for more than 70% of all MOASes observed on the global Internet. We identify this BGP engineering practice by extracting customer-provider relationships from the BGP AS paths collected by SpamTracer, from a daily AS-level Internet topology providing business relationships between ASes available at [18] and from the routing policies published in Internet Routing Registries (IRRs). In the BGP practices (ii) and (iii), the conflicting ASes usually belong to the same organisation, *e.g.,* AS20940 "Akamai Technologies European AS" and AS21342 "Akamai Technologies AS". We detect such BGP practices by measuring the similarity between the owner name of conflicting ASes. We use the *Levenshtein distance* [6], also called the *edit distance*, between the AS owner names to assess their similarity. This distance metric is commonly used to the assess the similarity between two sequences, in particular character strings. Examples of the similarity between AS owner names are given in Table 3.4 (note that we normalize the computed Levenshtein distance to obtain a value in $[0, 1]$).

ASN$_1$	**ASN$_2$**	**Norm. Leven. dist.** ($\in [0,1]$)	**Siblings?** (< 0.5)
AMAZON-02 - Amazon.com, Inc.	AMAZON-AS-AP Amazon.com Tech Telecom	0.44	Yes
Akamai Technologies European AS	Akamai Technologies AS	0.29	Yes
Canadian Research Network	Network Research Belgium	0.56	No

Table 3.4 – Attributing ASes to a single organisation based on the similarity between the owner name of the ASes.

(2.1.B) BGP AS path deviation: this anomaly measures the amount of overlap between BGP AS paths collected from a given BGP collector to a monitored network. Instead of trying to assess the legitimacy of AS paths changes by looking at the inter-AS relationships like in [144], our approach leverages the AS paths from topologically distributed BGP collectors to detect major routing changes. We detect the anomaly by measuring the similarity between any two consecutive AS paths in the sequence of daily AS paths from each BGP collector individually. We use the Jaccard index[7]

[6] http://people.cs.pitt.edu/~kirk/cs1501/Pruhs/Fall2006/Assignments/editdistance/Levenshtein%20Distance.htm

[7] The Jaccard index J of two sets S_1 and S_2 measures the amount of overlap between the two sets and is defined as $J = \frac{|S_1 \cap S_2|}{|S_1 \cup S_2|}$.

between two sets of elements to compute the amount of overlap between two AS paths. Formally, the BGP AS path anomaly for a network IP prefix p monitored on day d ($d = 1, \ldots, n$) from a BGP collector c ($c = 1, \ldots, 6$) can be expressed as follows:

$$a_{bgpasp}(p, c, d) = \frac{|asp_{c,d-1} \cap asp_{c,d}|}{|asp_{c,d-1} \cup asp_{c,d}|}, d = 2, \ldots, n$$

where $asp_{c,d}$ is the d^{th} AS path collected from the BGP collector c.

Traceroute anomaly detection.

The extraction of routing anomalies from the data plane aims at determining how a routing change observed in BGP impacted the forwarding paths towards a monitored network.

(2.1.C) Network/host reachability anomaly: this anomaly consists in observing a sudden and permanent change of reachability of the monitored network (AS)/host. As reported by Bush et al. [64] the data plane reachability of a network can be dissociated from its control plane visibility due to limited visibility of the BGP collectors used but also to default routing. This conclusion motivated us to measure the reachability of probed networks to characterise their routing behavior. Note that the network/host reachability anomaly consists of two values computed individually at the network and host levels. This anomaly suggests that a major routing change altered the configuration of the monitored network, *e.g.,* due to a blackholing hijack. Formally, the network/host reachability anomaly for a network IP prefix p monitored from a traceroute collector c ($c = 1, \ldots, 3$) is expressed as follows:

$$a_{reach}(p, c) = \begin{cases} 1 \text{ if } \exists d : r_{c,1} = \ldots = r_{c,d-1} \neq r_{c,d} = \ldots = r_{c,n} \\ 0 \text{ otherwise} \end{cases}$$

where n is the number of days in the monitoring period of the prefix p and $r_{c,d}$ is the network/host reachability of the d^{th} traceroute collected from the collector c, *i.e.,* $r_{c,d} = 1$ if we received a reply from the target network/host, and 0 otherwise. This reachability anomaly can result from a major routing change in the path from the source to the destination of the traceroute, which causes the destination host or AS to become (un)reachable. The AS reachability anomaly can indicate that the destination network has been blackholed as a result of a BGP hijack. This conclusion is reinforced if this anomaly happens with multiple hosts within a single AS.

(2.1.D) Hop count anomaly: this anomaly consists in observing an important and permanent change in the length (in number of IP-level hops) of the traceroutes. This situation suggests that a major routing change occurred that permanently changed the forwarding paths. Formally, the hop count anomaly for a network IP prefix p monitored from a traceroute collector c ($c = 1, \ldots, 3$) is expressed as follows:

$$a_{hop}(p, c) = \begin{cases} \frac{\min\{l_{c,1} \ldots l_{c,d-1}\} - \max\{l_{c,d} \ldots l_{c,n}\}}{30} & \text{if } \exists d : \min\{l_{c,1} \ldots l_{c,d-1}\} > \max\{l_{c,d} \ldots l_{c,n}\} \\ \frac{|\max\{l_{c,1} \ldots l_{c,d-1}\} - \min\{l_{c,d} \ldots l_{c,n}\}|}{30} & \text{if } \exists d : \max\{l_{c,1} \ldots l_{c,d-1}\} < \min\{l_{c,d} \ldots l_{c,n}\} \\ 0 & \text{otherwise} \end{cases}$$

where n is the number of consecutive traceroute instances to a given host, $l_{c,d}$ is the number of hops (length) in the d^{th} traceroute collected from the collector c.

(2.1.E) IP-level traceroute deviation: this anomaly measures the amount of overlap between IP-level traceroute paths collected from a given traceroute collector to a monitored network. The similarity between the IP-level paths is computed using the Jaccard index between each consecutive pair of IP-level paths. Formally, the IP-level traceroute path anomaly for a network IP prefix p monitored on day d ($d = 1, \ldots, n$) from a traceroute collector c ($c = 1, \ldots, 3$) can be expressed as follows:

$$a_{trip} = \frac{|ipp_{c,d-1} \cap ipp_{c,d}|}{|ipp_{c,d-1} \cup ipp_{c,d}|}, d = 2, \ldots, n$$

where $ipp_{c,d}$ is the d^{th} IP-level traceroute collected from the collector c.

(2.1.F) AS-level traceroute deviation: this anomaly measures the amount of overlap between AS-level traceroute paths collected from a given traceroute collector to a monitored network. The similarity between the AS-level paths is computed using the Jaccard index between each consecutive pair of AS-level paths. Formally, the AS-level traceroute path anomaly for a network IP prefix p monitored on day d ($d = 1, \ldots, n$) from a traceroute collector c ($c = 1, \ldots, 3$) can be expressed as follows:

$$a_{tras} = \frac{|asp_{c,d-1} \cap asp_{c,d}|}{|asp_{c,d-1} \cup asp_{c,d}|}, d = 2, \ldots, n$$

where $asp_{c,d}$ is the d^{th} AS-level traceroute collected from the collector c.

(2.1.G) Geographical deviation: this anomaly measures the amount of overlap between traceroute country-level paths collected from a given traceroute collector to a monitored network. The assumption behind this anomaly is that country-level traceroutes more likely remain constant even when routing changes occur at the IP or AS levels. The similarity between country-level paths is computed using the Jaccard index. Formally, the traceroute AS-level path anomaly for a network IP prefix p monitored on day d ($d = 1, \ldots, n$) from a traceroute collector c ($c = 1, \ldots, 3$) can be expressed as follows:

$$a_{geo} = \frac{|geop_{c,d-1} \cap geop_{c,d}|}{|geop_{c,d-1} \cup geop_{c,d}|}, d = 2, \ldots, n$$

where $geop_{c,d}$ is the d^{th} country-level traceroute collected from the collector c.

Summary. Every anomaly type is quantified with a score in $[0, 1]$. A BGP origin anomaly is defined by a triplet (IP, AS_1, AS_2) where IP is the monitored IP address block and AS_1 and AS_2 are the ASes announcing IP. In case an IP address block is announced by more than two ASes, several BGP origin anomalies can be produced. Path deviations are computed using the Jaccard index on the sets (p_d, p_{d+1}) where p_d is a path collected on day d and p_{d+1} is a path collected on day $d + 1$. Finally, network/host reachability anomalies and the hop count anomaly are computed once for all traceroutes collected for a network.

A network monitored for n days thus produces (i) zero or more BGP origin anomalies, (ii) $c \times (n - 1)$ path deviations for each anomaly type where c is the number of collectors ($c = 3$ for traceroutes and $c = 6$ for BGP AS paths) and (iii) zero or one network/host reachability and hop count anomalies.

Multi-stage aggregation and filtering

We introduce here a new approach to identify possibly hijacked networks from the set of anomalies extracted from our BGP and traceroute data. We use Multi-Criteria Decision Analysis (MCDA) to design a multi-stage decision-making process and identify the most interesting cases by ranking IP blocks according to their (anomalous) routing behavior. A typical MCDA problem consists to evaluate a set of alternatives with respect to different criteria using an *aggregation function* [62]. The outcome of this evaluation is a global score obtained with a well-defined aggregation model that incorporates a set of constraints reflecting the preferences and expectations of the decision-maker. We compute a global *suspiciousness score* for a given monitored IP address block by applying MCDA, in this case we use the Weighted Ordered Weighted Average (WOWA) operator [172] to perform the aggregation of the different individual anomaly scores. The advantage of this MCDA approach over a traditional decision tree is that it removes the need to define intermediate decision thresholds and allows to identify suspicious routing behaviors likely resulting from a BGP hijack in a more fuzzy way, since the final decision is only made at the end, based on the aggregate value obtained from the fusion of all anomaly scores. Moreover, MCDA provides flexible methods to help model complex decision schemes. It also simplifies the update of the model when removing or adding anomaly types or editing the parameters of the aggregation operator.

MCDA provides an extensive set of methods to model simple to very complex decision schemes, ranging from basic averaging functions to more advanced methods such as fuzzy integrals [62]. In our decision-making system, we rely mainly on the Weighted Ordered Weighted Average (WOWA) operator [172] to aggregate the different individual anomaly scores at various levels. The choice of using WOWA was motivated by a trade-off between flexibility and complexity of the decision model. In fact, WOWA combines the advantages of two types of averaging functions: the weighted mean (WM) and the ordered weighted average (OWA). This enables a decision maker to quantify, with a single operator, the reliability of the information sources (as WM does) but also to weight the individual scores according to their relative *ordering*. This sorting and weighted ordering aspects allow us to emphasize various distributions of scores (*e.g.,* eliminate outliers, emphasize mid-range values, ensure that "at least x" or "most of" the scores are significantly high, etc).

Aggregation functions are used in typical situations where we have several criteria of concern, with respect to which we assess different options. The objective consists in calculating a *combined* score for each option, from which decisions can be made. An aggregation function is defined as a monotonically increasing function of n arguments ($n > 1$) that maps the n-dimensional unit cube onto the unit interval: $f_{aggr} : [0, 1]^n \longrightarrow [0, 1]$. Typical examples of aggregation operators are the arithmetic and weighted mean, however these have very limited decision-modeling capabilities.

OWA extends averaging functions by combining two characteristics: (i) a weighting vector (like in a classical weighted mean), and (ii) *sorting* the inputs (usually in descending order), hence the name of *Ordered Weighted Averaging* [184]. OWA is

defined as:

$$OWA_{\mathbf{w}}(\mathbf{x}) = \sum_{i=1}^{n} w_i x_{(i)} =< \mathbf{w}, \mathbf{x}_{\searrow} >$$

where $\mathbf{x}_{\searrow}$ is used to represent the vector $\mathbf{x}$ arranged in decreasing order: $x_{(1)} \geq x_{(2)} \geq \ldots \geq x_{(n)}$. This allows a decision-maker to design more complex decision modeling schemes, in which we can ensure that only a portion of criteria is satisfied without any preference on which exactly (*e.g.,* "at least" c criteria satisfied out of n). For example, a decision maker might value input vectors that satisfy at least half of a set of criteria well (*i.e.,* have a significantly high score), regardless of how well they perform on the other half. OWA differs from a classical weighted means in that the weights are not associated with particular inputs, but rather with their *magnitude*, and it can thus emphasize the largest, smallest or mid-range values.

It might be useful also to take into account the *reliability* of each information source in the aggregation model, like in Weighted Mean (WM). Torra proposed thus a generalization of both WM and OWA, called *Weighted OWA* (WOWA) [172]. This aggregation function combines the advantages of both types of averaging functions by quantifying the *reliability* of the information sources with a vector $\mathbf{p}$ (as the weighted mean does), and at the same time, by weighting the values in relation to their relative *ordering* with a second vector $\mathbf{w}$ (as the OWA operator). *Weighted OWA* is defined by:

$$WOWA_{\mathbf{w},\mathbf{p}}(\mathbf{x}) = \sum_{i=1}^{n} u_i x_{(i)},$$

where $x_{(i)}$ is the i^{th} largest component of $\mathbf{x}$ and the weights u_i are defined as

$$u_i = G\left(\sum_{j \in H_i} p_j\right) - G\left(\sum_{j \in H_{i-1}} p_j\right)$$

where the set $H_i = \{j | x_j \geq x_i\}$ is the set of indices of the i largest elements of $\mathbf{x}$, and G is a monotone non-decreasing function that interpolates the points $(i/n, \sum_{j \leq i} w_j)$ together with the point $(0, 0)$. Moreover, G is required to have the two following properties:

1. $G(i/n) = \sum_{j \leq i} w_j$, $i = 0, \ldots, n$;

2. G is linear if the points $(i/n, \sum_{j \leq i} w_j)$ lie on a straight line.

When the number of criteria to be evaluated is large, it is generally considered a best practice to organise them in subgroups, which are then evaluated hierarchically. Figure 3.6 illustrates the design of our multi-stage anomaly scoring and aggregation system, in which we organise the aggregation of anomalies in different subgroups based on their semantics. The advantage of a multi-stage aggregation model is that intermediate decision thresholds are not needed. Intermediate aggregate scores are propagated up to the highest level where they can contribute to the overall score.

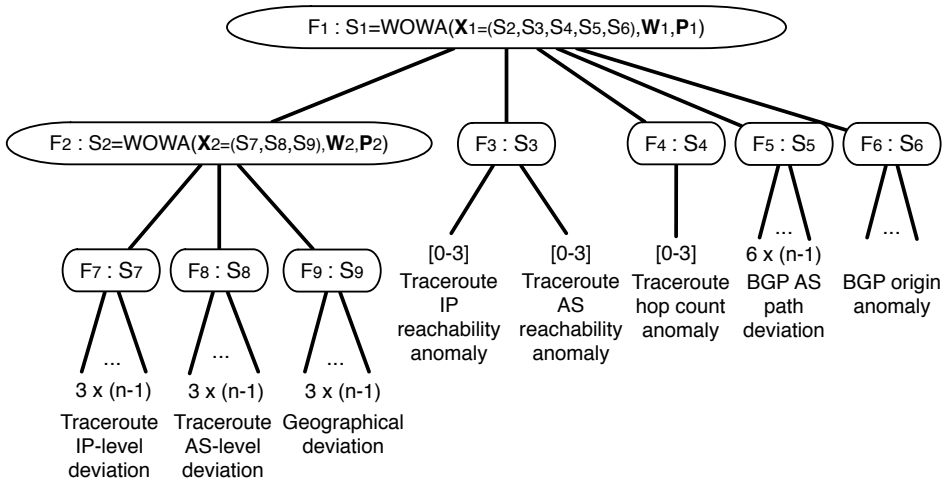


Figure 3.6 – Evaluation and ranking of candidate BGP hijacks: a multi-level aggregation model $\mathcal{M}$ for anomaly scores.

While the use of an aggregation operator like WOWA allows to remove intermediate thresholds, we still need to define the two weighting vectors $\mathbf{w}$ and $\mathbf{p}$ providing respectively the *degree of compensation* between large and small anomaly scores and the *importance* of the anomalies that supply the scores. Given the definitions here above, we define an aggregation function $\mathcal{F}_a$, with as output the aggregated score S_a given by WOWA calculated for the anomaly a as:

$$\mathcal{F}_a : S_a = WOWA(\mathbf{x}_a, \mathbf{w}_a, \mathbf{p}_a)$$

where $\mathbf{x}_a$ is the vector of scores to aggregate and $\mathbf{w}_a$ and $\mathbf{p}_a$ the WOWA weighting vectors. As shown in Figure 3.6, we can then define our multi-stage anomaly scoring and aggregation model $\mathcal{M}$, with as output the final score S_1, for a given monitored network, as the recursive function $\mathcal{F}_a$ where:

$$\begin{aligned}
&\mathcal{F}_1 : S_1 = WOWA(\mathbf{x}_1 = (S_2, S_3, S_4, S_5, S_6), \mathbf{w}_1, \mathbf{p}_1) \\
&\mathcal{F}_2 : S_2 = WOWA(\mathbf{x}_2 = (S_7, S_8, S_9), \mathbf{w}_2, \mathbf{p}_2) \\
&\quad \ldots \\
&\mathcal{F}_9 : S_9 = WOWA(\mathbf{x}_9 = (a_{geo_1}, \ldots, a_{geo_{n-1}}), \mathbf{w}_9, \mathbf{p}_9)
\end{aligned}$$

As an example, we define $\mathbf{w}_1 = (0.5, 0.3, 0.2, 0.0, 0.0)$ and $\mathbf{p}_1 = (0.2, 0.25, 0.15, 0.25, 0.15)$ to obtain the final score S_1 as outcome of the top-tier aggregation stage. Vector $\mathbf{w}_1$ translates here the intuition that a hijacked network does not always exhibit all anomalies (*e.g.,* a hijack does not necessarily involve a BGP origin anomaly) hence we require that "at least some" of the anomaly scores have a high score to contribute to a final aggregate score above a predefined decision threshold. The components of $\mathbf{p}_1$ translate the confidence we have in the different anomaly types to identify a suspicious routing change. The highest confidence score (0.25) is assigned to the traceroute reachability anomaly S_3 and BGP AS path deviation S_5, which by experience have proved being particularly reliable. On the other hand the traceroute hop

count anomaly S_4 and BGP origin anomaly S_6 are assigned a lower confidence score (0.15) because we observed them only in a few rare hijack scenarios. Finally, the traceroute path deviation S_2 is given a medium confidence (0.2) as it can be affected by inaccuracies in traceroute measurements. The model parameter definition is done similarly at the other intermediary stages (for w_i, p_i where $i = 2, \ldots, 9$) so as to include expert knowledge and model the preferences of a network analyst.

Obviously, like any other unsupervised technique (*i.e.,* in absence of reliable "ground-truth" data), a number of parameters must be defined – usually based on the acquired expertise and domain knowledge – to accurately model a decision scheme and ensure that the most relevant cases are ranked in the top tier, whereas truly benign cases are assigned very low scores. In the case of WOWA, we only have to specify two different weighting vectors, which already simplifies considerably the parameter selection phase. This said, it is important to stress that the primary goal of our multi-stage scoring and filtering approach is to narrow down, as much as possible, the number of cases to look at so that we can focus on a limited set of most promising BGP hijack candidates, which can be further validated through manual investigation.

Recall that the ultimate goal is to prove whether (i) "BGP spectrum agility" still exists and (ii) if the modus of BGP hijacking spammers has changed since 2006 [148]. In other words, we try to understand if this is a problem still worth of consideration in 2014, or not. Under these considerations, and without discrediting the importance of parameters selection, we argue that the determination of the *optimal* parameters for our decision model is, at this stage, not critical to achieving our goals. If our decision model proves to be effective at identifying malicious instances of BGP hijacks then we will run a rigorous sensitivity analysis to determine the optimal parameters and further improve our results.

3.2.3 Validation of candidate hijacks

Due to the lack of ground-truth information and the limitations of routing data alone to identify instances of BGP hijacks, an additional validation is required and consists in collecting additional evidence, usually involving some manual processing, about candidate hijacks to help confirm them or not. Moreover, we showed in [177] that, in this validation process, considering multiple and independent data sources, such as BGP and traceroute routing data, spam and other security-related data and IRR data, as well as feedback from network owners is primordial to avoid drawing conclusions biased by a limited set of evidence possibly skewed towards one verdict or the other. We (in)validate candidate hijacks using, besides the collected routing data, daily archives of the following external data sources:

- **Routing Information Base (RIB)** dumps from RIPE RIS [26] and Routeviews [43] consist of snapshots of routers routing table providing the list of announced IP address blocks and associated BGP AS paths.
- **Internet Routing Registry (IRR)** dumps [20] provide registration information on IP address and AS number holders as well as possible routing policies

established between interconnected networks (ASes) (*i.e.,* via BGP `import` and `export` rules as suggested in [61]).

- **Spamhaus Don't Route Or Peer (DROP)** [34] is a blacklist of IP address blocks allegedly controlled by cybercriminals, including some claimed to have been stolen from their legitimate owner.

- **Network operational mailing lists**, such as [22, 27], are sometimes used by network operators to report BGP hijack incidents (*e.g.,* the Link Telecom hijack [46]).

We examine the routing history related to candidate hijacked IP address ranges to study their routing characteristics including (i) when they were publicly announced, (ii) the BGP origin ASes used to advertise them, and (iii) the upstream provider ASes seen in the AS paths. Because our data collection system only collects routing information about IP address ranges for a limited period of time, we built the routing history of candidate hijacked IP address ranges from the archived dumps of routing information bases (RIBs). We use the decision tree depicted in Figure 3.7 to analyse the routing history of candidate hijacks. This decision tree is inspired by the characteristics of hijacks reported by Ramachandran et al. in [148], namely network blocks unannounced prior to being hijacked and only advertised by the hijacker for a limited period of time before being withdrawn.

(1) Not routed during the hijack (old target IP address, false positive): We filter out cases where a monitored network was **not** found to be routed **at the time** of the identified hijack. We found that some target IP addresses were present in our input IP address feeds although they had been involved in malicious activities several days to several months before, *e.g.,* IP address blocks listed in Spamhaus DROP remain listed until Spamhaus considers that the case is solved resulting in a large number of blocks being listed without actually being routed.

(2) Not routed after the hijack (disappearance of an old IP prefix, false positive): Networks routed **before** the hijack (for a period of at least three months), **during** the hijack but **not after** the hijack (for a period of at least three months), and exhibiting **no change** in the BGP origin AS are attributed to the disappearance of an old prefix. The routing history of IP address blocks in this category is illustrated in Figure 3.8. t_{hijack} corresponds to the time of the hijack as identified by SpamTracer. Recall from Section 3.2.1 that we start monitoring a network when we observe malicious activities from it so t_{hijack} corresponds to the routing change from the hijacked state of the network to the normal state of the network. The disappearance of an IP prefix from the routing tables occurs regularly as the amount of IPv4 address space advertised on the Internet constantly fluctuates[8]. Although the advertised IPv4 address space tends to grow with time, for instance, as of September 1^{st} 2014, there are an equivalent of 219,455 more /24's advertised than on January 1^{st} 2014 and 377,998 more than on January 1^{st} 2013. Nevertheless, IP address blocks appear and disappear from the routing tables regularly. For instance, over the course of

[8]Based on the archived IPv4 resource reports from http://resources.potaroo.net/iso3166/v4cc.html.

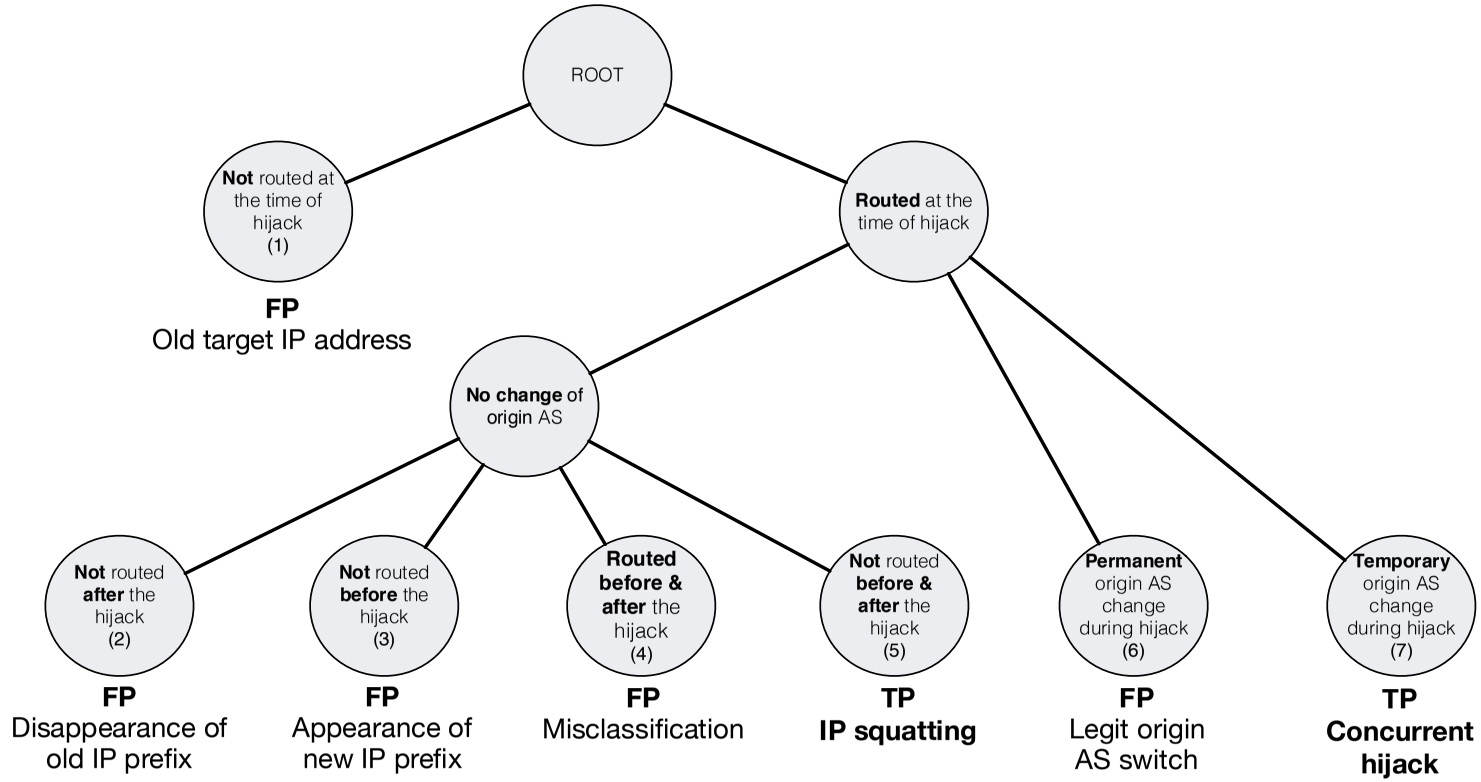


Figure 3.7 – SpamTracer: Validation of candidate hijacks via the routing history.

September 2014, although the amount of advertised IPv4 address space increased between the beginning and the end of the month by an equivalent of 28,899 /24's, on nine days we observed a decrease of the advertised IPv4 address space with a maximum of 1,126 less advertised /24's between September 15^{th} and 16^{th}. Moreover, a network in this category does not exhibit any change of BGP origin AS at the time of the hijack. Thus, for these reasons we flag these matching cases as false positives.

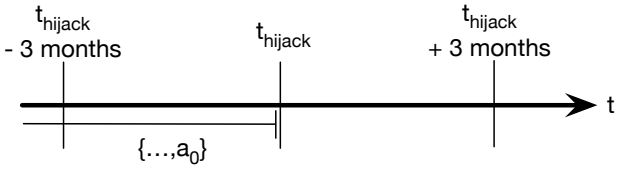


Figure 3.8 – Disappearance of an old IP prefix (false positive): routing history of IP address blocks **not** routed **after** the hijack, and exhibiting **no change** in the BGP origin AS (category (2)).

(3) Not routed before the hijack (appearance of a new IP prefix, false positive): Networks routed **during** and **after** the hijack but **not before** the hijack (for a period of at least three months), and exhibiting **no change** in the BGP origin AS (Figure 3.9) are attributed to the appearance of a new prefix. Similar to the category (2), as new IP prefixes regularly appear in the routing tables and provided they do not exhibit a change of BGP origin AS, we consider cases falling in this category to be false positives.

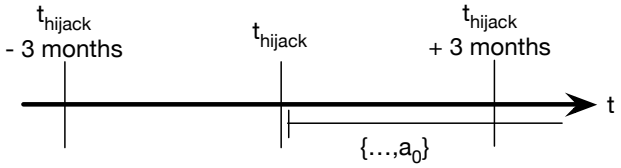


Figure 3.9 – Appearance of a new IP prefix (false positive): routing history of IP address blocks **not** routed **before** the hijack, and exhibiting **no change** in the BGP origin AS (category (3)).

(4) Routed before & after the hijack (misclassification, false positive): Networks routed for a period of at least three months **before** the hijack and three months **after** the hijack, and exhibiting **no change** in the BGP origin AS (Figure 3.10) are attributed to a misclassification by our system and filtered out as well.

(5) Not routed before & after the hijack (IP squatting, true positive): Networks that are found to be routed **only at the time** of the identified hijacks, and exhibiting **no change** in the BGP origin AS (Figure 3.11) are the most similar to the hijacks observed in [148] and are attributed to a hijack of type 7-8 (from our BGP hijack model in Section 2.3.4 of Chapter 2 on page 16). Unlike hijacks of already announced IP address space, *i.e.,* concurrent hijacks (category (7)), this type of hijack involves IP address space that is unannounced prior to being hijacked. In order to clearly differentiate the two phenomena, we refer to this type of hijack as "IP squatting".

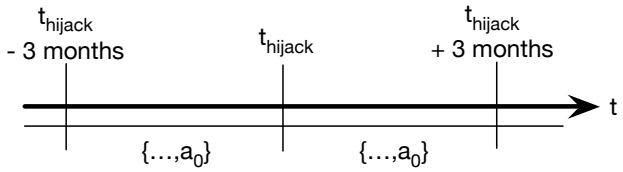


Figure 3.10 – Misclassification (false positive): routing history of IP address blocks routed **before** and **after** the hijack, and exhibiting **no change** in the BGP origin AS (category (4)).

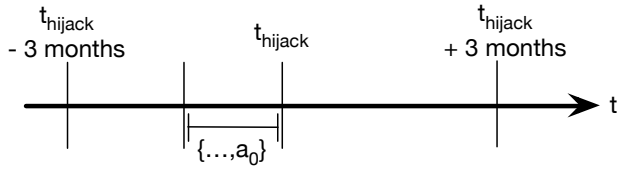


Figure 3.11 – IP Squatting (true positive): routing history of IP address blocks **not** routed **before** and **after** the hijack, and exhibiting **no change** in the BGP origin AS (category (5)).

(6) Permanent origin AS change during hijack (legit origin AS switch, false positive): Networks exhibiting a **permanent** BGP origin AS change (lasting for a period of at least three months) at the time of the hijack (Figure 3.12) are attributed to a legitimate origin AS change.

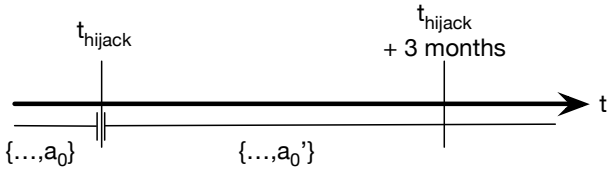


Figure 3.12 – Legit origin AS switch (false positive): routing history of IP address blocks exhibiting a **permanent** BGP origin AS change (category (6)).

(7) Temporary origin AS change during hijack (concurrent hijack, true positive): Networks exhibiting a **temporary** BGP origin AS change (lasting for a period of less than three months) during the hijack (Figure 3.13) are attributed to a hijack of type 1, 3 or 5 (from our BGP hijack model described in Section 2.3.4 of Chapter 2 on page 16). We refer to this type of hijack as "concurrent hijack" due to an IP address block being concurrently advertised in BGP by its legitimate owner and the hijacker.

We leverage IRR dumps to identify the country of registration and, the name and the contact details of the owner of IP address blocks and AS numbers involved in candidate hijacks. We use this information to assess the legitimacy of routing announcements and profile IP address block and AS number holders, *e.g.,* to determine whether the owner of an IP address block is also the owner of the originating AS or to determine whether the owner of an announced IP address block is still in

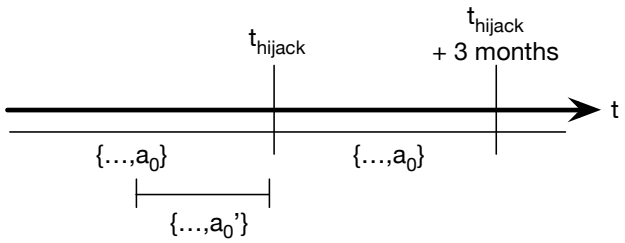


Figure 3.13 – CONCURRENT HIJACK (TRUE POSITIVE): routing history of IP address blocks exhibiting a **temporary** BGP origin AS change (category (7)).

business. As suggested in [159], we further assess the consistency of inter-AS links observed in BGP AS paths using the published routing policies when available. We consider an inter-AS link consistent if both AS refer to each other in their declared `import`/`export` rules.

We use feedback from the Spamhaus DROP list that is a subset of SBL consisting of "IP address blocks that are hijacked or leased by professional spam or cybercriminal operations" [34].

Finally, in order to facilitate the communication among network operators on the Internet, the operational community uses public mailing lists, such as the North American Network Operators' Group (NANOG) mailing list [22] or the RIPE Working Groups mailing lists [27]. We check our candidate hijack cases against reported routing incidents in the archives of these two mailing lists.

At the end of this stage, we should be left with a set of hijack cases that should allow us to confirm or not that BGP hijacks as described in [148] are still ongoing and, if yes, what their characteristics are.

3.2.4 Root cause analysis

While the external cross-validation of candidate hijacks described here above should increase our confidence in the existence of BGP spectrum agility spammers in the real world, we wanted to confirm our results by further investigating the root causes of the validated hijacks from a spam campaign perspective. Assuming we could identify good candidate hijacks that are perfectly matching the anomalous routing behavior of BGP spectrum agility spammers, one would expect that spam campaigns launched from these hijacked networks, by the same group of agile spammers, should intuitively share also a number of commonalities with respect to spam features (*e.g.,* advertised URI's, sender's address, etc).

We have thus used a multi-criteria clustering framework called TRIAGE [168] to identify series of spam emails sent from different hijacked IP address blocks that seem to be part of a campaign orchestrated by the same agile spammers. TRIAGE is a software framework for security data mining that relies on intelligent data fusion algorithms to reliably group events or entities likely linked to the same root cause. Thanks to a multi-criteria clustering approach, it can identify complex pat-

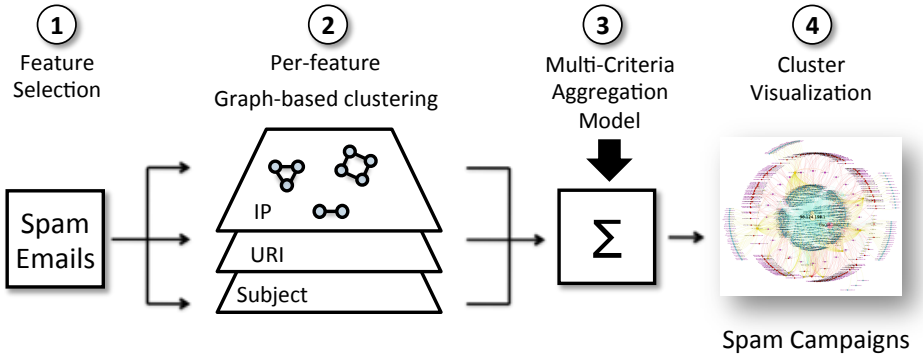


Figure 3.14 – Clustering spam emails sent from hijacked networks using TRIAGE.

terns and varying relationships among groups of events within a dataset. TRIAGE is best described as a security tool designed for intelligence extraction and attack investigation, helping analysts to determine the patterns and behaviors of the intruders and typically used to highlight how they operate. This novel clustering approach has demonstrated its utility in the context of other security investigations, *e.g.,* rogue AV campaigns [68], spam botnets [170] and targeted attacks [169].

Figure 6.2 illustrates the TRIAGE workflow, as applied to our spam dataset. In step ①, a number of email characteristics (or *features*) are selected and defined as decision criteria for linking related spam emails. Such characteristics include the sender IP address, the email subject, the sending date, the advertised URL's and associated domains and `whois` registration information. In step ②, TRIAGE builds relationships among all email samples with respect to selected features using appropriate similarity metrics. For text-based features (*e.g.,* subject, email addresses), we used string-oriented similarity measures commonly-used in information retrieval, such as the Levenshtein similarity and *N-gram* similarity [114]. However, other similarity metrics may be defined to match the feature type and be consistent to analyst expectations (*e.g.,* Jaccard to measure similarity between sets, or a custom IP addresses similarity metric that is based on their relative inter-distance in the binary space).

At step ③, the individual feature similarities are fused using an aggregation model reflecting a high-level behavior defined by the analyst, who can impose, *e.g.,* that some portion of highly similar email features (out of n available) must be satisfied to assign different samples to the same campaign (regardless of which ones). Similarly to the WOWA aggregation method explained here above, in TRIAGE we can assign different *weights* to individual features, so as to give higher or lower importance to certain features. For this analysis we gave more importance to the *source IP addresses*, *domain names* associated to spam URL's and `whois` registration names, since we anticipate that a combination of these features convey a sense and possible evidence of colluding spam activities.

As outcome (step ④), TRIAGE identifies *multi-dimensional clusters* (called MDCs), which in this analysis are clusters of spam emails in which any pair of emails is linked by a number of common traits, yet not necessarily always the same. As explained

in [168], a decision threshold can be chosen such that undesired linkage between attacks are eliminated, *i.e.,* to drop any irrelevant connection that could result from a combination of small values or an insufficient number of correlated features.

3.3 Conclusion

In the first part of this chapter we discussed the different types of data, *i.e.,* control plane and data plane routing data, and logs of malicious activities, that can be leveraged to achieve the first goal of this thesis, which is to determine whether malicious BGP hijacks are still occurring today on the Internet. We showed that using a single dataset is not enough to overcome the intrinsic limitations every dataset has, and is likely to produce a lot of false-positive cases. We thus decided to correlate routing data from the control plane and the data plane with malicious activities logs.

In the second part of this chapter, we presented SpamTracer, a data collection and analysis framework for the large-scale study of malicious BGP hijacks. Based upon the characteristics of the first BGP hijacking spammers involved in "BGP spectrum agility" uncovered by Ramachandran et al. [148], SpamTracer leverages live BGP data and traceroute measurements to characterise the routing-level behavior of networks originating malicious network traffic, in particular spam emails. Due to the large dataset collected by SpamTracer, we needed a system to automatically extract from our data the networks that exhibit an anomalous routing behavior. We thus designed a multi-stage scoring and data filtering scheme using a multi-criteria decision analysis (MCDA)-based method to aggregate routing anomalies scores extracted from the collected data into a global suspiciousness score. The key objective of this analytical step is to reduce the number of suspicious cases to a set of candidate hijacks that can then be manually validated. This extra validation step using external data sources is required by the lack of ground-truth data to confirm or not the maliciousness of the uncovered cases. Finally, spam emails issued by networks associated to validated BGP hijacks are clustered using an external tool called TRIAGE in order to reveal the modus operandi of the attackers. In the next chapter we present the results of our large scale study of malicious BGP hijacks over a period of 22 months.

4

The malicious BGP hijacks phenomenon

Contents

4.1 Routing data collection results 68
4.2 Multi-stage scoring and data filtering results 69
4.3 Validation of candidate hijack results 71
4.3.1 Long-lived hijacks 84
4.3.2 Short-lived hijacks 86
4.4 Root cause analysis results 94
4.5 Summary of findings 96
4.6 Effectiveness of current countermeasures 98
4.6.1 BGP hijack detection 98
4.6.2 BGP hijack prevention 99
4.7 Operationalizing SpamTracer 100
4.7.1 BGP hijack attacks characteristics 101
4.7.2 Real-time blacklist of hijacked networks 102
4.7.3 Real-world deployment 104
4.8 Conclusion 106

In this chapter[1] [2], we tackle the second and third research problems of this thesis and provide an answer to the two associated questions: (i) Do cybercriminals use BGP hijacking as a means to steal blocks of IP addresses and use them to perform other nefarious activities? In other words, we want to determine whether the "BGP spectrum agility" phenomenon introduced in 2006 is still a problem worth of consideration, as of 2014. (ii) If yes, what is the modus operandi used by these attackers to carry out such hijacks and what characterises this phenomenon? In particular,

[1] The experimental results described in this Chapter will be presented at the Network and Distributed System Security (NDSS) Symposium in February 2015 [179].

[2] Preliminary results have been presented at RIPE [175] and NANOG [176] meetings.

we seek to determine how attackers abuse the inter-domain routing infrastructure to hijack blocks of IP addresses so as to identify the appropriate countermeasures to prevent them. We further aim at assessing the frequency and prevalence of such attacks to determine the magnitude of the security threat posed by such a phenomenon.

In Chapter 3, we have introduced SpamTracer, a framework for collecting routing data related to networks having been seen originating malicious network traffic and extracting networks exhibiting an abnormal routing behavior likely resulting from a BGP hijack. We now turn to the description of our results, by detailing step-by-step the outcome of every component of our experimental environmental (see Figure 6.1 in Section 3.2 of Chapter 3). We continue with a thorough investigation and validation of the candidate malicious BGP hijacks we have identified. Finally, we explore the usage of uncovered hijack characteristics to build a real-time blacklist of hijacked IP address blocks to support the detection and mitigation of malicious BGP hijack attacks.

4.1 Routing data collection results

We consider a dataset of BGP and traceroute data collected between September 2012 and June 2014 (22 months). A summary of the dataset is provided in Table 4.1.

Statistic	**Sep 2012-Jun 2014**
Nr. of distinct IP address blocks (with one IP address monitored per block)	649,081
Nr. of distinct ASes	18,807
Nr. of traceroute viewpoints	(STv1) Sep 2012 - Nov 2013 1 (STv2) Dec 2013 - Jun 2014 3
Nr. of traceroutes	8,594,902
Nr. of BGP viewpoints	6
Nr. of BGP AS paths	28,679,725

Table 4.1 – Summary of the BGP and traceroute dataset.

Recall that SpamTracer was designed as an experimental environment to study malicious BGP hijacks. As such, its design evolved over time as we analysed the collected data and gained more insights into the malicious BGP hijacks phenomenon. In its original setup, SpamTracer (v1) used to run as a single traceroute measurement node running on a single host (hosted on the Renater academic network in France, Europe). Between September 2012 and November 2013 (15 months) traceroute measurements were thus performed from a single vantage point. The monitoring capacity was also limited to 8,000 network IP address blocks per day. Motivated by the results obtained from the initial deployment, we undertook the larger deployment of SpamTracer (v2) in the Internet cloud (as described in Section 3.2 in Chapter 3) in December 2013. Data collected between December 2013 and June

2014 (7 months) thus include traceroute measurements performed from three vantage points (hosted on the Amazon EC2 network in Newark - U.S., Dublin - Ireland, and Singapore). The new deployment also included an increased monitoring capacity of 40,000 networks per day. The collection of BGP data in the second deployment of SpamTracer remained unchanged from the first deployment, *i.e.,* BGP routes to each monitored network IP address block were collected from six RouteViews route servers distributed worldwide (located in Eugene - Oregon - U.S., Palo Alto - California - U.S., London - UK, Tokyo - Japan, Sydney - Australia and Sao Paulo - Brazil).

During 22 months we monitored a total of 649,081 distinct IP address blocks which sent spam to our spamtraps or were reported as originating malicious network traffic by one of the other feeds used and detailed in Section 3.2.1 of Chapter 3 on page 48. These networks were operated from 18,807 different ASes. Finally, more than 8.5M data plane measurements and about 28.6M BGP routes towards these networks were collected.

4.2 Multi-stage scoring and data filtering results

Figure 6.3 shows the distribution of scores as computed thanks to the procedure described in Section 3.2.2 of Chapter 3 on page 55 for the monitored networks. The first part of the curve between the score value 0 and approximately 0.25 corresponds to 31.29% of networks exhibiting almost no variability in collected BGP routes and traceroutes. For this reason we consider that they very likely only include benign cases. 68.642% of networks have a score between 0.25 and 0.75. Networks in that category usually exhibit a set of various anomalies, which makes them hard to attribute to a benign or malicious routing behavior. They may suffer from limitations of the aggregation model or from inaccuracies in the collected data [128], which, in the case of a benign routing behavior, mistakenly increases the suspiciousness score and, in the case of a malicious routing behavior, prevents it from being correctly extracted by our scoring system. Finally, 0.068% of monitored IP address blocks have a score higher than 0.75 and correspond to the most interesting candidates to find hijacked networks because they exhibit many midrange to high anomalies scores indicating a suspicious routing behavior.

There are 437 different IP address blocks which exhibit a score higher than 0.75. Each of them appeared in our spam feed or in the other feeds of malicious activities and was monitored only once during the 22 months of the experiment.

We now examine two cases from the extreme categories to determine whether they were correctly identified by our system as either a benign case or a real hijack. We also compare the score obtained using the WOWA aggregation operator and the score that would have been obtained using a naive arithmetic mean. We first consider a network that was assigned a score of 0.23. This network was not hijacked although it exhibited variations in the IP- and AS-level traceroute paths. These varying traceroute paths produced a score $S_2 = 0.67$ in our model. That high score was then mitigated by $S_{3-6} = 0$ giving a final score $S_1 = 0.23$. Using the arithmetic

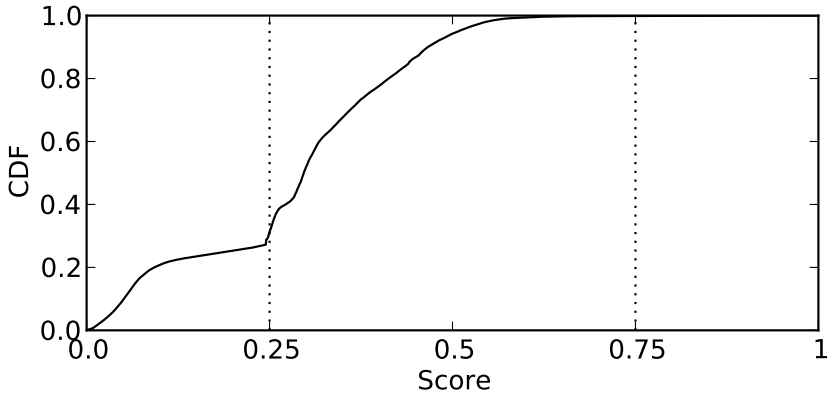


Figure 4.1 – BGP hijack identification: scores between September 2012 and June 2014.

mean instead of WOWA yields a score $S_1 = 0.04$. We now examine a network that scored 0.89. Manual investigation raised our suspicion about this case as the monitored IP address block was found to be routed only during the first four days in the monitoring period of seven days. Several routing anomalies including BGP and traceroute paths deviations were extracted as well as a traceroute hop count anomaly resulting from the significant difference in length between the traceroute path collected on the fourth day, when the network was still routed, and the almost[3] *empty* traceroute path collected on the fifth day, when the network was not routed anymore. The same IP address block would have scored 0.07 using the arithmetic mean as an aggregation function because of the property of this function to even distributions with a lot of low values and a few high values towards the low values, *i.e.,* low values swamp outlying high values.

The two examples above show that we could hardly differentiate suspicious cases from benign ones by naively averaging the scores of routing anomalies. Using a more sophisticated aggregation function such as WOWA we can include expert knowledge into the model by means of the weighting vectors $\mathbf{w}$ and $\mathbf{p}$. While the values of the weighting vectors in our model were set empirically, we leveraged our experience in analyzing routing incidents to ensure that our model can effectively differentiate suspicious from benign cases, even with more "borderline" or ambiguous cases. Moreover, it is noteworthy that the primary goal of our methodology is to narrow down the large number of cases so as to retain a set of interesting BGP hijack candidates. Under no circumstances aim we at designing a new full fledged BGP hijack detection system. Hence, we leave as future work a more extensive sensitivity analysis of the results for various combinations of $\mathbf{w}$ and $\mathbf{p}$, for example by running simulations on synthetic data corresponding to different benign and hijack routing behaviors.

While the number of candidate hijacks uncovered may seem rather low, we need

[3]Traceroute measurements to non routed IP address blocks typically produce *empty*, zero-length, paths. The configuration of the network of the probing host can however introduce a *few* permanent hops in traceroute paths due to intra- or inter-AS default routing.

only one validated case of malicious BGP hijack to confirm the existence of BGP spectrum agility. As we will show later, it is not necessary to be able to identify all hijacks in our dataset to confirm this phenomenon. As a cross-validation test, we did examine manually a number of cases belonging to the different categories (low, midrange and high suspiciousness scores) to be confident that our system correctly scored and ranked them in concordance with their routing behavior.

Additionally, the lack of ground-truth information about the monitored IP address blocks makes it very hard to leverage machine learning techniques to automatically generate a decision tree.

4.3 Validation of candidate hijack results

We leverage here the methodology presented in Section 3.2.3 to (in)validate uncovered candidate hijacks. Due to the large amount of time required to manually investigate cases, we focus our analysis on the 437 IP address blocks that scored above 0.75 in our multi-stage scoring and filtering system, *i.e.,* the upper quartile in the scoring distribution.

As a first step towards the validation, we look at the routing history of candidate hijacked IP address blocks and apply the decision tree described in Section 3.2.3 to the set of 437 cases to identify those most likely corresponding to BGP hijacks. The result is depicted in Figure 4.2.

(1) Not routed during the hijack (old target IP address, false positive): Looking at the 32 IP address blocks that were not routed during the suspicious hijacks we discovered they all come from the Spamhaus DROP list. It turns out that DROP contains several blocks, which are not routed. It appears this is due to the listing policy implemented at Spamhaus that keeps a block in the list as long as the case is not resolved even though the block is not announced anymore. We further investigated the reason why a non routed block could produce routing anomalies such that our system would consider it hijacked. It turns out that these cases result from *default BGP routes*: "a BGP engineering practice allowing routers not able to maintain a full routing table (*e.g.,* due to memory limitations) to have a gateway that is able to handle the default traffic in a more flexible way than with static routes" [88]. As explained in [100], default routes are commonly advertised by an ISP to its (single-homed) customers so that they can route the default traffic to the ISP and have the ISP taking on the advertisement of its customers' prefixes and the proper forwarding of customer traffic. Recall that the RouteViews BGP collectors queried in real time by SpamTracer consists of special-purpose routers collecting BGP routes from their peers without advertising anything. In their study of the effect of BGP route collector placement on the completeness of inferred Internet inter-AS topology maps, Gregori et al. [84] explains that in BGP routing data collection infrastructures such as RouteViews, RIPE RIS or Packet Clearing House it is common for BGP collectors to be treated as customers by the ASes they peer with, *i.e.,* BGP collectors and peering ASes have a customer-to-provider relationship (per the inter-AS economic relationship model introduced in [75]). The reason for this is that

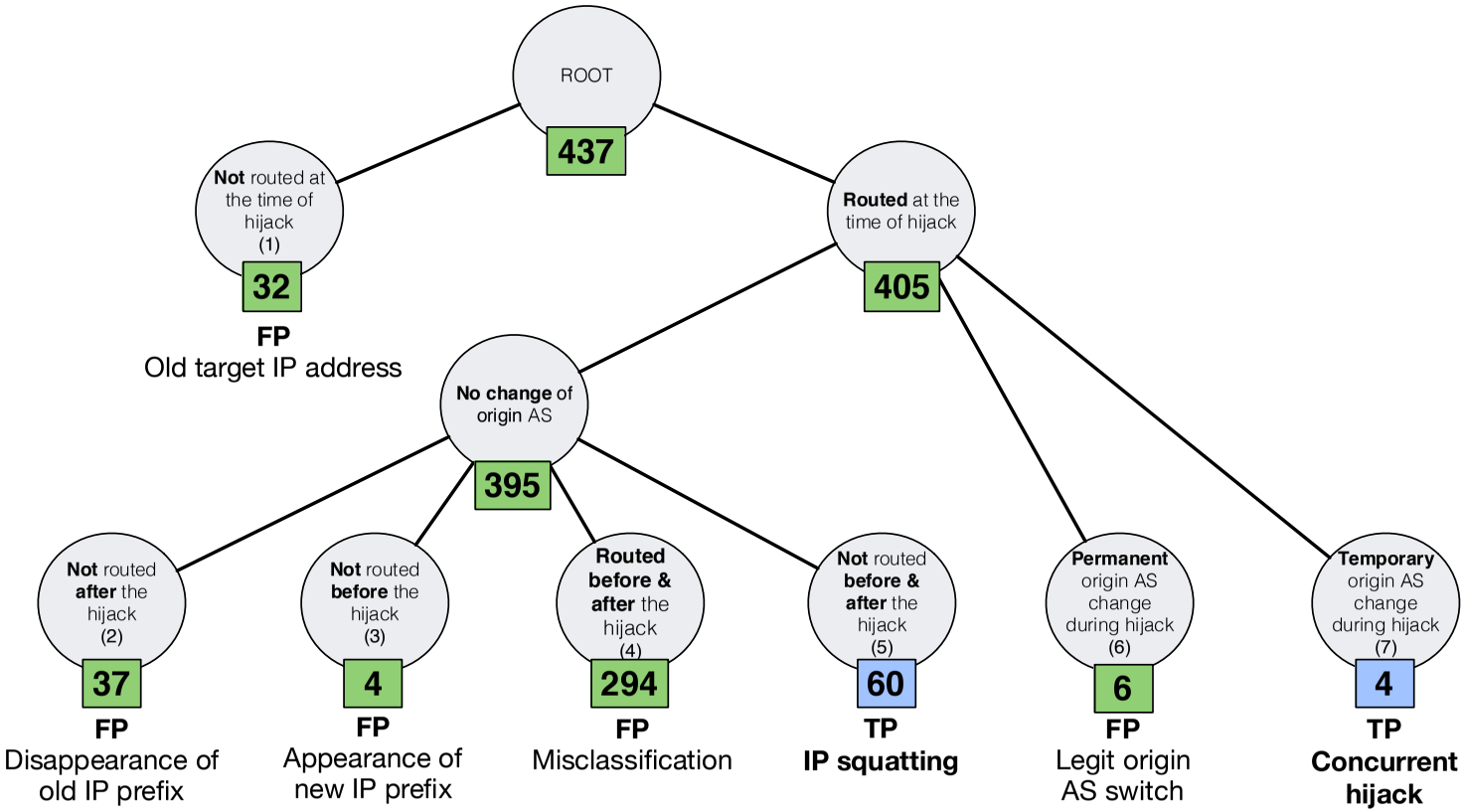


Figure 4.2 – SpamTracer: Results of the validation via the routing history of the 437 candidate hijacks with a score $S \in]0.75, 1]$.

providers typically advertise to their customers all routes received from their other customers, their providers and their peers, which is equivalent to a full version of the providers' routing table. When a BGP collector has a default route and is queried by SpamTracer with an IP address of a non routed block it returns the default route with the `PREFIX 0.0.0.0/0` and `AS PATH {<collector peer>,<collector peer's provider>}`. Each of the 32 IP address blocks that SpamTracer flagged as hijacked exhibited different default BGP routes, which in turn introduced high score BGP origin anomalies (MOAS conflicts) and BGP path anomalies resulting in a high final score. For example, the prefix 85.121.39.0/24, monitored though non routed between 2014-03-06 and 2014-04-04, was associated with default BGP routes from three out of six BGP collectors, which resulted in a final score

$$S_1 = WOWA(\mathbf{x_1}, \mathbf{w_1}, \mathbf{p_1}) = 0.7533$$

with $\mathbf{x_1} = (0, 0, 0, 0.89, 1.0)$, $\mathbf{w_1} = (0.2, 0.25, 0.15, 0.25, 0.15)$ and , $\mathbf{p_1} = (0.5, 0.3, 0.2, 0.0, 0.0)$. In this case the global high score was solely due to high score BGP anomalies since the fact that the block was not routed prevented any traceroute measurement to be performed[4].

(2) Not routed after the hijack (disappearance of an old IP prefix, false positive): The 37 IP address blocks in this group have been flagged as hijacked because they stopped being routed in the course of the monitoring period. From SpamTracer data only, the routing behavior of these IP address blocks is similar to the routing behavior of agile spammers exhibiting "BGP spectrum agility" as described in [148]. Indeed, they appear to be routed from the first day of the monitoring period until they disappear from the routing tables. Due to the lack of historical routing information SpamTracer is not able to distinguish the cases in (2) from those in (5), where the network appears not to be routed prior to being hijacked and monitored by SpamTracer. Using an a posteriori routing history-based validation allows to differentiate them. As explained in Section 3.2.3 of Chapter 3, the disappearance of a prefix from the routing tables occurs regularly, hence can be treated as a benign event. Since no change of BGP origin AS was observed related to these networks, we consider them to be false positives.

(3) Not routed before the hijack (appearance of a new IP prefix, false positive): We observed four IP address blocks that became routed at the beginning of the hijack and remained routed afterwards. We investigated these four cases and found out that they were flagged by our system as hijacked because the networks disappeared from the routing tables for a few hours to several days before the end of the monitoring period and then reappeared after we had stopped monitoring them. This situation resulted in different routing anomalies. We hypothesise this routing behavior is due to *route flap damping* [180]: a mechanism implemented in BGP routers aiming at preventing router overloading by suppressing unstable (flapping) BGP routes responsible for excessive updates. We have no firm evidence to confirm this assumption though. Since none of these networks exhibited a change of BGP origin AS at the time

[4]In SpamTracer traceroute measurements are not performed from the same ASes the BGP routing data is collected from, which explains why some BGP collectors can have default routes while traceroute collectors do not.

of the hijack, we thus consider them to be false positives, mostly to adopt a very conservative approach to decide what a validated BGP hijack should be.

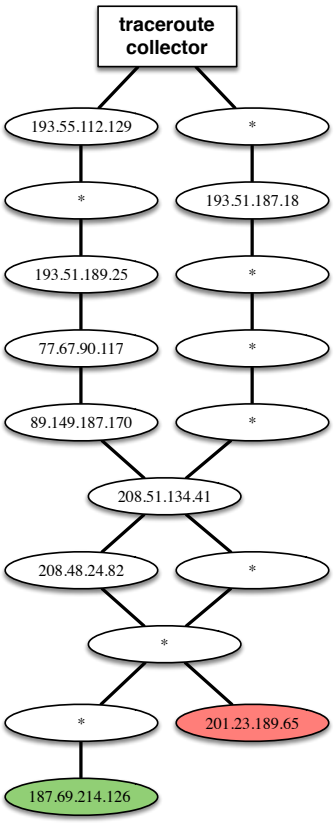
(4) Routed before & after the hijack (misclassification, false positive): With 294 IP address blocks (out of 437 scoring more than 0.75) this group contains the bulk of false-positive cases identified by our system. Based on their routing history, these networks do not have a suspicious routing behavior. They all have been routed for at least three months before and after the identified hijack and they did not exhibit any change in the BGP origin AS. They however all scored higher than 0.75. Manual in-depth analysis of traceroute and BGP data related to some of these cases revealed that they were mistakenly flagged as suspicious due to inaccuracies in traceroute measurements, such as traceroute paths cluttered with many non-responsive IP-level hops ("*"). We noticed that many of these IP address blocks exhibited traceroute paths of 10 or less IP hops and BGP AS paths of four or less AS hops (just above the average 3.7 according to [98]). Unfortunately, the Jaccard index we use to compute traceroute and BGP paths anomalies (by measuring the amount of overlap between paths) is greatly affected by the length of the compared paths and will produce a low similarity score for small changes in such paths. As an example, the prefix 187.69.192.0/18 was monitored between 2013-11-28 and 2013-12-04 and was assigned a final score

$$S_1 = WOWA(\mathbf{x_1}, \mathbf{w_1}, \mathbf{p_1}) = 0.7545$$

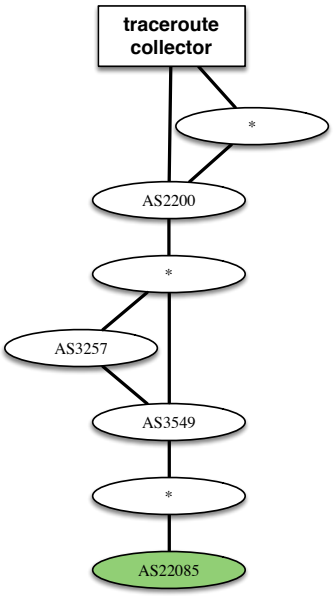
with $\mathbf{x_1} = (0.91, 0, 0.72, 0.5, 0), \mathbf{w_1} = (0.2, 0.25, 0.15, 0.25, 0.15)$ and , $\mathbf{p_1} = (0.5, 0.3, 0.2, 0.0, 0.0)$. The global high score was due to a high score in traceroute paths anomalies (0.91), a high score in the hop count anomaly (0.72) and a midrange score (0.5) in the BGP path anomalies. In this specific case, traceroute measurements were heavily cluttered with non-responsive IP-level hops (as depicted in Figure 4.3). We attribute these cases to a misclassification by our system.

(5) Not routed before & after the hijack (IP squatting, true positive): We uncovered 60 prefixes that appeared to be routed neither before nor after the alleged hijack. This routing behavior looks very similar to the type of hijacks carried out by spammers performing "BGP spectrum agility" reported by Ramachandran et al. [148].

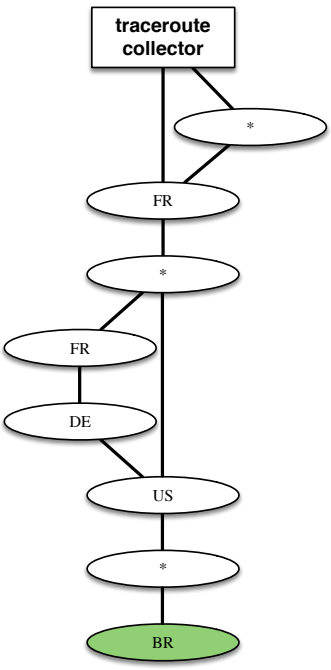
(6) Permanent origin AS change during hijack (legit origin AS switch, false positive): We observed 6 IP address blocks that exhibited a permanent BGP origin AS change at the time of the identified hijack, *i.e.,* the BGP origin AS changed at the time of the hijack and remained the same for a period of at least three months. The high score assigned to these cases result from the BGP origin AS change(s), which introduced subsequent changes in traceroute and BGP paths. In their BGP hijack prevention system pgPGP [108], Karlin et al. consider a BGP origin AS valid if it lasts at least 24 hours. In our case we take a larger margin and consider a change permanent and valid if it lasts at least a period of three months. Recall that the SpamTracer multi-stage scoring and data filtering module (described in Section 3.2.2 of Chapter 3) includes some heuristics to identify BGP origin AS changes that result from a legitimate BGP engineering practice, *e.g.,* sibling ASes announcing the same prefix, IP anycasting, an IP prefix originated by an AS and one of its provider ASes ("peering MOASes"). SpamTracer did not identify any legitimate



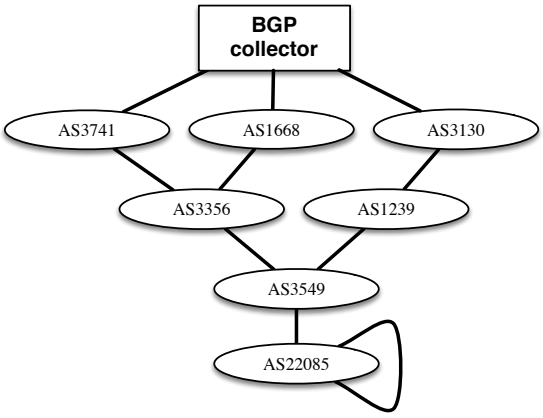
(a) IP-level traceroute paths



(b) AS-level traceroute paths



(c) country-level traceroute paths



(d) BGP AS paths

Figure 4.3 – IP-, AS- and country-level traceroute paths and BGP AS paths from one traceroute and BGP collector to the IP prefix 187.69.192.0/18.

explanation for the BGP origin change in the size cases in this category. After manually analysing these cases, we attribute them to a legitimate switch from one BGP origin AS to another for an IP prefix (*e.g.,* an AS switching from one upstream AS provider to another).

(7) Temporary origin AS change during hijack (concurrent hijack, true positive): A total of four IP address blocks scored more than 0.75 and exhibited a temporary BGP origin AS change observed at the time of the identified hijack. We consider an origin AS change temporary if it lasts for a period of less than three months. The high score assigned to these cases result from the BGP origin AS change(s), which introduced subsequent changes in traceroute and BGP paths. No legitimate explanation for the BGP origin AS change was identified at the SpamTracer multi-stage scoring and data filtering step, hence we attribute these cases to a highly suspicious hijack of already announced IP address space with a new origin AS (hijack type 1, 3 and 5 per the IP prefix hijacking attack model defined in Section 2.3.4 of Chapter 2). After looking closely at the routing history of the four alleged hijacked IP address blocks, they appeared to share a similar routing history pattern, as illustrated in Figure 4.4, *i.e.,* the IP address blocks were originated by an AS a_0 until the beginning of the hijack when they became concurrently originated by the original AS a_0 and the alleged hijacking AS a'_0 for a period of a few hours to several days. Furthermore, the advertisements from AS a'_0 were always for an IP prefix of the same length as the original advertisements from AS a_0 (hijack type 1), *i.e.,* there was no advertisement of more or less specific prefixes (hijack type 3 and 5).

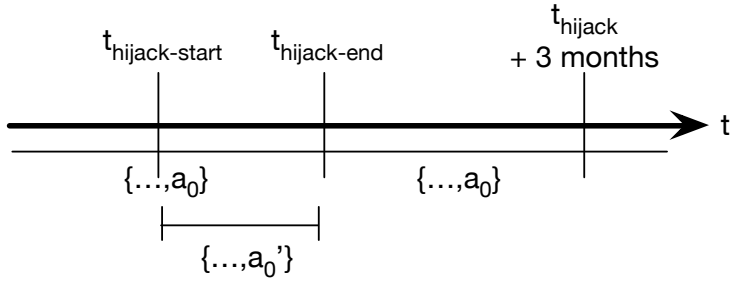


Figure 4.4 – Routing history pattern of the four suspicious hijacked IP address blocks exhibiting a **temporary** BGP origin AS change (category (7)).

After having reviewed all 437 candidate hijacks, it comes, from the previous discussion, that 60 fit the pattern of "BGP spectrum agility" we were looking for[5]. After examining the routing history of these blocks, we could classify them further into two different categories:

- Prefix hijack via valid upstream: In 90% of the hijacks, the IP address ranges were allocated but ***(1) unannounced*** by the time they were hijacked (*i.e.,* left idle by their valid owner), and the attacker forged part of the BGP

[5]Disclaimer: In the remainder of the paper, for the sake of conciseness, we talk about hijacks and attackers instead of candidate hijacks and likely attackers even though we have no bullet proof evidence of their wrong doing.

AS path to advertise the IP ranges using an ***(2) invalid BGP origin AS*** via a ***(3) valid direct upstream provider (first hop) AS***. This class of hijack is illustrated in Figure 4.5a.

- AS HIJACK VIA ROGUE UPSTREAM: In 10% of the hijacks, the IP address ranges were allocated but ***(1) unannounced*** and the attacker forged part of the BGP AS path to advertise the IP address ranges using the ***(4) valid BGP origin AS*** but via an ***(5) invalid direct upstream provider (first hop) AS***. This class of hijack is illustrated in Figure 4.5b.

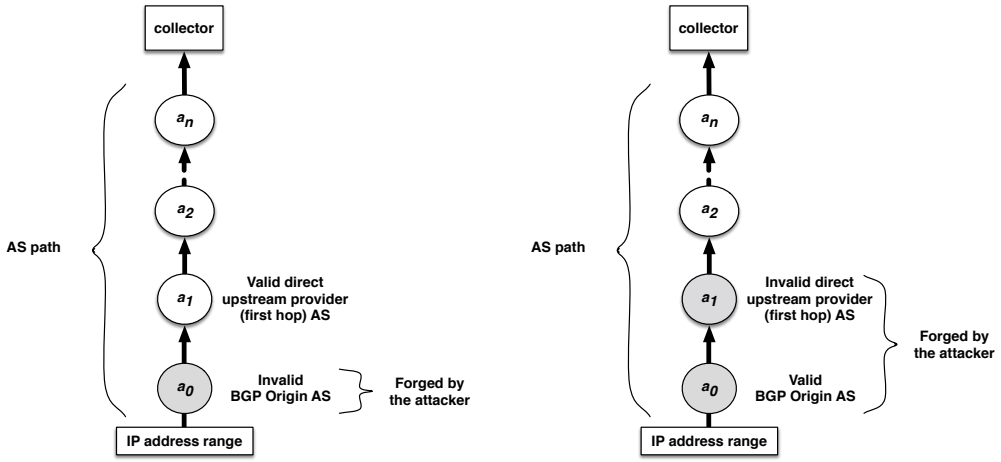
(1) Unannounced IP address space: The routing history revealed that all hijacked prefixes were unannounced before being hijacked.
(2)-(4) (In)valid BGP origin AS: In this work, we consider the origin AS for an IP address range *valid* if the IP address range is mapped to the origin AS in the IRRs (`whois`) and the IP address range owner is also the same as the origin AS owner.
(3)-(5) (In)valid direct upstream provider AS: In this work, we consider as *invalid* the AS a_1 appearing as the direct upstream provider of the origin AS a_0 in the BGP AS path $\{a_n, ..., a_1, a_0\}$ if all the following conditions are met: (1) it has never been used as a direct upstream provider AS for a_0 in the past, (2) it does not appear in the list of provider ASes of a_0 and does not have a_0 in the list of its customers (*i.e.,* `imports`/`exports`) published in the `whois` when such information was provided, (3) it is not used as an upstream provider to advertise any non hijacked IP address range at the time it is observed in the hijacks, (4) it is unused when it is observed for the first time in hijacks, (5) its holder refers to an inactive organisation, and (6) it has, at some point in time, been reported as suspicious by Spamhaus[6].

In the AS hijack cases, it thus appears that attackers actually forged part of the BGP AS path ($\{a_1, a_0\}$) by unauthorisedly using a_1 and a_0 in the BGP announcements for the different hijacked IP prefixes. BGP hijacking using a forged AS path is a stealthy BGP hijack technique [96, 154] and was probably used by the attackers in an effort not to raise suspicion.

We further observed that the 64 hijacked IP address blocks were advertised from only eight invalid distinct BGP origin ASes (prefix hijacks) and via only three invalid distinct upstream ASes (AS hijacks)[(7)]. Based on this observation we used the archived routing information bases (RIBs) from RouteViews and RIPE RIS to extract all IP address ranges originated by the same eight invalid BGP origin ASes or advertised via the same three invalid upstream provider ASes during the same 22 months time period. We applied the candidate hijack validation methodology as we did for the 437 initial cases uncovered by SpamTracer. Surprisingly no less than 2,649 additional hijacked IP ranges were uncovered! All of them have the exact same hijack signatures as the other 64 IP address blocks. The validation via the routing

[6]Spamhaus SBL records related to some identified hijacked IP address blocks are available at `http://www.spamhaus.org/sbl/query/SBL<record_id>`: 96354, 175835, 177177, 177452, 177570, 179312, 180606, 182044, 182223, 182351, 183715, 183836, 184596, 184865, 185726, 185728, 217199. Note that records are purged when the cases are considered to be solved.

[7]Disclaimer: IP address blocks and ASes were likely abused in hijacks between September 2012 and June 2014 and, therefore, might now be legitimately used.



(a) **Prefix hijack** via valid upstream. (b) **AS hijack** via rogue upstream

Figure 4.5 – Categories of uncovered hijacks

history revealed that 13 IP address blocks out of the total 2,713 (64 identified by SpamTracer + 2,649 additional) were involved in **concurrent prefix hijacks** and exhibited a change of BGP origin AS and were temporarily originated by one of the eight previously identified invalid BGP origin ASes. Recall from our BGP hijack attack model in Section 2.3.4 of Chapter 2 that the hijack of announced IP address space is a more devious attack than the hijack of unannounced IP address space due to the disruptions likely resulting from network traffic related to the hijacked addresses being diverted to the hijacker's network. Moreover, all 13 IP address blocks we observed in such hijacks involved concurrent BGP announcements for the *same* prefix (hijack type 1 in our BGP hijack attack model), *i.e.,* during the hijack each prefix was announced by both its legitimate owner and the hijacker resulting in a Multiple Origin AS (MOAS) conflict. The effectiveness of these hijacks was thus low due to the concurrency between the legitimate and the rogue BGP announcements. As an example, the IP address block 196.1.114.0/24 was observed as being hijacked between February 13 2014 05:22:26 and February 14 2014 16:21:51 using the invalid BGP origin AS57792 "PE Glazunov Yuriy Anatol'yevich" (Russia) via the valid direct upstream provider (first hop) AS9031 "EDPNet" while it was legitimately originated by AS55847 "NKM EDGE Network" (India). Based on data from RIPE RIS [26] the legitimate BGP route was seen before and after the hijack by more than 100 (90%) RIPE RIS collectors' peers. However during the hijack the visibility of the legitimate route drastically decreased so that the legitimate route was only seen by 66 (58%) peers while the hijacked route was visible through 38 (32%) peers. Other instances of suspicious concurrent hijacks have already been observed in the wild [48, 121], supposedly carried out to impersonate services hosted on the hijacked networks, a Spamhaus DNSBL server in [48] and cryptocurrency miners in [121]. Although the 13 hijacks of announced IP address space differ from the other 2,700 hijacks by the hijack type used, they share common traits with some of the other 2,700 hijacks with respect to the time they were carried out as well as

the BGP origin ASes and direct upstream provider ASes used to announce them. For instance, the concurrent hijack of the prefix 196.1.114.0/24 described here above started and ended at the exact same time as nine other non-concurrent hijacks (of unannounced IP address space) where the prefixes were also originated by AS57792 via AS9031. This suggests that the 13 prefix hijacks were likely performed by the same attackers responsible for the other hijacks. Finally, the small number of concurrent prefix hijacks observed suggests that they were either unintentional, possibly resulting from an error in the selection of unannounced IP address space by the attackers, or intentional and purposefully targeting the 13 hijacked IP address blocks. We do not have firm evidence to support either of theses conclusions though. The remaining 2,700 IP address blocks out of the total 2,713 were found to have been involved in **IP squatting** operations, *i.e.,* they were unadvertised by the time they were hijacked. Each of them appeared to have been advertised for a limited period of time by an invalid BGP origin AS or via an invalid direct upstream provider (first hop) AS. While we observed spam coming from the 64 hijacked IP address ranges identified by our system, we did not find any spam sent from the new 2,640 ranges in our spamtrap logs. In the remainder of this section we thus investigate **a total of 2,713 IP address ranges supposedly hijacked between September 2012 and June 2014**. In the remainder of this chapter, for the sake of conciseness and clarity, we refer to the identified cases of concurrent prefix hijacks and IP squatting operations as hijacks.

Figure 4.6 depicts the AS-level topology extracted from BGP announcements of the 2,713 hijacks observed between September 2012 and June 2014. In the context of the depicted AS-level topology, an invalid AS corresponds to the notion of invalid BGP origin AS or direct upstream provider (first hop) AS defined here above. A suspicious (resp. benign) AS refers to an AS that meets some (resp. none of the) conditions of an invalid AS. Additionally, Figure 4.7 shows the timeline of BGP announcements involving the 15 *invalid* or *suspicious* direct upstream provider (first hop) or BGP origin ASes. From Figure 4.6 we can see that three invalid direct upstream provider (first hop) ASes (AS42989, AS49473, AS57792) were involved in the hijack of 227 IP address blocks, mostly (224 out of 227) registered in the RIPE region and marginally (3 out of 227) in the APNIC region. The three upstream ASes also belonged to the RIPE region. All *AS hijacks* we observed were thus carried by advertising the prefixes via one of the three invalid upstream ASes. Figure 4.7 shows that AS hijacks were observed between October 4, 2012, and February 21, 2014. Investigating the AS-level topology related to *prefix hijacks* reveals that eight invalid BGP origin ASes were observed in the hijack of 2,486 IP address blocks distributed among the five Regional Internet Registries (RIRs). Interestingly, IP address space hijacked via invalid BGP origin ASes appears to belong to more diverse regions than address space involved in AS hijacks, *e.g.,* AS57792, AS33688 and AS28490 were involved in prefix hijacks of IP address blocks belonging to various organisations located all around the globe. Unlike AS hijacks, prefix hijacks are observed throughout the period of the experiment, between September 2012 and June 2014. Based on the observed hijacks, AS57792 was prominently used in 669 prefix hijacks of (un)announced IP address blocks between September 11, 2012, and March 15,

2014. It is noteworthy that AS57792, involved in both AS and prefix hijacks, was first observed connected to the transit upstream provider AS9031 "EDPNet" until March 15, 2014, when it disappeared from the routing tables. It then reappeared on June 28, 2014, connected to two new suspicious transit upstream providers AS57756 "PE Gaftkovich Irina Valer'evna" and AS58099 "Mikma Ltd.". Finally, from the AS-level topology, it also appears that AS42989 and AS49473 share a common upstream provider AS57954 "FOP Budko Dmutro Pavlovich". Moreover, AS57954 was found to be a suspicious AS so AS42989, AS49473 and AS57954 may be controlled by the same or by colluding hijacker(s).

Figure 6.4 shows the distribution of the 2,713 observed hijacks across time. We can see that 96.8% of the observed hijacks have occurred after July 2013. From that point the distribution becomes almost uniform, showing that hijacks were performed on a regular basis for more than one year. With an average of 4.06 hijacks per day, we note that BGP hijacks have been an ongoing and recurring threat in the past 22 months (and possibly before).

We now focus on another key characteristic of the identified hijacks: their duration. In [148] Ramachandran et al. reported on spam coming from IP prefixes involved in BGP routing announcements lasting less than one day. Figure 4.9 shows the distribution of hijack durations for the 2,713 validated cases. In our case, 75.6% lasted **less** than one day, from 29 minutes to 23 hours 47 minutes, 92.3% lasted no more than 2 days and 98.1% lasted no more than one week. A large fraction of hijacks are thus similar in duration to those reported in [148], *i.e.,* less than one day. All in all the majority of hijacks are rather **short-lived**. Taking a closer look at the duration of short-lived hijacks depicted in Figure 4.10, we can see that as much as 75.6% last less than one day, typically a few hours. Additionally, 1.9% of hijacks are **long-lived**, *i.e.,* lasted more than one week in our observations, with a maximum duration of 410 days (one year one month and 15 days). Looking at the duration of long-lived hijacks in Figure 4.11 reveals that no less than 66% last at most one month and only 3.8% last more than six months.

Though short-lived and long-lived hijacks share some characteristics, we consider them to be due to two distinct phenomena. In fact, short-lived hijacks can be used by an attacker to circumvent traceback and avoid blacklisting by hopping between IP addresses in a set until the set itself gets blacklisted and then move to another set. Long-lived hijacks however make it harder for the attacker to remain undetected. We show this later when checking the list of hijacked IP address ranges against several blacklists. Such long-lived hijacks have already been observed in the wild, for instance in 2011 a couple of IP prefixes belonging to the company Link Telecom were hijacked for 5 months and used to perform various malicious activities such as sending spam and hosting services but also exploiting remote hosts and originating suspicious IRC traffic [154].

In the remainder of this Section we investigate short-lived ($\leq$ 1 week) and long-lived ($>$ 1 week) hijacks separately to emphasize their similarities and differences. We consider the following characteristics of a hijack event:

- (*C.1*) Whether **spam emails** were received from the IP address range at our

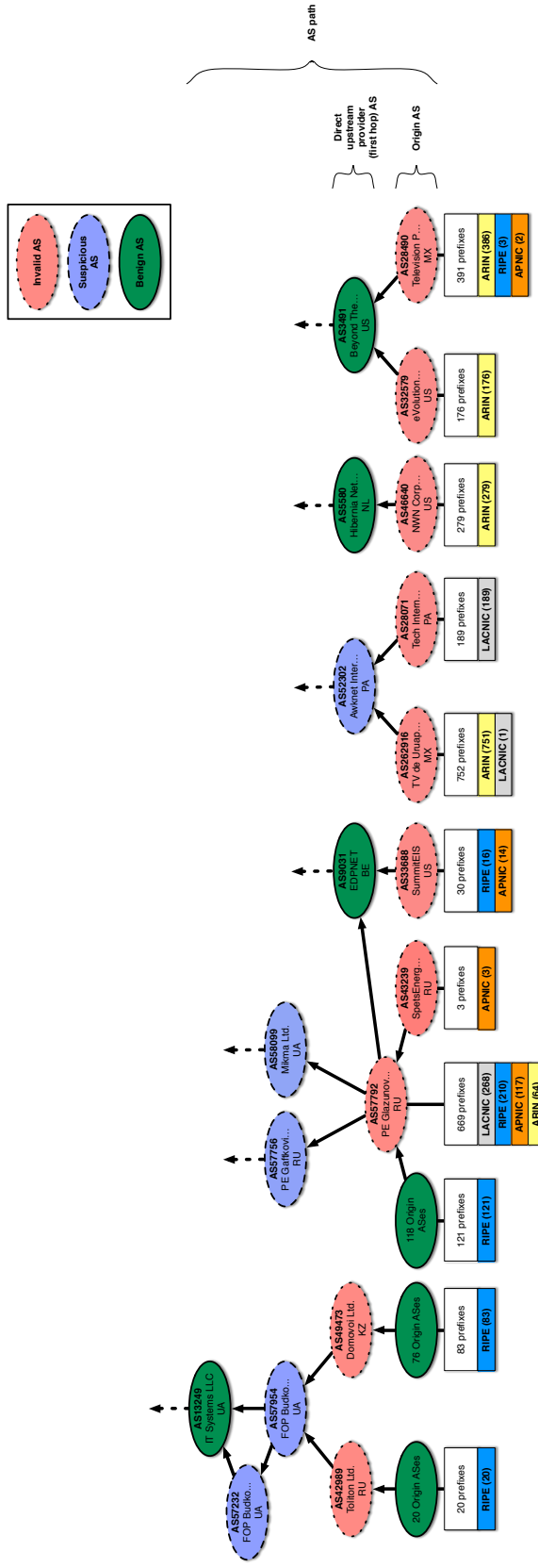


Figure 4.6 – AS-level topology extracted from BGP announcements of the 2,713 suspicious hijacks observed between September 2012 and June 2014. For the sake of readability only relevant ASes are depicted.

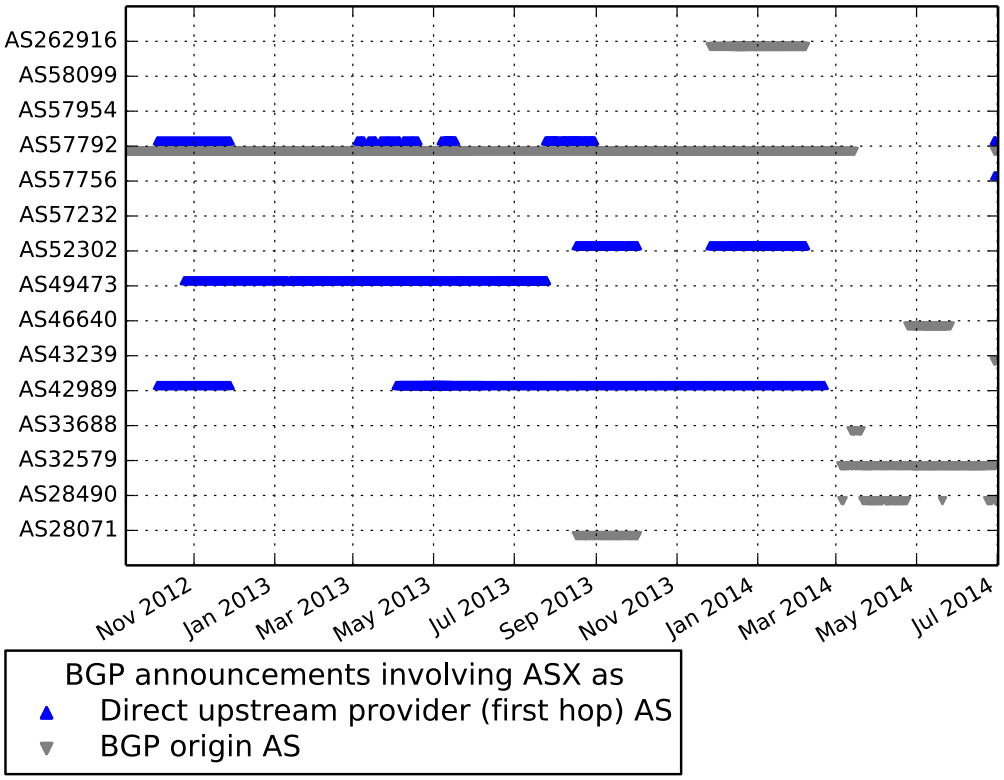


Figure 4.7 – Timeline of BGP announcements involving the 15 **invalid** or **suspicious** ASes as either direct upstream provider (first hop) AS or BGP origin AS. AS57232, AS57954 and AS58099 identified as suspicious (but not invalid) during the hijacks never appear as direct upstream provider (first hop) or BGP origin AS in the BGP announcements for the hijacked prefixes.

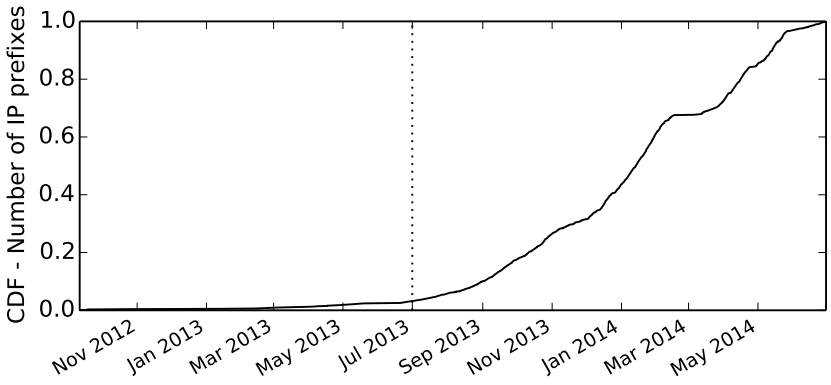


Figure 4.8 – The number of hijacked IP address ranges observed between September 2012 and June 2014. Most of the observed hijacks occurred after July 2013.

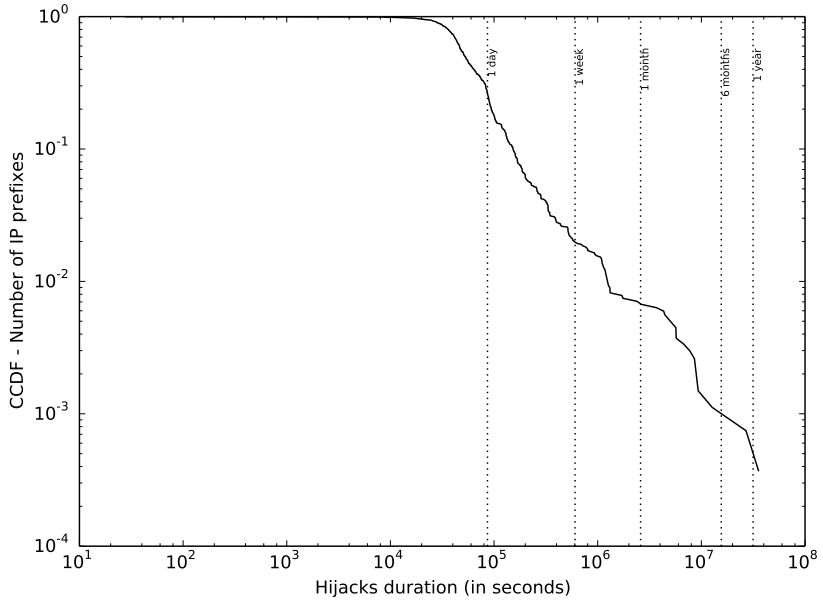


Figure 4.9 – The duration of hijacks identified between September 2012 and June 2014. Most of the observed hijacks are short-lived, *i.e.,* they last less than a week.

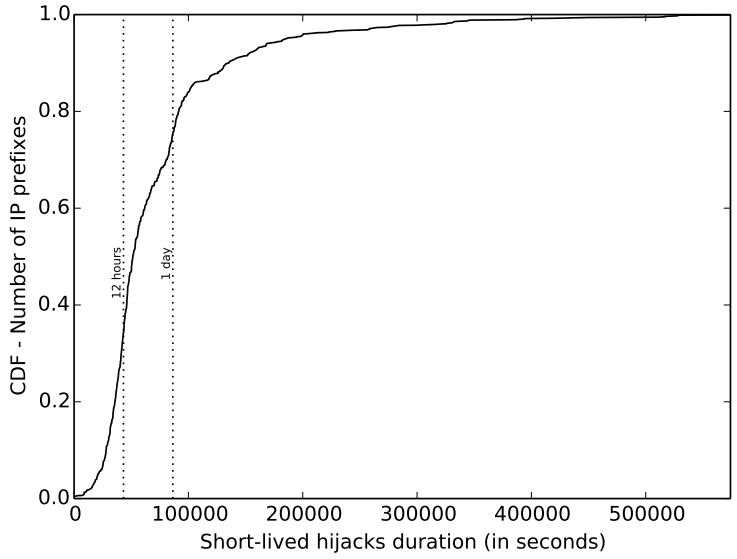


Figure 4.10 – The duration of **short-lived** ($\leq$ 1 week) hijacks.

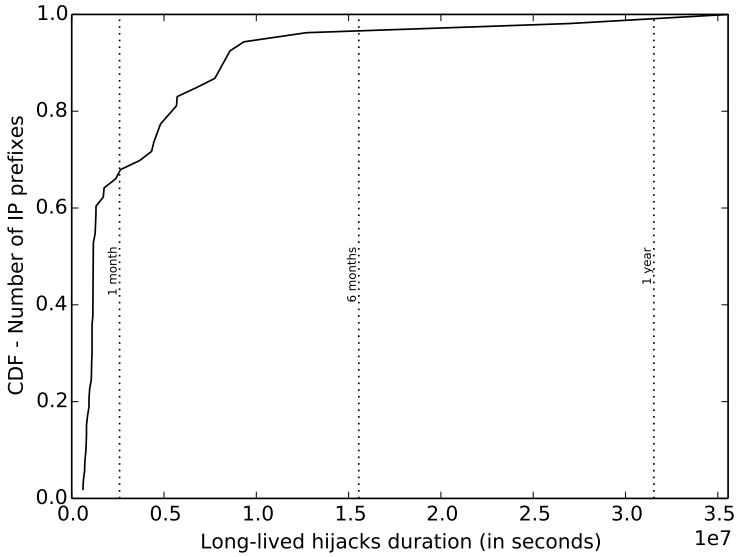


Figure 4.11 – The duration of **long-lived** (> 1 week) hijacks.

spamtraps and/or spam sources were **blacklisted** for the IP address range in Spamhaus SBL or DROP (Don't Route Or Peer) [34], Uceprotect [42] or Manitu [21].

- (*C.2*) The **duration of the unadvertised period** of the IP prefix, which corresponds to the amount of time elapsed between the last time it was announced and the moment it was hijacked.
- (*C.3*) The **registry** of the IP address block, which refers to the Regional Internet Registry (RIR) responsible for the allocation/assignment of the block.
- (*C.4*) The **registration date** of the IP address range, which is the date at which it was allocated or assigned by a Regional Internet Registry (RIR) to an Internet Service Provider (ISP) or end-user (*e.g.,* a company).
- (*C.5*) The **size** of the IP address range, which defines the number of individual IP addresses available in the range.
- (*C.6*) Whether the **owner** of the IP address block is still in business.

4.3.1 Long-lived hijacks

In this section we analyse more closely the 55 long-lived hijacks (out of the total 2,713 hijacks) we identified according to the *five* characteristics described above.

(C.1) Figure 4.12 shows the spam and blacklisted spam sources along with BGP announcements related to long-lived hijacked IP prefixes. Since those IP prefixes

were announced neither before nor after being hijacked, the BGP announcements shown here all relate to the time of the hijacks. We can see that seven out of 55 IP address ranges sent spam to our spamtraps. A total of 815 spam emails were sent from IP addresses scattered throughout each of the long-lived hijacked IP address blocks. Spam was mainly received at the start of the hijack period. No IP source was found to be blacklisted by the time the spam was received.

However, three networks (37.230.212.0/22, 193.138.172.0/22 and 91.198.40.0/24) out of the 55 became blacklisted by Spamhaus within two days after they became hijacked and we observed spam originating from them. In these cases, blacklists appear to have reacted quickly. Four additional networks which have not sent spam to our spamtraps also became blacklisted, although it took more time for them to appear on a blacklist. For two of them (61.45.251.0/24 and 115.85.133.0/24) it took 2 weeks and the hijack was over by the time they appeared on a blacklist. For the other two (91.220.63.0/24 and 192.12.131.0/24) it took one month and 2 months respectively before they appeared on a blacklist.

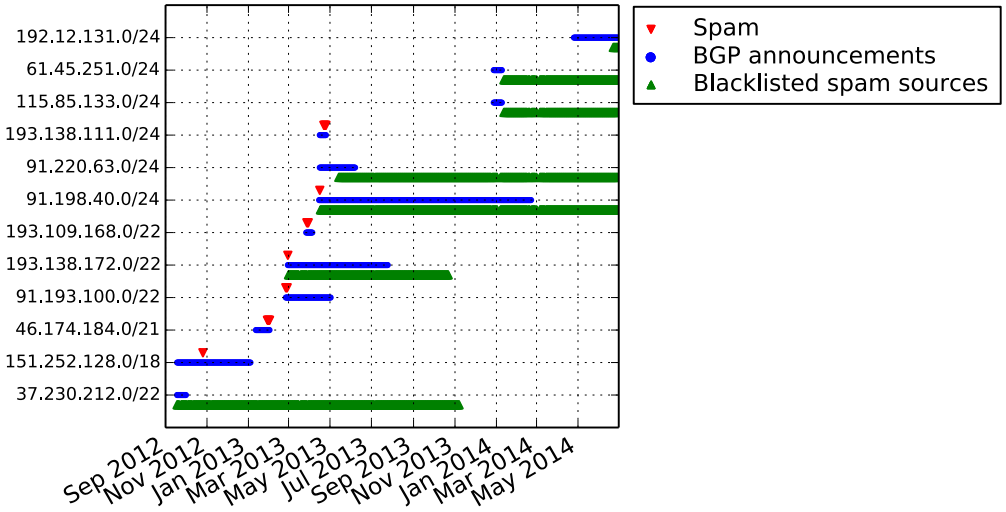


Figure 4.12 – BGP announcements, spam emails and blacklisted spam sources related to **long-lived** hijacked IP address ranges. For the sake of conciseness, only the 13 out of 55 IP address ranges that sent spam to our spamtraps or were blacklisted are depicted.

(C.2) 43 IP prefixes out of 55 were never publicly announced on the Internet before they were hijacked. The 12 others were hijacked on average one year after remaining unadvertised, with a minimum of one day and a maximum of three years and two months.

(C.3) The 55 long-lived hijacked IP address blocks were mostly registered in the RIPE (55%) and ARIN (28%) regions, which together accounts for about 68% of the IPv4 space allocated to the five Regional Internet Registries. The remainder of the blocks belonged to the APNIC (13%) and LANIC (4%) regions.

(C.4) The 55 long-lived hijacked IP address blocks were mostly registered after 2000. It is noteworthy that at the time they were hijacked, these ranges were all

properly registered IP address blocks assigned to various organisations. None of them was part of "bogon" IP address blocks, *i.e.,* IP addresses that should not be announced on the Internet [36].

(C.5) In [148], Ramachandran et al. claimed to have observed spam from large (*i.e.,* /8) hijacked IP address blocks. However such large blocks usually contain many smaller blocks (*e.g.,* /16's or /24's) which are allocated or assigned to various ISPs or end-user organisations. In our 55 long-lived hijack cases, the IP address blocks were smaller than what was claimed in previous studies, *i.e.,* the largest was a /19 and the smallest was a /24.

(C.6) The analysis of IRR (`whois`) records for long-lived hijacked IP address blocks revealed that most of the 55 blocks refer to organisations that are apparently out of business. This observation indicates that attackers might specifically target unannounced IP address space whose registrant does not exist anymore, for instance when a company is dissolved, acquired by or merged into another one. In some cases, its IP address blocks may be left unused.

Last but not least, we managed to get feedback from an ISP unwittingly involved in 23 out of the 55 long-lived hijacks carried out via the invalid AS57792. After investigation on their side, AS57792's ISP "EDPnet" (AS9031) confirmed these attacks had taken place and were performed by one of their customers, without them noticing it initially. The elements we provided corroborated their observations, and the ISP has since then terminated his peering contract with the misbehaving customer AS57792. This AS was observed connected to the provider AS9031 until March 2014 when it disappeared from the routing tables, likely due to the termination of its peering contract with AS9031. Interestingly, it reappeared in late June connected to two new providers AS57756 and AS58099.

4.3.2 Short-lived hijacks

In this Section we focus our analysis on the 2,658 short-lived hijacks (out of the total 2,713 hijacks). We further distinguish two episodes in these short-lived hijacks: (1) spam and blacklisted spam sources related to hijacked networks observed between February 2013 and May 2013 and (2) an apparently different hijack phenomenon observed between June 2013 and June 2014, showing a striking and unusual temporal pattern in the BGP announcements. We first present these two episodes and their differences with respect to characteristic C.1. Afterwards, we describe their commonalities in terms of the other characteristics C.2-6.

Episode 1: From February 2013 until May 2013

(C.1) Out of 2,658 short-lived hijacked IP prefixes, 57 have sent spam emails to our spamtraps between February and May 2013. Figure 6.5 shows the BGP announcements, spam and blacklisted spam sources related, for the readability, to a sample of 25 out of 57 short-lived hijacked IP prefixes. The figure highlights:

- the strong **temporal** correlation between BGP announcements and spam, and
- the **low** number of IP address blocks (7 out of 57) blacklisted by Spamhaus before the end of the hijack.

Note that the 32 remaining IP prefixes not depicted on Figure 6.5 exhibit the exact same pattern with respect to the BGP announcements, spam emails and blacklisted spam sources. A total of 4,149 spam emails were received from the short-lived hijacked IP address blocks. We extracted from this spam all advertised URLs that were pointing to 1,174 unique domain names, resolving to IP addresses belonging to the same hijacked IP address blocks, showing that some IP addresses were used in parallel to send spam and host the advertised scam websites. From `whois` information, we observed that these domain names were usually created within a few days before the networks being hijacked. This shows that attackers, very likely, control the entire IP address blocks and take full advantage of them.

Furthermore, spam emails collected by our spamtraps are enriched with the name of the spambot associated with the spam, by taking advantage of CBL signatures [11]. Spam emails sent from the supposedly hijacked IP address blocks we uncovered were not associated with any known spam botnet and thus must have been sent using another type of spamming infrastructure, instead of the traditional spam botnets. This is consistent with BGP spectrum agility where spammers need to set up a dedicated infrastructure with their own machines so that they can be assigned the hijacked IP addresses.

Episode 2: From June 2013 until June 2014

The hijacks performed during the second period, between June 2013 and June 2014, are more intriguing. This phenomenon is significant since it includes 2,601 short-lived hijacks representing 98% of all short-lived hijacks identified. Figure 6.6 depicts, for the sake of readability, a sample of 87 (out of 2,601) hijacks that occurred in June 2014[8] and shows that:

- all hijacks are actually performed by groups of two to four prefixes, starting and ending at the same time;
- during the one month period there are always, at any point in time, at least two IP prefixes hijacked.

Although only part of the phenomenon is depicted in Figure 6.6, it is recurrent and persistent over the complete 13 month period, between June 2013 and June 2014. This strongly indicates that they may have been performed with the same modus operandi. The fact that some groups of hijacks start only seconds after the end of previous groups further suggests that they might be carried out in an **automated**

[8]The figure depicting the complete phenomenon of episode 2 is available at `http://www.eurecom.fr/~vervier/public/bgphijacks_episode2_June2013_June2014.pdf`.

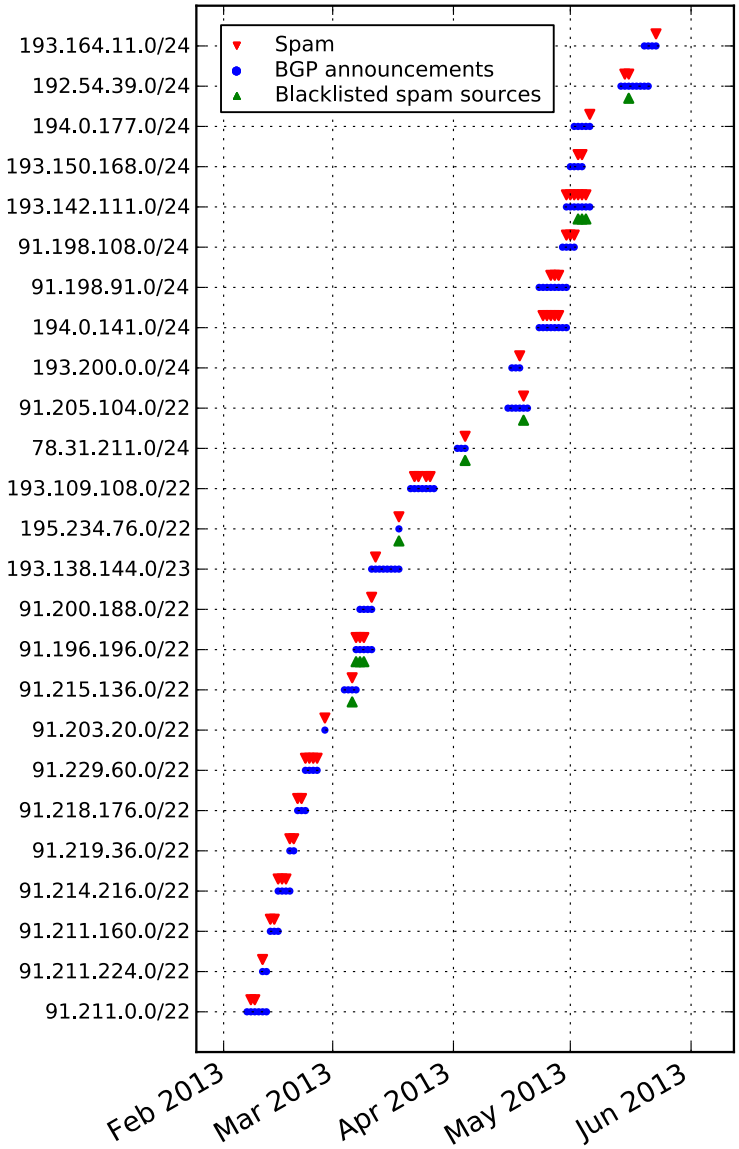


Figure 4.13 – Episode 1 of **short-lived** hijacks between February and May 2013: temporal correlation of BGP announcements, spam emails and blacklisted spam sources related to hijacked IP address ranges. For the sake of readability, only a sample of 25 out of 57 IP address ranges are depicted. The remaining 32 IP prefixes exhibit the same pattern.

way, possibly also relying on some automated process to find target network address blocks to hijack.

(C.1) Strangely enough, we have not been able to find any malicious traffic associated with those hijacked IP address blocks. The absence of spam and other scam-related traffic in our data may be due to incomplete visibility into malicious activities associated with these networks. To further investigate potential malicious network traffic originating from these networks at the time they were hijacked we mined various external data sources including NetFlow traces collected at the Munich's scientific network [177], DNS query logs from Symantec's Norton DNS infrastructure and Symantec's A/V telemetry [71] and found 557 additional scam-related domain names resolving to IP addresses hosted on only 15 of these 2,601 blocks. The absence of malicious network traffic related to most of these networks could also indicate that this is a moving infrastructure to host servers, *e.g.,* C&C servers. We have currently no conclusive evidence to validate this conjecture, though.

Common characteristics of episodes 1 and 2

In this section, we analyse the common characteristics of all 2,658 short-lived hijacks.

(C.2) Figure 4.15 presents the duration of the unadvertised period of all short-lived hijacked networks. 2,291 IP prefixes (86.2%) were never publicly announced before they were hijacked. From an informal discussion with a RIPE NCC executive [69] and discussions on the NANOG mailing list [52] it is apparently common practice for network operators to register and use publicly routable IP address blocks for internal network infrastructure only. This could explain why no route to such block can be found in our BGP feed. Apart from this reason we are not aware of any other reason why IP address blocks are registered but never actually publicly announced. The remaining 367 networks were last announced from 0 day, for the few cases of concurrent IP prefix hijack, as described in Section 4.3 on page 78, to 4 years before being hijacked, with a average of 24.6 months and a median of 24.5 months. With 72.4% of IP prefixes left unannounced for more than one year we can conclude that attackers mostly hijack networks left unannounced for a long period of time.

(C.3) Figure 4.16 depicts (i) the ratio of the number of short-lived hijacked IP prefixes per RIR, (ii) the ratio of short-lived hijacked IPv4 address space per the IPv4 address space allocated to each RIR, and (iii) the ratio of short-lived hijacked IPv4 address space per the complete IPv4 address space. The most affected region (RIR) appears to be ARIN with 61.9% of hijacked IP address blocks. Although the RIPE region is less affected than ARIN, they exhibit a similar ratio of hijacked address space (0.024%). Overall hijacked IP address blocks are mainly distributed among the ARIN, RIPE and LACNIC regions. APNIC and AfriNIC regions were seldom affected by hijacks with 5% and 0.4% of hijacked IP address blocks respectively. It is noteworthy that all hijacked IP address blocks were at least allocated to a RIR at the time they were hijacked, *i.e.,* no block was part of IPv4 address space reserved by

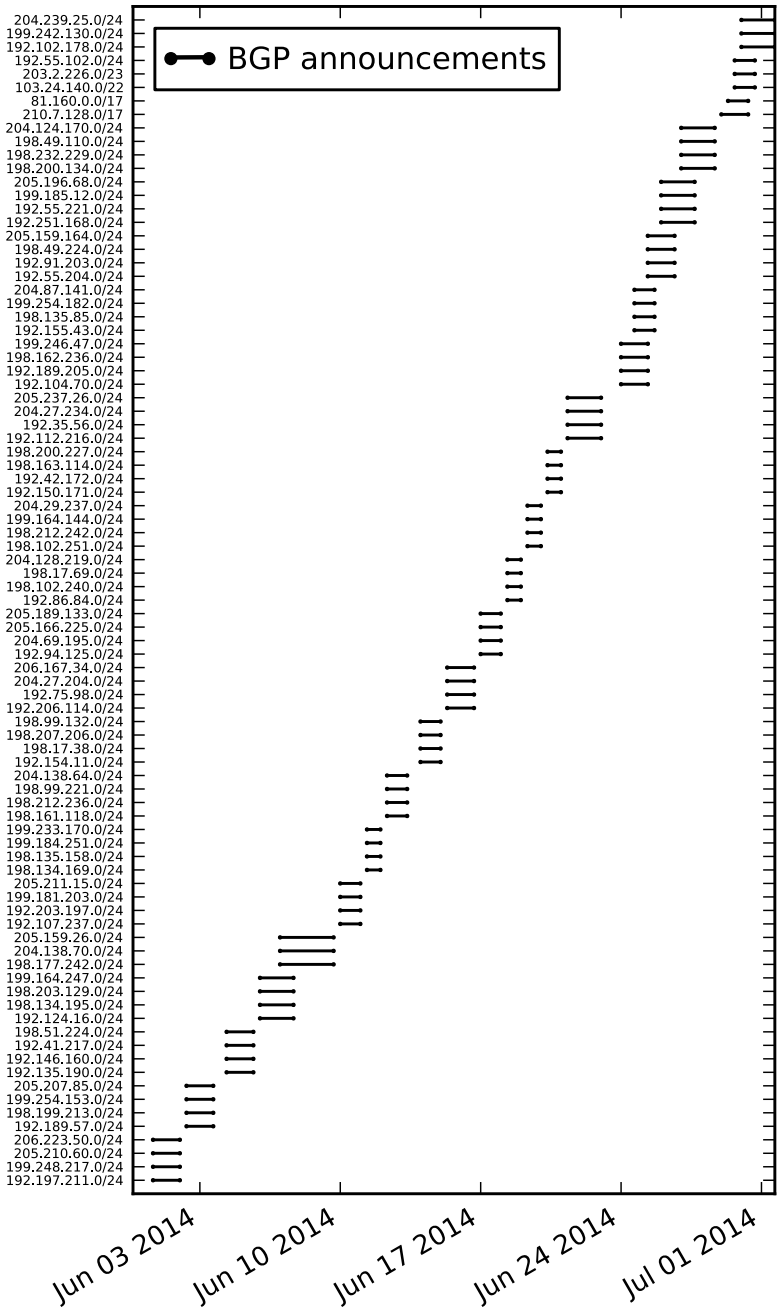


Figure 4.14 – Episode 2 of **short-lived** hijacks between June 2013 and June 2014: hijacks are always performed by groups of at least two IP prefixes. For the sake of readability, only a sample of 87 (out of 2,601) IP address ranges hijacked in June 2014 are depicted.

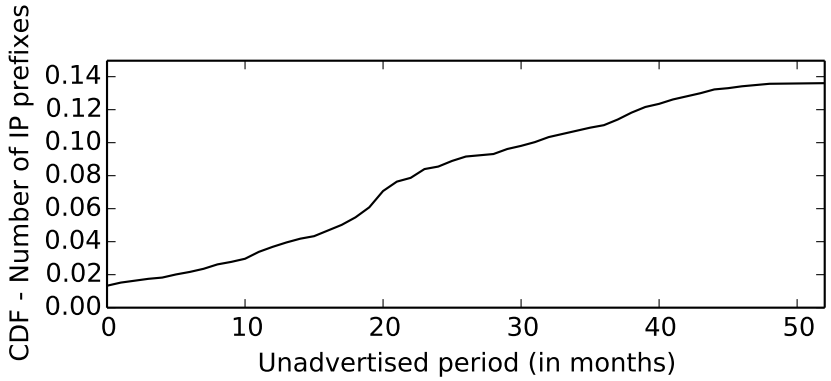


Figure 4.15 – The duration of the unadvertised period for 13.8% of the **short-lived** hijacked IP address ranges. The remaining 86.2% of IP address ranges were never announced before being hijacked.

the IANA for special use or not yet allocated by the IANA to a RIR. Finally, while ARIN space is involved in about three times more hijacks than RIPE and LACNIC space, the part of the complete IPv4 address space involved in ARIN space hijacks (0.01%) is only about twice that of RIPE (0.005%) and LACNIC (0.004%) indicating that hijacked ARIN's blocks are overall smaller than RIPE's and LACNIC's. This observation is later corroborated by the ratio of short-lived hijacked IP address block sizes per RIR in the characteristic C.5.

(C.4) Figure 4.17 shows the registration date of short-lived hijacked IP address blocks. It appears that 2,073 IP address ranges (76.43%) out of 2,658 were registered before 1997 when RIRs started taking on the registration of IP address resources and setting up the IRRs [109]. Network address ranges registered before 1997 can thus sometimes be poorly documented and, for that reason, have been assumed to be a target of choice for spammers to hijack them [93, 133]. This idea appears to be supported by our data.

(C.5) Short-lived hijacked IP address blocks include /17's, /21's, /22's, /23's and (92.6%) /24's, similar to the long-lived ones. Figure 4.18 depicts the distribution of IP address block sizes per RIR. Hijacked IP address blocks in ARIN and AfriNIC space are almost exclusively /24's whereas blocks in RIPE, APNIC and LACNIC include blocks the size of which varies between /17's and /24's. Although those hijacks look like the ones Ramachandran et al. reported in [148], the average size of hijacked address blocks is very different, namely /24, instead of /8.

(C.6) The analysis of `whois` records (from IRR databases) of short-lived hijacked networks revealed that all IP address blocks were, at the time they were hijacked, properly registered blocks assigned to an organisation with sometimes multiple blocks referring to the same organisation. Although we could not check all 2,658 IP address blocks, we looked at 100 of them and determined that 41% refer to organisations that are apparently out of business but, interestingly, 59% refer to organisations that appear to be still in business.

As a final validation step, we mined the archives of the NANOG [22] and RIPE Working Groups [27] mailing lists for public reports related to ASes or IP address

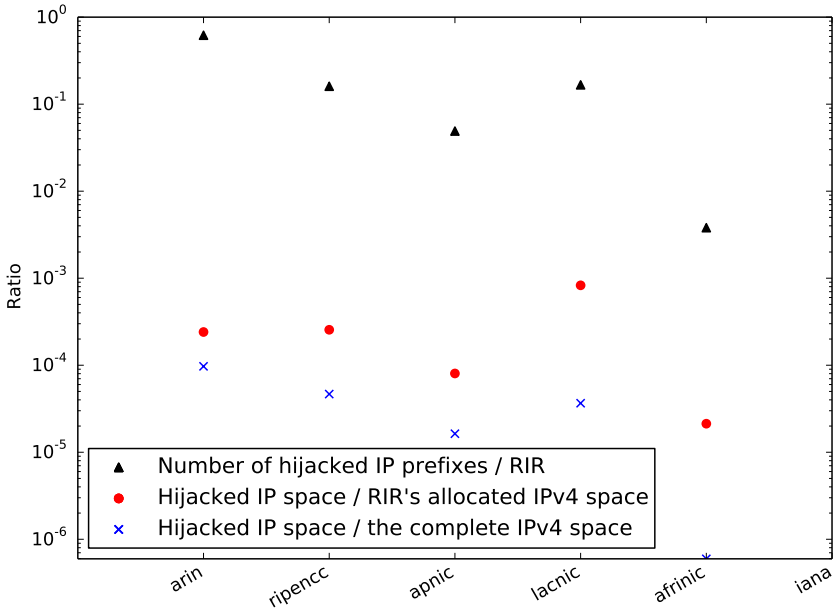


Figure 4.16 – (i) The ratio of the number of **short-lived** hijacked IP prefixes per RIR. (ii) The ratio of the **short-lived** hijacked IPv4 address space (*i.e.*, /32's) per the IPv4 address space allocated to each RIR. (iii) The ratio of the **short-lived** hijacked IPv4 address space per the complete IPv4 address space.

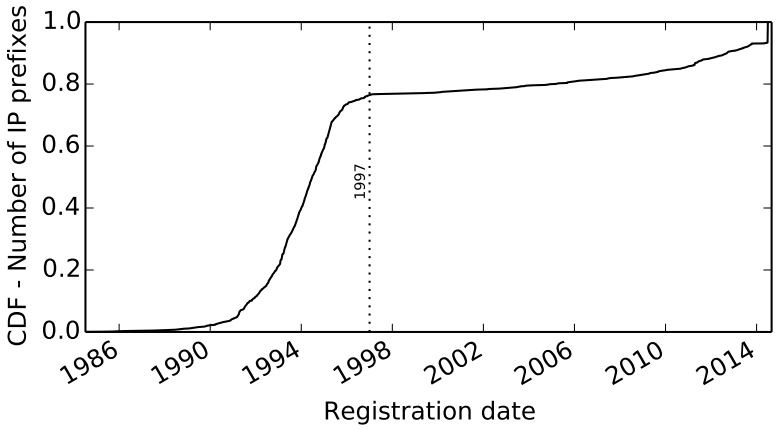


Figure 4.17 – The registration date of **short-lived** hijacked IP address ranges.

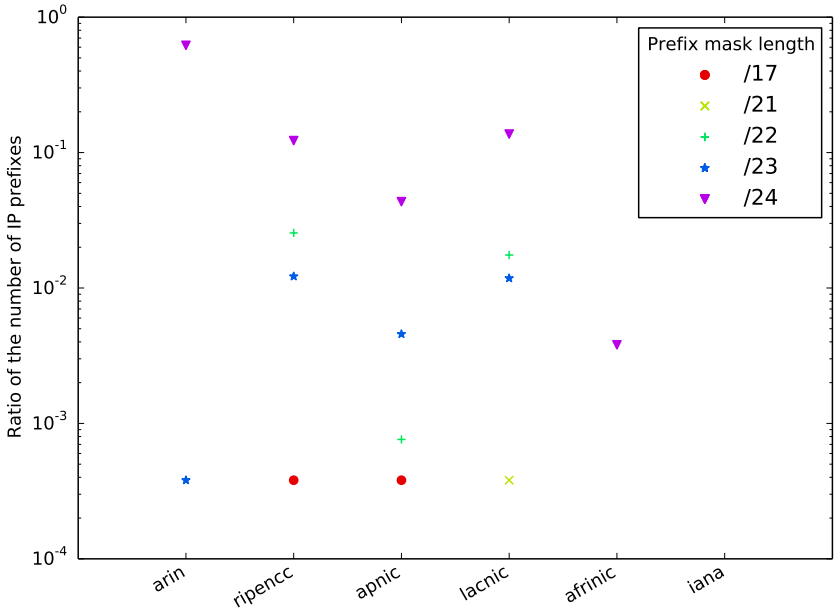


Figure 4.18 – The ratio of the number of **short-lived** hijacked IP prefixes per RIR and per IP address block size (prefix mask length).

blocks identified in our hijacks. We found one NANOG thread [49] reporting the hijack of the block 91.220.85.0/24 and its legitimate BGP origin AS51888 via the invalid direct upstream provider AS42989.

Early September 2014, BGPmon.net and Renesys published on their respective blog two analyses of "IP squatting for spam" operations [125, 171] involving some of the *invalid* ASes reported earlier in this document. These reports came out shortly after a network operator at AS132420 "E2E-Networks" in India complained, in late August 2014, on the NANOG mailing list that one of their IP prefixes had been hijacked by AS43239 "SpetsEnergo Ltd." located in Russia. According to the network operator of the victim network, the targeted IP address block was not routed by the time it was hijacked. The offending AS43239, identified in our traces, was also shown in BGPmon.net's and Renesys' analyses to have been used for several other IP address block hijacks associated with the emission of spam emails. BGPmon.net's article mentioned four additional AS numbers (AS28490, AS32579, AS57792 and AS262916) abused in other IP squatting campaigns, which were also identified in our traces. These analyses thus appear to corroborate our findings.

Finally, 770 out of the 2,658 short-lived hijacks carried out via the invalid AS57792 were later confirmed by AS57792's ISP "EDPnet" (AS9031), unwittingly involved in these attacks, that we interacted with via its national CERT organisation. After some investigation, the ISP discovered that one of their customers was indeed announcing IP address blocks he did not own, which were used for nefarious activities. The ISP has now terminated its contract with the misbehaving customer (who turned out to be colluding with a spammer apparently based in Russia).

4.4 Root cause analysis results

In the previous section we have uncovered strong evidence of BGP spectrum agility occurring on the Internet. However, we have not systematically analysed if the identified hijacks are isolated attacks or if some of them share a common root cause, as we would expect if they are part of campaigns orchestrated by the same spammers. This is why we have run the TRIAGE clustering tool against all spam emails coming from the 64 supposedly hijacked IP address blocks which have sent spam to our spamtraps.

Based on the 4,964 spam emails received, the multi-criteria clustering tool TRIAGE has identified only 30 multi-dimensional clusters (MDCs) in which spam emails are correlated by various combinations of features. Because of the way these clusters are generated, we anticipate they likely represent different campaigns organized by the same individuals - as spam emails within the same cluster share several common traits. More details on the methodology behind the creation of MDCs by TRIAGE can be found in Section 6.3.4 of Chapter 3 on page 130 and in [168]. Thus 64 prefixes were used to run 30 different spam campaigns. In the following we will show that some campaigns are rather short-lived and run from a single prefix whereas others last for several days relying on a number of different prefixes reinforcing the idea that the hijacks are repeatedly performed by a limited number of actors, in a coordinated fashion. Table 4.2 provides global statistics computed across all MDCs. Most spam campaigns seem to be short-lived (lasting on average only a couple of days), except two MDCs that existed for more than 30 days.

Statistic	**Avg**	**Med**	**Min**	**Max**
Nr of spam emails	141.8	11.5	2	1,178
Nr of IP prefixes	1.6	1	1	12
Nr of URL hosting server IP addresses	7.3	4	1	24
Nr of URL domain names	10.3	2	1	173
Nr of URL domain name whois registrants	44.5	6.5	1	556
Nr of spam subjects	47.7	7	2	455
Nr of active days	5.7	1	1	24
Lifetime in days	5.7	1	1	81
Compactness	0.43	0.43	0.27	0.74

Table 4.2 – Global statistics for the 30 MD-Clusters (spam campaigns).

By clustering spam emails into campaigns, we obtain new insights into hijacking spammers behavior. From the structure of MDCs, we uncover three key modus operandi of hijacking spammers: (1) 10 campaigns (out of 30) involve a single hijacked IP prefix that is not abused elsewhere in any other campaign, (2) 17 campaigns involve a single hijacked IP prefix, yet the hijacked prefix is abused concurrently in different spam campaigns, and (3) three campaigns were observed abusing *multiple hijacked IP prefixes* sequentially over a longer period of time. While the first two phenomena actually confirmed our intuition about the anticipated behavior of this class of spammers, the latter phenomenon is the most interesting one as it confirms

the existence of BGP spectrum agility in the form of campaigns of BGP hijacks orchestrated by the same spammers moving from one stolen block to the other to send spam emails. It highlights the existence of a more agile and sophisticated modus operandi of spammers capable of hijacking and abusing multiple IP prefixes, and subsequently hopping from one hijacked IP prefix to another to distribute spam. This agility enables them to send spam in a more stealthy manner and thus stay undetected "under the radar", but also to hinder their traceability and render IP blacklists ineffective.

The average MDC compactness (Cp) provides an objective metric to evaluate the consistency of the results. A Cp value close to 0.40 means that spam emails belonging to the same MDC have at least three strong features in common, thus reinforcing the intuition that the emails are sourced by the same gang of spammers.

Finally, from Table 4.2 we also observe a higher variability in spam email subjects and `whois` registrant addresses, suggesting that spammers have automated tools at their disposal to facilitate the creation of new email templates and automate the registration of new domains used for disposable "one-time URLs".

Figure 6.7[9] shows a graph visualization of one of the large-scale campaigns that involved multiple hijacked IP prefixes, which illustrates the typical modus operandi of agile spammers operating such stealthy campaigns. In this particular example, we can observe the following key points:

- over 662 spam emails have been sent from 12 different hijacked IP prefixes (yellow nodes), each of them used in turn by spammers to distribute spam using a bunch of one-time URL's, most of them including domain names (blue nodes) registered at ENOM (large pink node) using privacy-protected email addresses provided by *whoisprivacyprotect.com* (red nodes);
- spam advertised content (domain URLs) share the same server IP addresses (light grey nodes);
- the campaign has a lifetime of 84 days, yet only 24 active days (purple nodes laid out in a clockwise fashion), during which spammers are hopping from one hijacked IP prefix to another, which is an effective way of circumventing IP-based spam filters and reputation systems.

To the best of our knowledge, these results are completely novel and shed a new light on the behavior of agile BGP hijacking spammers. First, we observe that stealthy spam campaigns can be performed by exploiting multiple hijacked IP address blocks. Secondly, we observe that a very limited number of direct upstream providers were involved in these hijacks, which already gives some indication of possible countermeasures. Finally, all URLs advertised in spam emails are sharing a common hosting infrastructure and were registered in a similar way – suggesting that `whois` registration data can also be leveraged in prevention systems. The key take-away of this

[9]Disclaimer: IP addresses, domain names and email addresses were found in campaigns launched from likely hijacked networks only between September 2012 and June 2014. These may have been abused and stolen from their legitimate owners and, therefore, may now be legitimately used.

root cause analysis is that it enables us to link together different hijacked prefixes showing they are used by the same spamming actors for a long period of time in a very stealthy way.

4.5 Summary of findings

Finding 1 *We uncovered two types of hijack phenomena: long-lived and short-lived ones. Long-lived hijacks can last from a week to several months, whereas short-lived hijacks last from a few minutes to several days.*

Finding 2 *Uncovered hijacks targeted to a great extent IP address blocks that were not announced in BGP prior to being hijacked (*IP squatting*). A few cases however involved IP address space already announced by the legitimate owner of the network (*concurrent hijack*).*

Finding 3 *Attackers were found to stealthily hijack properly registered but unannounced IP address space by using two different hijacking techniques (as defined in Section 4.3 on page 71):* prefix hijacking *and* AS hijacking*. In prefix hijacking, the attacker announced an IP address block using an invalid BGP origin AS via a valid direct upstream provider (first hop) AS. In AS hijacking, the attacker announced an IP address block using its valid BGP origin AS but via an invalid direct upstream provider (first hop) AS.*

Finding 4 *In the 2,489* prefix hijacks *we found only* eight different *invalid BGP origin ASes. In the 224* AS hijacks *we found, for 214 different valid BGP origin ASes, only* three different *invalid upstream provider ASes. One AS, involved in the hijack of 793 IP address blocks over 22 months, was observed first as an invalid upstream provider AS, and then as an invalid BGP origin AS. These 793 hijacks were later confirmed by the ISP providing transit to that AS, who consequently terminated the contract with this customer abusing the routing infrastructure.*

Finding 5 *Spamming using hijacked IP prefixes appears to be an effective technique for defeating known protections, such as spam IP blacklists. Moreover, almost none of the IP address blocks were hijacked more than once meaning that in this case blacklisting those blocks after the hijack ends is useless. Finally, spammers also use the hijacked IP address blocks as a hosting infrastructure for spam advertised content.*

Finding 6 *Hijackers mostly hijack IP prefixes that have never been advertised or left unadvertised for a very long time, typically more than one year.*

Finding 7 *Hijackers seem to prefer IP address blocks that were properly registered, in contrast to "bogon" IP address blocks, such as the list from Team Cymru [36], whose announcements are commonly automatically filtered out.*

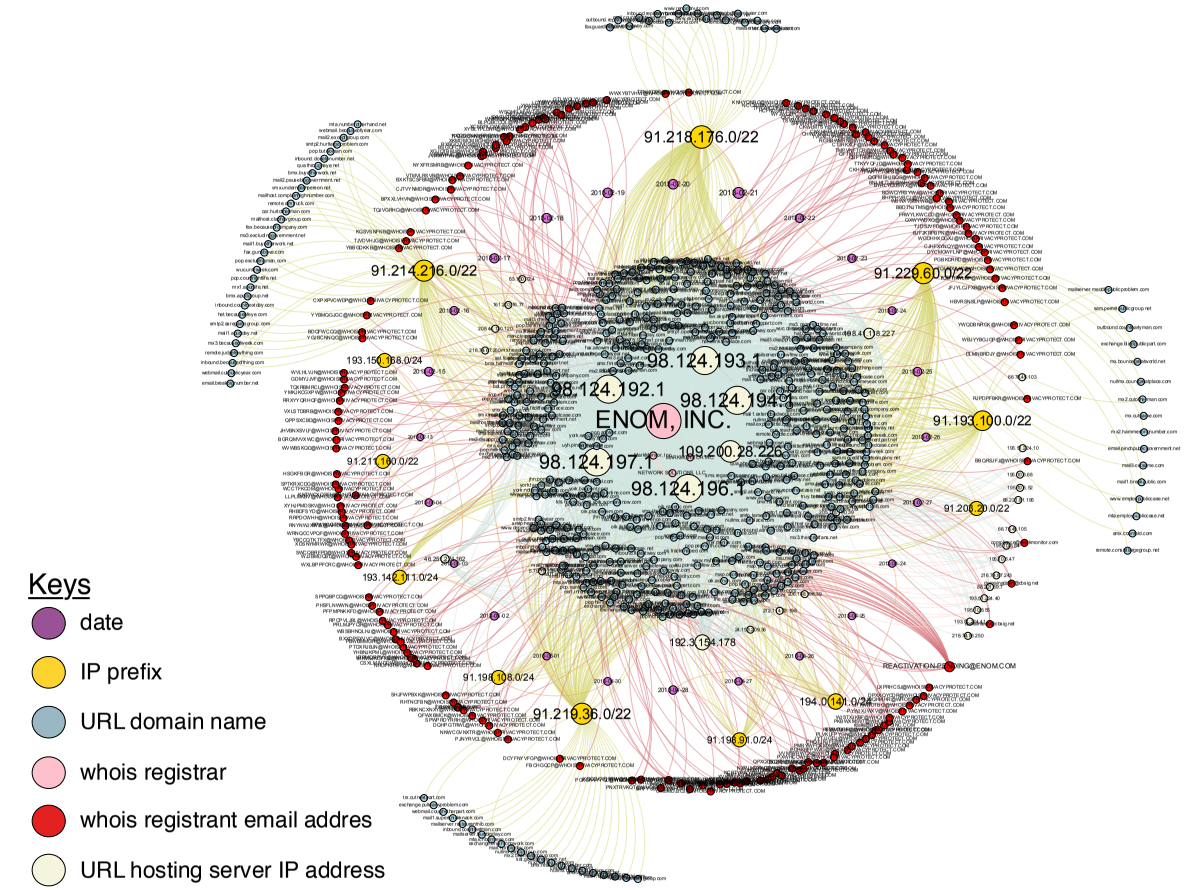


Figure 4.19 – An example of a large-scale spam campaign involving multiple hijacked IP prefixes. The nodes laid in clock-wise fashion reflect the timeline of the campaign.

Finding 8 *Many hijacked IP address blocks we identified refer to organisations that have ceased to exist. Orphan IP address blocks that are left behind then become targets of choice for hijackers as they can hijack them without being noticed. As of July 2014 as much as 20.26% of the whole IPv4 address space*[10] *is currently allocated or assigned without being publicly announced.*

Finding 9 *Some short-lived hijacks were clearly associated with spam activities, confirming the existence of the BGP hijacking spammers phenomenon as introduced in [148]. However, a large portion of them exhibited no spam and we conjecture that they would ideally serve as a moving infrastructure to host malicious servers, but as per page 93, after investigation an ISP unwittingly involved in many of these hijacks confirmed these blocks had sent spam emails, although we had not received any at our spamtraps.*

4.6 Effectiveness of current countermeasures

Different technologies and systems have been designed to detect and mitigate BGP hijacks. In this section we evaluate the effectiveness of two BGP hijack countermeasures: a state-of-the-art BGP hijack detection system called Argus [157] and the BGP security framework RPKI [101, 119, 120].

4.6.1 BGP hijack detection

There have been numerous systems and services, such as BGPmon.net [7], Renesys [25], PHAS [116] and Argus [157], developed to detect and mitigate BGP hijacks. One of them, Argus [157], aims at detecting BGP hijacks in real time by using a combination of BGP data and ping measurements to detect, upon a routing change related to a network, changes in the network's reachability indicating a possible hijack. In an effort to assess the security impact of the hijack incidents we uncovered, we decided to verify the effectiveness of the Argus system against these cases. We chose Argus for two reasons: (i) it is currently deployed and provides a publicly available historical feed of alerts, and (ii) it is also able to detect all types of hijacks, namely those where the attacker hijacks an IP address block by using an invalid BGP origin AS (prefix hijack) or by forging part of the BGP AS path (AS hijack).

It turns out that none of the 2,713 hijacks we identified were reported by Argus. The reason is that most BGP hijack detection systems [96, 116, 157], including Argus, work by building a model of the Internet AS-level topology and then using it to validate any routing change. However, because almost all hijacks we identified involve IP space that was unannounced prior to being hijacked, there is no state for the IP address blocks in the model resulting in any new route announcement to be accepted as legitimate. Although current BGP hijack detection techniques

[10] Based on statistics published by RIRs and available at http://bgp.potaroo.net/ipv4-stats/prefixes.txt

are valuable for network operators to monitor their own networks, their inability to currently detect hijacks like those we observed suggest that those techniques should integrate in the future in their detection scheme some characteristics of the hijacks we have identified.

As discussed in Section 4.3.2 on page 91, shortly after a network operator complained on the NANOG mailing list that one of his IP prefixes had been hijacked [51], two different reports describing "IP squatting operations for spam" were released on the blog of BGPmon.net [171] and Renesys [125]. These reports corroborated our findings about *five* (out of 10) invalid ASes we identified.

4.6.2 BGP hijack prevention

Besides BGP hijack detection techniques, the network operators have started to adopt and deploy a BGP hijack prevention framework commonly referred to as the RPKI system. Though many approaches have been proposed to bring security to BGP [102], this framework has been gaining more momentum than others in the last few years. This is likely due to the fact that it is the only framework being standardized by the IETF. We are not aware of any other framework that is ready and mature enough to go through that process. Recall from Chapter 2 that the framework relies on a Resource Public Key Infrastructure (RPKI), standardized in RFC 6480 [120], to prevent the injection of bogus routing announcements. The RPKI used in this scheme consists of a database of certificates of four types: (i) a type A called Route Origin Authorisation (ROA) binds an IP address block to its authorised BGP origin AS(es), (ii) a type B that binds a router to the AS number it belongs to, and (iii-iv) certificates C and D that binds respectively IP addresses and AS numbers to the public key of their respective owner. The certification chain follows the AS number and IP address delegation chain, with the IANA acting as the root certificate authority for RIRs' certificates, a RIR is then acting as the certificate authority for ISPs' certificates, etc. Each certificate is signed with the private key of its holder and also embeds its public key. The framework proposes two separate techniques to secure BGP: (i) secured *route origination* and (ii) secured *route propagation* (or BGPsec). (i) Secured route origination, standardized in RFC 6483 [101], uses ROAs (type A certificates) to verify that a given IP address block is originated by the authorised AS(es). A router is then able to verify the validity of a received BGP update for a given IP address block and BGP origin AS by (i-a) querying the RPKI for a ROA related to the IP address block and verifying its cryptographic validity, and, (i-b) if the ROA is valid, verifying that the origin AS and the length of IP prefix observed in the BGP update match the authorised origin AS(es) and prefix length in the ROA. This prevents an attacker from announcing a block he does not own. (ii) Secured route propagation [119] aims at preventing *AS path forgery* by ensuring that each AS in the AS path was not impersonated. This is done by having each router signing a BGP update it propagates so that subsequent routers can verify, using type B certificates from the RPKI, that all routers which have signed the update indeed belong to the ASes found in the path.

Secured route origination is progressively being deployed. According to the RIPE

NCC [152] there is currently 4.1% of the IPv4 address space covered by ROA's. Interestingly, none of the IP address blocks that we identified as having been hijacked were covered by a ROA at the time they were hijacked. In 90% of the hijacks we observed, the attacker announced the IP address blocks using an invalid BGP origin AS (*prefix hijacks* as defined in Section 4.3 on page 71). Providing a ROA had been issued for these blocks and their valid BGP origin AS, the RPKI would have invalidated the bogus announcements.

However, assuming ROAs would bind IP address blocks with their legitimate BGP origin AS, hijacks can still be successful if attackers forge the BGP AS path and prepend the valid BGP origin AS to it, which is exactly what we observed in 10% of the hijacks (*AS hijacks* as defined in Section 4.3 on page 71). Secured route propagation (BGPsec) is currently still at an early development stage and thus not yet being deployed. In the meanwhile, although BGP origin validation via ROAs does not intend to prevent BGP AS path forgery, as acknowledged in RFC 6483 [101], the RPKI and ROAs could nevertheless be leveraged to prevent unannounced IP address blocks from being hijacked by issuing a ROA for AS0 and each unannounced IP address block (such ROAs are already used to prevent the announcement of reserved/unallocated IP space as dictated in RFC 6483 [101]). Then, the RPKI will classify all routes for these IP address blocks as invalid. This solution is not perfect though as it requires a specific ROA to be issued when an IP address block becomes unannounced which, in the case of orphan blocks, is unlikely. Overall, the only proper solution to prevent BGP AS path forgery and the AS hijacks we identified is to have secured routed propagation, *i.e.,* BGPsec, deployed. Unfortunately, this solution is much more invasive and cannot be deployed without substantial software and hardware updates on all routers. Moreover, the standardization process of BGPsec is not yet completed and there is no router code available as of today. Some vendors are working on it, or intending to work on it, but some other vendors do not even list it on their roadmap yet.

4.7 Operationalizing SpamTracer

In the previous Sections of this Chapter, we presented the result of the analysis of 22 months of data collected using SpamTracer and exposed an ongoing "BGP spectrum agility" phenomenon where stealthy and persistent campaigns of BGP hijack attacks are taking place on the Internet on a regular basis. Although the phenomenon by itself was not new since the first evidence of "BGP spectrum agility" was reported in [148] in 2006, to the best of our knowledge, very little was known about the frequency and persistence of these attacks as well as about the modus operandi used by attackers to carry out malicious BGP hijack attacks. We also identified several differences between the hijack attacks reported in [148] and the ones we uncovered, such as the varying hijack duration (from minutes to months), the size of the hijacked IP address blocks (mostly /24), suggesting the phenomenon has evolved over the years. As explained in Chapter 3, SpamTracer was originally designed to determine whether, as of 2014, the "BGP spectrum agility" phenomenon was still a

problem worth of consideration. If yes, the collected data would allow us to characterise the observed BGP hijack attacks and gain further insights into the modus operandi of attackers in an effort to identify possible ways to detect and mitigate such attacks. Moreover, we have seen in Chapter 2 that besides the "Don't Route Or Peer (DROP)" list provided by Spamhaus [34] and the development and deployment of BGP hijack prevention techniques, such as the RPKI, little effort is being devoted to mitigate such attacks. Moreover, earlier in this Chapter, we saw that current BGP hijack detection and prevention techniques (*e.g.,* Argus, RPKI+ROA) were not very effective at detecting the hijacks we observed. In this Section, we propose to leverage some key characteristics of the BGP hijack attacks we uncovered to detect future attack instances in an effort to (i) continue to monitor the "BGP spectrum agility" phenomenon and (ii) provide a way to proactively mitigate these attacks by means of a (black-) list of hijacked IP address ranges. In the remainder of this section we first discuss the attack characteristics that can be leveraged in a real-time identification of BGP hijack. We then propose a technique to build a real-time blacklist of hijacked IP address blocks and discuss some possible applications.

4.7.1 BGP hijack attacks characteristics

Earlier in this Chapter we have exposed "BGP spectrum agility" we observed over a period of 22 months between September 2012 and June 2014 in the form of persistent campaigns of malicious BGP hijacks. We observed two major phenomena: *short-lived* and *long-lived* hijacks. Moreover, we witnessed two recurring modus operandi used by attackers to hijack blocks of IP addresses: *prefix hijacks* via invalid BGP origin ASes and *AS hijacks* via invalid direct upstream provider (first hop) ASes. Furthermore, almost all (99.52%) identified hijack cases involved IP address blocks that were not announced prior to being hijacked. Based on the observed attacks characteristics, we consider the following features in order to derive a signature for automatically detecting instances of IP address block hijacks:

1. the **BGP origin AS(es)** observed for an IP address block;
2. the **direct upstream provider (first hop) AS(es)** observed for an IP address block;
3. the **owner** of the IP address block, the BGP origin AS and the direct upstream provider AS, as published in the IRRs;
4. the **country of registration** of the IP address block, the BGP origin AS and the direct upstream provider (first hop) AS, as published in the IRRs;
5. the **Internet Routing Registry (RIR)** responsible for the allocation of the IP address block, the BGP origin AS and the direct upstream provider (first hop) AS;
6. possible **routing policies** related to the BGP origin AS and direct upstream provider AS published in the IRRs;

7. whether **malicious network traffic** was originated by the IP address block;
8. the **in-** and **out-degree** of the BGP origin AS and the direct upstream provider (first hop) AS, defined respectively as the number of downstream and upstream inter-AS links observed in BGP for an AS;
9. the **number of IP address blocks** originated by the BGP origin AS or advertised via the direct upstream provider (first hop) AS;
10. the **observation window** of the IP address block, the BGP origin AS and the direct upstream provider (first hop) AS, defined as the time period between its first and last appearance in the BGP routing tables;
11. the **uptime** of the IP address block, the BGP origin AS and the direct upstream provider (first hop) AS, defined as the time period during which it appeared in the BGP routing tables.

In the context of our system to build a blacklist of maliciously hijacked networks, features 1-8 can be used to detect the hijack of an IP address block as soon as the routing changes related to that block resulting from the hijack are visible by the BGP collectors used. The reason is that for features 1-8 the anomalous value can be recorded as soon as the hijack starts. Features 9-11 however only exhibit their anomalous value a posteriori, *i.e.,* after the hijack ends. We show how the different features are leveraged in the description of the algorithm for building a real-time blacklist of hijacked IP address blocks here below.

4.7.2 Real-time blacklist of hijacked networks

The input data to our system consists of (i) historical dumps of BGP routing tables (RIB's) from one RIPE RIS [26] collector (rrc00.ripe.net) and one RouteViews [43] collector (route-views2.routeviews.org) providing the BGP routes to network IP prefixes advertised on the Internet, and (ii) historical dumps of Internet Routing Registries (IRRs) providing registration information about IP and AS resources, such as the owner of an IP address block or AS, its country of registration, contact details and possible routing policies between ASes. The system architecture of our real-time blacklist of hijacked networks is depicted in Figure 6.8. Its workflow consists of the following steps:

1. ***Identification of unannounced IP prefixes.*** We historical dumps of BGP routing tables to first build a list of unannounced IP address blocks including (i) blocks that have been allocated or assigned to an Internet Service Provider (ISP) or an end-user but are currently unannounced (*e.g.,* a company terminated its business but did not return its IP prefixes), and (ii) blocks that are not supposed to be used at all because they are reserved for special purposes by the Internet Assigned Numbers Authority (IANA), not allocated to any Regional Internet Registry (RIR) or not allocated/assigned by any RIR to an ISP or end-user. All these network IP prefixes are considered more appropriate

targets for a BGP hijack than others because they do not host any machines and, as a result, will likely not create any visible disruption during a hijack.

2. ***Detection of resuming IP prefixes.*** The routing state of a network IP prefix corresponds to the BGP routes from different collectors in the Internet towards that prefix. We monitor the prefixes to detect those resuming with a suspicious routing state. Metrics used to identify a suspicious routing state include:

   - a new BGP origin AS for the prefix;
   - a new direct upstream provider (first hop) AS in BGP for the prefix;
   - reserved, unallocated or previously unadvertised BGP origin AS and/or upstream provider AS for the prefix;
   - suspicious BGP origin AS with respect to its country of registration and the country of registration of the prefix;
   - mismatch between the BGP origin AS owner and the prefix owner;
   - suspicious upstream provider ASes with respect to their country of registration and the country of registration of the prefix;
   - mismatch between the actual routing state as observed in BGP and the expected one from the registration information if available (*i.e.,* violation of published routing policies between the BGP origin AS and the direct upstream provider AS(es));
   - prefix, BGP origin AS or upstream provider AS involved in previous hijacks.

   We identify candidate hijacked prefixes at this step.

3. ***Correlation between hijack and malicious activities.*** Candidate hijacked prefixes identified at step ② are further monitored to determine whether malicious activities (*e.g.,* spam, scam websites hosting) are performed from these prefixes during the hijack.

The output of our system is thus a list of network IP prefixes that are considered hijacked from their legitimate owner together with a suspicion score computed based on the evidences gathered during the monitoring process. The suspicion score increases with the number of "indications of a suspicious routing state" extracted in step ② as well as if malicious activities are observed from the hijacked IP prefixes in step ③.

We can translate the suspicious routing state identification process in our system into the Algorithm 3. The algorithm takes as input (i) a set of resuming IP prefixes, (ii) the routing history of these prefixes, (iii) IRR records related to the prefixes, BGP origin ASes and direct upstream provider ASes, and (iv) the list of IP prefixes, BGP origin ASes and direct upstream provider ASes involved in previous hijacks. It outputs a set of suspicious hijacked networks with a score $s \in [0, 1]$. The score of a prefix can be further increased, up to 1, in step ③ if we observe malicious

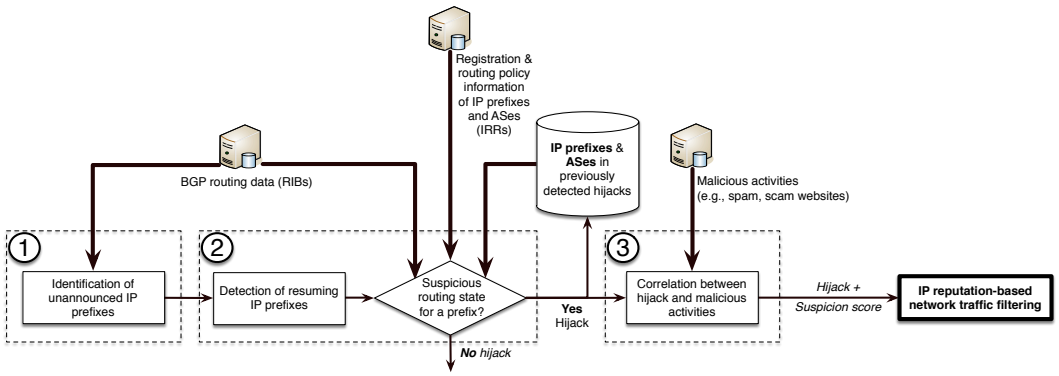


Figure 4.20 – Real-time blacklist of hijacked network IP address blocks: system architecture.

network traffic originating from the network. Finally, a BGP origin AS or direct upstream provider (first hop) AS exhibiting an average *uptime* of IP prefixes lower than seven days (as in short-lived hijacks described earlier in this Chapter) and an *observation window* significantly longer than the average uptime of the IP prefixes will be considered a suspicious hijacked AS, and will be included in the list of "IP prefixes and ASes in previously detected hijacks" in our system.

4.7.3 Real-world deployment

We now present the results obtained from the deployment of the system described here above in the real world. First of all, we leveraged as input data to our real-time blacklist the historical dumps of RIBs and IRRs as well as the spam dataset we introduced in Section 3.2.3 of Chapter 3 for the validation of candidate hijacks from SpamTracer.

Experimental validation

In an effort to validate our approach, we applied our system on historical data for the same period used in the analysis of SpamTracer data, *i.e.,* between September 2012 and June 2014. During that time period, our system blacklisted a total 4,262 IP address blocks among which 2,700 (99.5%) were part of the 2,713 hijack cases we uncovered and exposed in the first part of this Chapter. The 13 missed cases correspond to the 13 concurrent hijacks identified in our traces and described earlier in this Chapter on page 78. These missed cases result from the fact the our real-time blacklist system is specifically tailored to detect hijacks of prefixes not announced before the hijack. While some detected prefixes were associated with the emission of spam to our spamtraps, most of them could not be linked to any malicious traffic. These experimental results show that the features used in our system are able to extract relevant suspect hijack cases. As a final step, we applied our candidate hijack validation methodology described in Section 3.2.3 of Chapter 3 on the 1,562 additional suspicious prefixes that were not part of the 2,713 hijack attacks discussed earlier.

Algorithm 3 Real-time blacklist of hijacked network IP address blocks

Require: Sets of resuming IP prefixes P_r, their BGP origin ASes O_r and direct upstream ASes U_r
Require: History of BGP origin ASes H_{O_r} and direct upstream ASes H_{U_r} related to prefixes P_r
Require: IRR records I_r related to prefixes P_r, BGP origin ASes O_r and direct upstream ASes U_r
Require: Prefixes P'_h, BGP origin ASes O'_h and direct upstream ASes U'_h seen in previous hijacks
Ensure: Set of suspicious hijacked IP prefixes $P_h = \{(p_h, s) | p_h$ is a hijacked prefix with score $s\}$

```
procedure SuspiciousRoutingStateIdentification
    for p_r in P_r do
        s ← 0.0
        if p_r ∈ P'_h or O_r(p_r) ∈ O'_h or U_r(p_r) ∈ U'_h then
            #IP prefix, BGP origin AS or upstream AS involved in previous hijacks
            s ← 1.0
        else
            if I_r(p_r) = ∅ or I_r(O_r) = ∅ or I_r(U_r) = ∅ then
                #Unallocated/unassigned/reserved IP prefix, BGP origin or upstream AS
                s ← s + 0.2
            end if
            if O_r(p_r) ∉ H_{O_r}(p_r) then
                #New BGP origin AS
                if I_{r_owner}(p_r) ≠ I_{r_owner}(O_r(p_r)) then
                    #IP prefix and BGP origin AS owner mismatch
                    s ← s + 0.1
                end if
                if I_{r_country}(p_r) ≠ I_{r_country}(O_r(p_r)) then
                    #IP prefix and BGP origin AS country mismatch
                    s ← s + 0.1
                end if
                if outdegree(O_r(p_r)) ⋘ and count_prefix(O_r(p_r)) ⋙ then
                    #Prefix hijack: origin AS connected to a few upstream ASes
                    #and originating many prefixes
                    s ← s + 0.2
                end if
            end if
            if U_r(p_r) ∉ H_{U_r}(p_r) then
                #New upstream AS
                if O_r(p_r) ∉ I_{r_import/export}(U_r(p_r)) or U_r(p_r) ∉ I_{r_import/export}(O_r(p_r)) then
                    #Routing policy violation
                    s ← s + 0.1
                end if
                if I_{r_rir}(O_r(p_r)) ≠ I_{r_rir}(U_r(p_r)) then
                    #BGP origin AS and upstream AS registry (RIR) mismatch
                    s ← s + 0.1
                end if
                if indegree(U_r(p_r)) ⋙ and count_prefix(O_r(p_r)) ⋘ then
                    #AS hijack: upstream AS connected to many origin ASes,
                    #each of them originating a few prefixes
                    s ← s + 0.2
                end if
            end if
        end if
        if s ≥ 0.4 then
            P_h = P_h ∪ {(p_r, s)}
        end if
    end for
end procedure
```

Recall from Chapter 3 that the validation of candidate hijacks involves four different steps: (i) validate the suspicious routing history via the decision tree (of page 60), (ii) verify the consistency of the BGP origin AS and direct upstream provider ASes in BGP announcements with respect to information published in the IRRs, (iii) check the suspicious prefixes against the Spamhaus DROP list, and (iv) search archives of network operational mailing lists for external feedback and confirmation. While step (i) and (iii) are automated, steps (ii) and (iv) consists of a manual process and are thus time consuming. The validation of the remaining cases is something we are currently working on but we nevertheless have already been able to validate 45 hijacks out 1,562 suspicious cases and found no false positive so far. These cases exhibit the exact same characteristics as the corpus of 2,713 discussed earlier in this Chapter. Interestingly, the investigation of these hijacks allowed us to discover two new *invalid direct upstream provider ASes* (AS hijacks as described on page 132), namely AS198059 "C-media LLC" and AS58061 "Trade House BelRosResursu Ltd.", involved in the hijack of 45 IP address blocks. While we are confident that new malicious hijacks will likely be uncovered from the 1,517 cases left to validate, we also envision the presence of false positives. We leave as future work the investigation into these possible false-positive cases and a sensitivity analysis of the metrics used in the system.

Real-time blacklist to monitor "BGP spectrum agility"

We decided to let the real-time blacklist system run after the experimental validation period, *i.e.,* after June 2014. Figure 4.21 depicts the number of IP address blocks listed during the two months of July and August 2014. We can see that the number of entries can vary a lot from one day to the other, with an average of 17.3 IP address blocks listed per day. The peak at the beginning of July 2014 is due to the temporary tuning of parameters of Algorithm 3, which resulted in additional networks being abnormally flagged as suspicious hijacks.

In an effort to continue the study of malicious BGP hijacks and "BGP spectrum agility", we have initiated a collaboration with external partners to collect additional network traffic information related to suspicious hijacked IP address blocks. In the context of a collaboration with the Belgian CERT.be [9] and the BELgian national research NETwork (BELNET) [4], we have recently started to use our blacklist of suspicious hijacked networks to automatically collect NetFlow data related to these blocks of IP addresses and allow us to gain more insights into the usage, besides sending spam and hosting scam websites, behind malicious BGP hijack campaigns. The analysis of this additional data source and the potential new insights into the usage of hijacked IP address blocks represent an exciting future work avenue.

4.8 Conclusion

We conclude by providing concrete lessons that can be leveraged to improve existing spam and BGP hijack mitigation techniques and thwart these attacks.

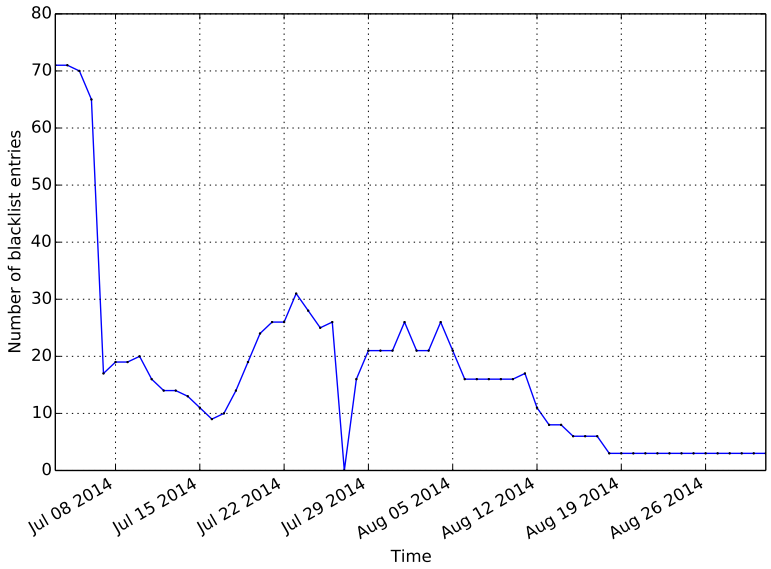


Figure 4.21 – Real-world deployment of the real-time blacklist of hijacked IP address blocks.

Lesson 1 *We have confirmed the existence, as of 2014, of BGP spectrum agility in the real-world in the form of stealthy and persistent campaigns of malicious BGP hijacks.*

Lesson 2 *Today's BGP hijack mitigation systems, such as [7, 96, 101, 116, 157], are **blind** to hijacks of registered though **unannounced** IP address space carried out by announcing an IP address block using its **valid** BGP origin AS but via an invalid upstream provider AS. The complete deployment of BGPsec and ROAs would prevent these attacks. In the meantime, we would suggest BGP hijack detection systems to include signatures for these hijacks based on the characteristics we uncovered.*

Lesson 3 *Owners of unannounced IP address blocks leave them vulnerable to hijacking. A best practice would be to announce all blocks even if they are unused.*

Lesson 4 *A worldwide hunt for orphan IP address blocks should be launched to prevent them from being hijacked and further used for malicious purposes. Additionally, IP address block owners that cease to exist or do not require the IP resources anymore should (be forced to) return them. Keeping IRRs and RPKI data fresh is therefore key to prevent hijacks of such IP address space.*

Lesson 5 *Uncovered hijacks involved many different IP address blocks and origin ASes but very few invalid BGP origin ASes and direct upstream provider ASes. This suggests that ASes identified as invalid or malicious in previous hijacks can be leveraged to identify subsequent hijacks or even block traffic from and to IP address blocks advertised via these ASes.*

Lesson 6 *We have built upon some characteristics of the uncovered hijacks to develop a real-time blacklist of hijacked IP address blocks to help detect and mitigate future instances of malicious BGP hijack attacks.*

As future work we plan to expand the collaboration we have recently initiated with CERTs, ISPs and the NANOG and RIPE communities at large. A concrete outcome of these ongoing discussions was the confirmation that one of the ASes found to be malicious by our system and responsible for the hijack of 793 IP prefixes has seen his peering contract terminated by its valid upstream ISP.

5

Conclusion and future challenges

The investigations carried out in this research thesis have brought new interesting insights into the little-known phenomenon of malicious BGP hijacks. They are summarized in this Chapter. We also present our solution to the research problems formulated in the introduction of this thesis. Finally, we conclude this work by presenting some interesting new avenues for future work.

5.1 Research contributions

The introduction Chapter to this thesis has raised several research questions regarding the possible existence, as of 2014, of malicious BGP hijack attacks on the Internet, the prevalence of these attacks, and the attackers' modus operandi. To answer these questions we have identified *three* research problems that this thesis needed to tackle. Based on the developments and experimental results described in the previous chapters, we now show how we managed to solve these research problems.

Problem 1 (P1): *correlation of security-related and routing data.*

We successfully addressed this first research challenge by developing a large-scale data collection environment dubbed SpamTracer, described in Chapter 3. The proposed system consists in monitoring networks generating malicious network traffic, such as spam emails, traffic related to C&C servers or malicious websites hosting, etc, using a combination of BGP data and traceroute measurements from *multiple vantage points* distributed around the globe. The network traces collected for each network are further enriched with information from the Internet Routing Registries (IRRs) about IP address blocks and ASes seen in the traces. This approach was motivated by the lack of technique or system readily available for collecting and analysing data for a large-scale study of real-world malicious BGP hijack attacks.

Inspired by the characteristics of the malicious BGP hijacks described in the first observations of "BGP spectrum agility" back in 2006 in [148], SpamTracer was

designed to monitor networks as soon as they are seen generating nefarious network traffic to observe the transition from the hijacked state to the normal state of the network. It also enables us to monitor the routes to suspicious networks from the *control* and *data plane* perspectives. When combined with registration information from IRRs, the collected data provides us with a *representative* set of features to characterise the observed routing behaviors and identify the ones most likely resulting from a BGP hijack.

Conclusion 1: we have shown that the SpamTracer data collection system successfully solved the first research problem **P1** introduced by this thesis. By running the experiment for almost two years, we were able to build a *comprehensive* dataset of enriched routing-related information related to offending networks. We now show how we managed to solve the next two research problems thanks to the data collected using SpamTracer.

Problem 2 (P2): *assessment of the existence of malicious BGP hijacks.*

The solution brought to the second research problem posed by this thesis is *twofold*.

First, in Chapter 3, we presented a novel approach to identify and validate possible instances of malicious BGP hijacks from the SpamTracer dataset. The developed methodology consists of a multi-stage scoring and filtering process: (i) it combines state-of-the-art as well as new heuristics to extract anomalous routing behaviours from the collected data, (ii) and then uses a multi-criteria decision analysis (MCDA)-based hijack identification model to score and retain the cases most likely resulting from a BGP hijack. These results were further enriched using external data sources and feedback from network owners in order to validate candidate hijack cases.

Second, in Chapter 4, we applied our methodology on a 22 months dataset, which enabled us to identify and validate over two thousand instances of malicious BGP hijack attacks, which have taken place on a regular basis over the whole period of the experiment. Moreover, we observed a specific class of "agile" spammers that appear to routinely, persistently and -quite likely- automatically make use of this attack to launch spam campaigns from stolen IP address blocks. Some of the observed hijacks were also confirmed by victim network owners and an ISP who was unwittingly involved in several hijack cases.

The observed attacks thus confirm the existence, as of 2014, of malicious BGP hijacks and the "BGP spectrum agility" phenomenon. These experimental results also enabled us to validate the two assumptions made in the introduction of this thesis, namely that (i) if "BGP spectrum agility" is still a problem worth of consideration in 2014 then it is likely to be observed used by spammers and (ii) if we monitor the data and control plane routes to networks generating malicious network traffic from sufficiently distributed vantage points then we should be able to uncover malicious BGP hijacks if there are any.

Conclusion 2: In the light of the findings related to the malicious BGP hijacks phenomenon exposed in Chapter 4, obtained using the methodology presented in Chapter 3, we can claim to have successfully solved the research problem **P2**.

Problem 3 (P3): *if existent, assessment of the prevalence of malicious BGP hijacks and characterisation of the attackers' behavior.*

In Chapter 4, we presented a forensic analysis of the identified cases of malicious BGP hijacks, which enabled us to assess the *prevalence* of such attacks and their *threat* on current security systems, characterise the *modus operandi* of attackers, and propose *improvements* to current BGP hijack detection and mitigation techniques.

First, we have shown that malicious BGP hijack attacks have taken place *frequently* and *persistently* on the Internet over the 22 months period of the experiment described in this thesis.

Second, we have unveiled a sophisticated modus operandi used by cybercriminals to stealthily hijack blocks of IP addresses that were left unannounced by their owner; actually squatting vacant and often orphan IP address space. Moreover, attackers were found to abuse different AS number resources in the bogus BGP announcements in an effort to hinder their traceability. We have also described two different types of hijack phenomena, namely short-lived and long-lived hijacks.

Third, we have shown that the identified hijacks were rather successful at circumventing traditional BGP hijack and spam protection techniques. We have proposed new directions to defend more effectively against this emerging threat posed by malicious BGP hijacks, in particular (i) pursue the development and deployment of security extensions to BGP like the RPKI-based secured BGP route origination and propagation, (ii) leverage IP and AS resources abused in previous hijacks to detect future hijack instances as attackers observed in our traces tend to extensively reuse these resources, and (iii) encourage network operators to return or publicly advertise their IP address ranges to dispose of the 20% of the global IPv4 address that is currently allocated but not publicly advertised, which makes it potentially vulnerable to such malicious BGP hijacks.

Fourth, we took advantage of characteristics of the unveiled hijack cases to build a real-time blacklist of hijacked network ranges to help detect and mitigate such attacks.

Conclusion 3: In the light of these last findings we can conclude that we have successfully solved the research problem **P3** posed by this thesis.

In conclusion, besides successfully answering the research problems posed, this thesis has brought completely new and interesting insights into the little-known phenomenon of *malicious BGP hijacks*. To the best of our knowledge, this work is the first to provide such a comprehensive analysis of validated real-world hijack attacks, some of which were observed generating spam emails and hosting scam-related websites, and to expose the modus operandi of the sophisticated cybercriminals using

such attacks. Finally, besides being an eye-opener to the fact that frequent, persistent and stealthy BGP hijack attacks have taken place on the Internet for months or even years, the results presented in this thesis have also opened new interesting and exciting avenues for future research, which should enable us to answer some questions left open or introduced by this work, for example determine the usage by cybercriminals of the identified hijacked IP address blocks from which we did not observe any malicious traffic.

5.2 Future research and perspectives

Besides shedding light on the security threat posed by these attacks, this work has uncovered different research challenges and future work perspectives. In this work we have confirmed the existence, as of 2014, of real-world stealthy and persistent malicious BGP hijacks and then characterised them along with the modus operandi of the attackers. As a first step we set up a data collection framework for the monitoring of the routing behavior of networks generating malicious network traffic, in particular spam emails. A *first* direction for future work thus goes towards augmenting the set of features collected to characterise the routing behavior of monitored networks. Thanks to the dataset built using our data collection prototype, we have been able to extract and score abnormal routing behaviors and then validate a set of candidate BGP hijack cases leading to the identification of real-world malicious BGP hijack instances. A *second* set of future challenges is thus related to the analysis of the collected data. Finally, while our analysis of these attacks allowed us to give a first answer to the questions posed by this thesis, the exploitation of these results as well as the refinement of the picture of the malicious BGP hijacks phenomenon depicted in this work bring some other new exciting research perspectives.

5.2.1 Expanding the scope of the security-related data

The primary security-related dataset used in this work was a live feed of spam emails. This choice was motivated by the first observations of "BGP spectrum agility" reported in [148] and describing spammers sending spam from hijacked networks. Nevertheless, we witnessed only a very small number of hijacked networks sending spam to our spamtraps or listed in one of the queried spam sender IP blacklists. We gathered evidence of scam websites hosted on some of the stolen IP address blocks but they as well were related to only a small percentage of all identified hijacked blocks. We did not find evidence of any other type of malicious activities performed from the identified hijacked prefixes. The absence of spam and other scam-related traffic in our data for most hijacked prefixes may be due to incomplete visibility into malicious activities associated with these networks. Pitsillidis et al. showed in [140] that spamtrap feeds are biased towards large spam campaigns and thus fail to capture outlying ones. According to the authors, a better coverage of the global spam activities could be achieved using a spam feed obtained from an ISP or operational spam blacklists. The absence of malicious network traffic may also indicate that hijacked networks

are used to carry out other types of nefarious activities, such as host C&C servers, launch DDoS attacks, distribute malware.

5.2.2 On the routing data collection

In the analysis of the data collected by SpamTracer, we noticed that, if BGP routes and traceroute paths to monitored networks are not collected from the same vantage points, this complicates the interpretation of discrepancies between such paths for a given network. This prevents us from effectively leveraging these discrepancies in our routing anomalies scoring system.

Furthermore, while the large scale deployment of SpamTracer in the Internet cloud enabled us to increase the number and geographical distribution of traceroute vantage points, we monitor each network from only six BGP collectors and three traceroute vantage points, which might affect our visibility of routing changes [84, 156]. The monitoring capacity of 40K networks per day, with one monitored IP address per network, limited by the available hardware resources of the measurement nodes and network connection latencies, allows to observe the routing behavior of two fifths of the spam networks seen in our spam feed. A higher monitoring capacity would enable a larger number of spam networks as well as networks from other security-related datasets to be monitored, and for a longer period of time.

Finally, a more serious limitation of our methodology is that neither BGP nor traceroute measurements were designed to infer the AS-level connectivity of the Internet and capture the complex inter-AS relationships, hence all results inferred from such data can only be as accurate as the data [153]. However, we tried to balance this limitation by setting up our own data collection process enabling us to collect the most appropriate data for studying the routing-level behavior of offending networks.

5.2.3 On the multi-stage scoring and data filtering

In the context of our MCDA-based approach to identify candidate BGP hijacks from the data collected with SpamTracer, using an aggregation function such as the weighted ordered weighted average (WOWA) we could include expert knowledge into the model by means of some parameter vectors. While the values of these parameter vectors in our model were set empirically, we leveraged our experience in analyzing routing incidents to ensure that our model could effectively differentiate suspicious from benign cases, even with more "borderline" or ambiguous cases. Hence, we leave as future work a more extensive sensitivity analysis of the results for various values of the few parameters used in our algorithms, for example by running simulations on synthetic data to discriminate benign from hijack routing behaviors.

Moreover, it is noteworthy that the primary goal of our methodology was to narrow down the large number of cases so as to retain a set of interesting BGP hijack candidates. It has never been our intent to design a new fullly fledged BGP hijack detection system. Along the previous point, we do not claim that we have

found all BGP hijacks that could have been found in our dataset. What matters is that our results must be seen as a proof of the existence of these recurring attacks.

Finally, we are also willing to explore new routing anomalies that could be extracted from the collected BGP and traceroute data to refine our identification of abnormal routing behaviors.

5.2.4 On the validation of candidate BGP hijacks

As we showed in [177], correlating routing anomalies with traces of malicious activities related to the same networks is insufficient to evidence harmful BGP hijacks. The methodology presented in Chapter 3 to validate candidate hijacks enabled us to gather contextual information to build a compelling case for the reality of these malicious hijacks. Feedback from an ISP who was unwittingly involved in several hijack cases apparently performed by a Russian spammer allowed us to confirm some of the identified hijacks. Consequently, we are planning to automate asking direct feedback from network operators at ISPs or from network owners involved in candidate hijack cases. We are also seeking to partner with organisations that have an interest in the investigation or mitigation of hijack events, such as RIRs, CERTs, governmental agencies, as well as communities of network operators, such as NANOG.

In an effort to improve the validation of the malicious nature of identified hijacks, we are also exploring using NetFlow data related to hijacked IP address blocks to assess the footprints left by attackers during the hijacks, in the same way that such data was used in the forensic analyses of former hijack cases in [154, 177]. The added value of using NetFlow traces to the validation of candidate hijacks and the study of confirmed cases is that it provides generic network traffic traces allowing to discover potential new types of harmful activities performed form such networks. As explained in Section 4.7.3 of Chapter 4, we have recently started to collaborate with the BELgian research NETwork (BELNET) and the Belgian CERT (CERT.be) to automatically collect NetFlow data related to the suspicious hijacked prefixes identified in our recently deployed real-time blacklist.

5.2.5 Exploitation of results

In Section 4.7.3 of Chapter 4 we explored the use of some characteristics of the hijack attacks identified in this work to build a real-time blacklist of suspicious hijacked networks to help with the detection and mitigation of future hijack instances. While we have experimentally validated our approach, we intend to evaluate the quality of the algorithm and the parameters used in the current deployment.

5.2.6 Monitor the IPv6 Internet

In this work we have focused our attention on the IPv4 Internet for one reason: we lack security-related data in IPv6, in particular logs of spam sent from IPv6 network ranges. However, the recent exhaustion of the Internet Assigned Numbers Authority

(IANA) unallocated IPv4 address pool in 2011 [135] and the reduction of available IP address space from Regional Internet Registries (RIRs)[1] is reducing the amount of unused IP address space, which appeared from our data to be specifically targeted by hijackers. As of today, as much as 20% of the IPv4 address space appears to be allocated yet not publicly announced[2]. Moreover, the progressive adoption of IPv6 by ISPs on the Internet[3] suggests that malicious BGP hijack attacks are or will occur in the IPv6 world similarly to IPv4. Since the version of BGP running in IPv6 is not different from the one used in IPv4, the BGP hijack attack model built for IPv4 can also be applied to the IPv6 world. As of the writing of this document, we are not aware of any reported cases of BGP hijack in the IPv6 Internet, though.

[1] http://www.potaroo.net/tools/ipv4/

[2] Based on statistics published by RIR's and available at http://bgp.potaroo.net/ipv4-stats/prefixes.txt

[3] http://bgp.potaroo.net/v6/as2.0/index.html

6

Résumé en français

Contents

6.1 Introduction 118
6.1.1 Exposé du problème 119
6.1.2 Objectifs de recherche 120
6.1.3 Structure de la thèse 122
6.2 Positionnement par rapport à l'état de l'art 122
6.2.1 Les attaques par détournement BGP malveillant 122
6.2.2 Le filtrage anti-spam 123
6.2.3 La sécurité du routage inter-domaine dans l'Internet 123
6.3 SpamTracer 124
6.3.1 La collecte de données de routage 125
6.3.2 Processus d'extraction et d'évaluation des anomalies de routage 126
6.3.3 Validation des cas suspects de détournements BGP 128
6.3.4 Analyse de l'origine des détournements BGP 130
6.4 Résultats 131
6.4.1 Résultats : la collecte des données de routage 131
6.4.2 Résultats : processus d'extraction et d'évaluation des anomalies de routage 131
6.4.3 Résultats : validation des cas suspects de détournements BGP 132
6.4.4 Résultats : analyse de l'origine des détournements BGP . . 139
6.4.5 Efficacité des contres-mesures 140
6.4.6 Opérationnalisation de SpamTracer 143
6.4.7 Enseignements 145
6.5 Conclusion et perspectives futures de recherches 146
6.5.1 Contributions scientifiques 146
6.5.2 Perspectives futures de recherche 148

6.1 Introduction

L'infrastructure de routage de l'Internet de connue pour être vulnérable aux attaques par détournement BGP. Cette attaque consiste en la prise de contrôle d'un bloc d'adresses IP sans le consentement de son propriétaire légitime. Cette attaque est rendue possible par le fait que dans BGP [150] - le protocole utilisé pour le routage inter-domaine dans le l'Internet - l'échange d'information de routage est basé sur la confiance mutuelle entre les réseaux (communément appelés « autonomous systems (ASes) » dans le jargon BGP) interconnectés. Des attaques par détournement BGP accidentels, donc pas nécessairement malveillant, surviennent régulièrement dans l'Internet. Ils sont généralement dus à des erreurs de configurations de la part des opérateurs réseau. Quelques cas ont été publiés sur des forums d'opérateurs réseau, comme NANOG, ou des sites web [40, 44, 46, 91]. Des techniques permettant de détecter ce type d'attaques ont été proposées afin d'aider les opérateurs réseaux à surveiller leur propre réseau et ainsi réagir rapidement en cas de possibles perturbations dues à ces attaques. Ces approches souffrent d'un taux de fausses alarmes très élevé [96, 116, 157, 194], ce qui est acceptable pour ces utilisateurs puisqu'ils connaissent leur réseau. D'autres techniques ont été proposées pour empêcher que ce type d'attaques puissent se produire [101, 102, 119] mais leur adoption à grande échelle est sérieusement freiné par le coût de leur déploiement.

En 2006, Ramachandran et al. [148] ont introduit une nouvelle menace pour la sécurité de l'Internet appelée « **BGP spectrum agility** ». Les auteurs expliquent avoir observé, sur une période de quelques mois, des courriers indésirables (ci-après désignés comme « spam ») envoyés depuis un ensemble de larges blocs d'adresses IP (*i.e.,* /8) ; chacun de ces blocs, précédemment non annoncés dans BGP, ayant été annoncés pendant une très courte période de temps (*i.e.,* moins de 24 heures). Plus tard, d'autres travaux de recherche ont rapporté avoir observé des courriers indésirables venant de préfixes IP détournés [70, 96]. En outre, complémentant le travail de Schlamp et al. [154], nous avons récemment exposé dans [47, 178] le cas d'une attaque par détournement BGP ayant affecté quelques blocs d'adresses IP utilisés par la suite pour envoyer du spam. Plus récemment, nous avons montré dans [177], grâce à un cas d'étude réel, que la corrélation d'anomalies de routage et de trafic réseau malveillant, comme par exemple du spam, est à elle seule insuffisante pour conclure qu'il s'agit d'une attaque par détournement BGP malveillant.

Hormis ces quelques cas et en dépit de la volonté apparente de certains propriétaires de réseaux d'être capables de détecter le vol de leurs blocs d'adresses IP, à notre connaissance, il n'existe à l'heure actuelle aucune preuve que ce type d'attaques vaut la peine qu'on s'en préoccupe ; puisque personne n'a montré jusqu'à présent que les cybercriminels utilisent ce type d'attaques de façon régulièrement pour commettre différentes activités malveillantes. S'ils en étaient capables, cela constituerait une menace sérieuse pour la sécurité de l'Internet, puisque cela leur permettrait non seulement d'envoyer du spam en déjouant les classiques listes noires (ou « blacklists »), mais cela leur permettrait surtout de lancer de larges attaques par déni de service distribué (ou « DDoS ») avec un coût presque nul, ou bien encore de lancer des attaques dites de « l'homme du milieu » (ou « man-in-the-middle ») contre

n'importe quelle cible. Ainsi, nous pensons qu'il est nécessaire d'évaluer de façon rigoureuse l'existence et l'importance de cette menace potentielle.

Ce phénomène soulève différentes questions actuellement sans réponse et auxquelles ce travail à pour objectif de répondre. A savoir, en 2014, les cybercriminels utilisent-ils les attaques par détournement BGP afin de réaliser d'autres activités malveillantes depuis les adresses IP volées ? Si non, quelles sont les hypothèses qui nous permettent de conclure que ce phénomène n'existe pas ? Si oui, à quel point cette attaque est-elle répandue dans l'Internet et quel est le mode opératoire des attaquants ?

6.1.1 Exposé du problème

Pour réponse aux précédentes questions, il est nécessaire de résoudre les trois problèmes suivants : (i) corréler des traces d'activités malveillantes avant des information de routage, (ii) démontrer l'existence ou l'absence d'attaques par détournement BGP malveillant, et, (iii) si ce phénomène existe réellement, évaluer à quel point il est répandu et caractériser le comportement des attaquants.

(P1) La corrélation de données sécurité et réseau

Afin de pouvoir corréler des traces d'activités malveillantes et des anomalies de routage dans l'Internet, il est nécessaire de collecter des données relatives à des événements pertinents pour la sécurité, par exemple la réception de spam, ainsi que des données relatives à l'état du routage des réseaux concernés au même moment dans le temps. La qualité des données doit être assurée pour nous permettre d'identifier de possibles attaques et par la suite de caractériser le comportement des acteurs responsables.

(P2) Evaluation de l'existence d'attaques par détournement BGP malveillant

Pour démontrer l'existence ou l'absence d'attaques par détournement BGP malveillant en utilisant les données précédemment collectées, il est nécessaire de pouvoir extraire des anomalies de routage résultantes de scénarios connus d'attaques par détournement BGP. Afin d'identifier de manière fiable les cas d'attaques par détournement BGP malveillant, il est nécessaire d'utiliser et d'étendre les techniques actuelles [59, 96, 116, 144, 167, 189, 194]. Etant donné l'absence de données permettant de confirmer nos résultats, il également important de trouver un mécanisme pour valider les cas suspects d'attaques identifiés.

(P3) Si elles existent, déterminer à quel point ce type d'attaques est répandu et caractériser le comportement des attaquants

Si le phénomène de « BGP spectrum agility » se trouve être toujours d'actualité, il est nécessaire d'identifier à quel point ce type d'attaques est répandu, et ce afin

de déterminer la menace réelle posée par ces attaques sur les mécanismes actuels de défenses contre diverses activités malveillantes et plus généralement sur sécurité de l'Internet.

6.1.2 Objectifs de recherche

Ce travail a pour but d'aborder les problèmes de recherche précédemment introduits en construisant un environnement de collecte et d'analyse de données pour l'étude des attaques par détournement BGP malveillant.

> **Affirmation :** *Bien qu'il existe des techniques visant à détecter ou empêcher les attaques par détournement BGP, il n'existe aucun environnement utilisable tel quel pour l'étude de ces attaques lorsqu'elle sont réalisées dans le but d'appuyer d'autres activités malveillantes, comme l'envoi de spam, l'hébergement de sites web d'hameçonnage, ou encore le lancement d'attaques DDoS.*

Cette affirmation, abordée en détail dans la Section 6.2, découle des trois principales limitations des méthodes actuelles de détection d'attaques par détournement BGP. Elle s'articule autour des trois points suivants.

1. La majorité des techniques existantes de *détection* d'attaques par détournement BGP sont basées uniquement sur la surveillance passive du plan de contrôle (BGP) [59, 116, 144, 189] et en conséquence souffrent d'un taux de fausses alarmes très élevé. D'autres méthodes [96, 157, 194] complète cette surveillance passive par des mesures réseau, comme par exemple « ping » ou « traceroute », afin d'affiner la précision de la détection. Cependant, ces techniques ne permettent de couvrir tous les types d'attaques par détournement qu'il est nécessaire d'étudier dans ce travail.

2. Des techniques de *prévention* d'attaques par détournement BGP, comme le système RPKI [120], visant à combler, habituellement moyen de la cryptographie, les lacunes de BGP en terme de sécurité, sont quant à elles freinées dans le déploiement par les changements importants qu'elles nécessitent à l'infrastructure de routage actuelle.

3. Enfin, dans leur première observation du phénomène de « BGP spectrum agility » en 2006, Ramachandran et al. ont rapporté que des spammeurs ont détourné de larges blocs d'adresses IP (*i.e.,* /8), précédemment non annoncés dans BGP, et ce pendant une courte période de temps (*i.e.,* moins d'un jour). Aussi, les techniques de détection et de prévention d'attaques par détournement BGP sont actuellement inefficaces contre ce type particulier d'attaques.

Pour aborder les problèmes de recherche décris ci-dessus, nous faisons les hypothèses suivantes.

Hypothèse 1 : *En nous basant sur les premières observations du phénomène de « BGP spectrum agility », si ce phénomène existe encore à l'heure actuelle, nous pensons qu'il doit être utilisé par les spammeurs.*

Hypothèse 2 : *Combiner une surveillance passive du plan de contrôle avec des mesures réseau actives depuis un grand nombre de points de vue devraient permettre de mettre au jour ce types d'attaques par détournement BGP.*

Sur base de ces hypothèses, nous envisageons d'aborder les problèmes de recherche de la façon suivante.

Pour aborder le premier **problème (P1)**, nous allons corréler des traces d'activités malveillantes avec des anomalies de routages. Pour cela, nous allons construire un système de collecte et d'analyse de données afin de démontrer rigoureusement et scientifiquement l'existence ou l'absence d'attaques par détournement BGP malveillant.

Pour aborder le second **problème (P2)**, nous allons étendre les technique actuelles de détection d'attaques par détournement BGP avec de nouveaux algorithmes. Pour valider les suspects identifiés, nous considérons également utiliser différentes sources de données externes ainsi que le retour certains opérateurs réseaux.

Enfin, pour aborder le **problème (P3)**, c'est-à-dire si les attaques par détournement BGP malveillant existent, nous souhaitons évaluer l'impact des ces attaques sur les mécanismes de défense actuels, comme par exemple les listes noires (ou « blacklists ») dans les filtres anti-spam. Nous souhaitons également utiliser ces attaques identifiées pour en apprendre plus sur le comportement et le mode opératoire des attaquants.

Thèse.

- La corrélation d'événements relatifs à la sécurité, en particulier des courriers électroniques indésirables, avec des informations de routage sur les réseaux ayant envoyés ces trafic réseau malveillant nous permettent d'identifier des cas suspects d'attaques par détournement BGP. L'utilisation de source de données externes ainsi que le retour d'opérateurs réseau permettent de confirmer la nature malveillante de ces cas.

- En 2014, le phénomène de *BGP spectrum agility* peut être observé dans l'Internet sous la forme de longues et furtives campagnes d'attaques par détournement BGP malveillant.

- Il existe une catégorie de spammeurs capables de détourner, de manière systématique, continuelle et, vraisemblablement, automatique un grand nombre de blocs d'adresses IP afin de les utiliser pour envoyer du spam et appuyer leurs diverses activités malveillantes.

- Les attaques par détournement BGP malveillants consistent en l'annonce dans BGP de blocs d'adresses IP, précédemment non annoncés, en utilisant des numéros d'ASs volés. Ces attaques semblent être très efficaces pour contourner les mécanismes traditionnels de défense contre le spam et les attaques par détournement BGP.

- Certaines caractéristiques des attaques identifiées, comme par exemple les numéros d'ASs volés par les attaquants, peuvent être utilisées afin de détecter ces attaques et ainsi atténuer leur impact sur la sécurité de l'infrastructure de routage et de l'Internet en général.

6.1.3 Structure de la thèse

Le reste de ce Chapitre est organisé de la façon suivante. Dans la Section 6.2, nous positionnons ce travail rapport à l'**état de l'art**. Nous discutons des cas connus d'attaques par détournement BGP malveillant, des technologies actuelles de filtrage anti-spam, ainsi que des limitations des techniques développées pour détecter ou empêcher les attaques par détournement BGP. Cette Section justifie ainsi le travail réalisé dans cette thèse. Dans la Section 6.3, nous décrivons SpamTracer, l'environnement de collecte et d'analyse de données que nous avons développer afin d'étudier l'existence des attaques par détournement BGP malveillant. La Section 6.4 rapporte les résultats expérimentaux obtenus en utilisant SpamTracer pendant presque deux ans. Premièrement, nous démontrer l'existence, à l'heure actuelle, de campagnes d'attaques par détournement BGP malveillant. Deuxièmement, nous proposons, sur base des attaques identifiées, des pistes afin de mieux se protéger contre cette menace.

Enfin, nous concluons cette thèse dans la Section 6.5 en présentant la solution apportée aux différents problèmes de recherche introduits ci-dessus. Nous discutons également des perpectives futures de recherche dans l'étude des attaques par détournement BGP malveillant.

6.2 Positionnement par rapport à l'état de l'art

6.2.1 Les attaques par détournement BGP malveillant

En 2006, Ramachandran et al. [148] ont introduit une nouvelle menace pour la sécurité de l'Internet appelée « **BGP spectrum agility** ». Les auteurs expliquent avoir observé, sur une période de quelques mois, des courriers indésirables (ci-après désignés comme « spam ») envoyés depuis un ensemble de larges blocs d'adresses IP (*i.e.,* /8) ; chacun de ces blocs, précédemment non annoncés dans BGP, ayant été annoncés pendant une très courte période de temps (*i.e.,* moins de 24 heures). Plus tard, Hu et al. [96] et Duan et al. [70] ont rapportés des observations similaires au phénomène de « BGP spectrum agility ». Cela étant dit, nous avons récemment montré dans [177], au moyen d'un cas d'étude pratique, que la corrélation entre des

anomalies de routage BGP et du traffic réseau malveillant (*e.g.,* du spam) venant des réseaux affectés est, à elle seule, insuffisante pour identifier de manière catégorique des cas d'attaque par détournement BGP malveillant. Entre temps, Schlamp et al. ont également décrit, dans [154], une attaque durant laquelle plusieurs blocs d'adresses IP ont été détournés via BGP, et ce pendant plusieurs mois, afin de les utiliser pour réaliser des activités malveillantes, comme envoyer du spam.

Ces quelques précédentes publications montre que le phénomène des attaques par détournement BGP malveillant est bien réel et est observable dans l'Internet. Cependant, la rareté avec laquelle ces attaques ont été observées depuis 2006 donne l'impression que cette menace reste très anecdotique et qu'il n'existe pas vraiment d'infrastructure mise en place par les cybercriminels pour automatiser efficacement et systématiquement le lancement de ce type d'attaque.

6.2.2 Le filtrage anti-spam

Les techniques utilisées pour filtrer les courriers indésirables (ou spam) peuvent être de deux types : pré-acceptation et post-acceptation. Les techniques dites de « pré-acceptation » tirent parti de caractéristiques réseau bas niveau pour identifier du trafic de spam avant qu'il n'arrive au niveau du serveur de courrier électronique. Ces techniques sont souvent peu gourmandes en ressources et sont donc utilisées comme une première couche de défense dans les filtres anti-spam. La plus populaire de ces techniques est de loin la liste noir (« blacklist ») d'adresses IP de machines envoyant du spam [11, 34, 42]. Malgré leur coût de maintenance, leur incomplétude [148] et leurs possibles inexactitudes, ces listes sont toujours très utilisées dans les systèmes de filtre anti-spam commerciaux [10, 39]. D'autre part, les technique de « post-acceptation » se basent sur des caractéristiques extraites des en-têtes et du contenu des courriers électroniques, comme par exemple des parties de texte ou encore des adresses web, pour différentier les courriers légitimes des indésirables. Ces techniques sont habituellement efficaces pour le filtrage anti-spam mais sont aussi beaucoup plus gourmandes en ressources.

6.2.3 La sécurité du routage inter-domaine dans l'Internet

On distingue deux grandes classes de défense contre les attaques par détournement BGP. (i) Les techniques dites « de détection » visent à surveiller l'état de l'infrastructure du routage dans l'Internet et à déclencher une alarme lorsqu'un changement anormal dans le routage est observé. (ii) Les autres techniques cherchent quant à elles à ajouter à BGP des mécanismes permettant d'empêcher que ces attaques par détournement puissent se produire. Contrairement à ces dernières, les techniques de détection ont l'avantage de ne nécessiter aucune modification du logiciel tournant sur les routeurs BGP, ce qui les rend beaucoup plus facile à déployer en situation réelle.

Plusieurs techniques visant à renforcer la sécurité de BGP en modifiant le protocole lui-même ont été proposées [101, 119, 120]. Ces techniques utilisent habituellement la cryptographie afin de signer différents éléments des messages BGP et

ainsi assurer l'authenticité et l'intégrité des informations de routage échangées. Récemment, un système de sécurisation de BGP reposant sur une RPKI (« Resource Public Key Infrastructure ») [101] et visant à empêcher n'importe qui d'annoncer des préfixes IP dans BGP sans y être autorisé a été adopté par la communauté réseau opérationnelle. Bien qu'encore peu répandu, ce système est progressivement déployé dans les différents réseaux de l'Internet.

En ce qui concerne les techniques de détection d'attaques par détournement, on distingue encore deux méthodes différentes. La première méthode [113, 115, 116, 144] propose de détecter les attaques par détournement en surveillant de manière passive l'état de l'infrastructure de routage. Cependant, cette méthode est connue pour produire un grand nombre de fausses alertes en raison de la forte similarité entre les symptômes d'une attaque par détournement BGP et certains changements de routage légitimes pouvant être observés dans BGP. Cette méthode est donc principalement utilisée par les opérateurs réseaux afin de surveiller leur propre réseau. En effet, étant donné qu'ils connaissent leur réseau ils sont capables d'écarter les fausses alarmes. Dans [113, 159], les auteurs utilisent un ensemble d'heuristiques [113] et les Registres du Routage dans l'Internet (« Internet Routing Registries » ou « IRRs ») éliminer certaines fausses alertes correspondant à des changement de routage bénins. Dans [115, 144], les auteurs utilisent certaines propriétés topologique et géographique du routage dans l'Internet [115] afin de détecter des changements anormaux dans le routage. Enfin, le système PHAS, proposé par [116], consiste à surveiller l'ensemble des ASs (« Autonomous Systems ») qui annoncent chaque préfixe IP dans l'Internet et à déclencher une alerte, accompagnée d'un courriel envoyé au propriétaire du préfixe IP concerné, au moindre changement observé dans cet ensemble d'ASs.

La deuxième méthode [96, 157, 189, 194] combine la surveillance passive de l'infrastructure de routage avec le sondage actif des réseaux surveillés afin d'améliorer la détection en complétant les alertes observées dans le « plan de contrôle (BGP) » par des mesures réseaux effectuées dans « plan de données ». Ainsi, Hu et al. ont proposé, dans [96], une technique consistant à surveiller de manière passive l'infrastructure de routage et à confirmer ou infirmer chaque alerte déclenchée en effectuant différentes vérifications au niveau du plan de données, par exemple en utilisant le « idle scan » ou encore la détection du système d'exploitation des machines se trouvant dans le réseau concerné. Dans [157, 189], les auteurs propose de détecter les attaques par détournement BGP complétant les alertes au niveau « plan de contrôle (BGP) » en effectuant des mesures « ping ». Shi et al. [157] propose de lancer des « ping » vers les réseaux ayant manifesté des changements de routage anormaux depuis un grand nombre de points de collecte. A l'opposé, Zhang et al. [189] propose de lancer des « ping » depuis les réseaux touchés par un changement de routage anormal. Dans les deux cas, le but final est d'utiliser les mesures « ping » afin de valider ou pas les alertes BGP.

6.3 SpamTracer

Rappelons que dans la Section 6.1 nous avons établi que l'objectif principal de cette thèse est de répondre à la question : *en 2014, les cybercriminels utilisent-ils les*

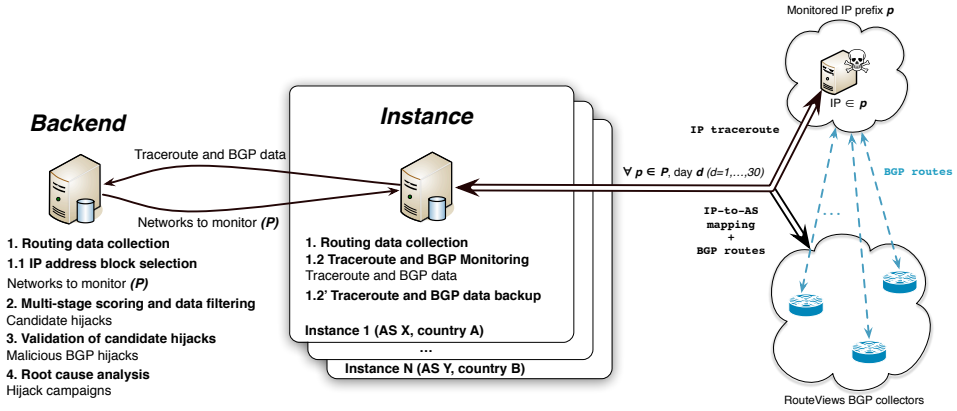


FIGURE 6.1 – Environnement expérimental.

attaques par détournement BGP malveillant pour utiliser l'espace IP volé afin de perpétrer d'autres activités malveillantes, comme envoyer du spam, distribuer des logiciels malveillants, etc ? Dans la Section 6.2 nous avons rappelé les quelques cas passés et isolés d'attaques par détournement BGP malveillant justifiant l'intérêt à l'heure actuelle d'étudier cette menace. Nous avons également montré qu'il existe un réel manque d'outil et de données pour répondre à cette question. Ces deux points justifient ainsi la construction d'une infrastructure de collecte et d'analyse de données spécifiquement pour l'étude des attaques par détournement BGP malveillant.

Nous avons donc construit un environnement expérimental, appelé SPAMTRACER, représenté schématiquement dans la Figure 6.1. Notre objectif est Ⓐ de collecter des informations sur l'état de routage de réseaux ayant envoyé du spam, Ⓑ d'extraire de ces données les blocs d'adresses IP ayant un comportement de routage anormal et ne conserver que les plus suspects d'entre eux, Ⓒ de (in)valider manuellement chaque cas suspect en utilisant des sources de données externes, et, enfin, Ⓓ d'analyser l'origine des cas validés afin d'en apprendre davantage sur le modus opérandi des pirates derrières ces attaques.

Notre approche se base sur l'hypothèse que lorsqu'un bloc d'adresses IP est détourné pour l'utiliser afin, par exemple, d'envoyer du spam, un changement dans l'état de routage du bloc sera observé lorsque l'attaque se terminera et donc que le bloc sera relâché par le spammeur. Ainsi, en surveillant l'état de routage d'un bloc aussitôt qu'on observe du spam depuis ce réseau, on cherchera un changement de routage du bloc indiquant la fin de l'attaque par détournement. Dans la suite de cette section nous décrivons en détail les différentes parties de notre environnement expérimental.

6.3.1 La collecte de données de routage

Sélection de blocs d'adresses IP à surveiller

Notre source de données principal consiste en un flux de spam alimenté par des pots de miel dédiés à la collecte de spam. Nous recevons tout les jours environ 3.500.000

spam venant d'environ 24.000 blocs d'adresses IP différents. En raison des ressources matérielles requises pour effectuer les mesures traceroute et collecter les données BGP depuis les points de collecte, notre système ne peut actuellement surveiller qu'environ 40.000 blocs d'adresses IP différents par jour. Ainsi, un échantillon de blocs d'adresses IP est extrait de notre flux de spam toutes les heures. Lors de cette sélection, nous prioritisons les blocs récemment annoncés car ils constituent de bons candidats pour des attaques par détournement BGP éphémères [148].

Collecte des données traceroute et BGP

Nous surveillons chaque bloc d'adresses IP sélectionné pendant une durée de 30 jours en collectant de façon journalière des données traceroute et BGP depuis différents points de vue dans l'Internet. Les mesures traceroute sont effectuées depuis trois points de vue et les données BGP sont collectées depuis six collecteurs RouteViews répartis dans le monde entier. De plus, pour chaque adresse IP observée dans un chemin traceroute, nous récupérons l'AS qui l'annonce dans BGP, ses coordonnées géographiques [15] ainsi que des informations sur son propriétaire [36].

En résumé, nous disposons, pour chaque bloc d'adresses IP sélectionné dans notre flux de spam et surveillé pendant une période de 30 jours consécutifs :

- d'un ensemble de chemins traceroute au niveau IP et AS vers le réseau et depuis chacun de nos trois points de vue ;
- d'un ensemble de chemins d'ASs BGP vers le réseau et depuis des six collecteurs RouteViews ;
- d'informations complémentaires sur la localisation et le propriétaire de chaque IP et AS traversé par les chemins traceroute.

6.3.2 Processus d'extraction et d'évaluation des anomalies de routage

Il nous est vite apparu qu'en raison du nombre important de réseaux surveillés, il était nécessaire de disposer d'un mécanisme pour les examiner de façon automatique. C'est pour cette raison que nous avons mis en place un procédé qui analyse les données brutes que nous collectons, en extrait des anomalies de routage, évalue ces dernières sur base d'un ensemble de critères que nous avons défini, et agrège les différents scores pour finalement mettre en avant les blocs d'adresses IP ayant vraisemblablement été victimes d'un détournement BGP. Nous présentons ci-dessous les différents composants de ce procédé d'évaluation des anomalies de routage.

Extraction des anomalies traceroute et BGP

(I) Les anomalies BGP permettent de caractériser le comportement de routage d'un réseau surveillé du point de vue du plan de contrôle. Elles sont extraites des chemins

d'AS BGP collectés chaque jour. (I.a) Une *anomalie de l'origine BGP* correspond à une situation où un bloc d'adresses IP est annoncé dans BGP par plus d'un AS. Ce type d'anomalie est communément appelé « conflit d'ASs d'origine » (« Multiple Origin AS (MOAS) conflict »). (I.b) Une *déviation de chemin d'ASs BGP* mesure la différence entre des chemins d'ASs vers un réseau donné et collectés depuis un collecteur BGP donné.

(II) Les anomalies traceroute permettent d'évaluer l'impact qu'un changement de routage survenant dans le plan de contrôle (BGP) a dans le plan de données. Elles sont extraites des chemins traceroute IP et AS. (II.a) Une *anomalie d'accessibilité au niveau IP (respectivement AS)* survient lorsque l'IP (respectivement l'AS) de destination de traceroute devient (in)accessible, et ce de façon permanente. (II.b) Une *anomalie de longueur de chemin* quantifie un possible changement permanent dans la longueur des chemins traceroute pour un réseau donné. (II.c) Une *déviation de chemin traceroute IP (respectivement AS)* mesure la différente entre les chemins traceroute IP (respectivement AS) collectés pour un réseau donné. (II.d) Une *déviation géographique* quantifie la différence observée entre les pays traversés par les différents chemins traceroute vers un réseau donné.

Chaque anomalie se voit assignée un score dans $[0, 1]$. Une anomalie de l'origine BGP est définie par un triplet (IP, AS_1, AS_2) où IP correspond au bloc d'adresses IP surveillé et, AS_1 et AS_2 sont les ASs annonçant IP. Dans le cas où un bloc d'adresses IP est annoncé par plus de deux ASs, plusieurs anomalies d'origine BGP sont crées. Les déviation de chemins sont calculées grâce à l'indice de Jaccard[1] des ensembles (p_j, p_{j+1}) où p_j est le chemin collecté le jour j et p_{j+1} est le chemin collecté le jour $j + 1$. Enfin, les anomalies d'accessibilité au niveau IP/AS et de longueur de chemin sont calculées une seule fois pour tous les chemins traceroute collectés pour un réseau donné. En résumé, un surveillé pendant une durée de n jours produit (i) zéro ou plusieurs anomalies de l'origine BGP, (ii) $c \times (n-1)$ déviations chemins pour chaque type d'anomalies où c est le nombre de collecteurs/points de vue ($c = 3$ pour les traceroutes et $c = 6$ pour les chemins d'ASs BGP), et (iii) zéro ou une anomalie d'accessibilité IP/AS et de longueur de chemin.

Evaluation des anomalies de routage

Grâce à l'expertise acquise par l'analyse manuelle d'un grand nombre de cas suspects d'attaques par détournement BGP, nous avons développé un procédé novateur d'évaluation des anomalies de routage basé sur l'analyse décisionnelle multicritittères (« Multi-Criteria Decision Analysis (MCDA) »). Cette approche permet d'allier flexibilité et simplicité d'implémentation et de maintenance de notre modèle d'identification des attaques par détournement BGP. En effet, elle constitue une bonne alternative aux arbres de décision du fait qu'elle permet d'assigner à n'importe quel bloc d'adresses IP un *score global de suspicion* sans devoir recourir à des seuils de décisions intermédiaires souvent définis de façon arbitraire et donc propices aux erreurs.

[1] L'indice de Jaccard J de deux ensembles E_1 et E_2 mesures le chevauchement entre les deux ensembles et est défini comme $J = \frac{|E_1 \cap E_2|}{|E_1 \cup E_2|}$.

L'analyse décisionnelle multicritères fournit un vaste ensemble de méthodes pour modéliser de très complexes schémas de décision, allant de fonctions basiques de moyenne à des méthodes plus avancées, telles que des intégrales floues [62]. Dans notre système de prise de décision, nous utilisons principalement l'opérateur Weighted Ordered Weighted Average (WOWA) [172] pour agréger le score des anomalies individuelles. La raison pour laquelle nous avons choisi l'opérateur WOWA est le compromis qu'il offre en la flexibilité et la complexité du modèle de décision construit. En effet, WOWA combine les avantages de deux types de fonctions d'agrégation : the moyenne pondérée (« weighted mean (WM) ») et la moyenne pondérée contrôlée (« ordered weighted average (OWA) »). Cela permet, lors de la prise de décision, de quantifier, avec un seul opérateur, la fiabilité des sources d'informations (comme la WM le fait) mais aussi de pondérer les scores individuels en fonction de leur *ordonnancement* relatif. Le triage et la pondération contrôlée nous permet de mettre en évidence différentes distributions de scores (*e.g.,* éliminer les observations aberrantes, mettre en évidence les valeurs de milieu d'intervalle, garantir que « au moins x » ou « la plupart » des scores sont significativement hauts, etc).

Logiquement, comme toute technique non-supervisée (*i.e.,* dans l'absence de données permettant de vérifier les résultats), elle nécessite la définition d'un nombre de paramètres – habituellement sur base de l'expertise et des connaissances acquises dans le domaine – afin de modéliser avec précision le schémas de décision et de garantir que les cas les plus pertinents termineront dans le haut du classement, alors que les cas réellement bénins se verront assignés un score global très bas. Dans le cas de WOWA, nous n'avons qu'à définir deux vecteurs de poids, ce qui simplifie déjà considérablement la phase de sélection des paramètres. Cela dit, il est important de préciser que l'objectif premier de notre procédé d'évaluation des anomalies de routage est de réduire, autant que possible, le nombre de cas correspondants vraisemblablement à attaques par détournement BGP afin de nous permettre de les examiner et de la (in)valider manuellement. Rappelons que l'objectif ultime est de démontrer que (i) le phénomène de « BGP spectrum agility » existe toujours et que (ii) le mode opératoire des spammeurs utilisant cette technique a changé depuis 2006 [148]. En d'autre termes, nous essayons de comprendre si, oui ou non, ce problème est encore d'actualité en 2014. Dans ce contexte, et sans discréditer l'importance de la sélection des paramètres, nous pensons que la définition des paramètres *optimaux* de notre modèle de décision n'est pas, à ce stade, critique pour l'accomplissement de nos objectifs.

Nous renvoyons le lecteur intéressé à la Section 3.2.2 du Chapitre 3 (en anglais) pour en apprendre davantage sur les aspects mathématiques et les différents paramètres de notre procédé d'évaluation des anomalies de routage basé sur l'analyse décisionnelle multicritères.

6.3.3 Validation des cas suspects de détournements BGP

En raison du manque d'informations permettant de valider (ou d'invalider) catégoriquement les cas suspects identifiés par notre système, il est nécessaire de passer par une étape additionnelle de *validation* qui consiste à collecter, parfois de façon ma-

nuelle, des éléments de preuve complémentaires via des sources de données externes et en interrogeant les propriétaires des réseaux concernés. Ainsi, pour (in)valider des cas suspects de détournements BGP, nous utilisons, outre les données de routage collectées au moyen de notre système, les sources de données suivantes :

- La **table de routage** des collecteurs BGP RIPE RIS [26] et RouteViews [43]. Elle permet de récupérer la liste des blocs d'adresses IP annoncés, à un moment donné, dans BGP, ainsi que les chemins d'ASs BGP (ou routes BGP) associés.
- Les **Registres du Routage dans l'Internet** [20] fournissant des informations sur le propriétaire des adresses IP et numéros d'AS alloués ou assignés. Ils renseignent également des possibles politiques de routage établies entre des réseaux qui sont interconnectés.
- La liste noire **Don't Route Or Peer (DROP)** de Spamhaus [34] renseignant des blocs d'adresses IP prétendument entièrement contrôlés par des cybercriminels, incluant certains ayant apparemment été volés à leur propriétaire.
- Les **forums d'opérateurs réseau**, comme NANOG [22] ou RIPE WG [27], parfois utilisés par les opérateurs réseaux pour rapporter des incidents d'attaque par détournement BGP (*e.g.,* le cas de l'attaque contre Link Telecom [46]).

Nous examinons, grâce aux tables de routage archivées par RIPE et RouteViews, l'historique du routage des blocs d'adresses IP identifiés comme suspects afin de déterminer (i) quand ils ont été annoncé publiquement dans l'Internet, (ii) les ASs d'origine utilisés pour les annoncer, et (iii) les ASs des différents FAIs observés en amont dans les chemins d'ASs.

Nous utilisons également des copies des Registres du Routage dans l'Internet (« Internet Routing Registries, IRRs ») afin de récupérer des informations sur le propriétaire de blocs d'adresses IP et de numéros d'AS impliqués dans des cas suspects de d'attaques par détournement BGP. Ces informations nous aident à déterminer la légitimité des annonces BGP observées. Comme Siganos et al. le suggèrent dans [159], il est également possible d'en apprendre davantage sur d'éventuelles politiques de routage entre des ASs lorsque les opérateurs réseau de ces derniers ont pris la peine d'en déclarer.

Nous vérifions également auprès de Spamhaus [34] la réputation des blocs d'adresses IP identifiés par notre système comme suspects.

Enfin, nous parcourons les archives des forums publiques d'opérateurs réseau, comme par exemple les forum NANOG [22] (« North American Network Operators' Group ») et RIPE Working Groups [27], afin de vérifier si certains des cas suspects que nous avons identifiés ont été discutés.

Nous devrions avoir, à la fin de cette étape, un ensemble de cas d'attaques par détournement BGP malveillant nous permettant de confirmer, ou non, l'existence de ce phénomène.

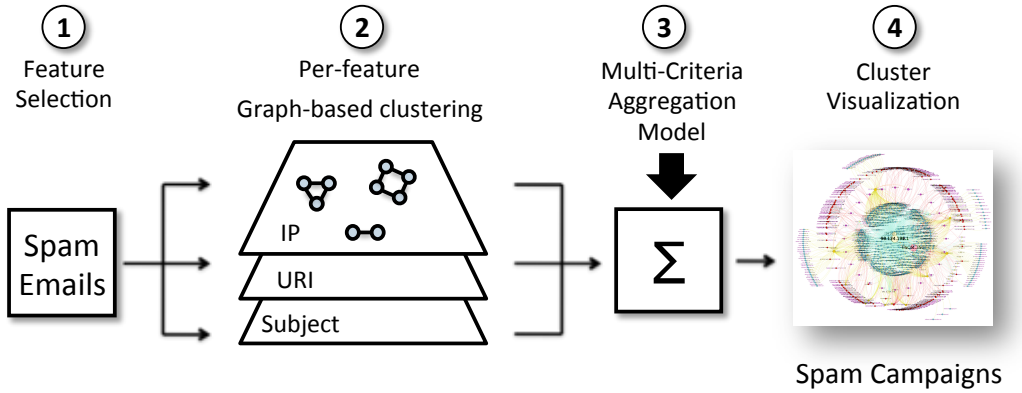


FIGURE 6.2 – Regroupement, grâce à TRIAGE, d'emails de spam envoyés depuis des réseaux ayant été détournés.

6.3.4 Analyse de l'origine des détournements BGP

Bien que l'étape de validation des cas suspects de détournement BGP décrites ci-dessus devrait nous permettre de déterminer avec plus de certitude s'il existe dans l'Internet des spammeurs « agiles » utilisant cette technique, nous souhaitions pouvoir confirmer ces résultats en analysant les cas de détournement BGP confirmés du point de vue des campagnes de spam qu'elles ont aidés à réaliser.

Ainsi, nous avons utilisé une technique de regroupement multicritères appelée TRIAGE [168] afin d'identifier des séries d'emails de spam envoyés depuis les différents blocs d'adresses IP ayant été détournés et qui feraient parties de campagnes orchestrées par les mêmes spammeurs « agiles ». TRIAGE peut être décrit comme un outil de sécurité aidant dans l'investigation d'attaques en permettant de reconstruire le mode opératoire des attaquants. Cette technologie a déjà démontré son utilité dans différentes applications [68, 170, 169].

La Figure 6.2 illustres le processus d'analyse de TRIAGE. A l'étape ①, nous définissons un certain nombres de caractéristiques décrivant des emails de spam nous permettant de les comparer entre eux. Ces caractéristiques incluent, par exemple, l'adresse IP de l'émetteur, la date d'envoi, les éventuelles URLs incluses dans le message. A l'étape ②, TRIAGE évalue les relations entre tous les emails analysés par rapport à chaque caractéristique définie séparément.

A l'étape ③, les valeurs de similarité pour chaque caractéristique sont fusionnées en utilisant un modèle d'agrégation défini par l'analyste, qui peut imposer, par exemple, que des emails groupés au sein d'une même campagne partagent un certains nombres de caractéristiques (sur les n disponibles) très similaires (peu importe lesquelles). De la même façon qu'avec notre méthode d'agrégation décrite ci-dessus utilisant WOWA, TRIAGE nous permet d'assigner des *poids* aux différentes caractéristiques utilisées et ce afin de donner plus ou moins d'importance à certaines d'entre elles.

Le résultat fourni par TRIAGE (étape ④) consiste en un ensemble de *groupes d'emails de spam* (appelés MDC's), chacun d'entre eux contenant des emails par-

tageant un certain nombre de traits communs, mais pas nécessairement les mêmes. Enfin, comme expliqué dans [168], il est nécessaire de définir un seuil de décision afin d'éliminer les cas résultants

As outcome (step ④), TRIAGE identifies *multi-dimensional clusters* (called MDC's), which in this analysis are clusters of spam emails in which any pair of emails is linked by a number of common traits, yet not necessarily always the same.

6.4 Résultats

Dans cette Section nous traitons les deuxième et troisième problèmes de recherche exposés dans l'introduction à cette thèse. En particulier nous décrivons les résultats obtenus en utilisant, pendant presque deux ans, notre environnement expérimental SPAMTRACER décrit ci-dessus. Nous complétons la description des résultats obtenus pour chaque composant de notre système par une analyse détaillée de cas confirmés d'attaques par détournement BGP malveillant que nous avons identifiés. Enfin, nous explorons l'usage de certaines caractéristiques des attaques observées pour construire un système de détection en temps réel d'attaques par détournements BGP afin d'aider à contrer les effets néfastes qu'elles produisent sur la sécurité de l'Internet.

6.4.1 Résultats : la collecte des données de routage

Nous considérons un ensemble de données BGP et traceroute collectées entre septembre 2012 et juin 2014 (22 mois). Durant ces 22 mois nous avons surveillé un total de 649.081 blocs d'adresses IP distincts depuis lesquels du spam ou diverses autres activités malveillantes ont été observées. Ces différents réseaux étaient opérés depuis 18.907 ASs différents. Enfin, nous avons collecté pas moins de 8,5 millions de mesures traceroute et environ 28,6 millions de routes BGP vers ces réseaux.

6.4.2 Résultats : processus d'extraction et d'évaluation des anomalies de routage

La figure 6.3 montre la distribution des scores obtenus pour chaque réseau surveillé grâce à la procédure décrite dans la Section 6.3.2. La première partie de courbe, entre les valeurs 0 et 0,25, correspond aux 31.29% des réseaux n'affichant presqu'aucune variation des les routes BGP et les traceroute. Pour cette raison, nous considérons qu'il s'agit très probablement de cas bénins. Ensuite, 68,642% des réseaux obtiennent un score compris entre 0,25 et 0,75. Ces réseaux affichent habituellement un ensemble varié d'anomalies de routage, ce qui les rend difficile à attribuer à un comportement bénin ou malveillant. Ces cas peuvent souffrir des limitations de notre modèle d'agrégation ou bien encore des inexactitudes observées dans les données collectées [128]. Enfin, 0,068% des blocs d'adresses IP surveillés obtiennent un score supérieur à 0,75 et correspondent aux cas les plus suspects d'attaques par détournement BGP.

Nous avons 437 blocs d'adresses IP différents qui ont obtenu un score supérieur à 0,75. Bien que ce nombre de cas suspects puissent paraître faible, nous n'avons,

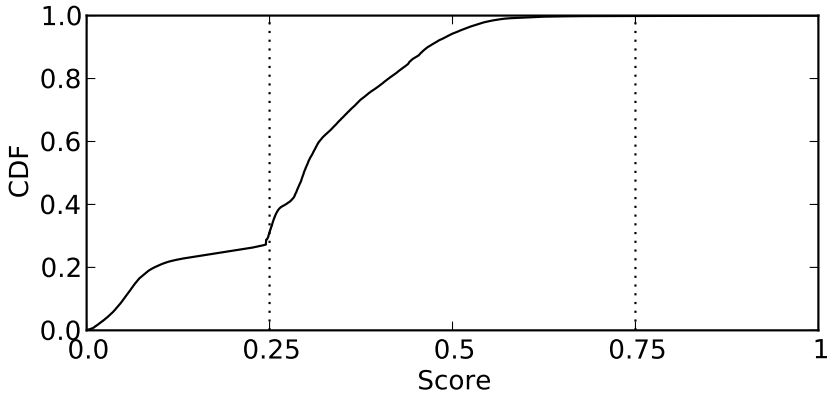


FIGURE 6.3 – Identification des attaques par détournement BGP : scores obtenus par les réseaux surveillés entre septembre 2012 et juin 2014.

en théorie, besoin que d'un seul cas validé de détournement BGP malveillant pour confirmer l'existence, à l'heure actuelle, du phénomène de « BGP spectrum agility ». Comme nous le montrerons plus tard, il n'est pas nécessaire de pouvoir identifier tous les cas de détournements BGP présents dans nos données pour pouvoir confirmer ce phénomène. Nous avons examiner manuellement un grand nombre de cas appartenant aux différentes catégories (scores faibles, moyens et élevés) afin d'être confiant dans la capacité de notre système à évaluer et classer ces cas en accord avec le comportement de routage qu'ils affichent.

6.4.3 Résultats : validation des cas suspects de détournements BGP

Nous avons utilisé la méthode présentée dans la Section 6.3.3 afin de valider (ou invalider) les cas suspects de détournements BGP identifiés. En raison du temps requis pour analyser manuellement l'entièreté les réseaux surveillés, nous nous somme concentrés sur les 437 blocs d'adresses IP les plus suspects et ayant obtenus un score supérieur à 0,75 dans notre système d'extraction et d'évaluation des anomalies de routage, *i.e.,* le quartile supérieur dans notre distribution de scores.

L'analyse des 437 cas suspects a révélé que 60 d'entre eux affichaient un comportement correspondant au phénomène de « BGP spectrum agility » que nous recherchions, et ce pour les différentes raisons que nous explicitons ci-dessous. Ainsi, en examinant l'historique de routage de ces 60 blocs, nous avons découverts qu'ils pouvaient être classés en deux catégories :

- DÉTOURNEMENT DE PRÉFIXE IP VIA UN FAI VALIDE : Dans 90% des cas, les blocs d'adresses IP étaient alloués mais ***(1) non annoncés*** au moment où ils ont été détournés (*i.e.,* ils avaient été laissés inactifs par leur propriétaire), et l'attaquant a forgé une partie du chemin d'ASs BGP afin d'annoncer les blocs depuis un ***(2) AS d'origine invalide*** mais via un ***(3) un FAI valide***.
- DÉTOURNEMENT D'AS VIA UN FAI INVALIDE : Dans 10% des cas, les blocs d'adresses IP étaient allouées mais ***(1) non annoncés*** et l'attaquant a forgé

une partie du chemin d'ASs BGP afin d'annoncer les blocs depuis un ***(4) AS d'origine valide*** mais via un ***(5) un FAI invalide***.

(1) Bloc d'adresses IP non annoncés : L'historique de routage révèle que tous les blocs d'adresses IP détournés n'étaient pas annoncés dans BGP avant qu'ils ne soient détournés.
(2)-(4) AS d'origine (in)valide : Dans ce travail, nous considérons l'AS d'origine pour un bloc d'adresses IP comme *valide* s'il existe une association entre cet AS et ce bloc dans les bases de données `whois` (IRRs).
(3)-(5) FAI (in)valide : Dans ce travail, nous considérons le FAI représenté par son AS a_1 apparaissant dans le chemin d'ASs BGP $\{a_n, ..., a_1, a_0\}$ (a_0 étant l'AS d'origine) comme *invalide* si l'ensemble des conditions suivantes sont satisfaites : (1) il n'est jamais apparu comme FAI pour a_0 auparavant, (2) il n'apparaît pas dans la liste des FAIs de a_0 et a_0 n'apparaît pas dans la liste de ses clients tel que renseigné dans les bases de données `whois`, (3) il n'est pas utilisé pour annoncer des blocs d'adresses IP non détournés au moment où les détournements se produisent, (4) il est inactif lorsqu'il est utilisé pour la première fois dans des détournements, (5) l'organisme officiellement propriétaire de l'AS apparaît être inactive, et (6) il a été, à un certain moment, marqué comme FAI suspect par Spamhaus [34].

La Figure 6.4 représente la distribution des 2.713 détournements observés au cours du temps.Nous pouvons voir que 96,8% des attaques se produisent après juillet 2013. Avec une moyenne de 4,06 détournements par jour, on peut noter que les attaques par détournement BGP ont été une menace continue et récurante pendant les 22 derniers mois (et potentiellement avant).

Figure 6.4 shows the distribution of the 2,713 observed hijacks across time. We can see that 96.8% of the observed hijacks have occurred after July 2013. From that point the distribution becomes almost uniform, showing that hijacks were performed on a regular basis for more than one year. With an average of 4.06 hijacks per day, we note that BGP hijacks have been an ongoing and recurring threat in the past 22 months (and possibly before).

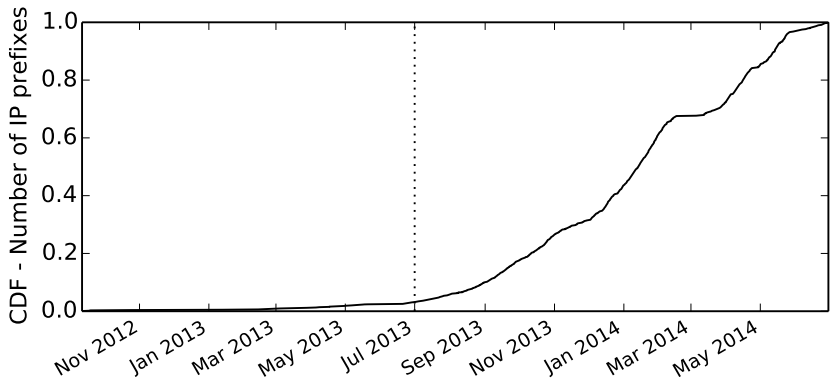


FIGURE 6.4 – Nombre the blocs d'adresses IP détournés identifiés entre septembre 2012 et juin 2014. La majorité d'entre eux est observée après juillet 2013.

Nous nous concentrons à présent sur une caractéristique des attaques identifiées : leur durée. Dans notre cas, 75,6% des détournements ont duré **moins** d'un jour et 98,1% n'ont pas duré plus d'une semaine. Une large fraction des cas que nous observés apparaissent donc être similaires, en terme de durée, aux cas rapportés dans [148].

Bien que les attaques par détournement de courte durée partagent quelques caractéristiques avec ceux de plus longue durée, nous les considérons comme appartenant à deux phénomènes différents. En fait, les détournements de courte durée peuvent être utilisés par les attaquants pour empêcher tout traçage d'attaques réalisées depuis les réseaux volés ainsi que pour éviter d'être bloqués en changeant régulièrement d'adresses IP dans un bloc jusqu'à que le bloc soit inscrit sur une liste noire pour ensuite passer à un autre bloc. D'autre part il est plus difficile, pour les attaquants réalisant des détournements de longue durée, d'éviter d'être détectés.

Dans le reste de cette Section nous analysons les détournements de courte ($\leq$ 1 semaine) et longue durée ($>$ 1 semaine) séparément afin de mettre en évidence leurs similitudes et leurs différences. Nous considérons les caractéristiques suivantes pour un bloc d'adresses IP détourné :

- (*C.1*) Si des **emails de spam** ont été reçus depuis ce bloc et/ou si des adresses IP de ce bloc ont été inscrites sur les listes noires de spammeurs comme Spamhaus SBL ou DROP (Don't Route Or Peer) [34], Uceprotect [42] ou encore Manitu [21].
- (*C.2*) La **durée de la période pendant laquelle le bloc est resté inactif avant d'être détourné**, ce qui correspond au temps entre la dernière fois où le bloc a été annoncé et le moment où il a été détourné.
- (*C.3*) La **taille** du bloc d'adresses IP, qui définit le nombre d'adresses IP disponibles dans le bloc.
- (*C.4*) Si le **propriétaire** du bloc est toujours en activité.

Détournements de longue durée

A présent nous analysons plus en détails les 55 détournements de longue durée identifiés (sur un total de 2.713 détournements) par rapport aux *cinq* caractéristiques décrites ci-dessus.

(C.1) Sept des 55 blocs d'adresses IP ont envoyés du spam. Un total de 815 emails de spam ont été envoyés depuis des adresses IP dispersées dans chacun des blocs. Nous avons principalement observé du spam au début de l'attaque par détournement. Aussi, aucune adresse IP ayant envoyé du spam n'était inscrite sur une liste noire de spammeurs au moment où elle a été observée pour la première fois. Enfin, sept des 55 blocs d'adresses IP sont apparus dans une liste noire, mais cela s'est produit plusieurs jours après le début du détournement voir après la fin de celui-ci.

(C.2) 43 blocs d'adresses IP sur 55 n'avaient jamais été annoncés publiquement dans l'Internet avant d'être détournés. Les 12 autres sont restés inactifs pendant une moyenne d'an avant d'être détournés.

(C.3) Dans [148], Ramachandran et al. déclarent avoir observé du spam venant de larges blocs d'adresses IP détournés (*i.e.,* /8). Dans les 55 cas de détournements de longue durée que nous avons observés, les blocs étaient plus petits, *i.e.,* le plus grand était un /19 et le plus petit un /24.

(C.4) L'analyse des base de données `whois` (IRRs) pour les 55 blocs concernés a révélé que le propriétaire de la plupart d'entre eux n'était plus en activité. Cette observation indique que les attaquants pourraient cibler spécifiquement les blocs d'adresses IP dont le propriétaire n'existe plus, ce qui est le cas par exemple lorsqu'une entreprise disparaît, est acquise par, ou bien encore fusionne avec une autre. Dans certaines de ces situations, il arrive que des blocs d'adresses IP soient laissés inactifs.

Détournements de courte durée

Nous nous concentrons maintenant sur les 2.658 détournements de courte durée observés (sur un total de 2.713 détournements). Nous distinguons deux épisodes dans ces attaques : (1) du spam et des adresses IP inscrites sur liste noire ont été observés entre février 2013 et mai 2013, et (2) un phénomène apparemment différent a été observé entre juin 2013 et juin 2014, révélant un schéma temporel marquant dans les annonces BGP. Nous commençons par présenter les deux épisodes et leurs différences par rapport à la caractéristique C.1. Ensuite, nous décrivons leurs ressemblances par rapport aux caractéristiques C.2-4.

Episode 1 : Entre février 2013 et mai 2013

(C.1) Des 2.658 blocs impliqués dans des détournements de courte durée que nous avons observés, 57 d'entre eux ont envoyé du spam entre février et mai 2013. La Figure 6.5 représente les annonces BGP, la réception de spam et la présence d'adresses IP inscrites sur liste noire pour ces blocs. Afin d'améliorer la lisibilité de la figure, nous n'avons représenté qu'un échantillon de 25 de ces blocs. La figure met en évidence :

- la forte corrélation **temporelle** entre les annonces BGP et le spam, et
- le **faible** nombre de blocs d'adresses IP (7 sur 57) inscrits sur une liste noire avant la fin du détournement.

Il est important de noter que les 32 blocs d'adresses IP non représentés sur la Figure 6.5 dépeignent exactement le même schéma temporel par rapport aux annonces BGP, la réception de spam et la présence d'adresses IP inscrites sur liste noire. Enfin, un total de 4.149 emails de spam ont été reçus depuis ces blocs d'adresses IP.

Episode 2 : Entre juin 2013 et juin 2014

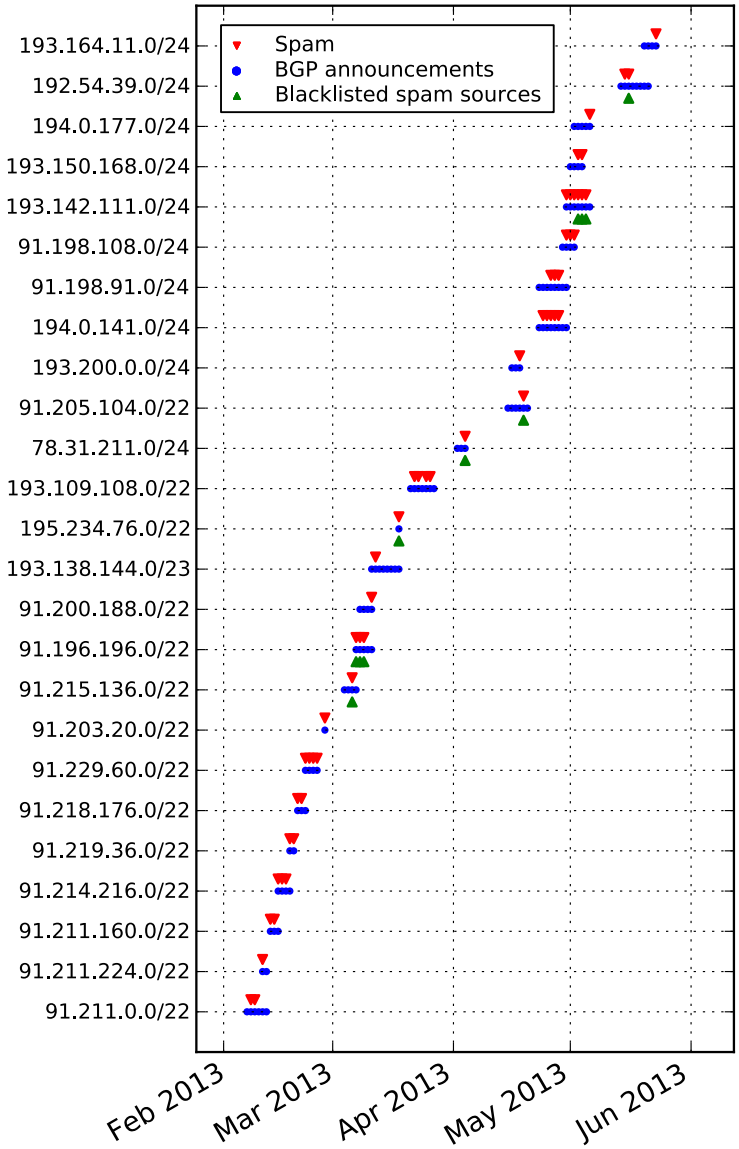


FIGURE 6.5 – Episode 1 des détournements de **courte durée** survenu entre février et mai 2013 : corrélation temporelle entre les annonces BGP, la réception de spam et la présence d'adresses IP inscrites sur liste noire pour les blocs d'adresses IP concernés. Afin d'améliorer la lisibilité de la figure, seuls 25 de ces blocs sont représentés. Les 32 autres dépeignent exactement le même phénomène.

Les détournements observés durant ce second épisode, entre juin 2013 et juin 2014, est encore plus intriguant. Ce phénomène est également important puisqu'il comprend pas moins de 2.601 attaques par détournement de courte durée, ce qui correspond à 98% des cas identifiés. La Figure 6.6 montre les annonces BGP pour les blocs concernés (pour améliorer la lisibilité de la figure, seuls 87 de ces blocs correspondant aux attaques observées pendant le mois de juin 2014 sont représentés)[(2)]. Elle permet de voir que :

- tous les détournements sont réalisés par groupe de deux à quatre blocs, tous commençant et se terminant au même moment ;
- durant la période d'un mois, il y a toujours, à tout moment, au moins deux blocs d'adresses IP détournés.

Bien que seul une partie du phénomène soit dépeint dans la Figure 6.6, celui-ci est récurrent et persistants durant la totalité des 13 mois, entre juin 2013 et juin 2014. Cela suggère fortement que ces attaques pourraient être réalisées suivant le même mode opératoire. Le fait chaque groupe d'attaques commence quelques secondes après la fin du précédant suggère qu'elles pourraient être réalisées de *façon automatique*, s'appuyant potentiellement sur un procédé automatique de sélection de bloc d'adresses IP à détourner.

(C.1) Curieusement, nous d'avons pu identifier aucun trafic de nature malveillante associé à ces blocs d'adresses IP. L'absence de traces de spam et autres menaces connexes dans nos données peut être attribuée à la visibilité limitée que nous pourrions avoir via nos sondes capturant les activités malveillantes dans l'Internet. Afin d'analyser plus en profondeur la nature des activités réalisées par les attaquants depuis les réseaux détournés, nous avons exploité différentes sources de données externes, comme des traces NetFlow collectées depuis le réseau scientifique de l'université de Munich [177], des archives de résolutions DNS collectées via l'infrastructure de Norton DNS de Symantec, ainsi que des données télémétriques issus de technologies anti-virus de Symantec [71]. Grâce à cette analyse nous avons identifié 557 noms de domaine supplémentaires résolvant vers des adresses IP dans 15 blocs (sur les 2.601 impliqués dans l'épisode 2). L'absence de trafic à caractère malveillant pour la majorité des réseaux impliqués dans l'épisode 2 pourrait également qu'il s'agit d'une infrastructure mouvante d'hébergement de serveurs, *e.g.,* de serveurs C&C. Nous n'avons, cependant, aucune preuve probante nous permettant de valider cette conjecture.

Caractéristiques communes aux épisodes 1 et 2

Nous analysons à présent les caractéristiques communes aux 2.658 attaques par détournement BGP de courte durée que nous avons identifiées.

(C.2) 2,291 blocs d'adresses IP (86.2%) n'avaient jamais été annoncés publiquement avant d'être détournés. Sur base d'une discussion avec un cadre du RIPE

[2]La figure montrant le phénomène de l'épisode 2 au complet est disponible à l'adresse suivante http://www.eurecom.fr/~vervier/public/bgphijacks_episode2_June2013_June2014.pdf.

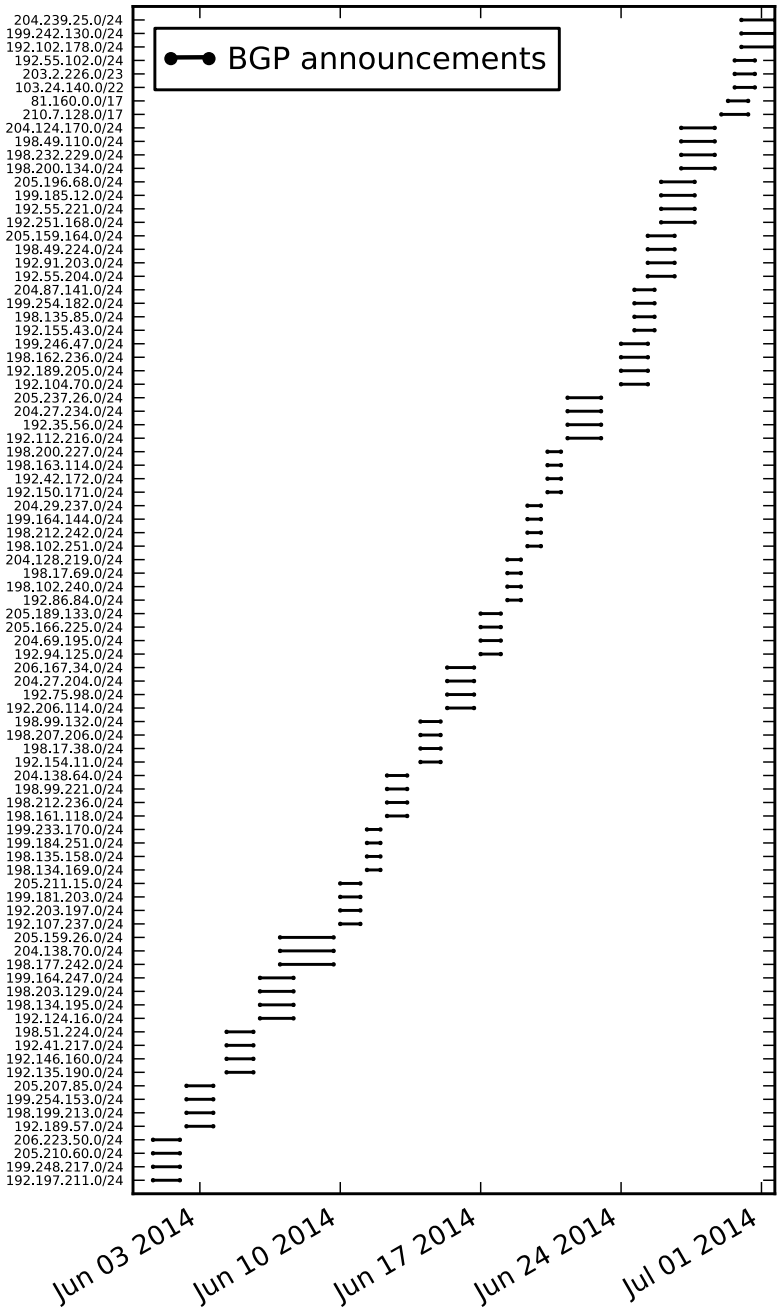


FIGURE 6.6 – Episode 2 des détournements de **courte durée** survenu entre juin 2013 et juin 2014 : les attaques sont toujours réalisées par groupe d'au moins deux blocs d'adresses IP. hijacks are always performed by groups of at least two IP prefixes. Afin d'améliorer la lisibilité de la figure, seuls 25 de ces blocs (sur 2.601) sont représentés.

NCC [69] ainsi que des messages échangés sur le forum NANOG [52], il serait pratique courante chez les opérateurs réseaux d'enregistrer des adresses IP publiques pour un usage strictement interne. Cela pourrait ainsi expliquer pourquoi nous n'observons aucune route vers ce type de bloc dans nos données BGP. Avec 72,4% des blocs d'adresses IP observés comme ayant été détournés laissés inactifs pendant plus d'un an, nous pouvons conclure que les attaquants cible principalement les blocs laissés inactifs depuis une longue période.

(C.3) Les blocs observés dans des détournements de courte durée incluent, de façon similaire aux détournement de longue durée, des /17, /21, /22, /23 et (pour 92.6%) des /24. Bien que ces attaques ressemblent, sur certains points, aux cas rapportés par Ramachandran et al. dans [148], la taille moyenne des blocs impliqués est très différente, à savoir /24, au lieu de /8.

(C.4) L'analyse des informations issues des bases de données `whois` (IRRs) concernant les blocs impliqués dans des détournements de courte durée a révélé que tous ces blocs étaient, au moment où ils ont été détournés, alloués à un organisme ou une entreprise. Bien que nous n'avons pu vérifier l'entièreté des 2.658 blocs, nous en avons analysé 100 et avons pu déterminer que 41% d'entre eux appartenaient à un propriétaire qui n'était plus en activité. Curieusement, 59% d'entre eux appartenaient à un propriétaire apparaissant comme étant toujours en activité.

6.4.4 Résultats : analyse de l'origine des détournements BGP

Dans la première partie de cette section nous avons exposé des preuves convaincantes que le phénomène de « BGP spectrum agility » existe actuellement dans l'Internet. Cependant, nous n'avons systématiquement déterminé si les attaques identifiées sont des cas isolés ou si certaines d'entre elles partagent une origine commune, ainsi qu'on pourrait s'y attendre si elles sont orchestrées par les mêmes spammeurs. C'est pourquoi nous avons appliqué TRIAGE sur l'ensemble des 4.964 emails de spam envoyés depuis 64 réseaux identifiés comme ayant été détournés.

TRIAGE n'a identifié que 30 groupes (ou multi-dimensional clusters, MDCs) dans chacun desquels les emails partagent des combinaisons variées de caractéristiques. En raison de la façon dont les groupes sont générés, nous prévoyons qu'ils représentent vraisemblablement différentes campagnes de spam organisées par les mêmes individus - puisque les emails de spam au sein d'un même groupe partagent plusieurs caractéristiques en commun. Ainsi 64 blocs d'adresses IP ont été utilisé pour réaliser 30 campagnes de spam différentes.

En regroupant les emails de spam en campagnes, nous obtenus de nouveaux éclairage sur le comportement des spammeurs. Sur base de la structure des MDCs, nous mettons en lumière trois modes opératoires clés des spammeurs réalisant des détournements BGP : (1) 10 campagnes (sur 30) impliquent un seul bloc d'adresses IP non impliqué dans aucune autre campagne, (2) 17 campagnes impliquent un seul bloc d'adresses IP impliqué simultanément dans différentes campagnes de spam, et (3) trois campagnes ont impliqué plusieurs blocs d'adresses IP séquentiellement pendant une longue période de temps. Alors que les deux premiers phénomènes confirment

notre intuition sur le comportement de ce type de spammeurs, le troisième phénomène est le plus intéressant car il confirme l'existence du phénomène de « BGP spectrum agility » sous la forme de campagnes de spam orchestrées par les mêmes spammeurs, allant d'un bloc volé à un autre pour envoyer du spam. Il met en évidence l'existence chez les spammeurs d'un mode opératoire plus agile et sophistiqué. Cette agilité leur permet d'envoyer du spam de façon plus furtive et ainsi de rester « sous le radar », mais aussi d'empêcher le traçage de leurs attaques et de rendre les listes noires inefficaces.

La Figure 6.7 présente une visualisation d'une des campagnes à grande échelle impliquant plusieurs blocs d'adresses IP détournés. Elle illustre le mode opératoire typique des spammeurs agiles opérant ce type de campagnes furtives.

A notre connaissance, ces résultats sont complètement novateurs et apportent un nouvel éclairage sur le comportement des spammeurs agiles. La leçon à retenir de cette analyse de l'origine des détournements BGP est que cela nous permet de relier entre eux différents blocs d'adresses IP en montrant qu'ils sont utilisés par les mêmes spammeurs, et ce, pendant une longue période de temps et de façon très furtive.

6.4.5 Efficacité des contres-mesures

Différentes technologies et systèmes ont été mis au point pour détecter et prévenir les attaques par détournement BGP. Nous allons à présent évaluer l'efficacité de deux de ces contre-mesures : un système avancé de détection de détournement BGP nommé Argus [157] et le système de sécurisation de BGP appelé RPKI [101, 119, 120].

Détection des détournements BGP

Au fil du temps de nombreux systèmes et services, comme par exemple BGPmon.net [7], Renesys [25], PHAS [116] ou encore Argus [157], ont été développé afin de détecter des attaques par détournement BGP. Un de ceux-ci, Argus [157], a pour objectif la détection en temps réel de ce type d'attaques. Pour ce faire ce système utilise une combinaison de données BGP et de mesures ping afin d'identifier des changements dans l'accessibilité d'un réseau consécutifs à un changement dans le routage de celui-ci, et pouvant indiquer que le réseau a été détourné. Afin d'évaluer l'impact sur la sécurité de l'Internet que représentent les attaques que nous identifiées, nous avons décidé de vérifier l'efficacité du système Argus contre ces attaques. Nous avons choisi Argus pour deux raisons : (i) le système est actuellement déployé et permet l'accès publiquement à l'historique de ses alertes, et (ii) il est supposé être capable de détecter n'importe quel type d'attaque par détournement BGP, c'est-à-dire aussi bien celles où l'attaquant utilise un AS d'origine invalide (détournement de préfixe IP via un FAI valide) que celles où l'attaquant utilise un FAI invalide (détournement d'AS via une FAI invalide).

Il s'avère qu'aucun des 2.713 attaques par détournement BGP que nous avons identifiées n'ont été détectées par Argus. La raison à cela est que la plupart des systèmes de détection [96, 116, 157], notamment Argus, fonctionne en construisant

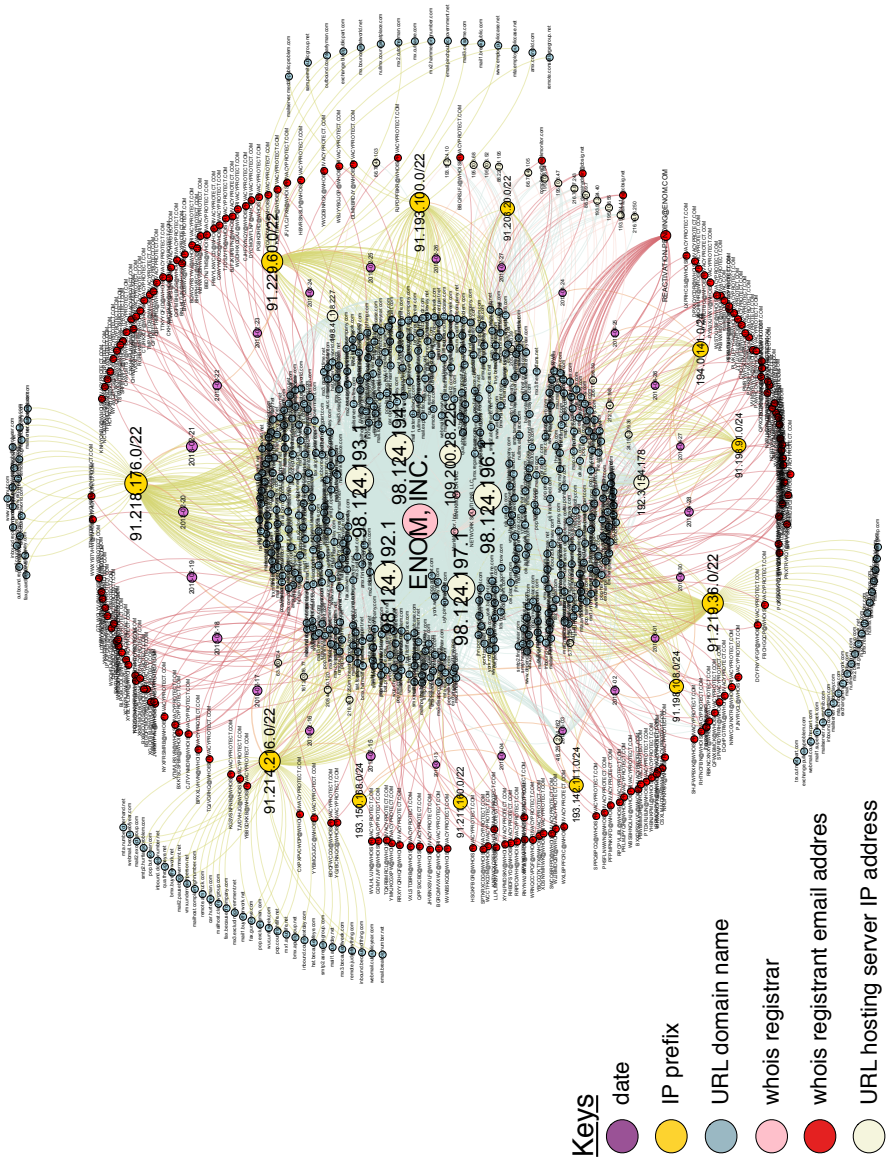


FIGURE 6.7 – Un exemple de campagne à grande échelle impliquant plusieurs blocs d'adresses IP. Les noeuds disposés dans le sens des aiguilles d'une montre reflètent la chronologie de la campagne.

un modèle de la topologie au niveau AS de l'Internet, qu'il utilise ensuite pour valider des changements de routage. Cependant, dans le cas d'un bloc d'adresses IP non annoncé avant d'être détourné, le système ne possède aucun état dans son modèle pour ce réseau. Il ne peut donc évaluer la légitimité d'un changement de routage pour ce réseau. Bien que les systèmes de détection sont très utiles pour les opérateurs réseau afin qu'il puisse surveiller l'état du routage pour leurs réseaux, leur incapacité à détecter le type d'attaques que nous avons observées suggère qu'il devraient, dans le futur, intégrer quelques unes des caractéristiques de ces attaques dans leurs signatures.

Peu de temps après qu'un opérateur réseau se soit plain sur le forum NANOG qu'un de ses blocs d'adresses IP ait été détourné [51], deux rapports différents décrivant des « opérations de squat d'adresses IP pour envoyer du spam » ont été publiés sur les blogs de BGPmon.net [171] et de Renesys [125]. Ces rapports ont corroboré nos résultats à propos de *cinq* (des 10) ASes invalides que nous avons identifiés.

Prévention des détournements BGP

Outre les techniques de détection de détournements BGP, les opérateurs réseau ont commencé à adopter et à déployer un système de sécurisation de BGP, communément appelé RPKI. Bien que de nombreuses approches différentes ont été proposées pour sécuriser BGP [102], ce système a particulièrement retenu l'attention de la communauté réseau ces dernières années. De plus, nous n'avons connaissance d'aucun autre système qui soit aussi prêt et mature que celui-ci.

Le système repose sur une infrastructure à clé publique de ressources (ou « Resource Public Key Infrastructure, RPKI ») standardisée dans le RFC 6480 [120] pour prévenir l'injection de routes BGP fallacieuses. La RPKI utilisée consiste en une base de données de certificats de quatre types : (i) un type A appelé « Route Origin Authorisation (ROA) » qui lie un bloc d'adresses IP aux ASs d'origine autorisés à l'annoncer dans BGP, (ii) un type B qui lie un routeur BGP à l'AS auquel il appartient, et (iii-iv) des certificats C et D qui lient, respectivement, des adresses IP et et des numéros d'AS à leur propriétaire. La chaîne de certification suit la chaîne de délégation des numéros d'AS et des adresses IP, avec l'IANA agissant comme l'autorité de certification pour les certificats des RIRs, chaque RIR agissant à leur tour comme l'autorité de certification pour les certificats des FAIs auxquels ils délèguent des adresses IP ou des numéros d'AS, etc. Chaque certificat est signé avec clé privée de son propriétaire et embarque également la clé publique de ce dernier. Le système propose ainsi deux techniques différentes pour sécuriser BGP : (i) la sécurisation de l'*origine des routes* et (ii) la sécurisation de la *propagation des routes* (ou BGPsec). (i) La sécurisation de l'origine des routes, standardisée dans le RFC 6483 [101], utilise des ROAs (certificats de type A) pour vérifier qu'un bloc d'adresses IP est annoncé par les ASs autorisés. Un routeur est alors capable de vérifier la validité d'un message BGP pour un bloc d'adresses IP et un AS d'origine donnés (i-a) en interrogeant la RPKI afin de récupérer un ROA associé au bloc et en vérifiant sa validité du point de vue cryptographique, et, (i-b) si le ROA est valide, en vérifiant que l'AS d'origine et la longueur du préfixe IP dans le message BGP correspondent

aux données renseignées dans le ROA. Cela permet d'empêcher un attaquant d'annoncer un bloc d'adresses IP dont il n'est pas le propriétaire. (ii) La sécurisation de la propagation des routes [119] a quant à elle pour but de prévenir la manipulation du chemin d'ASs par un attaquant en garantissant que l'identité d'aucun AS sur le chemin n'a été usurpée. Pour cela, chaque routeur recevant un message BGP se doit de le signer, au moyen d'un certificat de type B, avant de l'envoyer à ces routeurs voisins afin que ceux-ci puissent s'assurer que tous les routeurs par lesquels le message a transité appartiennent bien aux différents ASs renseignés dans le chemin d'ASs.

La sécurisation de l'origine des routes est actuellement en cours de déploiement par les FAIs et autres gestionnaires de réseaux à travers le monde. Selon le RIPE NCC [152], il y a actuellement 4,1% de l'espace IPv4 qui est sécurisé au moyen de ROAs. Curieusement, aucun des blocs que nous avons identifiés comme ayant été détournés n'était couvert par un ROA au moment de l'attaque. Dans 90% des attaques identifiées, l'attaquant a annoncé des blocs d'adresses IP en utilisant un AS d'origine invalide (*détournement de préfixe IP via un FAI valide*). En supposant qu'un ROA ait existé pour ces blocs avec leur AS d'origine valide, le système RPKI aurait été capable d'invalider les annonces BGP erronées.

Toutefois, en supposant que des ROAs lient l'entièreté des blocs d'adresses IP avec leurs ASs d'origine valides, il est toujours possible pour un attaquant de détourner ces blocs. Pour ce faire, il lui suffit de manipuler le chemin d'ASs en ajoutant à la fin de celui-ci l'AS d'origine valide. De cette façon, les annonces BGP fallacieuses passeraient les vérifications de l'origine des routes (via les ROAs) avec succès. C'est exactement la situation que nous avons observée dans 10% des détournements identifiés (*détournement d'AS via un FAI invalide*). Seule la sécurisation de la propagation des routes (BGPsec) peut empêcher ce type d'attaques. Cependant, BGPsec est toujours au début de son processus de développement et son déploiement n'a pas encore débuté.

Globalement, la seule solution définitive au problème des détournements BGP est le déploiement de la sécurisation de la propagation des routes, *i.e.,* BGPsec. Malheureusement, cette solution est beaucoup invasive (que par exemple la sécurisation de l'origine des routes via les ROAs) et ne peut être déployée sans changer de façon substantielle le logiciel et le matériel équipant actuellement les routeurs. De plus, la standardisation de BGPsec est n'est pas encore terminée et il n'existe actuellement pas encore d'implémentation disponible pour les routeurs. Certains fabricants de routeurs y travaillent, mais pour certains autres l'implémentation de BGPsec ne figure même pas sur leur feuille de route.

6.4.6 Opérationnalisation de SpamTracer

Nous avons, jusqu'à présent, présenté le résultat de l'analyse de 22 mois de données collectées grâce à SpamTracer. Cela nous a permis d'exposer un phénomène soutenu de « BGP spectrum agility » dans lequel des campagnes furtives et persistantes d'attaques par détournement BGP ont lieu régulièrement dans l'Internet. Nous avons également mis en lumière plusieurs différences entre les attaques que nous avons observées et celles rapportées dans [148], comme par exemple la durée

variable des détournements (de quelques minutes à plusieurs mois) ou encore la taille des blocs d'adresses IP détournés (principalement des préfixes IP /24), suggérant que le phénomène a évolué au fil des ans. De plus, nous avons vu que les techniques actuelles de détection et de prévention des attaques par détournement (*e.g.,* Argus, RPKI+ROA) ne sont pas très efficaces contre le type d'attaques que nous avons observées. Dans la suite de cette section, nous proposons de tirer parti des quelques caractéristiques clés des attaques par détournement BGP identifiées afin de détecter de futures instances de ces attaques. L'objectif poursuit ici est double : (i) continuer à surveiller le phénomène de « BGP spectrum agility » et (ii) fournir un moyen concret de se défendre contre ces attaques via la mise en place d'une liste (noire) de blocs d'adresses IP détournés.

Détection en temps réel de détournements BGP

Notre système utilise, en entrée, (i) un historique des routes BGP vers les différents réseaux dans l'Internet issues des tables de routage BGP de deux collecteurs (un RIPE RIS [26] et un RouteViews [43]), et (ii) un historique des bases de données `whois` (IRRs) fournissant des informations sur l'enregistrement (assignations/allocations) des adresses IP et numéros d'ASs, comme par exemple le propriétaire d'une adresse IP ou d'un numéro d'AS, le pays dans lequel une adresse IP ou un numéro d'AS a été enregistré, des informations permettant de contacter un propriétaire ou encore d'éventuelles politiques de routage déclarées entre des ASs. L'architecture de notre système de détecter est dépeint dans la Figure 6.8. Le processus de détection consiste en les étapes suivantes :

1. ***L'identification de blocs d'adresses IP inactifs.*** Nous utilisons l'historique des routes BGP que nous possédons pour tout d'abord identifier les blocs d'adresses IP inactifs, c'est-à-dire les blocs qui ne sont pas annoncés publiquement dans l'Internet. Comme nous avons plus tôt dans notre analyse des attaques par détournement identifiées, les attaquants semblent cibler particulièrement les blocs d'adresses IP laissés inactifs par leur propriétaire. De plus, l'absence, à priori, de machine dans ces réseaux réduit la probabilité pour l'attaquant de créer des perturbations lorsqu'il détourne ces blocs.

2. ***Détection de blocs d'adresses IP revenants.*** Nous surveillons l'état de routage des blocs d'adresses IP inactifs et, lorsqu'un de ceux-ci réapparaît dans l'Internet, nous vérifions la validité de l'annonce BGP en utilisant le même procédé que celui-ci développé pour la validation des cas suspects de détournements BGP décrit dans la Section 6.3.3. A ce stade nous disposons d'un ensemble de cas réels de détournements BGP.

3. ***Corrélation entre les détournements BGP et des traces d'activités malveillantes.*** Afin d'identifier d'éventuelles activités malveillantes réalisées depuis les réseaux détournés, nous surveillons la présence de l'ensemble des blocs identifiés à l'étape ② comme ayant été détournés dans différents ensembles de données contenant des traces de divers types d'attaques survenant dans l'Internet (*e.g.,* du spam, des sites web malveillants, etc).

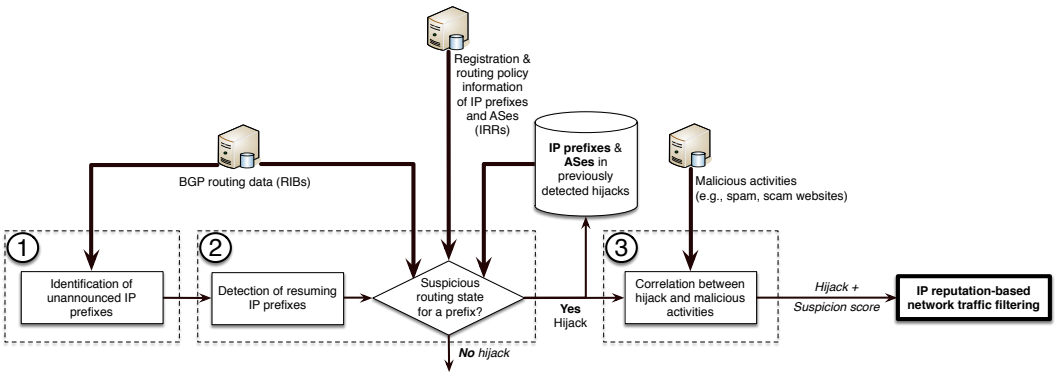


FIGURE 6.8 – Détection en temps réel de détournements BGP : architecture du système.

Afin de valider notre approche, nous avons appliqué notre système sur des données de la même période que celle utilisée dans l'analyse des données de SPAMTRACER, *i.e.,* entre septembre 2012 et juin 2014. Durant cette période, notre système de détection a identifié un total de 4.262 blocs d'adresses IP détournés dont 2.700 (99,5%) faisaient partie des 2.713 cas identifiés au moyen des données de SPAMTRACER. Bien que nous somme confiant dans le fait que nous allons vraisemblablement identifier de nouveaux cas de détournements dans les 1.517 cas supplémentaires fournis par notre système de détection et restants à vérifier, nous rencontrerons probablement également des faux positifs. Nous planifions d'analyser ces possibles faux positifs et de réaliser une étude de sensibilité des différents paramètres utilisés dans notre système.

6.4.7 Enseignements

Nous concluons cette section en donnant quelques enseignements concrets tirés de notre analyse des attaques par détournement identifiées et qui peuvent être utilisées afin d'améliorer les contre-mesures actuelles et, de façon générale, d'aider à mieux se défendre contre la menace que pose ces attaques sur la sécurité de l'Internet.

Enseignement 1 *Nous avons confirmé l'existence, à l'heure actuelle, du phénomène de « BGP spectrum agility » dans l'Internet sous la forme de campagnes d'attaques par détournement BGP malveillant furtives et persistances.*

Enseignement 2 *Les contre-mesures actuelles pour les attaques par détournement BGP telles que [7, 96, 101, 116, 157] sont inefficaces contre des attaques impliquant des blocs d'adresses IP préalablement* non annoncés *et, de façon générale, lorsque l'attaquant utilise un AS d'origine* ***valide*** *et un FAI invalide pour détourner les blocs. Seul le déploiement complet de BGPsec et du système de ROAs pourraient empêcher efficacement ces attaques. En attendant, nous suggérons que les systèmes de détection d'attaques par détournement BGP incluent dans leurs signatures des caractéristiques des attaques que nous avons observées.*

Enseignement 3 *Les propriétaires de blocs d'adresses IP non annoncés publiquement dans l'Internet laissent leurs ressources vulnérables aux détournements. Une pratique exemplaire serait de systématiquement annoncer publiquement tous les blocs alloués, même si ceux-ci ne sont pas utilisés.*

Enseignement 4 *Une chasse mondiale des blocs d'adresses IP orphelins devraient être lancées afin d'empêcher que ces blocs soient détournés à des fins malveillantes. De plus, les propriétaires de blocs disparaissant ou ne requérant plus leurs ressources IP devraient (être forcés de) les retourner. Garder les informations contenues dans les bases de données* `whois` *(IRRs) et RPKI à jour est primordiale afin d'aider à empêcher le détournement de blocs d'adresses IP.*

Enseignement 5 *Les attaques identifiées ont révélé qu'un grand nombre de blocs d'adresses IP différents ont été détournés au moyen d'un très faible nombre d'ASs (d'origine ou de FAIs) invalides ou malveillants différents. Cela suggère qu'on peut utiliser les ASs identifiés comme invalides ou malveillants afin de détecter de prochains détournements ou même bloquer tout trafic réseau en provenance ou à destination de blocs d'adresses IP annoncés via ces ASs.*

Enseignement 6 *Nous avons utilisé quelques caractéristiques des attaques identifiées dans ce travail pour développer une système de détection en temps réel de détournements BGP. Le but de ce système est non seulement d'aider à contrer ces attaques mais aussi à permettre de surveillé sur le long terme le phénomène de « BGP spectrum agility ».*

Enseignement 7 *Nous avons récemment entrepris une collaboration avec des CERTs, FAIs et les communautés NANOG et RIPE au sens large. Cette collaboration nous a récemment permis de confirmer qu'un des ASs identifié comme invalide et malveillant par notre système, et qui était impliqué dans pas moins de 793 attaques, avait vu son contrat d'appairage résilié par son FAI.*

6.5 Conclusion et perspectives futures de recherches

6.5.1 Contributions scientifiques

L'introduction à cette thèse a posé plusieurs questions de recherche concernant l'existence, en 2014, d'attaques par détournement BGP malveillant dans l'Internet, l'importance de ces attaques, ainsi que le mode opératoire des attaquants. Pour répondre à ces questions, nous avons défini *trois* problèmes de recherche que cette thèse se devait d'aborder. Sur base des résultats expérimentaux obtenus et décris dans les sections précédentes, nous allons maintenant montrer comment nous avons résolu ces problèmes de recherche.

Problème 1 (P1) : *corrélation de données de sécurité et de routage.*

Nous avons résolu le premier problème de recherche en construisant un environnement de collecte de données à grande échelle, SpamTracer, décrit dans la Section 6.3. Motivée par le manque de données ou de systèmes permettant d'étudier les attaques par détournement BGP malveillant dans l'Internet, nous avons développé une méthode qui combine des alertes au niveau BGP (plan de contrôle) et des mesures réseau traceroute (plan de données) relatives à des réseaux ayant émis du trafic réseau malveillant, comme par exemple du spam ou des communications avec des serveurs C&C. Ces traces, collectées depuis différents points de vue dans le monde, sont également enrichies d'informations issues des Registres du Routage de l'Internet (« Internet Routing Registries (IRRs) »). Elle nous fournissent ainsi un large ensemble de caractéristiques sur les anomalies de routage observées afin d'identifier celles qui résultent très probablement d'une attaque par détournement BGP.

Problème 2 (P2) : *évaluation de l'existence d'attaques par détournement BGP malveillant.*

Le solution au second problème posé par cette thèse est *double*.

Premièrement, dans la Section 6.3, nous avons présenté une nouvelle approche pour l'identification et la validation de possible attaques par détournement BGP malveillant depuis les données collectées par SpamTracer. Il s'agit d'un un procédé novateur de filtrage et d'évaluation d'anomalies de routage : (i) il combine des heuristiques d'extraction d'anomalies de routage avec (ii) un modèle d'identification des attaques par détournement BGP basé sur l'analyse décisionnelle multicritittères. Les cas sélectionnés par le système sont ensuite validés en utilisant des sources de données externes et le retour éventuel des propriétaires des réseaux concernés.

Deuxièmement, dans la Section 6.4, nous avons analysé presque deux ans de données et dévoilons plus de 2.000 attaques par détournement BGP malveillant qui ont eu lieu de façon régulière pendant toute la période de l'expérimentation. Un grand nombre de ces attaques ont été confirmées par des victimes ou par des FAIs (Fournisseurs d'Accès à Internet) impliqués involontairement.

Problème 3 (P3) : *si elles existent, déterminer à quel point ce type d'attaques est répandu et caractériser le comportement des attaquants.*

Dans la Section 6.4, nous révélons un mode opératoire sophistiqué utilisé par les cyber-criminels afin de subrepticement prendre le contrôle de blocs d'adresses IP sans l'autorisation de leur propriétaire. Nos résultats montrent que les attaques identifiées ont réussi à mettre en échec des mesures de prévention contre le spam et les attaques par détournement BGP. À la lumière de ces résultats, nous proposons des pistes afin de mieux se protéger contre cette menace émergente. Nous tirons également parti des caractéristiques des attaques observées pour concevoir un système de détection en temps réel de détournements de blocs IP.

En conclusion, en plus d'avoir résolu les problèmes de recherche posés, cette thèse

a mis en lumière des éléments encore inconnus jusque là à propos des *attaques par détournement BGP malveillant*. A notre connaissance, ce travail est le premier à fournir une description aussi détaillée de cas confirmés d'attaques par détournement BGP malveillant réalisés dans le but d'appuyer des activités de spam ou d'hébergement de sites web frauduleux. Il est aussi le premier à exposer au grand jour le mode opératoire des pirates reponsables de ces attaques. Enfin, cette thèse invite la communauté réseau à prendre conscience que des attaques par détournement BGP malveillant ont eu lieu dans l'Internet, et ce de façon récurante et persistante, pendant des mois, voire des années. Nous espérons que ce travail inspirera de nouvelles recherches afin de comprendre mieux encore la motivation des pirates derrières ces attaques ainsi que le moyen de s'en prémunir.

6.5.2 Perspectives futures de recherche

Hormis la mise en lumière de menace que pose les attaques par détournement BGP malveillant sur l'Internet, ce travail a permis d'identifier de futures perspectives de recherche. Tout d'abord, nous avons mis en place une infrastructure de collecte de données afin d'étudier le comportement de routage de réseaux émettant du trafic réseau malveillant. Une *première* perspective de recherche se situe donc dans la collecte d'un ensemble accrus de caractéristiques permettant d'encore mieux décrire les réseaux surveillés. Ensuite, nous avons utilisé les données collectées afin d'extraire des cas suspects d'attaques par détournement BGP. Nous avons pu en valider un sous-ensemble menant ainsi à l'identification de cas réels d'attaques par détournement BGP malveillant. Une *seconde* perspective de recherche future se situe donc dans la méthodologie d'analyse des données. Enfin, bien que notre analyse de ces attaques nous a permis de répondre aux questions posées par cette thèse, les découvertes rapportées sur le phénomène des attaques par détournement BGP malveillant ouvrent également la voie à de nouvelles perspectives de recherche sur ce phénomène.

Les données de sécurité

Dans ce travail, nous nous sommes concentré sur l'étude du comportement de routage de réseaux émettant du spam. Cela nous a permis de confirmer l'existence, en 2014, du phénomène de « BGP spectrum agility ». Nous avons également vu que certains des réseaux identifiés comme ayant été détournés ont hébergés des sites web malveillants. Cependant, nous avons également observé qu'une grande partie des réseaux identifiés comme ayant été détournés n'ont exposé aucun trafic réseau malveillant. Nous pensons que cela peut résulter de la couverture limitée des données de sécurité utilisée. Nous pensons également que cela peut indiquer que les réseaux détournés sont utilisés pour perpétrer d'autres types d'activités malveillantes que du spam, par exemple héberger des serveurs C&C, lancer des attaques de déni de service distribuées ou encore distribuer des logiciels malveillants.

Les données de routage

Grâce au déploiement à grande échelle de SpamTracer, nous collectons pour chaque réseau surveillé des données BGP depuis six collecteurs et des données traceroute depuis trois collecteurs. Cela dit, cela ne donne pas une visibilité complète sur tous les changements dans le routage pouvant survenir dans l'internet [84, 156]. Les ressources matérielles disponibles sur les machines sur lesquelles tournent le système a également des répercussions sur le nombre maximum de réseaux pouvant être surveillés par jour (40.000) ainsi sur la période pendant laquelle ces réseaux sont surveillés.

Enfin, une limitation intrinsèque de notre méthodologie vient du fait que ni les données BGP ni les mesures traceroute n'ont été conçues pour capturer les relations complexes qui peuvent exister en les ASs dans l'Internet. Ainsi, la précision des résultats ne peut être qu'égale à la précision des données [153]. Nous avons essayé de contrebalancer cette limitation par la mise en place d'un système de collecte de données spécifiques à l'étude du phénomène de « BGP spectrum agility ».

Le procédé de filtrage et d'évaluation d'anomalies de routage

Premièrement, dans le procédé de filtrage d'évaluation d'anomalies de routage que nous avons développé pour identifier des suspects d'attaques par détournement BGP malveillant, nous avons utilisé une approche basée sur l'analyse décisionnelle multi-critères (MCDA). Cette approche nous a permis d'inclure de l'expertise dans notre modèle de détection, notamment en utilisant une fonction d'agrégation telle que WOWA (Weighted Ordered Weighted Average). Aussi, les paramètres utilisés dans notre modèle ont été déterminés de façon empirique. Dans le futur, nous souhaiterions réaliser une analyse de sensibilité pour différentes valeurs de ces paramètres.

Deuxièmement, il faut noter que l'objectif de notre méthode d'identification d'attaques par détournement BGP malveillant est d'identifier, parmi le grand nombre d'anomalies de routage observées dans nos données, un ensemble réduit correspondant aux plus suspectes et donc à celles résultants le plus vraisemblablement d'une vraie attaque. A aucun moment nous n'avons eu pour objectif de développer un système pour identifier l'entièreté des cas possibles d'attaques dans nos données ou bien encore un nouveau système générique de détection d'attaques par détournement BGP.

Enfin, nous souhaiterions également, dans le futur, explorer de nouveaux types d'anomalies de routage pouvant être extraites des données BGP et traceroute que nous collectons.

La validation des cas suspects d'attaques par détournement BGP malveillant

Ainsi que nous l'avons montré dans [177], corréler des anomalies de routage et des traces d'activités malveillantes est insuffisant pour identifier de façon catégorique des attaques par détournement BGP malveillant. Notre méthode de validation des cas

suspects d'attaques décrite dans la Section 6.3 nous a déjà permis de rassembler un ensemble de preuves pour valider un certain nombre de cas. Nous avons également vu que la collaboration avec des Fournisseurs d'Accès à Internet (FAIs) permet d'aller encore plus loin dans la validation de cas suspects. Dans le futur, nous avons donc l'intention d'automatiser cette collaboration autant que possible pour les différents cas suspects que nous identifions.

L'exploitation des résultats

Dans la Section 6.4.6 nous avons exploré l'utilisation de certaines caractéristiques des attaques identifiées dans ce travail afin de développer un système de détection en temps-réel de ce types d'attaques et ainsi pouvoir endiguer ce phénomème. Nous avons pu valider notre approche de façon expérimentale. Nous cherchons à présent à évaluer la qualité de notre algorithme et des paramètres utilisés.

La surveillance de l'Internet IPv6

Dans ce travail, nous avons concentré notre attention sur l'Internet IPv4 pour une bonne raison : nous manquons actuellement de données de sécurités relatives à l'IPv6. Même si les adresses IPv4 s'épuisent[3], il reste encore à l'heure actuelle environ 20% d'adresses IPv4 qui ont été alloués à un propriétaire mais qui ne sont pas annoncées publiquement dans l'Internet[4]. Comme nous l'avons vu, ces adresses constituent des cibles de choix pour les pirates qui peuvent les détourner sans que cela soit trop visible. Nous pensons qu'avec l'adoption progressive d'IPv6 par les FAIs[5] et le fait que la version de BGP utilisée pour IPv6 fonctionne de manière identique à celle utilisée pour IPv4, des attaques par détournement BGP malveillant toucherons l'Internet IPv6 de la même manière qu'elles touchent l'Internet IPv4. A l'heure où nous écrivons ce document, nous n'avons cependant eu vent d'aucun cas d'attaque par détournement BGP ayant touché IPv6.

[3] http://www.potaroo.net/tools/ipv4/

[4] Sur base de statistiques publiées par les RIRs et disponibles à http://bgp.potaroo.net/ipv4-stats/prefixes.txt

[5] http://bgp.potaroo.net/v6/as2.0/index.html

Bibliography

[1] Abuse.ch. http://www.abuse.ch/. 12

[2] Amazon Elastic Compute Cloud (EC2). http://aws.amazon.com/ec2/. 46

[3] [atlas] BGP hijacking. http://www.ripe.net/ripe/mail/archives/ripe-atlas/2013-December/001233.html. 10

[4] Belnet. http://www.belnet.be/. 106

[5] BGP Monitoring System. http://bgpmon.netsec.colostate.edu/. 32, 36

[6] BGP: the Border Gateway Protocol: Advanced Internet Routing Resources. http://www.bgp4.as/. 31

[7] BGPmon.net. http://www.bgpmon.net/. 33, 98, 107, 140, 145

[8] bgpTables: A global BGP visibility analysis tool. http://bgpinspect.merit.edu/. 33

[9] CERT.be. https://www.cert.be/. 106

[10] Cisco SenderBase. http://www.senderbase.org/. 11, 123

[11] Composite Blocking List. http://cbl.abuseat.org/. 11, 87, 123

[12] Cyclops. http://cyclops.cs.ucla.edu. 32

[13] Dshield: Cooperative Network Security Community. http://www.dshield.org. 12, 48

[14] Emerging Threats. http://www.emergingthreats.net. 12, 48

[15] GeoIP API: MaxMind. http://www.maxmind.com/. 48, 50, 126

[16] Huricane Electric BGP Toolkit. http://bgp.he.net. 32

[17] Internet Alert Registry. http://www.cs.unm.edu/~karlinjf/IAR/. 28

[18] Internet Topology Collection. http://irl.cs.ucla.edu/topology/. 52

[19] Introduction to ARIN's database. https://www.arin.net/knowledge/database_text.html. 33

[20] IRR.net. http://www.irr.net/. 33, 50, 58, 129

[21] Manitu.net DNSBL. http://www.dnsbl.manitu.net/. 11, 84, 134

[22] North American Network Operators' Group (NANOG) mailing list. https://www.nanog.org/list/. 10, 59, 63, 91, 129

[23] Peeringdb. https://www.peeringdb.com/. 32

[24] PlanetLab: An open platform for developing, deploying and accessing planetary-scale services. https://www.planet-lab.org/. 46

[25] Renesys - the internet intelligence authority. http://www.renesys.com/. 33, 98, 140

[26] RIPE Routing Information Service (RIS). http://www.ripe.net/data-tools/stats/ris/. 32, 36, 58, 78, 102, 129, 144

[27] RIPE Working Group Mailing Lists. https://www.ripe.net/ripe/mail/wg-lists/. 10, 59, 63, 91, 129

[28] RIPEstat: Internet Measurements and Analysis. https://stat.ripe.net/. 32

[29] Robtex Swiss Army Knife Internet Tool. https://www.robtex.com/. 32

[30] ROVER Testbed: BGP Route Origin Verification. https://rover.secure64.com. 24

[31] RPKI Dashboard. http://rpki.surfnet.nl/. 25

[32] Shadowserver. http://www.shadowserver.org/. 12, 48

[33] SpamCop.net. http://www.spamcop.net/. 11

[34] Spamhaus. http://www.spamhaus.org/. 9, 11, 59, 63, 84, 101, 123, 129, 133, 134

[35] Symantec.cloud: Email Security, Web Security, Endpoint Protection, Archiving, Continuity, Instant Messaging Security. http://www.symanteccloud.com/. 11, 40, 47

[36] Team Cymru Community Services. http://www.team-cymru.org/. 86, 96, 126

[37] Team Cymru IP to ASN. https://asn.cymru.com/. 48

[38] Team Cymru IPv4 Fullbogons. http://www.team-cymru.org/Services/Bogons/. 13, 20, 48, 50

[39] The Apache SpamAssassin Project. http://spamassassin.apache.org/. 11, 123

[40] The New Threat: Targeted Internet Traffic Misdirection. http://www.renesys.com/2013/11/mitm-internet-hijacking/. 1, 21, 118

[41] The Spamhaus Don't Route Or Peer (DROP) List. http://www.spamhaus.org/drop/. 12, 48, 50

[42] Uceprotect. http://www.uceprotect.net/. 11, 84, 123, 134

[43] University of Oregon RouteViews Project. http://www.routeviews.org/. 32, 36, 47, 49, 58, 102, 129, 144

[44] YouTube IP Hijacking. http://www.nanog.org/mailinglist/mailarchives/old_archive/2008-02/msg00453.html. 1, 21, 118

[45] Route Hygiene: The Dirt on the Internet. http://www.renesys.com/2009/03/compliance-scoring-by-country/, March 2009. 28, 33

[46] Prefix hijacking by Michael Lindsay via Internap. http://mailman.nanog.org/pipermail/nanog/2011-August/039381.html, August 2011. 1, 10, 59, 118, 129

[47] Symantec Internet Security Threat Report: Future Spam Trends: BGP Hijacking. Case Study - Beware of "Fly-by Spammers". http://www.symantec.com/threatreport/, April 2012. 1, 118

[48] BGP hijack of Spamhaus? http://mailman.nanog.org/pipermail/nanog/2013-March/057340.html, March 2013. 10, 37, 78

[49] Illegal usage of AS51888 (and PI 91.220.85.0/24) from AS42989 and AS57954 (in ukraine). http://mailman.nanog.org/pipermail/nanog/2013-May/058230.html, May 2013. 10, 93

[50] *IP Routing: BGP Configuration Guide, Cisco IOS XE Release 3S*. Cisco Systems, Inc., August 2013. 25

[51] Prefix hijacking, how to prevent and fix currently. http://seclists.org/nanog/2014/Aug/479, August 2014. 10, 99, 142

[52] Re: Prefix hijacking, how to prevent and fix currently. http://seclists.org/nanog/2014/Sep/30, September 2014. 89, 139

[53] C. Alaettinoglu, C. Villamizar, E. Gerich, D. Kessens, D. Meyer, T. Bates, D. Karrenberg, and M. Terpstra. Routing Policy Specification Language (RPSL). RFC2622, June 1999. 33

[54] Y. H. Andre, Y. Hyun, A. Broido, and K. Claffy. On Third-party Addresses in Traceroute Paths, 2003. 39

[55] Y. H. Andre, Y. Hyun, A. Broido, and K. Claffy. Traceroute and BGP AS Path Incongruities. Technical report, CAIDA, 2003. 38, 39

[56] APNIC. Internet Routing Registry Tutorial. http://training.apnic.net/docs/TROU08.pdf. 33

[57] B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira. Avoiding traceroute anomalies with Paris traceroute. In *IMC*, pages 153–158. ACM, 2006. 39

[58] F. Baker and P. Savola. Ingress Filtering for Multihomed Networks. RFC3704, March 2004. 16

[59] H. Ballani, P. Francis, and X. Zhang. A study of prefix hijacking and interception in the Internet. In *SIGCOMM*, pages 265–276. ACM, 2007. 2, 3, 21, 30, 119, 120

[60] A. Barbir, S. Murphy, and Y. Yang. Generic Threats to Routing Protocols. RFC4593, October 2006. 15

[61] G. D. Battista, T. Refice, and M. Rimondini. How to extract BGP peering information from the internet routing registry. In *Proceedings of the 2nd Annual ACM Workshop on Mining Network Data, MineNet 2006, Pisa, Italy, September 15, 2006*, pages 317–322, 2006. 59

[62] G. Beliakov, A. Pradera, and T. Calvo. *Aggregation Functions: A Guide for Practitioners*. Springer, Berlin, New York, 2007. 55, 128

[63] L. Blunk, M. Karir, and C. Labovitz. Multi-Threaded Routing Toolkit (MRT) Routing Information Export Format. RFC6396, October 2011. 36

[64] R. Bush, O. Maennel, M. Roughan, and S. Uhlig. Internet optometry: assessing the broken glasses in internet reachability. In *IMC*, pages 242–253. ACM, 2009. 38, 49, 53

[65] K. Butler, T. R. Farley, P. McDaniel, and J. Rexford. A Survey of BGP Security Issues and Solutions. *Proceedings of the IEEE*, 98(1):100–122, Jan. 2010. 15, 17, 23

[66] CAIDA. Archipelago Measurement Infrastructure. http://www.caida.org/projects/ark/. 39

[67] Y.-J. Chi, R. Oliveira, and L. Zhang. Cyclops: the AS-level connectivity observatory. *SIGCOMM CCR*, 38(5):5–16, 2008. 32

[68] M. Cova, C. Leita, O. Thonnard, A. D. Keromytis, and M. Dacier. An analysis of rogue AV campaigns. In *RAID*, pages 442–463. Springer-Verlag, 2010. 64, 130

[69] A. de la Haye. Chief Operations Officer at RIPE NCC. RIPE67, October 2013. 89, 139

[70] Z. Duan, K. Gopalan, and X. Yuan. An Empirical Study of Behavioral Characteristics of Spammers: Findings and Implications. *Computer Communications*, 34(14):1764–1776, Sept. 2011. 1, 9, 118, 122

[71] T. Dumitras and D. Shou. Toward a standard benchmark for computer security research: the worldwide intelligence network environment (WINE). In *First Workshop on Building Analysis Datasets and Gathering Experience Returns for Security*, pages 89–96. ACM, 2011. 89, 137

[72] J. Durand, I. Pepelnjak, and G. Doering. BGP operations and security. draft-ietf-opsec-bgp-security-03.txt, April 2014. 15

[73] Z. Durumeric, E. Wustrow, and J. A. Halderman. ZMap: Fast Internet-wide Scanning and Its Security Applications. In *22nd USENIX Security Symposium (USENIX Security 13)*, pages 605–620, Washington, D.C., 2013. USENIX. 39

[74] V. Fuller and T. Li. Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan. RFC4632, August 2006. 12

[75] L. Gao. On Inferring Autonomous System Relationships in the Internet. *IEEE/ACM Trans. Netw.*, 9(6):733–745, Dec. 2001. 52, 71

[76] J. Gersch and D. Massey. Characterizing vulnerability to IP hijack attempts. In *Technologies for Homeland Security (HST), 2013 IEEE International Conference on*, pages 328–333, Nov 2013. 31

[77] J. Gersch and D. Massey. ROVER: Route Origin Verification Using DNS. In *ICCCN*, pages 1–9, July 2013. 23

[78] J. Gersch, D. Massey, C. Olschanowsky, and L. Zhang. DNS Resource Records for Authorized Routing Information. draft-gersch-grow-revdns-bgp-02, February 2013. 23

[79] J. Gersch, D. Massey, E. Osterweil, and C. Olschanowsky. Reverse DNS Naming Convention for CIDR Address Blocks. draft-gersch-dnsop-revdns-cidr-04.txt, February 2013. 23

[80] J. Gersch, D. Massey, and C. Papadopoulos. Incremental Deployment Strategies for Effective Detection and Prevention of BGP Origin Hijacks. In *Distributed Computing Systems (ICDCS), 2014 IEEE 34th International Conference on*, pages 670–679, June 2014. 31

[81] V. Gill, J. Heasley, D. Meyer, E. P. Savola, and C. Pignataro. The Generalized TTL Security Mechanism (GTSM). RFC5082, October 2007. 16

[82] S. Goldberg, M. Schapira, P. Hummon, and J. Rexford. How Secure Are Secure Interdomain Routing Protocols. In *Proceedings of the ACM SIGCOMM 2010 Conference*, pages 87–98, New York, NY, USA, 2010. ACM. 31

[83] G. Goodell, W. Aiello, T. Griffin, J. Ioannidis, P. McDanial, and A. Rubin. Working Around BGP: An Incremental Approach to Improving Security and Accuracy of Interdomain Routing. In *NDSS*, 2003. 23

[84] E. Gregori, A. Improta, L. Lenzini, L. Rossi, and L. Sani. On the incompleteness of the AS-level graph: a novel methodology for BGP route collector placement. In *IMC*, pages 253–264, 2012. 37, 71, 113, 149

[85] S. Hagen. *IPv6 Essentials*. "O'Reilly Media, Inc.", 2006. 12

[86] Y. Hamada. Case Study: Trouble with Messy IRR. `http://archive.apnic.net/meetings/25/program/routing/hamada-messy-irr.pdf`, February 2008. 28, 33

[87] S. Hao, N. A. Syed, N. Feamster, A. G. Gray, and S. Krasser. Detecting spammers with SNARE: spatio-temporal network-level automatic reputation engine. In *Proceedings of the 18th conference on USENIX security symposium*, SSYM'09, pages 101–118, Berkeley, CA, USA, 2009. USENIX Association. 11

[88] D. Haskin. Default Route Advertisement In BGP2 And BGP3 Versions Of The Border Gateway Protocol. RFC1397, January 1993. 71

[89] A. Heffernan. Protection of BGP Sessions via the TCP MD5 Signature Option. RFC2385, August 1998. 16

[90] R. Hinden and B. Haberman. Unique Local IPv6 Unicast Addresses. RFC4193, October 2005. 13

[91] R. Hiran, N. Carlsson, and P. Gill. Characterizing large-scale routing anomalies: a case study of the china telecom incident. In *PAM*, pages 229–238. Springer-Verlag, 2013. 1, 118

[92] R. Hiran, N. Carlsson, and N. Shahmehri. PrefiSec: A Distributed Alliance Framework for Collaborative BGP Monitoring and Prefix-based Security. In ACM, editor, *To appear in CCS Workshop on Information Sharing and Collaborative Security*, Scottsdale, AZ, Nov 2014. 31

[93] M. Hogewoning. IP Hijacking: Secure Internet Routing, March 2012. 21, 91

[94] S.-C. Hong, J.-K. Hong, and H. Ju. Ip prefix hijacking detection using the collection of as characteristics. In *Network Operations and Management Symposium (APNOMS), 2011 13th Asia-Pacific*, pages 1 –7, sept. 2011. 31

[95] S.-C. Hong, H.-T. Ju, and J. W. Hong. IP prefix hijacking detection using idle scan. In *APNOMS'09: Proceedings of the 12th Asia-Pacific network operations and management conference on Management enabling the future internet for changing business and new computing services*, pages 395–404. Springer-Verlag, 2009. 29, 31, 38

[96] X. Hu and Z. M. Mao. Accurate Real-Time Identification of IP Prefix Hijacking. In *Security and Privacy*, pages 3–17. IEEE, 2007. 1, 2, 3, 8, 11, 17, 26, 28, 37, 38, 51, 52, 77, 98, 107, 118, 119, 120, 122, 124, 140, 145

[97] Y.-C. Hu, A. Perrig, and M. Sirbu. SPV: Secure Path Vector Routing for Securing BGP. In *SIGCOMM*, pages 179–192. ACM, 2004. 23

[98] G. Huston. AS65000 BGP Routing Table Analysis Report. http://bgp.potaroo.net/as2.0/bgp-active.html. 74

[99] G. Huston. CIDR report. http://www.cidr-report.org/as2.0/. 12, 39

[100] G. Huston. *ISP Survival Guide: Strategies for Running a Competitive ISP*. John Wiley & Sons, Inc., 1998. 12, 13, 14, 71

[101] G. Huston and G. Michaelson. Validation of Route Origination Using the Resource Certificate Public Key Infrastructure (PKI) and Route Origin Authorizations (ROAs). RFC6483, February 2012. 1, 24, 25, 98, 99, 100, 107, 118, 123, 124, 140, 142, 145

[102] G. Huston, M. Rossi, and G. Armitage. Securing BGP: A Literature Survey. *Communications Surveys Tutorials, IEEE*, 13(2):199–222, 2011. 1, 16, 22, 99, 118, 142

[103] I. A. N. A. (IANA). Internet Protocol Version 6 Address Space. http://www.iana.org/assignments/ipv6-address-space/ipv6-address-space.xhtml, February 2013. 13

[104] I. A. N. A. (IANA). IANA IPv4 Address Space Registry. http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xhtml, October 2014. 13

[105] Q. Jacquemart, G. Urvoy Keller, and E. W. Biersack. A longitudinal study of BGP MOAS prefixes. In *International Traffic Monitoring and Analysis Workshop*, 04 2014. 37, 52

[106] E. Jones and O. L. Moigne. OSPF Security Vulnerabilities Analysis. http://tools.ietf.org/html/draft-ietf-rpsec-ospf-vuln-02, June 2006. 14

[107] Juniper Networks, Inc. *Example: Configuring Origin Validation for BGP*, May 2013. 25

[108] J. Karlin. Pretty Good BGP: Improving BGP by cautiously adopting routes. In *ICNP*. IEEE, 2006. 28, 74

[109] D. Karrenberg, G. Ross, P. Wilson, and L. Nobile. Development of the Regional Internet Registry System. http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_4-4/regional_internet_registries.html, December 2001. 91

[110] S. Kent, C. Lynn, J. Mikkelson, and K. Seo. Secure border gateway protocol (S-BGP) – real world performance and deployment issues. In *NDSS*, pages 103–116, February 2000. 22

[111] S. Kent and K. Seo. Security Architecture for the Internet Protocol. RFC4301, December 2005. 16

[112] A. Khan, H.-c. Kim, T. Kwon, and Y. Choi. A Comparative Study on IP Prefixes and Their Origin Ases in BGP and the IRR. *SIGCOMM CCR*, 43(3):16–24, July 2013. 28, 33

[113] V. Khare, Q. Ju, and B. Zhang. Concurrent prefix hijacks: occurrence and impacts. In *IMC*, pages 29–36. ACM, 2012. 26, 31, 124

[114] G. Kondrak. N-gram similarity and distance. In *Conf. on String Processing and Information Retrieval*, pages 115–126, 2005. 64

[115] C. Kruegel, D. Mutz, W. Robertson, and F. Valeur. Topology-Based Detection of Anomalous BGP Messages. In *RAID*, pages 17–35, 2003. 26, 124

[116] M. Lad, D. Massey, D. Pei, Y. Wu, B. Zhang, and L. Zhang. PHAS: A prefix hijack alert system. In *USENIX Security Symposium*, 2006. 1, 2, 3, 26, 36, 98, 107, 118, 119, 120, 124, 140, 145

[117] M. Lad, R. Oliveira, B. Zhang, and L. Zhang. Understanding the impact of bgp prefix hijacks. *ACM SIGCOMM Poster*, 2006. 17

[118] M. Lad, R. V. Oliveira, B. Zhang, and L. Z. 0001. Understanding Resiliency of Internet Topology against Prefix Hijack Attacks. In *DSN*, pages 368–377. IEEE Computer Society, 2007. 17, 31

[119] M. Lepinski. BGPSEC Protocol Specification. Internet-Draft, February 2013. 1, 24, 25, 98, 99, 118, 123, 140, 143

[120] M. Lepinski and S. Kent. An Infrastructure to Support Secure Internet Routing. RFC6480, February 2012. 4, 24, 98, 99, 120, 123, 140, 142

[121] P. Litke and J. Stewart. BGP Hijacking for Cryptocurrency Profit. `http://www.secureworks.com/cyber-threat-intelligence/threats/bgp-hijacking-for-cryptocurrency-profit/`, August 2014. 37, 78

[122] M. Luckie, Y. Hyun, and B. Huffaker. Traceroute probe method and forward IP path inference. In *IMC*, pages 311–324, New York, NY, USA, 2008. ACM. 39, 49

[123] A. Lutu, M. Bagnulo, and O. Maennel. The BGP Visibility Scanner. In *Computer Communications Workshop (CCW)*, pages 115–120. IEEE, 2013. 36, 38

[124] H. V. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and A. Venkataramani. iPlane: An Information Plane for Distributed Services. In *UNENIX OSDI*, pages 367–380, Berkeley, CA, USA, 2006. USENIX Association. 39

[125] D. Madory. Sprint, Windstream: Latest ISPs to hijack foreign networks. `http://renesys.com/2014/09/latest-isps-to-hijack/`, September 2014. 93, 99, 142

[126] R. Mahajan, D. Wetherall, and T. Anderson. Understanding BGP Misconfiguration. In *SIGCOMM*, pages 3–16. ACM, 2002. 2, 8

[127] Malware Domains. Malware Domain Block List. http://www.malwaredomainlist.com/. 12, 48

[128] Z. M. Mao, J. Rexford, J. Wang, and R. H. Katz. Towards an accurate AS-level traceroute tool. In *SIGCOMM*, pages 365–378. ACM, 2003. 31, 39, 69, 131

[129] C. Mcarthur and M. Guirguis. Stealthy IP prefix hijacking: don't bite off more than you can chew. In *GLOBECOM*, pages 2480–2485. IEEE Press, 2009. 31, 36

[130] Merit Network, Inc. Routing Registry Tutorial. https://www.nanog.org/meetings/nanog51/presentations/Sunday/NANOG51.Talk34.NANOG51IRRTutorial.pdf, January 2011. 33

[131] J. Mitchell. Autonomous System (AS) Reservation for Private Use. RFC6996, July 2013. 12

[132] G. Nakibly, A. Kirshon, D. Gonikman, and D. Boneh. Persistent OSPF Attacks. In *NDSS*. The Internet Society, 2012. 14

[133] L. Nobile and L. Vegoda. Address Space and AS Hijacking. http://meetings.ripe.net/ripe-48/presentations/ripe48-eof-nobile-vegoda.pdf, May 2004. 21, 91

[134] O. Nordstrom and C. Dovrolis. Beware of BGP attacks. *SIGCOMM CCR*, 34(2):1–8, Apr. 2004. 17

[135] Number Resource Organization (NRO). Free Pool of IPv4 Address Space Depleted. http://www.nro.net/news/ipv4-free-pool-depleted, February 2011. 115

[136] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang. The (in)Completeness of the Observed Internet AS-level Structure. *IEEE/ACM Transactions on Networking*, January 2010. 37

[137] P. v. Oorschot, T. Wan, and E. Kranakis. On interdomain routing security and pretty secure BGP (psBGP). *ACM Trans. Inf. Syst. Secur.*, 10, July 2007. 22

[138] V. N. Padmanabhan and D. R. Simon. Secure traceroute to detect faulty or malicious routing. *SIGCOMM CCR*, 33:77–82, 2002. 24, 39

[139] A. Pilosov and T. Kapela. Stealing The Internet: An Internet-Scale Man In The Middle Attack. http://www.defcon.org/images/defcon-16/dc16-presentations/defcon-16-pilosov-kapela.pdf. 17, 21, 31

[140] A. Pitsillidis, C. Kanich, G. M. Voelker, K. Levchenko, and S. Savage. Taster's choice: a comparative analysis of spam feeds. In *IMC*, pages 427–440. ACM, 2012. 112

[141] J. Porenta and M. Ciglarič. Empirical Comparison of IP Reputation Databases. In *CEAS*, pages 220–226. ACM, 2011. 11

[142] J. Posluns and S. Sjouwerman. *Inside the SPAM Cartel: By Spammer-X*. Syngress, 1 edition, November 2004. 11, 16, 21

[143] K. Poulsen. Cracking down on cyberspace land grabs. `http://www.theregister.co.uk/2003/06/11/cracking_down_on_cyberspace_land/`, June 2003. 11

[144] J. Qiu and L. Gao. Detecting Bogus BGP Route Information: Going Beyond Prefix Hijacking. In *SecureComm*, pages 381–390. IEEE, 2007. 2, 3, 26, 27, 52, 119, 120, 124

[145] S. Qiu, F. Monrose, A. Terzis, and P. McDaniel. Efficient techniques for detecting false origin advertisements in inter-domain routing. In *Proceedings of the 2006 2nd IEEE Workshop on Secure Network Protocols*, pages 12–19, Washington, DC, USA, 2006. IEEE Computer Society. 31

[146] T. Qiu, L. Ji, D. Pei, J. Wang, and J. Xu. TowerDefense: Deployment strategies for battling against IP prefix hijacking. In *Network Protocols (ICNP), 2010 18th IEEE International Conference on*, pages 134–143, Oct 2010. 31

[147] T. Qiu, L. Ji, D. Pei, J. Wang, J. Xu, and H. Ballani. Locating Prefix Hijackers Using LOCK. In *Proceedings of the 18th Conference on USENIX Security Symposium*, pages 135–150, Berkeley, CA, USA, 2009. USENIX Association. 31

[148] A. Ramachandran and N. Feamster. Understanding the Network-Level Behavior of Spammers. In *SIGCOMM*, pages 291–302. ACM, 2006. 1, 4, 8, 10, 11, 21, 37, 47, 48, 58, 59, 61, 63, 65, 73, 74, 80, 86, 91, 98, 100, 109, 112, 118, 122, 123, 126, 128, 134, 135, 139, 143

[149] A. Ramachandran, N. Feamster, and S. Vempala. Filtering spam with behavioral blacklisting. In *CCS*, pages 342–351. ACM, 2007. 11

[150] Y. Rekhter, T. Li, and S. Hares. A Border Gateway Protocol 4. RFC4271, January 2006. 1, 14, 15, 36, 118

[151] Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot, and E. Lear. Address Allocation for Private Internets. RFC1918, February 1996. 13

[152] RIPE NCC. RPKI ROA certification statistics. `http://certification-stats.ripe.net/`. 25, 100, 143

[153] M. Roughan, W. Willinger, O. Maennel, D. Perouli, and R. Bush. 10 Lessons from 10 Years of Measuring and Modeling the Internet's Autonomous Systems. *IEEE Journal on Selected Areas in Communications*, 29(9):1810–1821, 2011. 36, 38, 40, 113, 149

[154] J. Schlamp, G. Carle, and E. W. Biersack. A forensic case study on AS hijacking: the attacker's perspective. *SIGCOMM CCR*, pages 5–12, 2013. 1, 9, 11, 77, 80, 114, 118, 123

[155] Y. Shavitt and E. Shir. DIMES: Let the Internet measure itself. *SIGCOMM CCR*, 35(5):71–74, 2005. 39

[156] Y. Shavitt and U. Weinsberg. Quantifying the Importance of Vantage Points Distribution in Internet Topology Measurements. In *INFOCOM*, pages 792–800. IEEE, 2009. 39, 113, 149

[157] X. Shi, Y. Xiang, Z. Wang, X. Yin, and J. Wu. Detecting prefix hijackings in the internet with argus. In *IMC*, pages 15–28. ACM, 2012. 1, 3, 28, 29, 36, 37, 38, 98, 107, 118, 120, 124, 140, 145

[158] G. Siganos and M. Faloutsos. Analyzing BGP policies: methodology and tool. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 3, pages 1640–1651 vol.3, March 2004. 28, 31

[159] G. Siganos and M. Faloutsos. Neighborhood watch for internet routing: Can we improve the robustness of internet routing today? In *INFOCOM*, pages 1271–1279. IEEE, 2007. 26, 27, 31, 63, 124, 129

[160] S. Sinha, M. Bailey, and F. Jahanian. Shades of grey: On the effectiveness of reputation-based blacklists. In *Malicious and Unwanted Software, 2008. MALWARE 2008. 3rd International Conference on*, pages 57–64, Oct 2008. 11

[161] S. Sinha, M. Bailey, and F. Jahanian. Improving Spam Blacklisting Through Dynamic Thresholding and Speculative Aggregation. In *NDSS*. The Internet Society, 2010. 11

[162] J. Snijders. PeeringDB Accuracy: Is blind faith reasonable? NANOG58, June 2013. 33

[163] K. Sriram, O. Borchert, O. Kim, P. Gleichmann, and D. Montgomery. A Comparative Analysis of BGP Anomaly Detection and Robustness Algorithms. In *Conference For Homeland Security, 2009. CATCH '09. Cybersecurity Applications Technology*, pages 25–38, March 2009. 28, 31

[164] R. Steenbergen. Examining the validity of IRR data. `https://www.nanog.org/meetings/nanog44/presentations/Tuesday/RAS_irrdata_N44.pdf`, October 2008. 28, 33

[165] L. Subramanian, S. Agarwal, J. Rexford, and R. H. Katz. Characterizing the Internet hierarchy from multiple vantage points. In *INFOCOM*, 2002. 2

[166] L. Subramanian, V. Roth, I. Stoica, S. Shenker, and R. H. Katz. Listen and whisper: security mechanisms for BGP. In *NSDI*, pages 10–10. USENIX Association, 2004. 24

[167] M. Tahara, N. Tateishi, T. Oimatsu, and S. Majima. A Method to Detect Prefix Hijacking by Using Ping Tests. In *APNOMS*, pages 390–398. Springer-Verlag, 2008. 2, 28, 31, 119

[168] O. Thonnard. *A multi-criteria clustering approach to support attack attribution in cyberspace.* PhD thesis, École Doctorale d'Informatique, Télécommunications et Électronique de Paris, March 2010. 63, 65, 94, 130, 131

[169] O. Thonnard, L. Bilge, G. O'Gorman, S. Kiernan, and M. Lee. Industrial espionage and targeted attacks: Understanding the characteristics of an escalating threat. In *RAID*, pages 64–85. Springer, 2012. 64, 130

[170] O. Thonnard and M. Dacier. A strategic analysis of spam botnets operations. In *CEAS*, pages 162–171. ACM, 2011. 64, 130

[171] A. Took. Using BGP data to find Spammers. `http://www.bgpmon.net/using-bgp-data-to-find-spammers/`, September 2014. 93, 99, 142

[172] V. Torra. The weighted OWA operator. *Int. Journal of Intelligent Systems*, 12(2):153–166, 1997. 55, 56, 128

[173] J. Touch, A. Mankin, and R. Bonica. The TCP Authentication Option. RFC5925, June 2010. 16

[174] I. van Beijnum. *BGP*. O'Reilly, 2002. 13, 14

[175] P.-A. Vervier. SpamTracer: Tracking Fly-By Spammers. RIPE67, October 2013. 67

[176] P.-A. Vervier. SpamTracer: Tracking Fly-By Spammers. NANOG60, February 2014. 67

[177] P.-A. Vervier, Q. Jacquemart, J. Schlamp, O. Thonnard, G. Carle, G. Urvoy-Keller, E. W. Biersack, and M. Dacier. Malicious BGP Hijacks: Appearances Can Be Deceiving. In *International Conference on Communications (ICC) Communications and Information Systems Security (CISS) Symposium*, pages 884–889, Sydney, June 2014. IEEE. 1, 58, 89, 114, 118, 122, 137, 149

[178] P.-A. Vervier and O. Thonnard. SpamTracer: How Stealthy Are Spammers? In *5th International Traffic Monitoring and Analysis (TMA) Workshop (INFOCOM workshops)*, pages 453–458, Turin, April 2013. IEEE. 1, 35, 118

[179] P.-A. Vervier, O. Thonnard, and M. Dacier. Mind Your Blocks: On the Stealthiness of Malicious BGP Hijacks. To appear in the Network and Distributed System Security (NDSS) Symposium. IEEE, 2015. 67

[180] C. Villamizar, R. Chandra, and R. Govindan. BGP Route Flap Damping. RFC2439, November 1998. 16, 73

[181] Q. Vohra and E. Chen. BGP Support for Four-Octet Autonomous System (AS) Number Space. RFC6793, December 2012. 12

[182] D. A. Wheeler and G. N. Larsen. Techniques for cyber attack attribution. Technical report, DTIC Document, 2003. 10

[183] R. White. Securing BGP Through Secure Origin BGP. *The Internet Protocol Journal*, Volume 6(Number 3), Sept. 2003. 22

[184] R. Yager. On ordered weighted averaging aggregation operators in multicriteria decision-making. *IEEE Trans. Syst. Man Cybern.*, 18(1):183–190, 1988. 55

[185] H. Yan, R. Oliveira, K. Burnett, D. Matthews, L. Zhang, and D. Massey. BGPmon: A Real-Time, Scalable, Extensible Monitoring System. In *CATCH*, pages 212–223, Washington, DC, USA, 2009. IEEE Computer Society. 32

[186] B. Zhang, R. Liu, D. Massey, and L. Zhang. Collecting the internet AS-level topology. *SIGCOMM CCR*, 35:53–61, January 2005. 37

[187] Y. Zhang, Z. Zhang, Z. M. Mao, C. Hu, and B. MacDowell Maggs. On the Impact of Route Monitor Selection. In *IMC*, pages 215–220. ACM, 2007. 37

[188] Z. Zhang, Y. Zhang, Y. C. Hu, and Z. M. Mao. Practical defenses against BGP prefix hijacking. In *Proceedings of the 2007 ACM CoNEXT conference*, page 3. ACM, 2007. 31

[189] Z. Zhang, Y. Zhang, Y. C. Hu, Z. M. Mao, and R. Bush. iSPY: Detecting Ip Prefix Hijacking on My Own. In *SIGCOMM*, pages 327–338. ACM, 2008. 2, 3, 28, 29, 38, 119, 120, 124

[190] J. Zhao and Y. Wen. Emulation on the Internet Prefix Hijacking Attack Impaction. In K. Mustofa, E. Neuhold, A. Tjoa, E. Weippl, and I. You, editors, *Information and Communication Technology*, volume 7804 of *Lecture Notes in Computer Science*, pages 485–489. Springer Berlin Heidelberg, 2013. 17

[191] J. Zhao and Y. Wen. Evaluation on the influence of internet prefix hijacking events. *Comput. Sci. Inf. Syst.*, 10(2):611–631, 2013. 17

[192] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, S. F. Wu, and L. Zhang. An analysis of BGP multiple origin AS (MOAS) conflicts. In *SIGCOMM Workshop on Internet Measurement (IMW)*, pages 31–35. ACM, 2001. 19, 26, 36, 51, 52

[193] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, S. F. Wu, and L. Zhang. Detection of invalid routing announcement in the internet. In *DSN*, pages 59–68. IEEE Computer Society, 2002. 31

[194] C. Zheng, L. Ji, D. Pei, J. Wang, and P. Francis. A light-weight distributed scheme for detecting IP prefix hijacks in real-time. In *SIGCOMM*, pages 277–288. ACM, 2007. 1, 2, 3, 28, 30, 36, 38, 118, 119, 120, 124