The Blind Man and the Elephant: Measuring Economic Impacts of DDoS Attacks

Abhishta Abhishta

This page is intentionally left blank.

THE BLIND MAN AND THE ELEPHANT: MEASURING ECONOMIC IMPACTS OF DDOS ATTACKS

DISSERTATION

to obtain
the degree of doctor at the Universiteit Twente,
on the authority of the rector magnificus,
prof. dr. T.T.M. Palstra,
on account of the decision of the graduation committee,
to be publicly defended
on Thursday 5 December 2019 at 14.45 uur

by

Abhishta Abhishta

born on 4 December 1991
in Meerut, India

This dissertation has been approved by:

supervisor
Prof. dr. ir. L.J.M. Nieuwenhuis
co-supervisor
Dr. R.A.M.G. Joosten

Type set with LaTeX. Printed by IPSKAMP printing.
Cover design: Design Crowd
ISBN: 978-90-365-4912-7
DOI: 10.3990/1.9789036549127

Graduation committee:

Chairman/secretary:	prof. dr. T.A.J. Toonen
Supervisor:	Prof. dr. ir. L.J.M. Nieuwenhuis
Co-supervisor:	Dr. R.A.M.G. Joosten

Committee Members:
Prof. dr. ir. D. Hiemstra, Radboud University, Nijmegen, The Netherlands
Prof. dr. J. van Hillegersberg, University of Twente, The Netherlands
Prof. dr. M. Junger, University of Twente, The Netherlands
Dr. M Korczyński, Grenoble Institute of Technology, Grenoble, France
Prof. dr. ir. A. Pras, University of Twente, The Netherlands
Dr. A. Sperotto, University of Twente, The Netherlands

Funding source:

D3 - Distributed Denial-of-Service Defense – NWO project (628.001.018)

Contents

List of Tables

List of Figures

Acknowledgements

Even for a moment I don't fool myself into believing that this work would have been possible without the help of many wonderful people who have influenced me over the years. That being said this is also supposed to be the most read portion of my thesis. So, if you are reading this, I would like to inform you that this thesis is about measuring the economic impact of DDoS attacks and now I will present the findings of my thesis (just kidding!). I hope I am able to thank everyone who has been part of this journey in the next few lines. If I forget someone I am truly sorry.

First, I would like to thank my promoter, Bart and daily supervisor, Reinoud for all support they provided me during this period. Thank you for dealing with my last-minute change of ideas and annoying stories about food. You both helped me with several decisions that were not only related to work but also to personal life. I cannot express my gratitude in words. I don't think I could have asked for better supervisors. Bart, thank you for giving me this opportunity and teaching me how to deal with volatile situations. Reinoud, thank you for all the weekly discussions and pushing me (figuratively) when I needed some extra incentive to write things down. I will always be grateful to both of you for giving me this opportunity and providing necessary foundation for this work. Thank you!

I am thankful to my paranymphs, Wouter and Mattijs. Thank you for being part of this day! Wouter, a special thanks to you for being my longest standing coffee buddy. I think without your help I would have been dependent on Google for all my translations of Dutch documents. Over the last four and a half years you have been a cherished friend. Mattijs, it was a privilege to work besides you on the D3 project. Without your help it would have been difficult for me to position my work for the measurements community. I am really happy

that we made the trip to Budapest together, otherwise I would have missed the opportunity of knowing you.

My Ph.D. life was spent in University of Twente but its foundations were laid by all teachers that supported my enthusiasm during my school, under-graduate and post-graduate life. I wanted to become an academician, I cannot be grateful enough to Prof. Padmakumar Nair for mentoring me during a critical period of my life for making this possible for me. I was fortunate to have great teachers and I owe a lot to them for keeping me motivated as a youngster.

I would like to express my gratitude towards my colleagues at DACS whose inputs were crucial for my research. Aiko, I think your comments during my qualifier have a lot to do with how this thesis turned out, thank you. Anna, the list of things I have to thank you for grows each time we meet. Thank you for all your support during my Ph.D. Roland, a special thanks to you for being forthcoming to talk about topics. Each time we talk, I learn something new. I am happy to find two friends (Menno and you) with whom I can share stories about good food. Jair, thank you for all the coffee walks and our discussions about life. I take this opportunity to thank SURFnet (Xander and Bart), without whose help Chapter 8 of this thesis would not have been possible.

Table tennis (Thibats) has been a big component of my life in Enschede. Wesley, thank you so much for introducing me to the club and being a great friend over the years. I would also like to thank my doubles partners Tiago, Niels and Annelies for having faith in my table tennis skills. I would not be able to name all, but I am truly grateful to all of Thibats for providing me with a perfect atmosphere for learning the sport and finding timeless friends.

I am thankful to my graduation committee members for agreeing to be part of the ceremony. Jos, thank you for always being up for a quick chat in the hallway and giving me opportunities to grow as a teacher during my Ph.D. Marianne, none of the components related to criminology would have been possible without you. Thank you for all your support during my doctoral period. Maciej, thank you for all the interesting conversations when I travelled for WTMC. I look forward to our collaborations in the near future. Djoerd, if you wouldn't have organised the data science workshop, Chapter 6 of this thesis would not have been possible. Thank you for finding time to discuss the data collection when I needed your help.

A big thank you to all my colleagues at IEBIS. Berend, thank you for all the discussions we had related to finance and philosophy. Nils, thank you being a wonderful friend during this journey. Laura and you made my Ph.D. life a lot more exciting. Luca, thank you for your friendship and all the car rides that we shared. Hope there are a lot more to come. I would like to thank my office mates

Andreij, Arturo, Gréanne, Guido (for teaching me how to brew beer), Lucas, Sajjad, Sina, Sjoerd (for promoting LaTeX), Vahid and Wenyi for all the coffee breaks and moments that we shared. Elke and Hilde, thank you for continuing to make my life easy and all the chats that we had over the years.

I first came to Enschede in 2014 for my masters thesis. Raja, thank you for introducing Enschede to me in 2014 and for being a guide to me around Europe. Michel, thank you for guiding me at various moments before and during my Ph.D. Vishal, without mentioning you I don't think this section would be complete. I found a great friend in you. Thank you for your advice and giving me a patient ear whenever I needed it. You were a big support for me during my life in Enschede. Thank you for sharing your taste of music and comedians with me. It has left an impact on my life. Akshita, Kriti, Monika, Vishal (Ahuja) and Sahana thank you for filling the first year of my Ph.D. with activities and all the after-work discussions. Deepak and Sugandh thanks for all the celebrations you made me a part of. I hope there are more to come. I also express my gratitude to all my friends and family in India, Aarsheya, Abhinav, Ankush, Anupam, Chayank, Garima, Innayat, Kartik, Konark, Pavleen, Samarth and Utsav for understanding why I couldn't be part of some of the important events in their lives. I will try and make up for it. Zinzy, thank you for all the positivity that you bring in with you. Everyone needs a friend like you.

Letizia, it is difficult for me to write something about you here, not because I have nothing to say but because I have enough for a book in itself. Everyone who knows me understands how important you are for every step I take. There are probably a few truly selfless people I have met in my life. Would it be selfish of me to say that I will be married to the best of them all? I am thankful to Arturo and Maria Luisa for all their support during the last few years. Thank you for making me part of your family. This is probably the toughest part for me: How can I thank my own parents? Hmm... a step in this direction is *dedication of this work to you.* I am happy that you made me independent enough to build and maintain a life far away from home. Mom, I know I am not the most thankful son on the planet, but I will always be thankful that you understand and don't give up. Dad, I chose the same profession as you, hopefully I am able to influence as many lives as you did. Thank you!

Abstract

Internet has become an important part of our everyday life. We use services like Netflix, Skype, online banking and Scopus etc. daily. We even use Internet for filing our tax returns and communicating with municipalities. This dependency on network-based technologies provides an opportunity to malicious actors in our society to remotely attack IT infrastructure. One type of cyberattack that may lead to unavailability of network resources is known as distributed denial of service (DDoS) attack. A DDoS attack leverages many computers to launch a coordinated Denial of Service attack against one or more targets.

These attacks cause damages to victim businesses. According to reports published by several consultancies and security companies these attacks lead to millions of dollars in losses every year. One might ponder: are the damages caused by temporary unavailability of network services really this large? One of the points of criticism for these reports has been that they often base their findings on victim surveys and expert opinions. Now, as cost accounting/book keeping methods are not focused on measuring the impact of cyber security incidents, it is highly likely that surveys are unable to capture the true impact of an attack. A troubling fact is that most C-level managers make budgetary decisions for security based on the losses reported in these surveys. Several inputs for security investment decision models such as *return on security investment* (ROSI) also depend on these figures. This makes the situation very similar to the parable of the *blind men and the elephant*, in which several blind men try to conceptualise how the elephant looks like by touching it. Hence, it is important to develop methodologies that capture the true impact of DDoS attacks. In this thesis, we study the economic impact of DDoS attacks on public/private organisations by using an empirical approach.

In Chapter 1 we explain the motivation for our work and illustrate the problems associated with measuring the economic impacts of DDoS attacks. We then formulate our main research question and break it down into sub-questions that we investigate in later chapters. We state our main research question as follows:

> *What are the economic impacts of DDoS attacks on public/private organisations?*

Our first contribution is *identifying the main stakeholders* in a DDoS attack. In Chapter 2, we discuss the evolution of DDoS attacks in the last decade and briefly describe the strategies adopted by attackers and defenders. By studying the business model of a botnet, we also analyse how DDoS attacks can be used by attackers for monetary gains.

Our second contribution is to *develop methodologies to capture the direct impact of DDoS attacks.* In Chapters 3 and 4 we measure the direct consequences of DDoS attacks on large managed domain name service (DNS) providers and a cryptocurrency exchange respectively. We find that a successful DDoS attack on a managed DNS service provider, changes the security behaviour of its customers. In the case of cryptocurrency exchange we find that the losses are recovered very quickly, on most instances even within a single day. We show how longitudinal datasets can be used to asses the impacts.

The third contribution of this thesis is to *develop methodologies to measure the indirect consequences of DDoS attacks.* In Chapter 5, we propose a more robust event study approach and use it to analyse the impact of DDoS attack announcements on victims' stock prices. We find that in most cases this impact is short lived (5-10 days). In Chapter 6, we introduce a dataset based on web articles on DDoS attacks which captures the social context of an attack. We show how machine learning algorithms can be used to filter news articles that are reporting a DDoS attack from the dataset.

We recognise that it is not possible to measure the true impact of DDoS attacks on the victim without learning about the aims of attackers. In Chapter 7, we propose a model based on Routine Activity Theory (RAT) *to study attacker's aims* by using the information about the attack reported in the news articles. Later in Chapter 8, we show *how postulates of RAT may be used to explain DDoS attack trends* on educational institutions.

Our results show that DDoS attacks are not a random phenomenon and attackers are instigated by the circumstances surrounding them. We observe that measuring the true economic impact of these attacks is complex and requires us to consider the context of an attack. Some of the consequences of short duration IT unavailability are temporary and they are recovered rather quickly. Hence,

to take this work forward we propose to give economic meaning to the empirical data that is presently available and collect more data at employee level to measure the resilience of firms towards IT unavailability.

Samenvatting

Internet is een belangrijk onderdeel van ons dagelijks leven geworden. We maken dagelijks gebruik van diensten zoals Netflix, Skype, online bankieren, en Scopus etc. We gebruiken internet zelfs voor het indienen van onze belastingaangiftes en het communiceren met de gemeente. Deze afhankelijkheid van netwerkgebaseerde technologieën biedt kwaadwillende agenten in onze samenleving de mogelijkheid om op afstand een IT-infrastructuur aan te vallen. Een cyberaanval, die kan leiden tot onbeschikbaarheid van netwerkbronnen, staat bekend als Distributed Denial of Service-aanval (DDoS). Een DDoS-aanval maakt gebruik van een groot aantal computers om een gecoördineerde Denial of Service-aanval tegen een of meer doelen te starten.

Deze aanvallen veroorzaken schade aan bedrijven die slachtoffer zijn. Volgens rapporten van verschillende adviesbureaus en beveiligingsbedrijven leiden deze aanvallen elk jaar tot miljoenen dollars aan verliezen. Je zou kunnen denken: is de schade veroorzaakt door tijdelijke onbeschikbaarheid van netwerkdiensten echt zo groot? Een van de kritiekpunten aangaande deze rapporten is dat ze hun bevindingen vaak baseren op enquêtes onder slachtoffers en meningen van deskundigen. Aangezien kostenberekening / boekhoudmethoden niet zijn gericht op het meten van de impact van cyberveiligheidsincidenten, is het zeer waarschijnlijk dat enquêtes niet in staat zijn om de ware impact van een aanval vast te leggen. Een zorgwekkend feit is dat de meeste top-level managers budgettaire beslissingen voor beveiligingsmaatregelen nemen op basis van de verliezen die in deze enquêtes worden gerapporteerd. Verschillende variabelen voor beslissingsmodellen voor beveiligingsinvesteringen, zoals *rendement op beveiligingsinvesteringen* (ROSI), zijn ook afhankelijk van deze cijfers. Dit maakt de situatie erg vergelijkbaar met de parabel van de *blinde mannen en de olifant*, waarin blinde mannen proberen te bedenken hoe de olifant eruit ziet door hem

aan te raken. Daarom is het belangrijk om methodologieën te ontwikkelen die de ware impact van DDoS-aanvallen vastleggen. In dit proefschrift bestuderen we de economische impact van DDoS-aanvallen op publieke / private organisaties met behulp van een empirische aanpak.

In Hoofdstuk 1 lichten we de motivatie voor ons werk toe en illustreren we de problemen bij het meten van de economische impact van DDoS-aanvallen. Vervolgens formuleren we onze belangrijkste onderzoeksvraag en splitsen deze op in deelvragen die we in latere hoofdstukken onderzoeken. We folmuleren onze hoofdvraag als volgt:

> *Wat zijn de economische implicaties van DDoS-aanvallen op publieke / private organisaties?*

Onze eerste bijdrage is *de identificatie van de belangrijkste belanghebbenden* in een DDoS-aanval. In Hoofdstuk 2 bespreken we de evolutie van DDoS-aanvallen in het afgelopen decennium en beschrijven we kort de strategieën die aanvallers en verdedigers volgen. Door het bedrijfsmodel van een botnet te bestuderen, analyseren we ook hoe DDoS-aanvallen door aanvallers kunnen worden gebruikt voor geldwinsten.

Onze tweede bijdrage is het *ontwikkelen van methodologieën om de directe impact van DDoS-aanvallen vast te leggen.* In Hoofdstuk 3 en 4 meten we de directe gevolgen van DDoS-aanvallen op respectievelijk grote beheerde domain name service (DNS) providers en een cryptocurrency-uitwisseling. We zien dat een succesvolle DDoS-aanval op een beheerde DNS-serviceprovider het beveiligingsgedrag van zijn klanten verandert. In het geval van cryptocurrency-uitwisseling zien we dat de verliezen zeer snel worden teniet gedaan, in de meeste gevallen zelfs binnen een enkele dag. We laten zien hoe longitudinale datasets kunnen worden gebruikt om de impact te beoordelen.

De derde bijdrage van dit proefschrift is *methodologieën te ontwikkelen teneinde de indirecte gevolgen van DDoS-aanvallen te meten.* In Hoofdstuk 5 stellen we een robuustere benadering van zogenaamde gebeurtenisstudies (event studies) voor en gebruiken deze om de impact van aankondigingen van DDoS-aanvallen op de aandelenkoersen van het slachtoffer te analyseren. We merken dat deze impact in de meeste gevallen van korte duur is (5-10 dagen). In Hoofdstuk 6 introduceren we een dataset op basis van webartikelen over DDoS-aanvallen die de sociale context van een aanval weergeeft. We laten zien hoe machine learning-algoritmen kunnen worden gebruikt om nieuwsartikelen die DDoS-aanvallen rapporteren uit de dataset te filteren.

We stellen dat het niet mogelijk is om de ware impact van DDoS-aanvallen op het slachtoffer te meten zonder de doelen van aanvallers te kennen. In Hoofd-

stuk 7 stellen we een model voor op basis van Routine-ActiviteitsTheorie (RAT) *teneinde de doelen van de aanvaller te bestuderen* met behulp van de informatie over de aanval die in de nieuwsartikelen wordt gerapporteerd. Later in Hoofdstuk 8 laten we *zien hoe postulaten van RAT kunnen worden gebruikt om DDoS-aanvalstrends* op onderwijsinstellingen te verklaren.

Onze resultaten laten zien dat DDoS-aanvallen geen willekeurig verschijnsel zijn en aanvallers worden gemotiveerd door externe omstandigheden. We stellen vast dat het meten van de werkelijke economische impact van deze aanvallen complex is en dat we de context van een aanval moeten meenemen. Sommige van de gevolgen van korte onbeschikbaarheid van IT zijn tijdelijk en worden vrij snel teniet gedaan. Daarom adviseren we om de empirische gegevens die momenteel beschikbaar zijn economische inhoud te belang te geven, en meer gegevens op werknemersniveau te verzamelen om de weerbaarheid van bedrijven tegen IT-onbeschikbaarheid te meten.

This page is intentionally left blank.

Chapter 1

Introduction

We introduce the topic and motivation of this Ph.D. thesis. We describe the main research question and formulate the sub-questions. We also describe the research methodology used to answer the research questions. We end the chapter by giving an overview of this thesis and listing the main contributions of each chapter.

Many believe that Internet is going to be one of the basic needs for homo sapiens just like food, clothing and shelter. Since the implementation of the world wide web on the 6^{th} of August 1991, internet has increasingly become part of our everyday life. We use the services based on it for communication, research, financial transactions, entertainment etc. Information and communication technology (ICT) has helped organisations belonging to all possible sectors in improving efficiency and achieving economies of scale [171]. The use of ICT has not only provided economic benefits to businesses, but also better and more customised facilities to their customers. Today we can buy gadgets (e.g., Google Home) that can identify the owner by his/her voice, and perform a given task as efficiently as any human. Students around the world can learn from the best teachers and even surgeons can perform operations remotely, all thanks to quick and reliable internet based technologies and services.

1.1 Cyber attacks

The discussion above shows that we have become highly dependent on network based technologies in today's world. This however also gives an opportunity to nefarious actors in the society to plan malicious activities using the Internet. These actors have an opportunity to attack IT infrastructure remotely. These

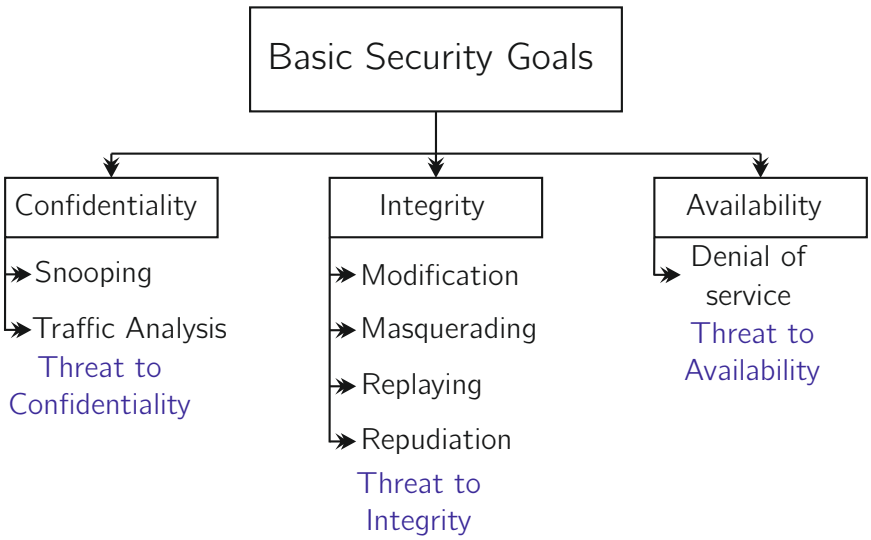


Figure 1.1: Taxonomy of cyber attacks categorised by basic security goals [71].

attacks that intend to damage or destroy a computer network/system are known as cyber attacks. They can be broadly classified with respect to the threat posed by them into three categories:

- Attacks that are a threat to *confidentiality*.
- Attacks that are a threat to *integrity*.
- Attacks that are a threat to *availability*.

Figure 1.1 shows a taxonomy of illicit actions classified according to the type of threat posed by them. The first category of attacks target the confidentiality of digitally stored data. With the help of malicious software (malware), actors can infiltrate the IT infrastructure of a company or an individual. This can provide them with the access to confidential information, which they can use to derive economic/non-economic gains. For example, in August 2015 user data of *Ashley Madison* (a commercial website known for enabling extra-marital affairs) were leaked [34]. Taking into account the business model of Ashley Madison, the confidentiality of user information was critical to its clients and data leakage led to public shaming of the clients.

The next category of attacks pose a threat to the assurance of the accuracy and consistency of data. On some occasions, nefarious actors can also make

use of vulnerabilities in the infrastructure to manipulate critical information. A well known example of an integrity attack from the past is the use of the Stuxnet worm to make changes to Iran's nuclear reactor settings in an attempt to destroy it [112].

The third category of attacks are aimed at making the infrastructures connected to the network unavailable for intended users. In a recent attack event, Github was targeted with a 1.35 terabits per second DDoS attack which led to short unavailability of its services [144].

1.2 Economic impact of cyber attacks and cyber security investments

These cyber attacks can impose a heavy cost on the victims. Organisations can suffer damages due to loss of productivity or bad publicity and can also be forced to pay reparations if attacks lead to violation of service level agreements (SLAs). Anderson et al. [13] provides a framework for measuring these costs. They decompose these costs into three components: 1) defence costs 2) direct losses and 3) indirect losses. The first component measures the amount of money already invested by the company to defend itself against cyber attacks. The second component takes into account monetary losses as a consequence of an attack and other immediate damages such as loss of intellectual property, distress suffered by victims etc. The last component of the framework considers indirect consequences such as loss of trust among customers, missed business opportunities etc. The reported damages due to cyber attacks runs in millions of dollars per company per year. In 2017, a study by Accenture estimated the average cost of malware attacks to be $ 2.4 million [2]. In 2018, a study by Ponemon Institute estimated the annual cost of data breaches at $ 3.9 million [156].

In order to protect against these attacks, organisations need to invest in cyber security. Security investments, unlike other investments such as buildings and machines, do not generate monetary returns [53]. Instead, their benefits are a result of cost savings by preventing or reducing the damage due to security breaches. Just like all other investments, cyber security investments should be managed by analysing the cost-benefit trade-offs. Several models used for supporting decisions with regards to these investments take into account financial measures based on the impact of past attacks. Gordon and Loeb [77] and Huang, Hu and Behara [92] suggest models based on expected losses, threat and vulnerability to calculate optimal investment. Butler [35] proposes a comparative approach known as Security Attribute Evaluation Method (SAEM), which

is a stepwise quantitative cost-benefit analysis for security investment decisions. Several researchers have suggested models following the return on investment (ROI) approach e.g., [48, 187, 93]. While some have also made use of other multi-criteria decision making approaches such as the analytic hierarchy process (AHP), value at risk (VaR) and balanced scorecard [26, 205, 194]. All methods mentioned here for evaluating security investments depend on the financial value of damages in case an asset is breached. Hence, for reaping the maximum benefits from security investments (by investing optimally), it is imperative that we have reliable methods for measuring the economic damage.

1.3 Problems with measuring economic impact

Measuring the economic impact of cyber attacks is challenging. Publicly available empirical data for calculating the damages are scarce due to the lack of willingness of organisations to share information [39]. Cashell, Jackson, Jicklin and Webel [39] even suggests that there are strong incentives for companies that discourage sharing of information. They argue that there can be high costs of public disclosure for organisations that choose to share information on security events. Hence, very few studies have been successful in empirically evaluating damages due to cyber attacks. Most of these studies have analysed the impact of cyber security breaches on stock prices of publicly traded companies [38, 40, 78, 64].

The studies that do report the economic damage done by cyber attacks do so on the basis of surveys [2, 3, 155, 99, 156, 51]. As shown in Section 1.2, these studies report the damages done by cyber attacks in millions. But are the damages due to cyber attacks really as high as reported by these studies? Here are a few reasons, why the numbers reported might be inflated:

- Cost accounting/book keeping methods used by companies are not focussed on measuring the impact of cyber security incidents, organisations are often unable to quantify the risks of cyber attacks [39]. Hence, it is almost impossible for survey takers to answer with numbers that capture the true impact of an attack.

- The estimated damages reported in surveys are based on inaccurate guesstimates of security experts [59]. Another problem with the losses reported by these studies is that the majority of these studies calculate average losses based on inputs provided by large companies. Florêncio and Herley [70] find evidence that most of these surveys are dominated by a minority of responses in the upper tail leading to over/estimation of losses.

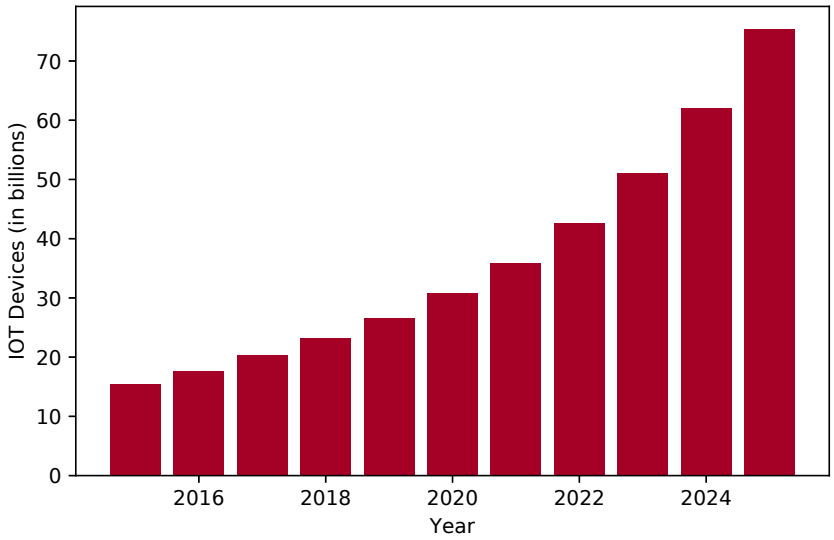


Figure 1.2: Expected increase in the number of IoT devices between 2015-2025 [97].

- Most of theses studies are organised by cyber security companies who have a clear incentive to inflate the losses due to cyber attacks.

In this thesis, we focus on analysing the direct and indirect damage caused by one particular attack known as distributed denial of service (DDoS) attack. Akamai (an organisation that provides DDoS protection services) based on a survey conducted by Ponemon Institute estimates the financial damage caused by DDoS attacks at $ 1.7 million per year per organisation [99]. The already discussed limitations of studies that use surveys as an instrument to measure economic impact are also applicable to this report. Also, unlike most other cyber attacks, DDoS attacks in isolation only affect the availability of network based services and do not lead to loss of intellectual property [57]. However, these days with the advent of internet of things (IoT) devices we can remotely control manufacturing equipments, household gadgets etc., only if network resources are available. Looking at the estimated growth in the number of IoT devices (Figure 1.2), it is clear that it is only matter of time before DDoS attacks lead to substantial financial damages to individuals as well.

For correctly estimating the losses due to DDoS attacks we need to consider the circumstances that form the setting of an attack. In other words we need

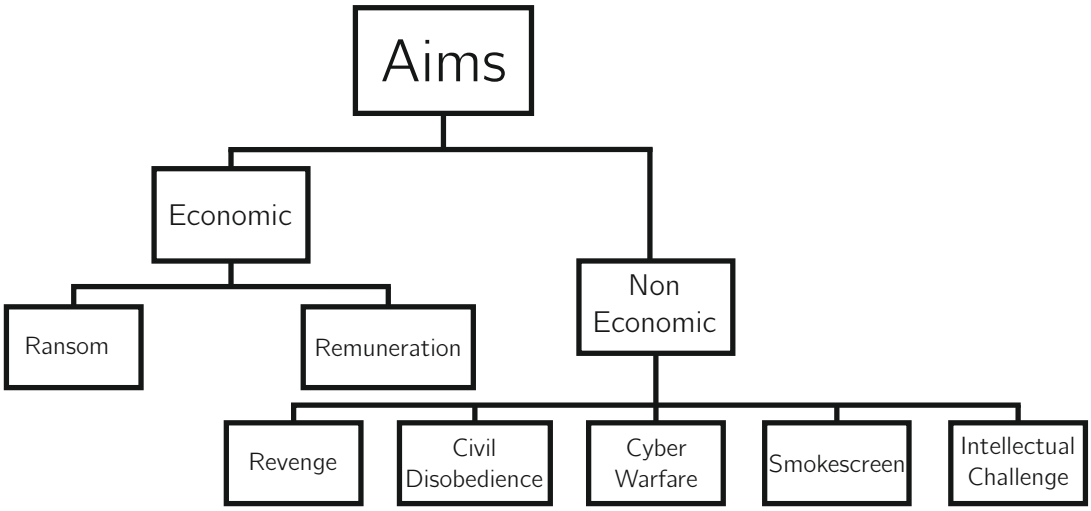


Figure 1.3: Typology of *aims* for attackers to use DDoS attacks.

to analyse the attacks while taking into account their *context*. A DDoS attack on a static website (website providing general information) of a multi-national company leads to smaller losses compared to an attack on the gaming servers of an online gaming company. We also need to examine the fact that some of the damages caused by temporary unavailability of services are reversible (i.e., the losses are recovered within a few days of an attack). We can investigate the true damage only by using an empirical approach to measure the economic impact of DDoS attacks.

1.4 Attacker aims

The way people think and behave is just as important to study as the malicious code used to exploit vulnerabilities in technology [115]. The aim of an attacker influences the amount and nature of the damage they are hoping to inflict on the victim. Attackers are not always looking for economic damages. Sauter [180] discussed the role of DDoS attacks in portraying civil disobedience. When attackers wish to portray civil disobedience their primary aim is to get the attention of concerned authorities (e.g., governments). In Figure 1.3 we show a typology of aims for attackers to use DDoS attacks. We classify the aims as economic (i.e., when the primary aim of an attacker is to inflict financial damages) and non-economic. Economic aims include ransom and remuneration. In 2015, a cyber criminal group called Armada Collective launched a DDoS based

ransom campaign known as DDoS for bitcoin (DD4BC), where the clear aim of the group was to collect ransom. At the same time Booters form a compelling case for when attackers launch DDoS attacks and get paid for it [177, 94]. Non-economic aims include using DDoS for revenge, portraying civil disobedience, cyber warfare, smokescreen and intellectual challenge. In case of non-economic aims the primary goal is not to inflict economic damage. According to a report by Kaspersky, several businesses believe that DDoS is being used as a revenge tactic [95].

Several studies in the field of classical criminology have analysed the aims of perpetrators [33, 186]. These studies have resulted in a better understanding of attacker behaviour and have helped law enforcement in making educated strategies [45]. In order to prepare ourselves for future attacks we need to improve our understanding of attackers that make use of DDoS attacks. A step in this direction can be to use theories from classical criminology to understand DDoS attacker behaviour.

1.5 Research question

In the previous sections of this chapter we discussed the difficulties in empirically measuring the economic impact of cyber attacks. In this thesis we focus on one attack in particular i.e., distributed denial of service (DDoS) attack. Consequently, the main research question investigated in this thesis is as follows:

> ***Research Question:*** *What are the economic impacts of DDoS attacks on public/private organisations?*

To answer the main research question, we divide it into five sub-questions. The first three sub-questions are related to empirically measuring the economic impact of DDoS attacks and they are answered in Part 1 of this thesis. In Section 1.4 we argued that the aims of attackers are not always economic. It is important for us to understand the aims of attackers as many a times they might be looking for attention of specific stakeholders rather than causing huge damages to the victim (e.g., in case of an act of civil disobedience). Hence, the topic of economic impact of DDoS attacks cannot be addressed without evaluating attacker aims. The last two sub-questions deal with analysing aims of attackers, and they are answered in Part 2 of this thesis.

1.6 Sub-questions and approach

In this section, we formulate sub-questions that will in turn help us to answer the main research question. We also provide an overview of the approach used to answer these questions.

1.6.1 Part I: Economic impact of DDoS attacks

DDoS attacks not only have consequences for the victim organisation but also for other stakeholders involved in an attack. To accurately measure the economic impact of DDoS attacks on the victim, we need to consider the role of the major parties involved. A few studies have analysed the consequences of DDoS attacks on one of the stakeholders [16, 40]. However, a holistic view of all the agents involved in an attack is often absent. Hence, in the first part of this thesis we focus on identifying the major stakeholders of a DDoS attack and then measuring the economic impact of DDoS attacks on them. Thus, our first sub-question is about identifying the stakeholders:

> ***SQ 1:*** *Who are the major stakeholders in a DDoS attack? How are they affected by a DDoS attack?*

We address *SQ 1* in the Chapter 2 of this thesis.

Once we have identified the major stakeholders in a DDoS attack and studied how they are affected, we proceed towards measuring the consequences. Based on the framework provided by Anderson et al.[13], we divide them into *direct* and *indirect* consequences. Therefore, the second sub-question is as follows:

> ***SQ 2:*** *How can we measure the direct consequences of a DDoS attack?*

We provide answers to *SQ 2* in Chapters 3 and 4 of this thesis.

Then, we want to evaluate the indirect consequences of a DDoS attack on an organisation. Our third sub-question is as follows:

> ***SQ 3:*** *How can we measure the indirect consequences of a DDoS attack?*

We answers *SQ 3* in Chapters 5 and 6 of this thesis.

1.6.2 Part II: Attacker aims

Attacker aims may influence the amount of damages that the attacker is hoping to inflict on the victim. The second part of this thesis deals with studying the aims of attackers with the help of classical theories in criminology and to evaluate whether their postulates can be used to explain DDoS attack trends. Our fourth sub-question is as follows:

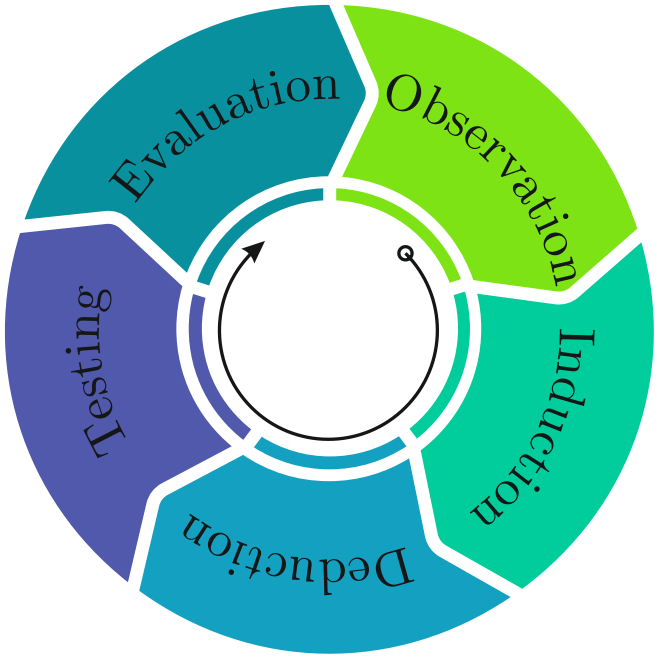


Figure 1.4: Groot's cycle of empirical scientific enquiry[81].

> ***SQ 4:*** *What are the various aims of attackers to use DDoS attacks? How can classical theories in criminology be used to explain the aims of attackers?*

We answer *SQ 4* in Chapter 7 of this thesis.

After explaining how classical theories in criminology are able to justify the aims of attackers to target an organisation's IT infrastructure with a DDoS attack, our final sub-question deals with validating this theory with the help of data collected in the real world. We would like to find out if their postulates can be used to explain DDoS attack trends on an organisation. Hence, our fifth sub-question is as follows:

> ***SQ 5:*** *How can we use the postulates of classical theories in criminology to explain DDoS attack trends?*

We answer *SQ 5* in Chapter 8 of this thesis.

1.6.3 Approach

To address the research questions that will be answered in this thesis, we make use of an empirical approach suggested by Groot [81]. Groot proposed the cycle

of empirical scientific inquiry that serves as a basic construct in our logico-methodological approach and is shown in Figure 1.4. This cycle has the following five phases:

Phase 1: This is the *observation* phase. It involves the collection and grouping of empirical materials and (tentative) formation of hypotheses.

Phase 2: This is the *induction* phase. In this phase one formulates hypotheses.

Phase 3: It is the *deduction* phase. Here one derives specific consequences from hypotheses, in the form of testable predictions.

Phase 4: The *testing* phase; here one tests the hypotheses against new empirical materials, by way of checking whether or not the predictions are fulfilled.

Phase 5: This is the *evaluation* phase. Now one evaluates the outcome of the testing procedure with respect to the hypotheses or theories stated, as well as with a view to subsequent, continued or related investigations.

For examining each of the research questions, we base our observations on previously established theories in finance and criminology. Considering these observations we formulate our hypotheses and deduce statistically testable hypotheses. We make use of a variety of datasets from different vantage points to test our hypotheses and evaluate results. In this thesis, we also introduce a novel dataset that can be helpful in collecting contextual information regarding DDoS attacks. We utilise the content change detection and notification service, called *Google Alerts*, i.e., provided by Google in order to collect this dataset. Such a dataset can be very helpful for researchers to track online news articles related to an attack. Not only do these articles provide technical insights on the methods used by attackers but also provide information on the socio-cultural, political and economic circumstances of the victim firm at the time of an attack. We explain in detail the collection methodology and show two case studies based on the dataset in Chapter 6. Furthermore, if necessary, we develop methods to analyse the data and validate them using the available datasets. To enable reproducibility as well as future research we release the data we collect (if not restricted by confidentiality clause) for use by other researchers.

1.7 Thesis organisation

Figure 1.5 shows how this thesis is organised. The figure shows the relationship between the chapters and serves as a map for readers. Below we provide a brief

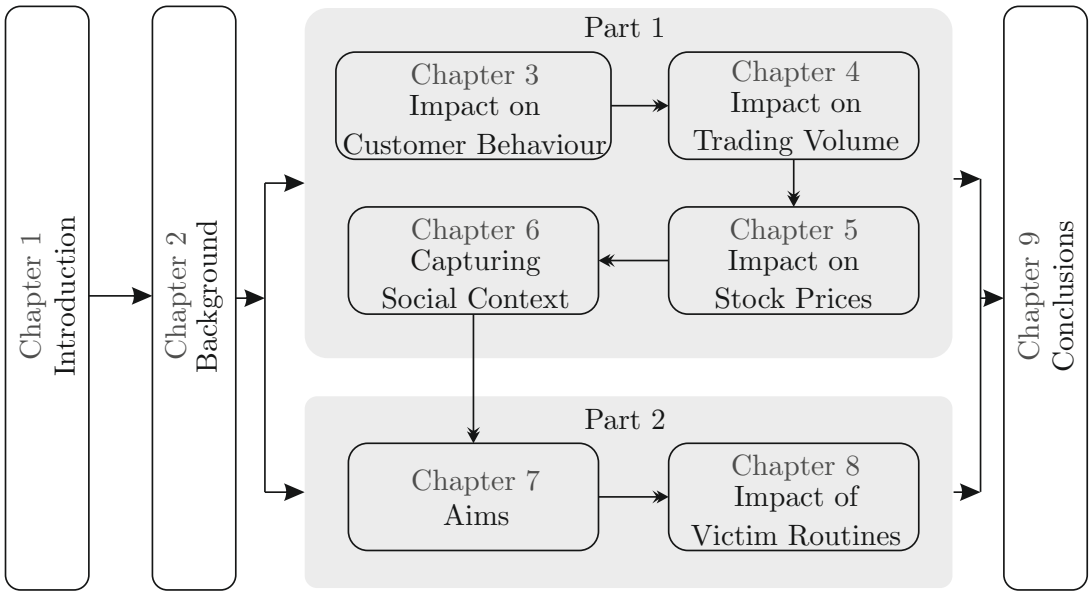


Figure 1.5: Organisation of this thesis.

summary of each chapter and provide references for the publications on which the chapter is based.

Chapter 2: Background

In this chapter, we introduce DDoS attacks and explain the evolution of these attacks with respect to strategies used and vulnerabilities exploited by attackers. We also track the increase in maximum attack intensity over the years and explain the mitigation strategies used by organisations. After gathering all the information needed, we answer the first sub-question. We identify the main stakeholders of a DDoS attack. Thereafter, we present the business model of a botnet using a business model canvas and explain the framework proposed by Anderson et al. [13] for measuring the cost of cybercrime. We end the chapter by explaining the usefulness of routine activity theory (RAT) in analysing the aims of attackers.

Parts of this chapter are based on the following peer-reviewed publication:

- C. Putman, Abhishta and L. J. Nieuwenhuis. 'Business Model of a Botnet'. *2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*. IEEE. 2018, pp. 441–445 [160].

Part I

The following four chapters form Part I of this thesis and focus on developing and validating methods for measuring the direct (Ch. 3 & 4) and indirect (Ch. 5 & 6) consequences of DDoS attacks.

Chapter 3: Impact on Customer Behaviour

Here, we develop and validate a method for analysing one of the direct consequences of DDoS attacks, i.e., loss of customers. The Domain Name System (DNS) is one of the core services that forms a crucial factor in successful delivery of internet services. Because of the importance of DNS, specialist service providers have come up in the market , that provide managed DNS services. One of their key selling points is that they protect DNS for a domain against DDoS attacks. We analyse two major DDoS attack events on managed DNS (MDNS) service providers (NS1 and Dyn). For our analysis we leverage data from OpenINTEL active DNS measurement system, which covers large parts of the global DNS over time [166]. The main contributions of this chapter are as follows:

- We develop a framework for measuring the behaviour of domains that use a MDNS service provider.
- We use this framework to analyse the impact of two DDoS attack events on the victims.
- We observe statistically significant changes in customer behaviour after the attacks (e.g., addition of a second DNS provider for a domain).
- Our results show that, even though it leads to higher costs, using a second DNS/MDNS provider is a good strategy to guarantee availability at all times.

This chapter is based on the following peer-reviewed publications:

- Abhishta, R. van Rijswijk-Deij and L. Nieuwenhuis. 'Measuring the Impact of a Successful DDoS Attack on the Customer Behaviour of Managed DNS Service Providers'. *WTMC '18*. ACM Press, 2018, pp. 1–7 [11].
- A. Abhishta, R. van Rijswijk-Deij and L. Nieuwenhuis. 'Measuring the Impact of a Successful DDoS Attack on the Customer Behaviour of Managed DNS Service Providers'. *Computer Communication Review* 48.5, 2018, pp. 70–76 [7].

Chapter 4: Impact on Trading Volume

This chapter focusses on analysing another direct consequence of a DDoS attack. DDoS attacks have become an effective tool to target the availability of any online platform. As a consequence, these businesses may lose sales volume during the attack period. We analyse the impact of DDoS attacks on the trading volume of a major cryptocurrency exchange. In order to do so we use an event analysis methodology to analyse the daily volume traded on the exchange on attack days. The key contributions of this chapter are as follows:

- We utilise a few concepts of behavioural finance. We develop an estimation model to predict the volume of cryptocurrency traded on the basis of price change.
- We modify the event analysis methodology to measure the impact of DDoS attacks on the volume traded on a major cryptocurrency exchange.
- We show that on most occasions the negative impact of a DDoS attack was recovered on the same day.
- Finally, with the help of hourly trading volumes we discuss the cause for delayed recovery by the exchange in 4 cases.

This chapter is based on the following peer-reviewed publication:

- A. Abhishta, R. Joosten, S. Dragomiretskiy and L. Nieuwenhuis. 'Impact of Successful DDoS Attacks on a Major Crypto-currency Exchange'. *2019 27th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*. United States: IEEE, 2019, pp. 379–384 [4].

Chapter 5: Impact on Stock Prices

In this chapter, we analyse one of the indirect consequences of DDoS attacks. If an organisation's stock is publicly traded, it is possible to measure the reaction of investors to the events that are reported in media. We analyse the impact of a DDoS attack announcement on a victim firm's stock price. We select 45 different DDoS attack events over a period of 5 years and apply a more robust and less naive event analysis methodology to measure the impact on stock price. We avoid the wide-spread assumption about short term returns being normally distributed and use the empirical distribution for testing our hypotheses.
This chapter is based on the following peer-reviewed publications:

- Abhishta, R. Joosten and L. J. M. Nieuwenhuis. 'Analysing the Impact of a DDoS Attack Announcement on Victim Stock Prices'. *Proc. of 25th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP'17), St. Petersburg,Russia*. United States: IEEE, 2017, pp. 354–362 [9].

- Abhishta, R. Joosten and L. J. Nieuwenhuis. 'Comparing Alternatives to Measure the Impact of DDoS Attack Announcements on Target Stock Prices'. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)* 8.4, 2017, pp. 1–18 [10].

Chapter 6: Capturing Social Context

We discuss a method to collect data to provide context to DDoS attack events. We utilise the content change detection and notification service provided by Google called *Google Alerts* to collect articles related to DDoS attacks for more than 3 years. We show the breadth and benefits of this data collection with the help two case studies. In the first case study, we compare the data collected by us with the DDoS attack related articles available on LexisNexis. We show that, all the news articles available on LexisNexis also appear in our dataset. In the second case study, we successfully use supervised machine learning algorithms to filter attack reporting news articles. We test the efficiency of 8 different machine learning algorithms with the help of an annotated sample of 1000 articles. We select the best performing algorithm to filter the entire collected data for attack reporting news articles and show how it can be used for tracking DDoS attack events.

Parts of this chapter were presented as following poster at IEEE S&P, 2019:

- A. Abhishta, R. Joosten, M. Jonker, W. Kamerman and L. Nieuwenhuis. 'Poster: Collecting Contextual Information About a DDoS Attack Event Using Google Alerts'. 2019. Poster presented at 40th IEEE Symposium on Security and Privacy, San Francisco, CA [5].

Part II

The next two chapters form Part II of this thesis and focus on studying the aims of attackers for the use of DDoS attacks. In Chapter 7 we identify the attacker aims and then in Chapter 8 we evaluate the impact of daily routines of a victim on DDoS attack pattern.

Chapter 7: Aims

This chapter focusses on analysing the various aims for which attackers might use DDoS attacks. With the help of the dataset presented in Chapter 6, we study the aims of the most reported DDoS attacks in 2016. Taking into account the socio-cultural, political and economic (SPEC) dimensions of DDoS attacks and the postulates of routine activity theory (RAT) we propose a methodology to analyse news articles reporting an attack event to explain probable aims of attackers. We then evaluate 27 different attack events using the proposed methodology. The main contributions of this chapter are as follows:

- We observe that news articles are able to explain the *context* of a DDoS attack. Using the proposed model it is possible to explain probable aims of attackers.
- Organisations can become a target because of their socio-cultural and political environment.
- Organisations can also become a target just because they are virtually invincible.

This chapter is based on the following peer-reviewed publication:

- A. Abhishta, M. Junger, R. Joosten and L. J. Nieuwenhuis. 'A Note on Analysing the Attacker Aims Behind DDoS Attacks'. *International Symposium on Intelligent and Distributed Computing*. Springer. 2019, pp. 255–265 [8].

Chapter 8: Impact of Victim Routines

In this chapter, we study the impact of daily routines of a victim on DDoS attack trends. Routine activity theory (RAT) suggests that changes in crime rates should be associated with days that affect daily routines. Holidays not only have an impact on attacker routines but also on the routines of the victim. We analyse the impact of academic routines on Dutch educational institutions using data collected at SURFnet*. The main contributions of this chapter are as follows:

- We show how routine activity theory can be used to evaluate the influence of victim routines on attack patterns.

*SURFnet is the primary supplier of advanced networking to Colleges, universities and research institutions in the Netherlands

- We formulate and test multiple hypotheses on the basis of RAT to analyse the impact of academic routines on Dutch educational institutions.
- Our results show that the number of denial of service attacks targeting academic institutions in the Netherlands are higher during business hours.

This chapter is based on the following peer-reviewed publication:

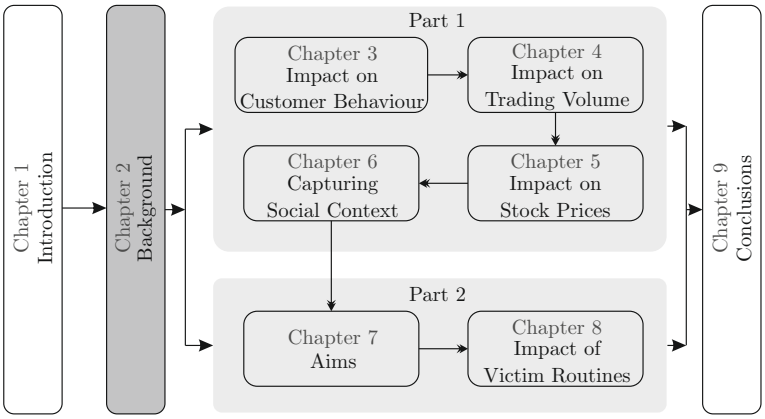
- A. Abhishta, M. Junger, R. Joosten and L. Nieuwenhuis. 'Victim Routine Influences the Number of DDoS Attacks: Evidence from Dutch Educational Network'. *2019 IEEE Security and Privacy Workshops (SPW).* 2019, pp. 242–247 [6].

Chapter 9: Conclusions and Future Works

In the final chapter of this thesis we draw the overall conclusions and answer the research questions formulated in Chapter 1. We also discuss the limitations of our work and suggest future directions for research.

Chapter 2

Background

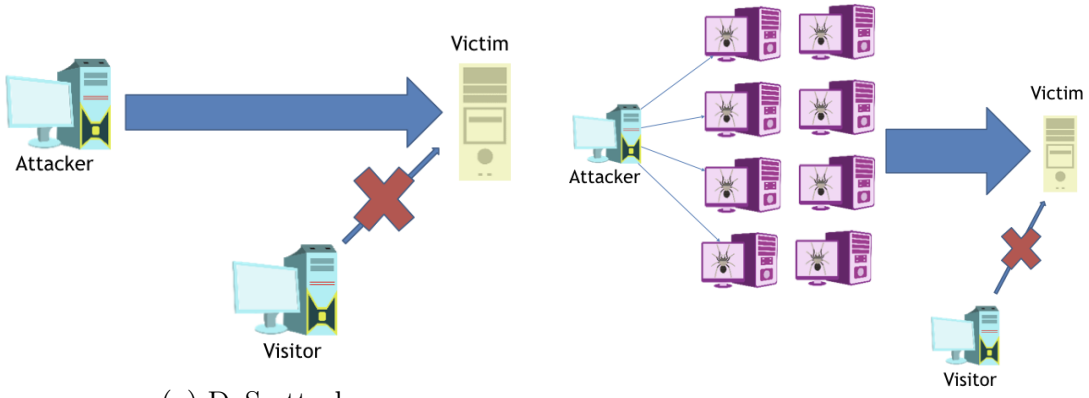


Here, we give the background information on distributed denial of service (DDoS) attacks. The chapter provides a peek in the evolution of DDoS attacks over the last 2 decades. It also gives a brief description of the various DDoS protection strategies available to organisations and explains the factors involved in selecting the most suitable strategy. It then identifies the various stakeholders involved in a DDoS attack and describes the interactions between them. Finally, the chapter ends by discussing the framework used for measuring cost of cybercrime.

2.1 Reading guidelines

We give an overview of the theories and frameworks used in our analysis to measure the impact of DDoS attacks. We begin by discussing in brief the history and evolution of DDoS attacks in Section 2.2. Then in Section 2.3 we review DDoS mitigation strategies available to organisations and the most popular tools to evaluate security investments. On the basis of this information we determine the main stakeholders in a DDoS attack in Section 2.4. In Section 2.5 we evaluate the profits made by attackers and the framework used by us to measure victim losses. We end the chapter by discussing how Routine Activity Theory (RAT) can be used to evaluate attacker aims and the impact of victim routines on attack trends.

2.2 DDoS attacks and their evolution



(a) DoS attack.

(b) DDoS attack.

Figure 2.1: DoS and DDoS attacks.

There phenomenon of Denial of Service (DoS) attacks has been known to the network research community since early 1980s [218]. According to WWW Security FAQ [192] a DoS attack can be defined as an attack designed to render a computer or network incapable of providing normal services. In the summer of 1999, the Computer Incident Advisory Capability (CAIC), now known as the original computer security incident response team at the Department of Energy (United States) reported the first Distributed DoS attack incident. According to WWW Security FAQ [192], "A DDoS attack uses many computers to launch

a coordinated Denial of Service attack against one or more targets. Using client/server technology, the perpetrator is able to multiply the effectiveness of the Denial of Service significantly by harnessing the resources of multiple unwitting accomplice computers, which serve as attack platforms." Just like a DoS attack, this results in the unavailability of network resources for the intended user. Figure 2.1 shows the difference between a DoS and DDoS attack.

Since 1999, distributed denial of service (DDoS) attacks have been used on numerous occasions by actors wanting to make the network services of the victim unavailable. Over the years attackers have used various different strategies and exploited several vulnerabilities in the network infrastructure to carry out DDoS attacks. Specht & Lee [189] and Mirkovic & Reiher [136] provided one of the first taxonomies for DDoS attacks. Specht & Lee [189] categorised DDoS attacks on the basis of attack model and techniques used by a perpetrator. They broadly classified these assaults based on following attack models: 1) *Agent-Handler* attacks 2) *internet relay chat (IRC)* based attacks.

An *Agent-Handler* model comprises of clients, handlers and agents (*a.k.a.* bots). A client is used by an attacker to communicate with rest of the attack system. Depending on the configuration, agents can be instructed by a single or multiple handlers. The handlers are software packages that are located throughout the Internet that attacker's client uses to communicate with the agents. The agents are compromised systems that will eventually carry out the attack. The attacker communicates with handlers to identify the active agents and also to carry out an attack. In such a model owners and users of an agent system have no knowledge of that their systems are compromised and take part in a DDoS attack. Specht and Lee [189] also propose the use of internet relay chat (IRC) by attackers as a substitute to handler program installed on a network server.

On the basis of the techniques used by attackers, Specht and Lee [189] classify the attacks as bandwidth depletion attacks and resource depletion attacks. When an attack involves agents sending large volumes of traffic to a victim system, to congest the victim system's bandwidth, it is called a bandwidth depletion attack. Flooding attacks and amplification attacks fall under the category of bandwidth depletion attacks. On the other hand, when an attack exploits the capacity of a network protocol, it is known as a resource depletion attack. Protocol exploitation attacks and malformed packet attack are categorised as resource depletion attacks.

Mirkovic and Reiher [136] further classified these attacks degree of automation, degree of attack rate dynamics and degree of impact. Attacks can be launched manually by the use of (D)DoS tools (e.g., low orbit ion cannon (LOIC), trinoo, tribe flood network (TFN) and Shaft) that are freely available

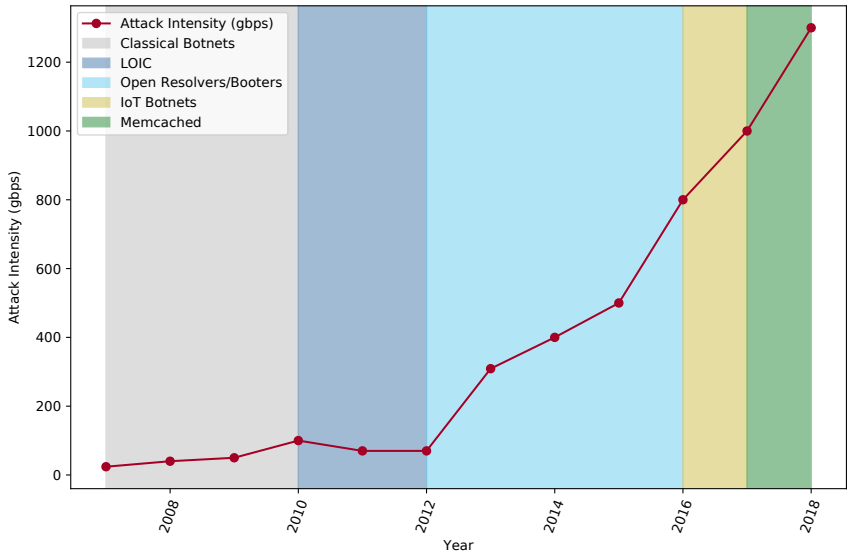


Figure 2.2: Historical increase in the intensity of DDoS attacks (2007-2018) [138].

online or automatically by using a botnet or booter. The rate at which packets flood the victim can be continuous or variable. Finally, attacks can also be classified on the basis of their intensity as disruptive and degrading. Zargar, Joshi and Tipper [218] provide a more recent overview of the types of DDoS attacks by updating the sub-classes of above mentioned attacks.

Practically, attack intensity has also risen in the last two decades. Attacks have evolved and various different strategies have been used by attackers to achieve higher intensities. In 2018, we observed a 1.3 terabits per second (Tbps) attack on Github [144]. Figure 2.2 shows the increase in attack intensities between 2007 and 2018. It also shows the prominent attack strategy used by attackers in that period. Between 2007 and 2009, the prominent attack strategy was the use of computer based botnets. The highest attack intensities recorded were under 100 gigabits per second (Gbps). A large number attacks reported in this period were politically motivated [176]. A report by Radware [161] has described the attacks in the period 2009 to 2012 to be driven by hacktivist groups such as *Anonymous*. Sauter [180] has shown that these attacks were launched primarily with the help of DDoS attack tools such as *LOIC*. The period from 2012 to 2016 was dominated by amplification and booter attacks. First, open DNS resolvers were used to amplify attacks and later NTP servers were used

for this purpose. In this period, the attacks started becoming a threat for high capacity networks as the intensity of attacks peaked above 500 Gbps. With the rise of IoT based botnets in 2016, we saw a number of high intensity attacks in latter half of 2016. The attack on managed domain name service (DNS) provider peaked at 1 Tbps. In 2018 attackers used memcached servers to amplify the attacks and were able to achieve 1.3 Tbps. This shows that by leveraging various vulnerabilities in the Internet, attackers have been able to target organisations with higher intensity attacks that are much harder to mitigate.

2.3 Defence and security investment

There are many different strategies that organisations can adopt in order to defend themselves against DDoS attacks. The choice of strategy depends upon multiple variables such as the location (with respect to network stream) and techniques of DDoS detection and response [218]. When DDoSed, a victim is flooded by network packets, to defend its infrastructure from becoming unavailable a victim can choose to deploy packet filtering based on IP traceback mechanisms, management information base (MIB), packet marking and filtering mechanisms, history based mechanisms, hop count mechanisms and path identifier (PI) mechanism. On the other hand, the victim can also distribute the traffic on multiple servers, such that none of the servers are overwhelmed.

In the last decade several organisations have come up that offer DDoS protection services. These organisations allow the victim to either host the service on their platform or direct traffic towards their traffic cleaning systems during an attack. A common problem organisations face is the decision to outsource DDoS protection or to have an in-house stand alone DDoS mitigation system. This decision depends upon the prospective benefit of DDoS protection, paying capacity of organisations and privacy laws applicable to an organisation. Over the years, researchers have proposed a number of methods to calculate optimal investment in security. We mentioned some of these models in Chapter 1. Most of these models consider parameters such as expected loss due to cyber attacks, probability of attack (i.e., threat) and probability of success of an attack (i.e., vulnerability) for determining optimal investment in security [78, 92]. According to Gordon & Loeb, in a one-period economic model and risk neutral setting if λ represents the monetary loss conditioned on the breach occurring, t represents the threat probability, v represents vulnerability (i.e., defined as a conditional probability that a threat once realised would be successful), z represents the investment in security and $S(z, v)$ denotes the security breach probability function

then the expected benefit from this investment in information security $EBIS(z)$ can be calculated as:

$$EBIS(z) = \{v - S(z, v)\}t\lambda \tag{2.1}$$

Hence, expected net benefits from investment in security (ENBIS) can be modelled as:

$$ENBIS(z) = \{v - S(z, v)\}t\lambda - z \tag{2.2}$$

The strategy suggested by [78] is derived by maximising $ENBIS(z)$. Another notable metric used for judging investments in security is known as return on (security) investment ($ROSI/ROI$). Equation 2.3 shows the expression used to calculate the $ROSI$. Higher values of $ROSI$ denote more efficient investments.

$$ROSI = \frac{\text{benefit of security} - \text{cost of security}}{\text{cost of security}} \tag{2.3}$$

As we can see that these metrics rely on variables such as expected loss and benefits of security, it is important for organisations to be able to measure these variables for efficient decision making. Measuring the value of these variables in the real world can be complex and challenging. In this thesis, we develop and apply methods for measuring the expected loss on an organisation due to a DDoS attack. This measurement can also be used to calculate the benefit of security.

2.4 Attack stakeholders

Till now we have discussed the various ways adopted by attackers to carry out DDoS attacks, we have also discussed the options available to organisations to protect themselves against these attacks. In this section, we identify the main stakeholders of a DDoS attack. We define main stakeholders as actors on whom a DDoS attack has an impact. We will also explain the interactions between these stakeholders.

Based on the previous discussion, we identify four main stakeholders in a DDoS attack. These are:

- The attacker.
- The victim.
- Customers of the victim.

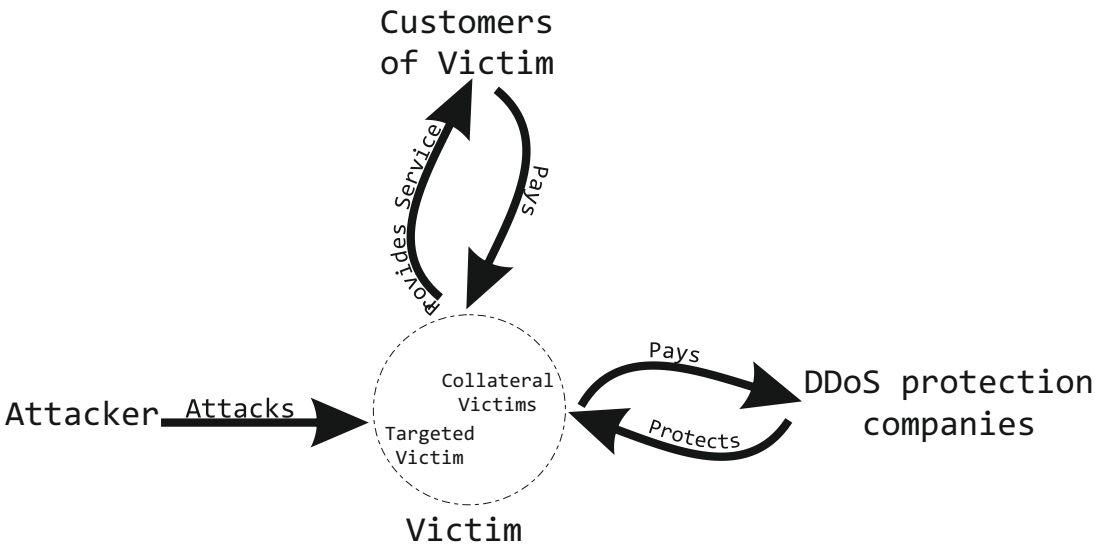


Figure 2.3: Interdependencies between the actors.

- DDoS protection companies.

The actor who initiates the attack is referred to as an *attacker*. The number of attackers can vary as per the model of the strike. Some strikes can be organised by a single malicious actor however it has been noticed that during most online protests a number of attackers collaborate to achieve higher attack intensities [180]. The intended target of a DDoS attack is referred to as the *victim*. In case of shared hosting an organisation shares the hosting platform with other organisations, in such a situation if any one of these organisations is targetted with a DDoS attack then all the other organisations will also suffer the consequences. We divide victims as targeted and collateral. An attack on an organisation may lead to unavailability of services on the side of its customers therefore the attack would result in losses for them. Thus, they form the third stakeholder in a DDoS attack. A report by Arbor Networks [214] suggests that 66% of the times *customers of the victim* are the real target. To protect themselves from the ever growing threat of these attacks, many a times firms outsource the security to *DDoS protection companies*. They form the fourth category of stakeholders. In Figure 2.3 we show the interactions between the stakeholders.

We focus on measuring the consequences of a DDoS attack on the victim. However, in the latter part of this thesis we also analyse the aims of attackers

for the use of DDoS attacks. In Chapter 3 and Chapter 4, we analyse the direct impacts of DDoS attacks on the victim. In Chapter 5 we analyse the impact of a DDoS attack on the market values of the firms and in Chapter 6 we present a dataset that captures the circumstances surrounding a DDoS attack and can be used to track articles reporting an attack. Later in Chapter 7 and Chapter 8, we analyse attack aims and the impact of victim routines on DDoS attack trends.

2.5 Impact of DDoS attack on attacker and victim

A DDoS attack event has implications on each of the stakeholders discussed above. In this section, we briefly discuss the impact of a DDoS attack on two these stakeholders, i.e., attacker and victim. In order to show the possible costs and benefits of DDoS attacks on an attacker, we analyse the business model of a botnet. For most DDoS attacks, botnets form the core infrastructure needed to launch an attack [125]. It not only allows attackers to access multiple IP addresses at the same time but also provides them anonymity [88]. We then based on previous research explain the possible impact of DDoS attacks on a victim.

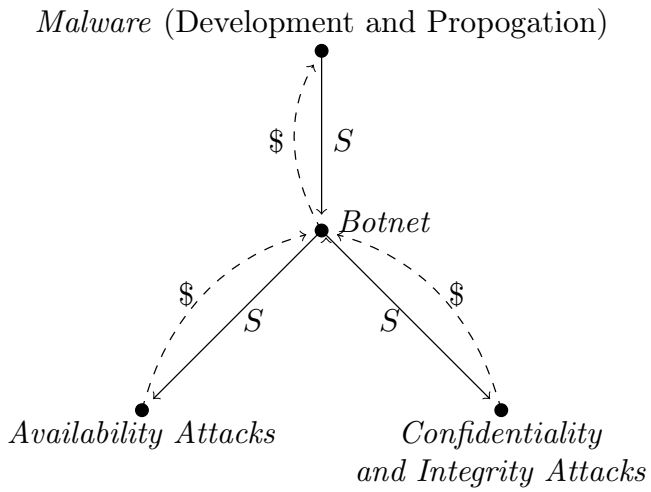
2.5.1 Business model of a botnet

Botnets are one of the basic infrastructures that help an attacker to launch a DDoS attack. A *botnet* is a network of various computers that can be controlled by attackers. The controller of the network is known as the *botmaster*. They give commands to the network by making use of various communication channels. The malicious software that is used to control this network of computers is called the *malware*. It comes as no surprise that these malicious networks (botnets) can provide economic and other benefits to their owners. Results as presented by Miller [135] and Bottazzi & Me [28] provide an insight on the actors involved in deployment and handling of a botnet. In this section, we present a brief overview of the development and life cycle of a botnet and finally based on this information analyse business model of a botnet.

2.5.1.1 Botnet ecosystem

Several practitioners are involved in development of a botnet from scratch, and can be divided in four tiers as follows[79]:

Tier 1: Practitioners who rely on others to develop malicious code, delivery mechanism and execution strategy.



S: Service Provided

Figure 2.4: Botnet Ecosystem.

Tier 2: Practitioners who have a great depth of experience, with the ability to develop their own tools.

Tier 3: Practitioners who focus on the discovery and use of unknown malicious code.

Tier 4: Practitioners who are organised, highly technical, proficient and well funded to discover new vulnerabilities and develop exploits.

Practitioners who belong to *Tier 3* and *Tier 4* are involved in development and propagation of malware. *Tier 2* practitioners use the botnet developed by *Tier 3* and *Tier 4* practitioners, while *Tier 1* practitioners use the services provided by *Tier 2* practitioners to initiate attacks. On the basis of this tier distribution we develop a botnet ecosystem as shown in Figure 2.4. The ecosystem gives a snapshot of how the botnet economy functions at a macro level. It also shows that botnets not only can be used for DDoS attacks (availability attacks) but can also be used for confidentiality and integrity attacks. A botnet owner can provide these attacks as a service, which can serve as a revenue stream. At the same time, the owner has to pay malware developers and propagators in order to keep the bots up and running. For gathering further knowledge on the business of botnets, we need to understand how a botnet can be built from scratch (botnet assembly chain) and maintained (botnet life cycle).

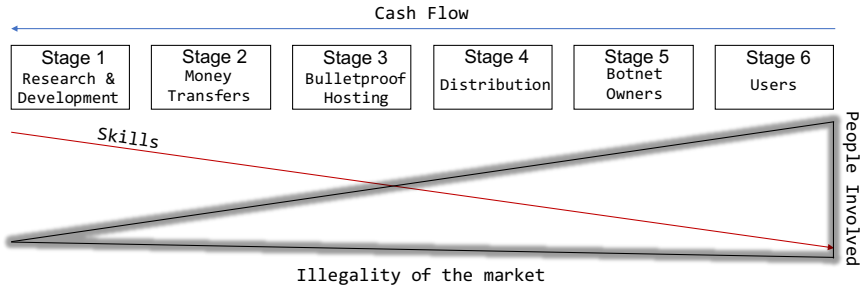


Figure 2.5: Botnet assembly chain [28].

2.5.1.2 Botnet assembly chain

Bottazzi and Me [28] has defined a botnet assembly chain as a six stage process based on activities varying from development to utilisation. Furthermore, this assembly chain is evaluated in terms level of skill necessary to be able to successfully complete the stage, and darkness (illegality) of the market it is operating in. Gosler and Von Thaer [79] propose a similar division, however their model is more based on the involved actors. It has already been discussed in Section 2.5.1.1.

The first stage of the botnet assembly chain is research & development and setting up of money transfers. These are arguably legal businesses. Research and development involves the continuous search for exploits in software, the development of new malicious software and selling knowledge of computer systems and software. The actors behind this process are mostly IT professionals and offer customer support. They can even customise the software to the wishes of the customer. The actors in this stage of development can be linked to tiers two, three and four in the hierarchy proposed by Gosler et al. [79]. The second stage involves setting up of payment methods for money transfers. The botnet developer can choose to use an online payment system such as Paypal or a crypto currency based payment system.

The third and fourth stage involves building a C&C (command and control) and distributing the malware to create bots respectively. As C&C is a critical component of the botnet infrastructure, botnet developers try to use bulletproof hosting services to host the command and control. Bulletproof hosting may also serve as web-based storage for the botnet end-user to store stolen information like banking credentials, passwords etc.. The control centre can consist of one or multiple computers for building redundancies [170]. Many malware developers do not have enough resources to spread the malware to a large number of com-

Article	Activity	Malware	Context	Finances
[109]	Spam advertised pharmaceuticals	Unknown	360 Million emails per hour 10,000 bots	$100 per sale, $3.5 million annually
[28]	Robbing bank credentials	Euro grabber, ZeuS based	30,000 Targets across Europe	$47 million over 2.5 months
[32]	Booter, botnet-for-hire	Unknown	800,000 Attacks over a 1 year period	$26k monthly revenue, median of 24 months
[28]	Advertisement click fraud	Zero-Access	140,000 hosts	$900k of daily ad revenue losses

Table 2.1: An overview of botnet case studies.

puters across the world. To solve this problem they make use of the so called PPI (pay-per-install) Distribution model. In essence this involves the owner of the malware paying affiliates to spread the malware, providing a commission to these malware spreaders per infected device. The client making use of PPI distribution to spread the malware usually collects the funds to be able to afford this by selling regular botnet related services. Taking a look at the tier-based hierarchy, it is likely that the practitioners which are mentioned in tier 1 make use of the PPI Distribution model. Caballero *et al.* [36] indicates the PPI model is one of the most used ways of distributing malware. Estimates are that, of the twenty most prevalent families of malware, twelve made use of the PPI distribution model [121].

Lastly, stage 5 and stage 6 lie in the highly illegal spectrum. Actors involved in this stage are the owners of botnets, the ones who actually perform the attacks. Several of these actors use the so called dark web for carrying out their activities.

2.5.1.3 Botnet life cycle and attacker revenue

Like any legitimate business, a botnet business passes through multiple phases during its life time. The first phase of the botnet life cycle is conception, that is all about motivation: why does one want to setup a botnet? On this subject, Rodriguez-Gomez *et al.* [167] argues that there are five motives for a botnet developer to setup a botnet. These are money, entertainment, ego, cause and social status. Of these five, it is argued that the primary motive is financial gain. Source code of the botnet malware can be sold or rented out to multiple buyers from around the world.

The second is the recruitment phase. Infecting computers (or paying others to infect computers for the botnet developer) with botnet malware resulting in the botmaster being able to control the computer. Usually, larger the botnet the better it is, as the power of a botnet is highly dependent on its size. Depending on the size, renting a botnet for DDoS attacks can cost up to several thousands of dollars a day.

Next, a botnet owner can decide to use the botnet himself or rent the services based on botnet. Booters are an example of renting DDoS attack services based on botnets [176]. It often takes place by making use of underground online marketplaces or forums, which can be found and accessed via the dark web. In United States of America, the law that prohibits the user to create a botnet (amongst other fraudulent computer activities) is known as the Computer Fraud and Abuse Act [1]. Other countries have similar laws in regard to fraudulent computer use, which include botnet use and ownership.

However, by using a botnet criminals can generate hefty revenues. Bottazzi and Me [28] states that spamming and DDoS-attacks can be considered least profitable among the activities mentioned in Table 2.1, since the operation is too noisy and hence more bots are to be replaced frequently. A summary of previous findings on the cases analysed by various researchers can be found in Table 2.1.

2.5.1.4 Business model canvas

Now that we have discussed the process of botnet development and have a sense of revenue streams of botnet owners, we can now design the business model of a botnet. According to Zott and Amit [221], a business model design is the purposeful weaving together of interdependent activities, performed by the firm itself or by its suppliers, partners and/or customers. We use the Osterwalder Business Model Canvas [146] framework to depict the *business* of developing, using and maintaining a botnet.

Osterwalder & Pigneur [146] propose nine building blocks as the basis of a business model, the logic of how a company intends to generate profit. These building blocks are customer segments, value propositions, channels, customer relationships, revenue streams, key resources, key activities, key partnerships and cost structure each have their own core questions that can be used to characterise every business.

Figure 2.6 shows the business model canvas for a botnet business. It provides a snapshot of the key value propositions of a botnet business. It lists the infrastructure and collaborations required to create one's own botnet and also provides an overview of the revenue stream of a botnet owner.

Key Partners	*Key Activities*	*Value Propositions*	*Customer Relationships*	*Customer Segments*
Malware Developers [28, 135] Money Handlers [28, 32] Bulletproof Hosting Provider [28, 32] Malware Distributors [28, 135, 36] Dark Web Market Places [161] Government Intelligence Agencies [147]	Botnet Maintenance [134, 135] Managing Botnet Infrastructure [135] Perform Attacks [28, 79]	Advertising products [109] Stealing Money [28] Creating Fraudulent ad-clicks [28] Shutting Down Websites [32, 147] Cryptocurrency Mining [147]	Automation Forums [87] Internet Relay Chat [123] Email [32]	Governments [147] Malicious Parties Website/Server Administrators [32, 122]
	Key Resources		*Channels*	
	Bots Identity Protection Software Network Connectivity		Dark Web Marketplaces [161] Hacker Forums [87] Emails [32] Websites [32]	
Cost Structure			*Revenue Streams*	
Malware Development [28, 135] Infections [28, 135, 36] Hosting [32] Bandwidth Transaction Fees [32] Customer Service [32]			Stolen Bank Account Money [28] Click Fraud [28] Sale of Booter Services [32, 147] Sale of Spam Services [109]	

Figure 2.6: Business Model Canvas for a botnet owner.

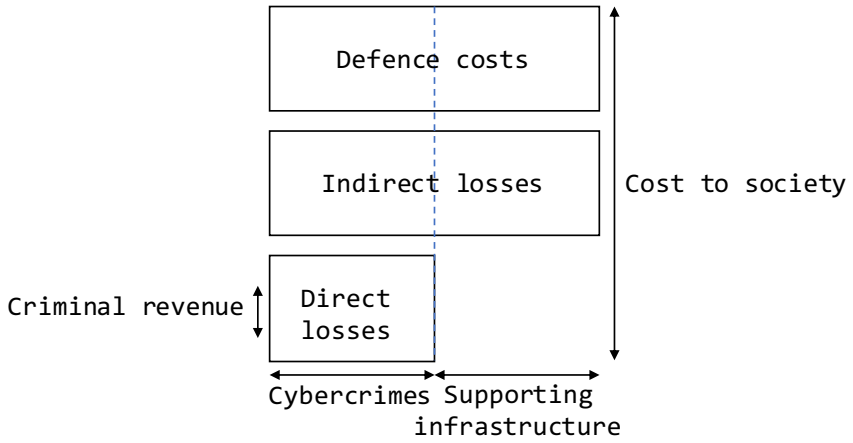


Figure 2.7: Framework for analysing the cost of cybercrime [13].

2.5.2 Measuring victim losses

Just like any other cyber attack, DDoS attacks can cause damages to the victim. Anderson *et al.* [13] have provided an elaborate framework for analysing the cost of cybercrimes. Figure 2.7 shows the proposed framework. They define the cost categories as follows:

Defence Costs: Defence costs are defined as the monetary equivalent of prevention efforts. These are costs that are incurred in anticipation of an attack. They include investment in security products, security services etc.. If an organisation invests in DDoS protection strategies and still the DDoS attack is successful then these costs can be considered as losses. In Chapter 3 we briefly discuss the increase in defence costs for an organisation once a service becomes unavailable to them due to a DDoS attack.

Indirect losses: Indirect loss is the monetary equivalent of the losses and opportunity costs imposed on society by the fact that certain cybercrime is carried out. One of the indirect losses due to a DDoS attack can be change in investor perception about the market value of a victim firm. In Chapter 5 we make use of event study methodology to measure these losses. Another method for measuring this indirect loss can be to measure the popularity of the attack in news media. In Chapter 6 we introduce a dataset that can be used to measure popularity of DDoS attacks and present a simple metric to measure popularity of attacks.

Direct losses: Direct loss is defined as the monetary equivalent of losses, damages or other suffering felt by the victim as a consequence of a cybercrime. In case of internet based service providers (e.g., hosting providers, DNS service providers etc.) customers of the victim might be forced to move to an alternative provider due to service unavailability. In Chapter 3 we analyse the impact of a DDoS attack on the customers of DNS service providers. Also, web based businesses may lose online sales due to a successful DDoS attack on their platform. In Chapter 4 we analyse the impact of DDoS attacks on crypto currency trading volume over a large crypto currency exchange.

Cost to society: According to the framework, the summation of direct losses, indirect losses and defence costs is the cost to society.

Criminal Revenue: Criminal revenue is defined as the monetary equivalent of the gross receipts from a crime. There have been instances when organisations are forced to pay ransoms to stop attackers from DDoSing them [148]. In Section 2.5.1 of this chapter we discuss the profits attackers can make by launching DDoS attacks.

Dubendorfer, Wagner and Plattner [57] use a systems engineering approach to suggest a framework to calculate the economic impact of DDoS attacks. They classify the damage due to DDoS in four types namely: downtime loss, disaster recovery, liability and customer loss. In this thesis we follow the framework proposed by Anderson et al. [13].

2.6 Routine activity theory

As discussed in Chapter 1, not all attackers are looking to inflict economic damage to the victim organisation. Hence, analysing attacker aims can help us in understanding the type of damage an attacker is looking to inflict. Also, routines of victims can have a considerable effect on the economic impact of short-term IT unavailability. Depending on the business model of a victim organisation, IT unavailability can cause varied damages at different times. Several studies in criminology have analysed the influence of attacker and victim routines on the occurrence of crime. Cohen and Felson [43] proposed a *routine activity approach* for analysing crime rate trends and cycles. Unlike most studies at that time, they concentrated on the circumstances in which offenders carry out criminal acts rather than emphasising their characteristics. They hypothesized that the dispersion of activities away from households and families leads to higher crime rates and presented a variety of data in support of their hypothesis.

Several studies in cybercrime have also made use of this routine activity theory (RAT) to evaluate the activities of cyber criminals. Yar [217] explored the

extent to which this theory's concepts and aetiological schema can be transposed to cybercrimes. Yar concluded that the theory's core concepts can indeed be applied to cybercrime, however there are important differences between *virtual* and *terrestrial* worlds. Yar [217] suggested that suitability of a target for an attack can be estimated according to its following four characteristics: Value, Inertia, Visibility and Accessibility. We use these characteristics in Chapter 7 of this thesis to propose a model that can be used to analyse attacker aims.

Pratt, Holtfreter and Reisig [158] based on RAT and consumer behaviour research analyse how personal characteristics and online routines increase people's exposure to motivated cyber criminals. Maimon, Kamerdze, Cukier and Sobesto [133] analysed if daily routines of a university has an impact on DDoS attack trend using the Intrusion Prevention System (IPS) data of a single university and showed that attacks on university are more likely to happen during business hours. In Chapter 8 of this thesis we analyse the DDoS attack trends on all the educational institutions in the Netherlands and show that daily routines of educational institutions have a significant impact on the DDoS attack trends even in this more generalised context.

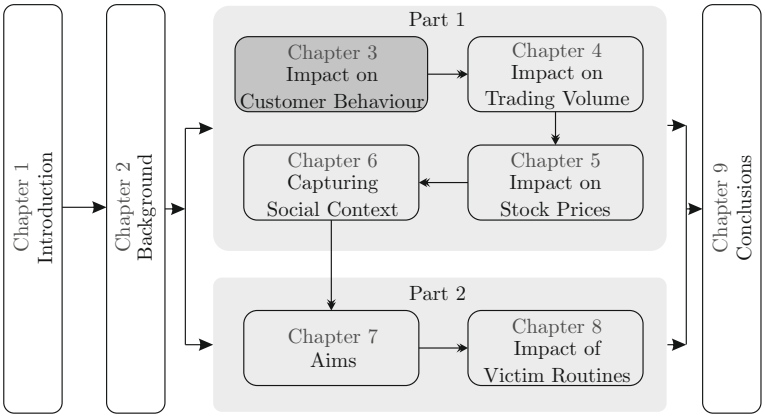
Part I

Economic Impact of DDoS Attacks

This page is intentionally left blank.

Chapter 3

Impact on Customer Behaviour



In this chapter, we analyse one of the direct impacts of DDoS attacks, i.e., change in customer behaviour of managed DNS service providers. Our analysis leverages data from the OpenINTEL active DNS measurement system, which covers large parts of the global DNS over time. Results show an almost immediate and statistically significant change in the behaviour of domains that use NS1 or Dyn as a DNS service provider. We observe a decline in the number of domains that exclusively use NS1 or Dyn as a managed DNS service provider, and see a shift toward risk spreading by using multiple providers. While a large managed DNS provider may be better equipped to protect against attacks, these two case studies show they are not impervious to them. This calls into question the wisdom of using a single provider for managed DNS. Our results show that spreading risk by using multiple providers is an effective countermeasure, albeit probably at a higher cost.

3.1 Introduction

An immediate impact of a successful DDoS attack is the unavailability of services provided by the victim to its customers. For instance, for an e-commerce firm this unavailability might result in decrease of sales during the attack and can also cause damage to the reputation of the victim [40].

These attacks also threaten the availability of services that support the Internet usage for an everyday user. One of the core services on which the Internet is built is the Domain Name System (DNS). DNS is responsible for translating easy to remember domain names into machine readable IP addresses. Thus, unavailability of the DNS leads to unavailability of web services for most users. On several occasions, attackers have targeted the DNS with a DDoS attack to bring down web services. Hence, it is important for firms that prioritise availability to choose a DNS provider that is resilient in the face of DDoS attacks. There are several managed DNS providers that provide DDoS resilient services. NS1 and Dyn are two such managed DNS (MDNS) service providers. On May 16$^{\text{th}}$, 2016 and October 21$^{\text{st}}$, 2016, DDoS attacks targeted NS1 [22] and Dyn [86] respectively. The attacks were successful in hindering the services provided by NS1 and Dyn for the better part of a day.

While much has been said about the impact of especially the Dyn attack, one aspect of these attacks has received far less attention, namely: *What is the impact of such an attack on the behaviour of customers of affected MDNS providers?* In this chapter, we study this impact by looking at the DNS configuration of domains in a large DNS dataset. This allows us to answer questions such as: do customers continue to use the services of the attacked MDNS after the attack or not? If they remain a customer, do they change their behaviour?

Our contributions are as follows:

- We provide a framework for measuring the behaviour of domains using an MDNS service provider.
- We use this framework to analyse the impact of successful DDoS attacks on NS1 and Dyn on the behaviour of domains that use their services.
- We show statistically significant changes in customer behaviour after the attacks, such as, e.g., adding a second DNS provider for a domain.
- We show that most customers that start using a second provider use another MDNS service provider as a secondary DNS to further reduce the risk of downtime.

3.2 DNS as a resource

In order to understand the behaviour of customers after a DDoS attack, it is important to first understand the additional benefits of the service provided by NS1 & Dyn to its customers. In this section we look at the Domain Name System (DNS) as a resource [20] and explain its benefits [195] with the help of a so-called value network. DNS is one of the core services that supports the Internet. It translates human readable *domain names* (e.g. `www.example.com`) into machine readable *IP addresses* (e.g. `93.184.216.34`) [165]. Hence, it is safe to categorise DNS as a resource that facilitates the delivery of other web-based services (e.g. e-banking) to the customers of a bank.

The DNS itself is hierarchically organised: *root level, top-level domains, public suffixes, second-level domains, third level domain and so on.* The data (a tuple of domain name and IP address) on a domain name server is distributed according to zones and is stored in zone files. The authority for the records related to a domain name is delegated to a so-called authoritative name server (ANS) with the help of so-called `NS` records.

Value of managed authoritative name servers for firms: A *value network* – with related concepts such as actors, roles and value adding activities – can be used to describe and analyse a specific product or service offering in a detailed way [61]. A value network shows the value adding actors involved in the service delivery process and their relationships. Such a network helps in understanding the benefits and roles of each of the actors in the process. A value network is defined as *"a spontaneous sensing and responding spatial and temporal structure of largely coupled value proposing social and economic actors interacting through institutions and technology, to: (1) co-produce service offerings, (2) exchange service offerings, and (3) co-create value"* [130] . Figure 3.1 shows the value network of a web service delivery.

We can understand the value network shown in Figure 3.1 by considering an example of a customer who wishes to transfer money to another account without physically visiting a bank. In this case the customer first needs to log on to the e-banking website of their bank using a web browser. Once the customer requests the e-banking website, the web browser then queries the DNS resolver of its network for the *IP address* associated with this *domain name*. In case the response to this query is not present in the cache of the DNS resolver, it retrieves the *IP address* from the authoritative name server (*ANS*) associated with this domain name and forwards the response to the web browser. The web browser then connects to the server located at the *IP address* and in-turn provides the web service to the customer. With the help of this example, it is

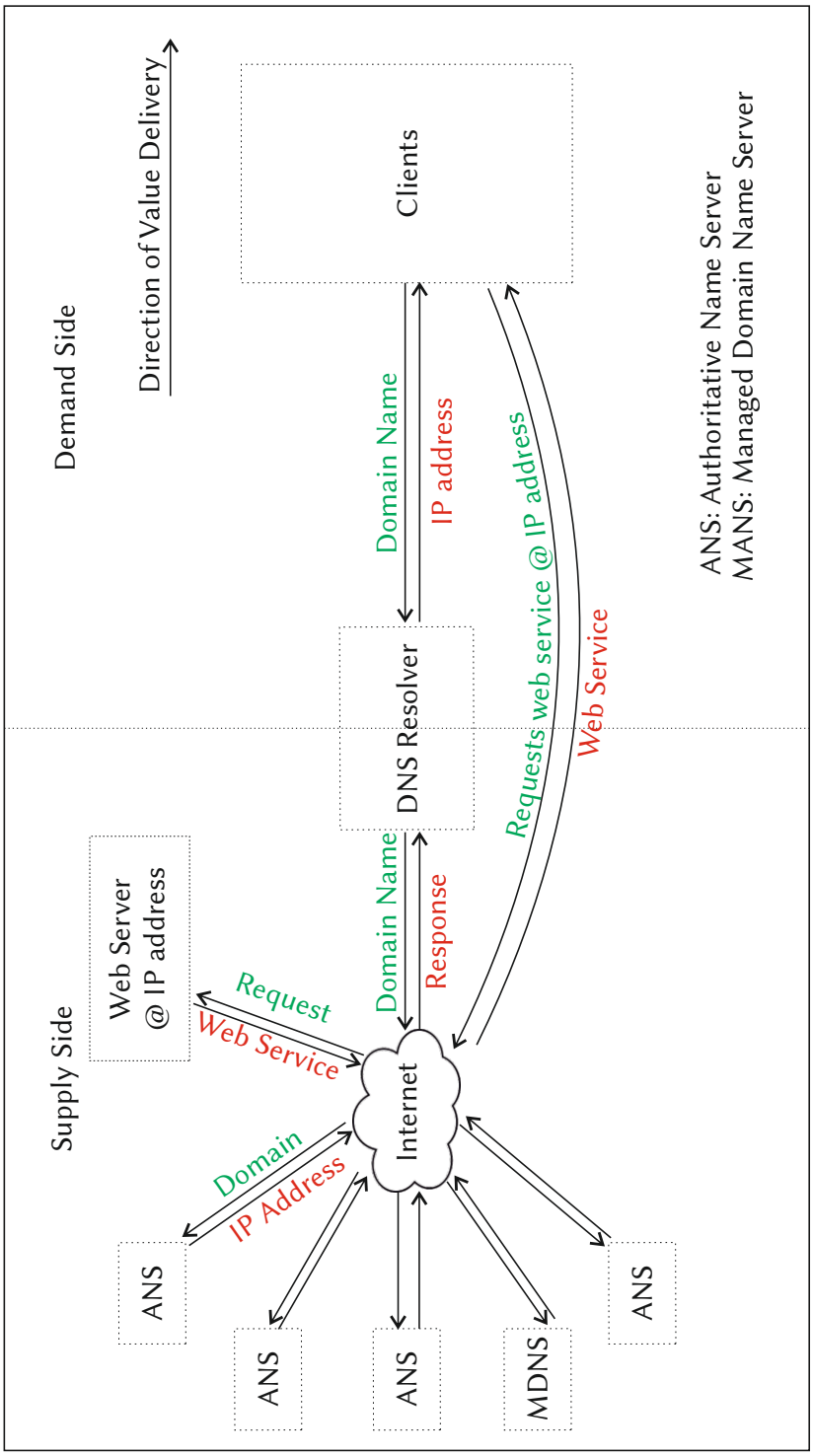


Figure 3.1: Value network of web service delivery showing the role of various components of the DNS.

evident that unavailability of *ANS* can lead to potential unavailability of the e-banking service.

On multiple occasions [152, 120, 148], criminals have targeted the availability of various components of the value network as described above using DDoS attacks. The need for availability of *ANS* has created a market for managed domain name service providers (*MDNS*). These *MDNS* provide the following benefits in addition to the features of an *ANS* [204]:

- faster response times;
- load balancing;
- and DDoS protection.

Hence, for a *domain* that forms a source of revenue for a company, a *MDNS* promises greater availability and helps the company in efficiently catering to the needs of its consumer.

It is common practice for domain owners to specify multiple ANS for their domain. DNS resolvers may then query each of these authoritative name servers, although they will have a preference on the basis of metrics (e.g. round trip time) [141]. In the context of the use of MDNS, this practice has additional consequences. A domain owner can choose to *exclusively* use multiple ANS from a single provider. If this provider then somehow goes down, the domain owner will suffer unavailability as a consequence. Another option is for the domain owner to procure services from multiple MDNS providers and thus to *non-exclusively* use MDNS services. While this makes DNS management a bit more complex for the domain owner, and potentially comes at a higher cost, it has one significant benefit: if one MDNS provider goes down, the domain will still be available under the assumption that the other MDNS provider(s) are still operational.

3.3 Impact of a DDoS attack

A successful DDoS attack hampers the availability of an MDNS provider. As the added value of using MDNS is DDoS protection, a successful attack can lead to loss of customers in a market where availability is of great importance [220, 50]. In this section we introduce a framework that can capture the behaviour of domains using an MDNS provider. We use this framework to study two DDoS attack events: (1) on NS1 on 16$^{\text{th}}$ May 2016 and (2) on Dyn on 21$^{\text{st}}$ October 2016. For our analysis, we make use of a large longitudinal dataset that is introduced in the following section.

MDNS	Dataset	Start Date	End Date
NS1	OpenINTEL	29th October 2015	5th June 2016
Dyn	OpenINTEL	4th of April 2016	11th November 2016

Table 3.1: Details of dataset.

3.3.1 Dataset

We use the OpenINTEL dataset as source data to analyse the impact of DDoS attacks on the behaviour of domains using an MDNS provider. The OpenINTEL project collects unique long-term datasets with daily DNS measurements for all domains under the main top-level domains on the Internet (including .com, .net and .org). Currently, OpenINTEL covers over 60% of the global DNS name space every 24 hours. Van Rijswijk-Deij *et al.* [166] explain the data collection method in detail.

We use data for the domains in three generic top-level domains (gTLDs) .com, .net and .org. In order to get a list of domains that use Dyn/NS1 on a given day we query the dataset for all domains that use Dyn/NS1 name server addresses in their NS records on that day. We use the measurements in the OpenINTEL dataset for time intervals as shown in Table 3.1.

3.3.2 Type of domains

On the basis of the number of different service providers found in `NS` records for a domain, we categorise domains into two types:

Exclusive: A domain is categorised as exclusive if it uses only Dyn/NS1 name server addresses in its `NS` records.

Non-exclusive: A domain is categorised as non-exclusive if it uses name server addresses from multiple providers including Dyn/NS1.

This categorisation is of great importance for this study as a non-exclusive domain will not experience an inferior service quality during an attack. Now, in order to measure the change in the behaviour of domains we need to first define *behaviour*.

3.3.3 Measuring the impact

We define a step-by-step procedure that we use to perform our analysis. In order to measure the impact of the DDoS attack on the domains using NS1/Dyn as

an MDNS provider, we use an approach similar to event studies [132]. We use a five step approach to measure the impact as described below:

Step 1: Define variables representing the behaviour of domains.

Step 2: Define a trend period and an event period.

Step 3: Measure behaviour in the trend period.

Step 4: Measure behaviour in the event period.

Step 5: Analyse any changes in behaviour.

Step 1 is discussed in Section 3.3.3.1, Step 2 in Section 3.3.3.2, Steps 3 and 4 in Section 3.3.3.3, and Step 5, the analysis, is discussed in Section 3.4.

3.3.3.1 Behaviour of domains

In this study we define the behaviour of domains that use NS1/Dyn's MDNS infrastructure on the basis of the following variables:

$Domains_n$ Total number of domains using NS1/Dyn on day n.

$Exclusive\_Domains_n$ Total number of domains exclusively using NS1/Dyn on day n.

$Nonexclusive\_Domains_n$ Total number of domains that are non-exclusively using NS1/Dyn on day n.

$To\_Exclusive_n$ Total number of domains that move from being non-exclusive to exclusive users of NS1/Dyn on day n.

$To\_Nonexclusive_n$ Total number of domains that move from being exclusive to non-exclusive users of NS1/Dyn on day n.

$New\_Exclusive_n$ Total number of *new* domains that became a new exclusive users of NS1/Dyn on day n (did not use NS1/Dyn on day $n-1$).

$New\_Nonexclusive_n$ Total number of *new* domains that became a new non-exclusive users of NS1/Dyn on day n (did not use NS1/Dyn on day $n-1$).

$Ex\_Exclusive_n$ Total number of exclusive domains that stopped using NS1/Dyn on day n.

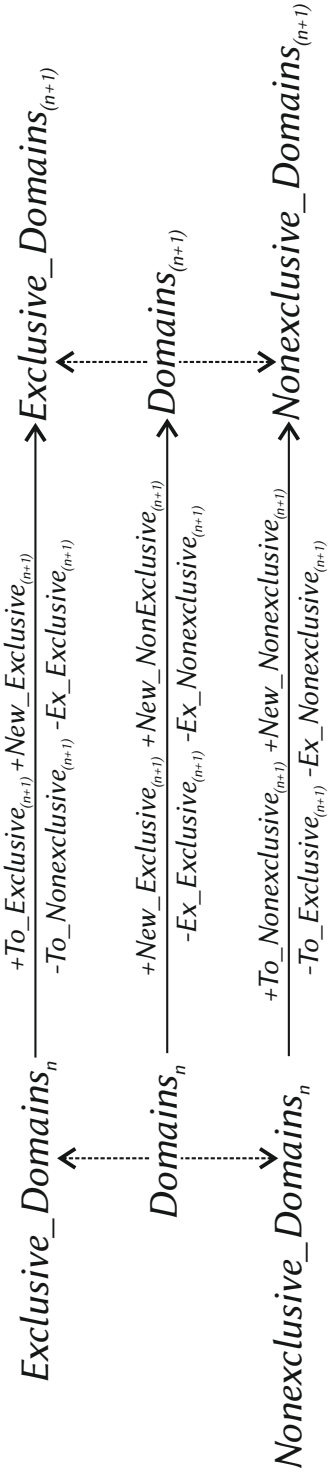


Figure 3.2: Relationship between the behaviour variables showing the changes in variable from day n to day $n+1$.

$Ex\_Nonexclusive_n$ Total number of non-exclusive domains that stopped using NS1/Dyn on day n.

$Ex\_Domains_n$ Total number of domains that stopped using NS1/Dyn on day n.

Figure 3.2 shows the relationship between the behaviour variables. Daily measurements of each of the behavioural variables provide us with a time series. In order to analyse this time series we calculate the daily change and 10-day cumulative average of the behavioural variables. For example, the change in variable $Domains_n$ represented by variable $\Delta Domains_n$ can be calculated with the help of Equation 3.1.

$$\Delta Domains_n = Domains_n - Domains_{(n-1)} \tag{3.1}$$

Calculating a 10-day cumulative average for the behavioural variables helps us to measure the net behaviour over a 10 day period [181]. It also filters any short term effects of random events from the time series. A 10-day cumulative average variable will not show changes due to an event whose effects disappear in less than 10 days. The use of cumulative averaging of time series is common practice in statistics to filter out noise. We can calculate the net cumulative average of a behaviour variable for day i as shown in Equation 3.2.

$$Cumulative\_Variable_i = \frac{1}{10}\sum_{n=0}^{9} Behaviour\_Variable_{i-n} \tag{3.2}$$

3.3.3.2 Trend and event period

The trend period is the interval before the attack date that we analyse to study the usual tendency of behaviour variables. This gives us a measure of behavioural variables without the influence of a large DDoS attack event. In this chapter, we study the usual behaviour of the behaviour variables for 200 days before the DDoS attacks. The trend period considered by us is consistent with the studies that analyse the impact of events on stock prices of the event stakeholders [9]. Similarly, the event period is the interval after the attack date that we analyse to study the deviations from the usual tendency (measure of behavioural variables under the influence of a large DDoS attack event). For this study we chose an event period of 20 days. Relatively soon after the second attack event we are analysing (i.e. the Dyn attack on 21st October 2016) a news article regarding the sale of Dyn to Oracle was published (on 11th November 2016). As this is another major event that may influence customers of Dyn, we

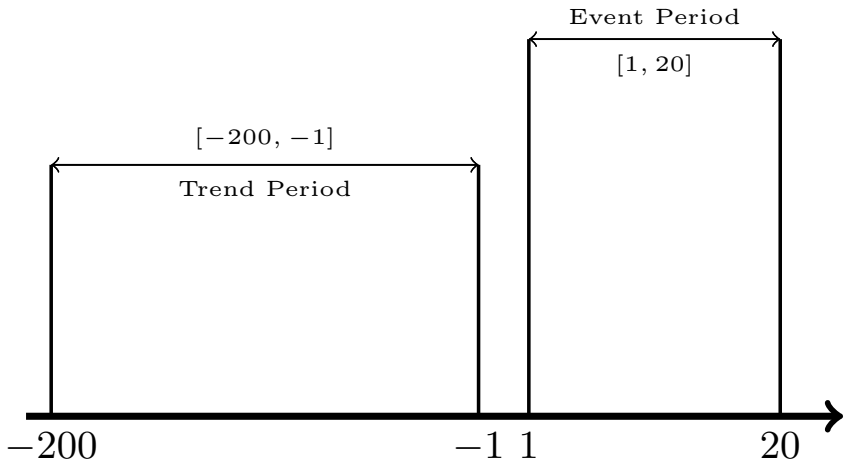


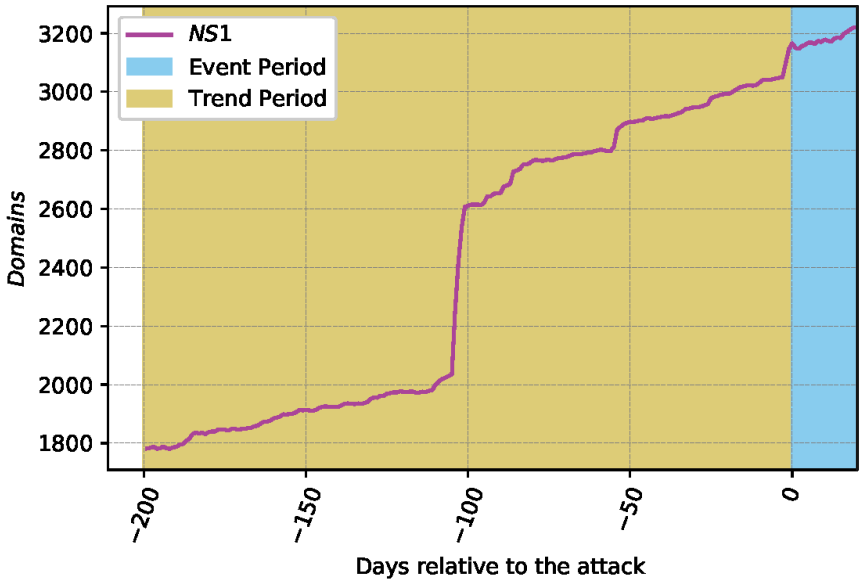
Figure 3.3: Trend and event periods.

run into the risk that any analysis of the behaviour variables beyond the 20 day window will be biased as it will also show effects that are a consequence of the takeover by Oracle. In order to keep this event window consistent for both the measurements we consider a 20 day event period for NS1 as well.

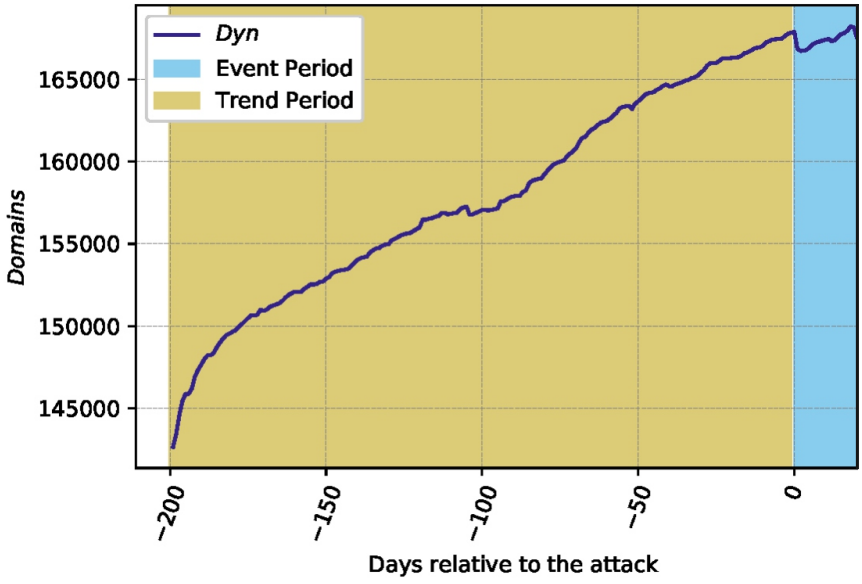
3.3.3.3 Measurement of behaviour variables

We measure the behaviour of domains by calculating the daily values for the behaviour variables that are described before. We do this with the help of the OpenINTEL dataset. $Domains_n$ is computed for each day of both the trend and the event period on the basis of the number of domains having Dyn or NS1 name server addresses in their `NS` records. If a domain had only Dyn/NS1 `NS` addresses then it was counted in variable $Exclusive\_Domain_n$, else it was counted in variable $Nonexclusive\_Domain_n$. We also calculated the daily changes in these variables as explained previously. We plot a time series of each of these variables in Figure 3.5. We discuss the interpretation of these plots in Section 3.4.1.

Next, we measure the activity of these domains on the basis of the difference in the domains using Dyn/NS1 on two consecutive days. If a Domain was a user of Dyn or NS1 on day $n-1$ but not a user on day n we count it in variable $Ex\_Exclusive_n$ or $Ex\_Nonexclusive_n$ depending on the state of the domain on day $n-1$. For example, if a domain `www.example.com` is an exclusive user of Dyn on day $n-1$ but does not use the services of Dyn on day n, then it

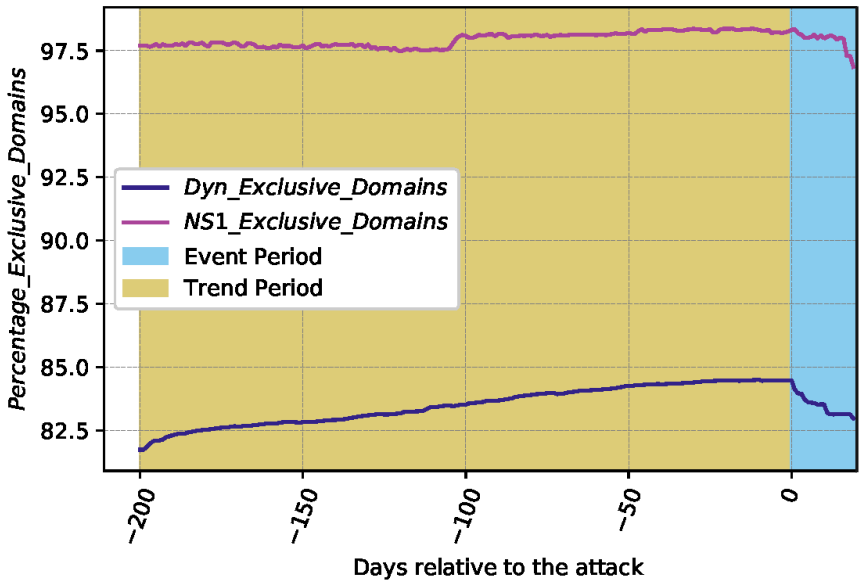


(a) Total domains NS1.

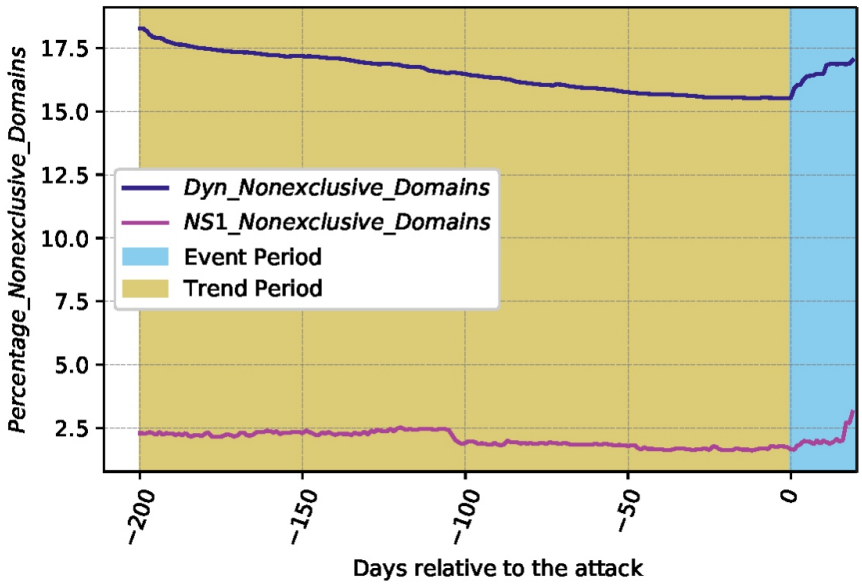


(b) Total domains Dyn

Figure 3.4: Total domains using NS1 and Dyn.

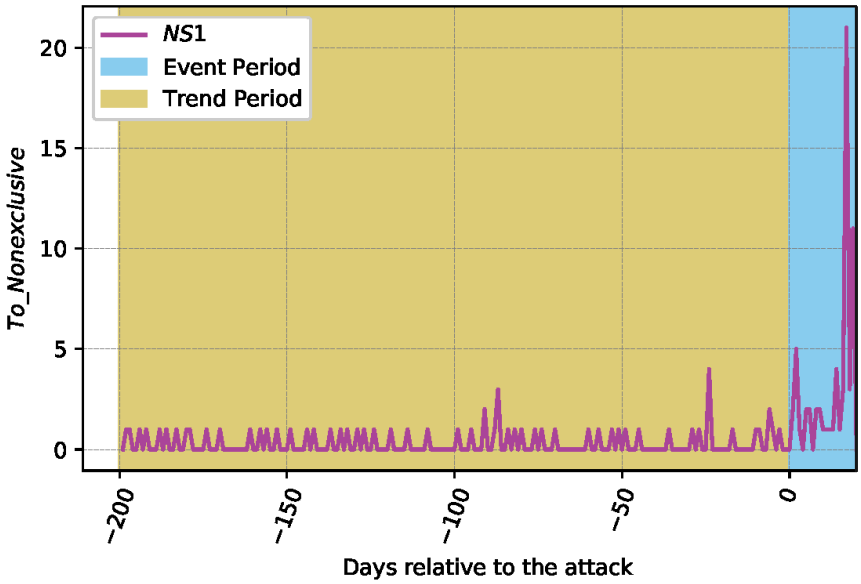


(a) Exclusive domains.

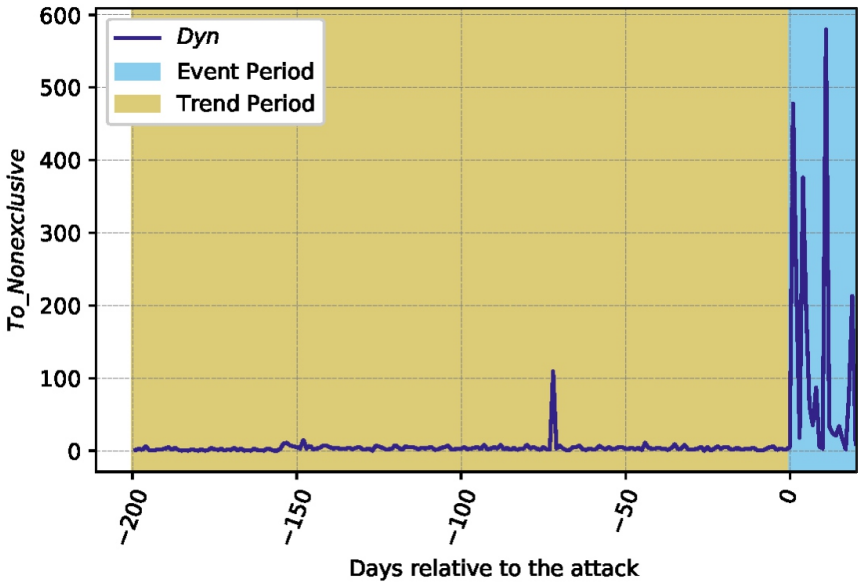


(b) Non-exclusive domains.

Figure 3.5: Time-series of behaviour variables.

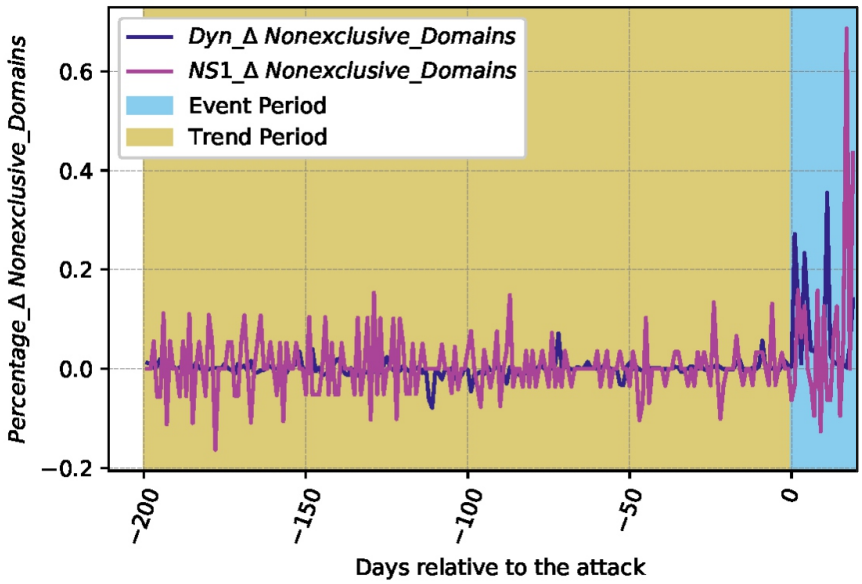


(a) Domains that become Non-exclusive for NS1.

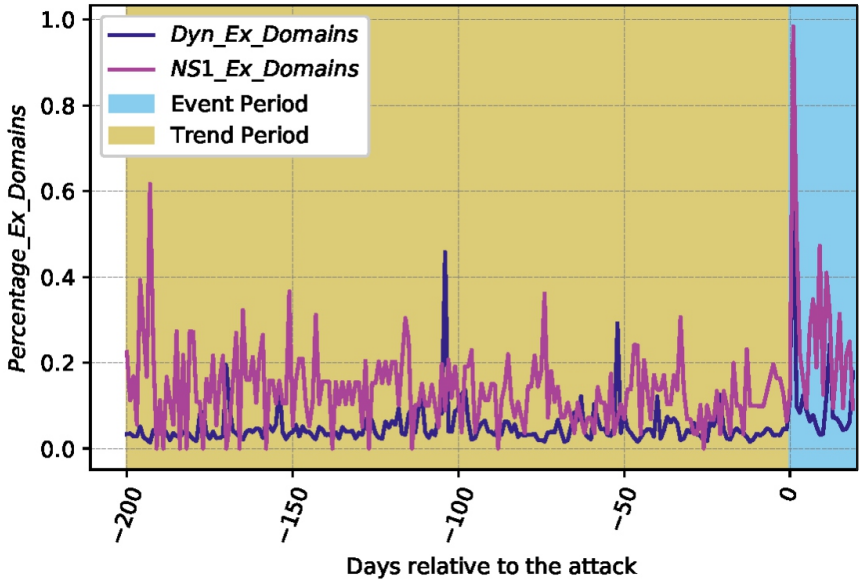


(b) Domains that become Non-exclusive for Dyn.

Figure 3.6: Time-series of behaviour variables (cont.).

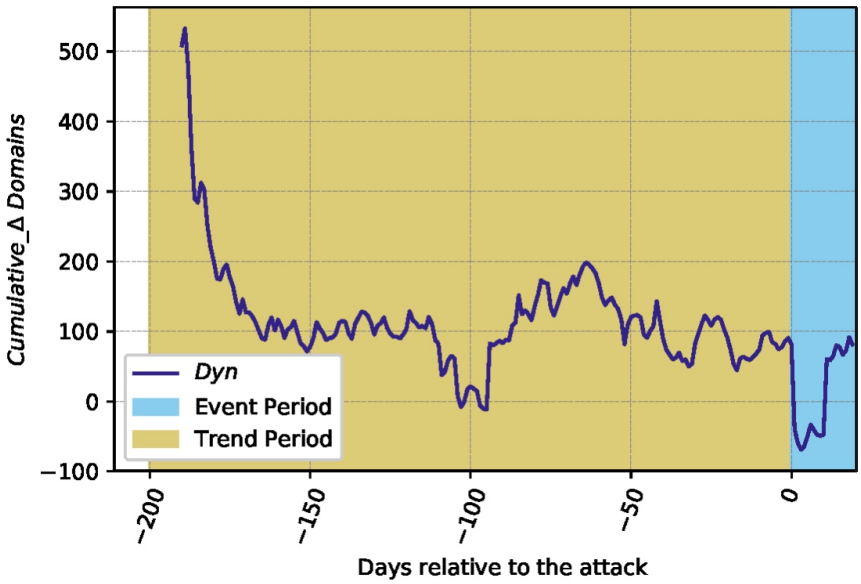


(a) Change in Non-exclusive Domains.

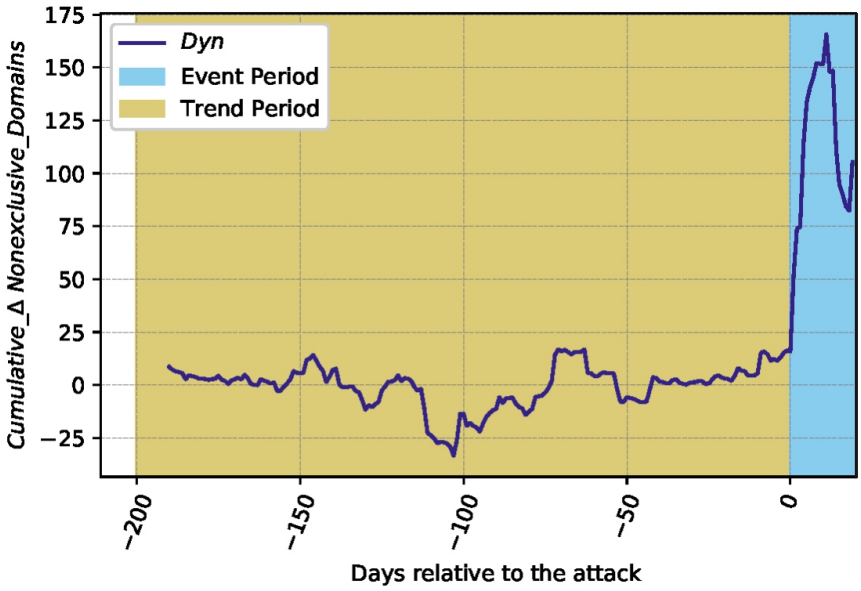


(b) Total Ex-Domains.

Figure 3.7: Time-series of behaviour variables (cont.).



(a) Cumulative change in Domains for Dyn.



(b) Cumulative change in Non-exclusive Domains for Dyn.

Figure 3.8: Time-series of behaviour variables.

will be counted in variable $Ex\_Exclusive_n$. In another case, if a domain was exclusive on day $n-1$ and non-exclusive on day n, we count it in variable $To\_Nonexclusive_n$. If a domain moved from being non-exclusive to exclusive on the next day we count it in variable $To\_Exclusive_n$. Some new domains also start using services of Dyn or NS1 each day, we count them in variable $New\_Exclusive_n$ or $New\_Nonexclusive_n$ depending on their joining status.

3.4 Analysis and results

We study the change in behaviour of the domains in three stages. First, we present the time series analysis of behaviour variables in Section 3.4.1. Then, we examine the statistical significance of the changes observed, in Section 3.4.2. Finally, we study the choice of secondary DNS service provider for the domains that become non-exclusive in Section 3.4.3.

3.4.1 Observations

Due to the DDoS attacks on May 16$^{\text{th}}$ and October 21$^{\text{st}}$, 2016 the service provided by NS1 and Dyn respectively was interrupted and the availability of the domains that used NS1/Dyn was threatened. In this section we discuss the interpretation of the changes in time series of the behaviour variables observed during the event period for both NS1 and Dyn.

Figures 3.4a and 3.4b show a sudden drop in number of domains using NS1 and Dyn just after the DDoS attack. These figures also show that NS1 was a much smaller MDNS (in terms of number of domains) than Dyn. The drop in the case of Dyn is much more observable than in the case of NS1. The drop shows that some domains stopped using the services of NS1 and Dyn after the attack and moved to a different MDNS. However, we observe a recovery pattern after the attack as the total number of domains starts increasing again a day after the attack. This indicates that some of the domains that stopped using NS1/Dyn, return when the services provided by NS1/Dyn are no longer affected due to a DDoS attack. At this point we would also like the reader to consider the analysis shown in Appendix 3.A, which shows that the return behaviour of the domains which stopped using Dyn over a longer period of 40 days after the attack. As all the domains who leave Dyn on the day of attack do not return together on a single day, we believe that the decision to use a different DNS service provider was that of the customer of Dyn (owner of the domain) rather than that of Dyn itself.

On the other hand, a similar recovery pattern is not seen in the case of *Exclusive_Domains* (Figure 3.5a). The lack of recovery pattern for domains

that use NS1 and Dyn exclusively can be attributed to the sudden and continuous rise in the number of domains using NS1 and Dyn non-exclusively. This sudden rise in the number of non-exclusive domains can be seen in Figure 3.5b. This shows that exclusive customers start using services from additional providers (become non-exclusive) in order to diversify the risk posed by DDoS attacks on their MDNS provider. The increase in the number of domains using NS1 and Dyn non-exclusively can be more clearly observed with the help of Figures 3.6a and 3.6b respectively. The notable change in preference of domains from using NS1/Dyn exclusively to non-exclusive use of their services after the attack can be clearly observed in Figure 3.7a. The percentage of total domains that choose to be non-exclusive in a single day in the event period is considerably higher than the trend period for both attacked MDNS providers.

Figure 3.7b shows a large number of domains leaving NS1/Dyn after the attack. During the event period, in case of NS1, 63.5% of the total domains that left using its services were exclusive users. In case of Dyn, 96.7% of the total users that stopped using its services during the event period were exclusive.

Zooming in on the larger of the two attack events, on Dyn, we can see the severity of the impact of the DDoS attack on Dyn with the help of the time series of cumulative variables. In Figure 3.8a we observe a strong negative cumulative impact on the total number of domains using Dyn in the event period (relative to the trend period). The only negative dip in the trend period can be attributed to a large number of non-exclusive domains leaving Dyn in the period 80 to 120 days before the attack (July-August 2016) as seen in Figure 3.8b. Contrastingly, in the event period we observe a sharp increase in the number of non-exclusive domains in Dyn. This behaviour is consistent and helps us re-emphasise the fact that domains tend to become non-exclusive users of an MDNS provider after the attack.

3.4.2 Statistical significance of the change in behaviour variables

With the help of the time-series plots we can observe the changes in the behavioural variables. In this section we test for statistical significance of the changes observed in the time series for both MDNS providers. The null hypothesis considered to examine the change in behaviour of the domains is as follows:

H_{a1}: There is no change in the behaviour of domains that use an MDNS provider after a DDoS attack.

In context with the measurement variables considered in this study we can reformulate the null hypothesis as follows:

Variable	Trend Period Mean		Event Period Mean		t-statistic	
	Dyn	NS1	Dyn	NS1	Dyn	NS1
Δ *Domains*	127.05	6.87	-9.545	3.42	2.229*	1.45
Δ *Exclusive_Domains*	126.985	6.80	-127.82	1.42	3.16*	2.18*
Δ *Nonexclusive_Domains*	0.065	0.07	118.27	2	−3.341*	−1.42
Ex_Exclusive	66.63	2.85	212.59	5.47	−2.595*	−2.02*
Ex_Nonexclusive	10.68	0.24	7.682	3.19	1.93	−7.32*
New_Exclusive	194.29	9.68	195.4	8.90	-0.057	0.40
New_Nonexclusive	10.07	0.29	15.32	3.19	−2.49*	−8.1*
To_Nonexclusive	3.8	0.3	114	3	−3.12*	−2.57*
To_Exclusive	3.1	0.27	3.36	1	-0.44	-5.1

*p -value ≤ 0.05

Table 3.2: Results of T-test on behavioural variables.

H_{a2}: There is no change in the mean of behaviour variables in the trend and the event periods.

We evaluate the null hypothesis by comparing the mean values of behavioural variables for both MDNS providers in the trend and the event period with the help of a t-test [169]. We consider the change in variables with a p -value ≤ 0.05 to be statistically significant. Table 3.2 shows the test statistics for each variable.

We find that the mean values for the change in total domains and change in exclusive domains during the trend period were significantly (statistically) higher than during the event period. The negative mean values for daily change in domains and daily change in exclusive domains shows that domains leave Dyn after the attack in the event period. On the other hand, the number of domains using Dyn non-exclusively witness a significant growth in the event period. We notice a similar statistically significant increase in the non-exclusive users on NS1.

We also find the change in variable *Ex_Exclusive* to be statistically significant for both Dyn and NS1. This demonstrates that an abnormally large number of exclusive domains stopped using their services in the event period.

We do not observe any change in the average number of new exclusive domains joining the attacked MDNS in the event period. However, we notice an abnormally large number of new non-exclusive domains joining the MDNS. This can be an indication that a number of exclusive domains that leave Dyn after the attack returned as non-exclusive. Looking at the results of the t-test for variables *To_Nonexclusive* and *To_Exclusive* we can say that a large num-

ber of exclusive domains became non-exclusive in the event period but the trend of non-exclusive domains becoming exclusive did not really change.

3.4.3 Choice of secondary DNS

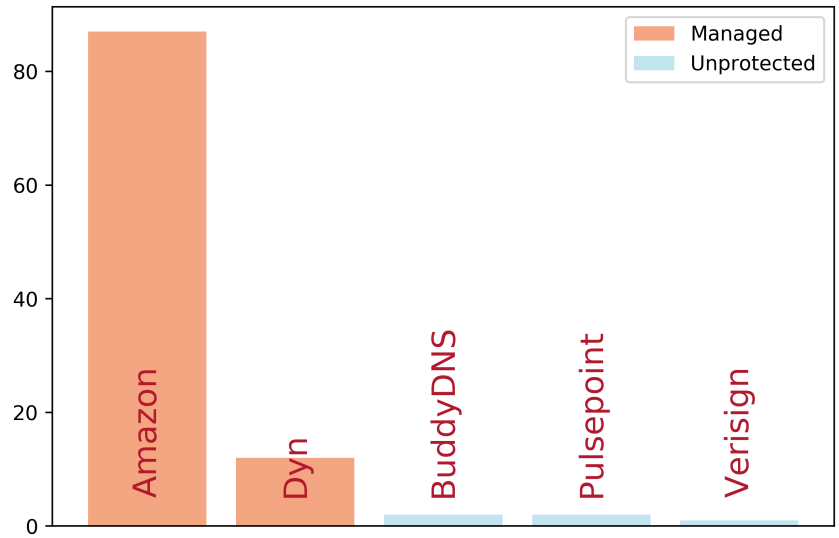
Given that a significant number of Dyn and NS1 customers become non-exclusive users, we also analysed which secondary providers they choose. In order to understand the choices made by non-exclusive domains using NS1 and Dyn before the attack we evaluate the secondary `NS` addresses in the `NS` records of these domains one day before the attack. Figures 3.9a and 3.9b show the top secondary DNS choices for non-exclusive domains using NS1/Dyn one day before each attack. Most of the very few non-exclusive domains of NS1 used another MDNS provider for secondary DNS. However, in the case of Dyn we see that a remarkable number of domains used non-managed DNS service providers as a secondary choice.

After the DDoS attacks, we can observe with the help of Figures 3.10a and 3.10b that most of the users of NS1 and Dyn that became no-exclusive over a period of 20 days after the attacks added another MDNS service provider. Since, it is highly unlikely that two MDNS service providers fail due to a DDoS attack at the same time it underlines the fact that in terms of risk management, using multiple providers is a good strategy.

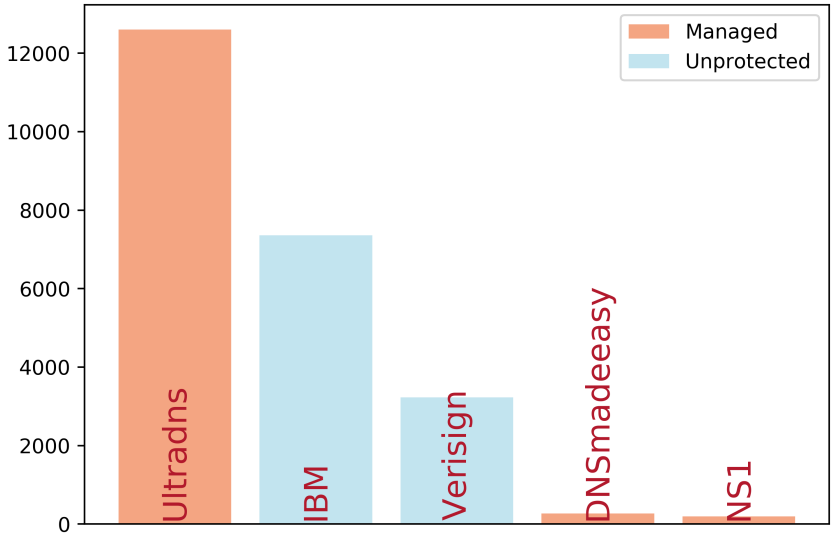
3.5 Related work

Distributed Denial-of-Service (DDoS) attacks have been the subject of intense study. Studies of the technical aspects of DDoS attacks have shown that there are myriad strategies for conducting an attack. The booter phenomenon has made DDoS attacks accessible to every one [175]. Studies have also shown reflection and botnet based attacks to be extremely effective [209]. Characterisation of DDoS attacks has been done by studies on the basis of intensity, source and event ports [105, 137, 119, 199]. At the same time, various DDoS mitigation techniques have been suggested by multiple researchers [136, 219]. Studies have also been conducted in order to evaluate the effectiveness of mitigation techniques [54].

Focusing specifically on the DNS, Moura *et al.* [139] evaluate the Nov. 30 and Dec. 1, 2015 events on the Root of the DNS. They show that large attacks can overwhelm some sites of some root letters. In addition, they also provide evidence that high traffic on one service can result in collateral damage to other services, possibly in the same data centre. In the event analysed in that study the overall DNS service was resilient to the DDoS attack. In case of the events

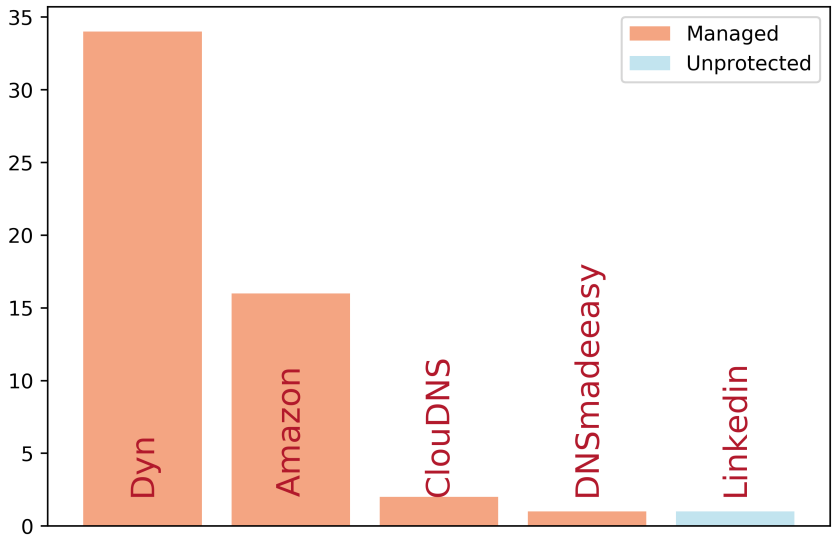


(a) Top secondary DNS choices for NS1.

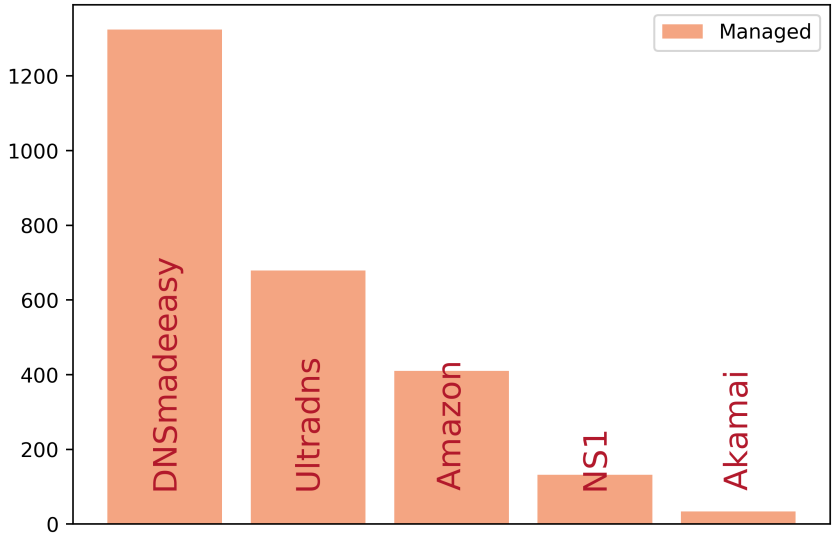


(b) Top secondary DNS choices for Dyn.

Figure 3.9: Secondary DNS choices for attacked MDNS (before attack).



(a) Top changes after attack NS1.



(b) Top changes after attack Dyn.

Figure 3.10: Secondary DNS choices for attacked MDNS (after attack).

evaluated by us in this chapter (Dyn and NS1 attack), the overall DNS service provided by Dyn and NS1 was not able to absorb the attack.

Jonker *et al.* [107] study the adoption of DDoS protection services in general, using active DNS measurements. They observe that there are generally three strategies for mitigation in the face of an attack, two that use redirection via the DNS and one that redirects traffic using the Border Gateway Protocol (BGP). Furthermore, Jonker *et al.* observe that there are two general types of customer behaviour: one group of customers uses on-demand DDoS protection, only switching it on in case of an actual attack. The other group chooses to enable DDoS protection permanently, always routing traffic via the DDoS protection service. In this chapter, we study a particular case of the latter, in which the DNS for a customer is supposed to always be protected against DDoS attack by making use of Dyn or NS1's managed DNS service.

Finally, industry reports [206, 27] from DDoS protection firms have studied the impact of DDoS attacks on the customers of Dyn. But in contrast to the framework used by us in this chapter, they did not consider the domain segmentation (exclusive and non-exclusive) or the return behaviour of domains. These measurements form an integral part of such an analysis.

To the best of our knowledge, our chapter is the first to empirically measure the direct impact of a successful DDoS attack on the behaviour of the victim's customers.

3.6 Conclusions

In this chapter, we study the effects of a successful DDoS attack on a managed DNS provider. Using data from the OpenINTEL platform, we test if the fallout of a successful attack results in changes in customer behaviour. We introduce a novel framework that measures the decisions a customer of an MDNS provider can take. We then use this model to analyse the change in customer behaviour after a successful DDoS attack.

According to the observations from our datasets, we can identify two types of customer behaviour. Most Dyn and NS1 customers use the MDNS *exclusively*, that is: they only configure authoritative name servers provided by Dyn or NS1 for their domains. A small, but non-trivial fraction of customers use the MDNS services *non-exclusively*. This means that they configure some of the authoritative name servers for a domain to be from Dyn/NS1, and some from other providers, or managed by themselves. In the period leading up to the attack, we observe a gradual growth in the use of services provided by both Dyn and NS1. Furthermore, we observe no significant changes in customer

behaviour from using Dyn/NS1 exclusively to non-exclusively for both existing and new customers. If we then focus on the aftermath of the attack, we observe a number of statistically significant changes:

- A significant number of MDNS customers that were using Dyn's or NS1's service exclusively switch to non-exclusive use in the aftermath of the attack. Furthermore, our analysis shows that in most cases this change is lasting, that is: in the period analysed the majority of domains that switch from exclusive to non-exclusive remain in that configuration.
- We observe no significant changes in the behaviour of Dyn customers that were already non-exclusive users. While this result was to be expected – since they were likely not affected by the attack – it underlines the fact that in terms of risk management, using multiple providers is a good strategy.
- Lastly, we observe that most of the newly non-exclusive customers after the attack on Dyn and NS1 use an MDNS service provider as a secondary DNS to further reduce the risk of downtime.

Summarising, our study shows that our model captures significant changes in customer behaviour in the wake of a large, successful DDoS attack on a provider whose business model includes protecting customers against such attacks. Furthermore, these changes in behaviour are not just temporary, but we observe lasting changes in customer behaviour and permanent loss of customers.

3.7 Future work

In this chapter we showed *that* there is a change in customer behaviour, and especially that customers choose to hedge their bets by starting to use multiple managed DNS service providers. The next step is to understand *why* customers change their behaviour, and especially why they make specific choices, such as starting to use multiple providers. Intuitively, one might assume that using more than one provider leads to a cost increase, so it would be valuable to understand if this is the case, and if so, what rationale customers have to make this choice, and whether they have an upper bound on an increase in cost. To study this, we believe it is necessary to conduct a qualitative study, where decision makers at organisations affected by an attack are consulted about their decision-making process. The outcome of such a study may also be valuable for future decision making when organisations plan to outsource DNS to a managed DNS provider, and may even have wider applicability in cloud outsourcing strategies. It is clear

from the examples of NS1 and Dyn that when taking into account that even large providers may be taken down by DDoS attacks that there are serious risks when outsourcing to a single provider.

Appendix 3.A More on Dyn

In the appendix to this chapter, we zoom into the analysis of the attack on Dyn. In the following section we analyse the return behaviour of customers that stopped using the services of Dyn in the trend and event periods. Then in Section 3.A.2 we estimate the effective number of domains that Dyn might have lost due to the attack.

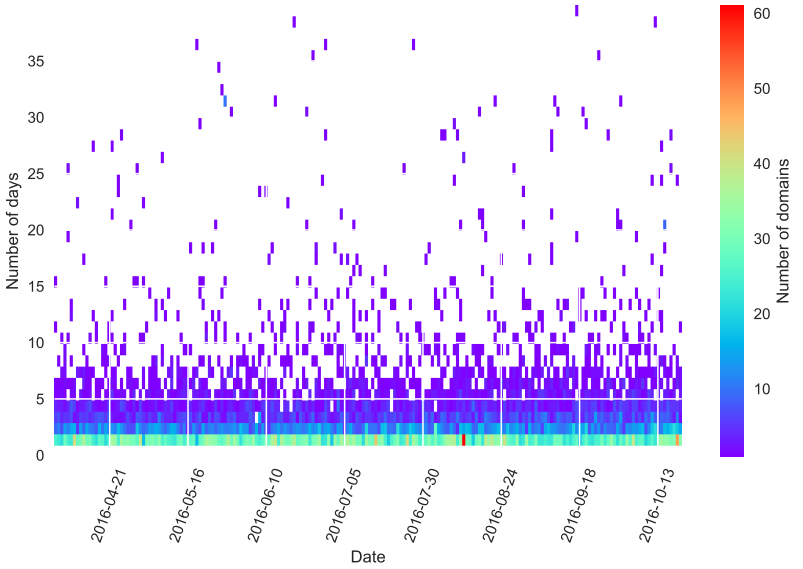
3.A.1 Return behaviour of domains

In this section we analyse the return behaviour of domains. We do this in order to determine whether the domains that stop using Dyn services return or not. We study the comeback behaviour both in the trend and in the event period in order to evaluate the change in behaviour. We demonstrate the return behaviour of domains with the help of heat-maps as shown in Figure 3.A.1. In these heat plots on the x-axis we have the dates on which the domain leaves Dyn and on the y-axis the number of days after which the domain starts using Dyn services again. The colour on the plot shows the number of domains that come back n days after leaving. We count the number of domains returning within 40 days of leaving. We choose a 40 day window to account for short term contracts that ex-domains might have with other ANS service providers.

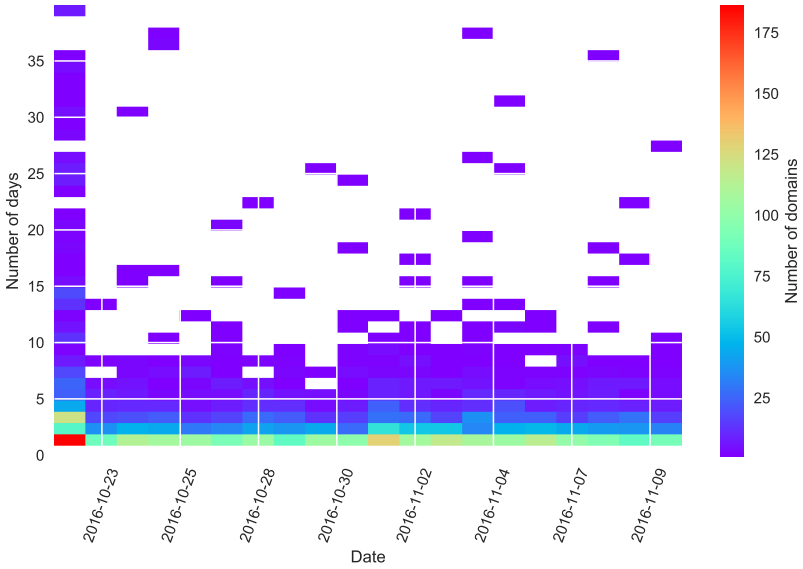
Figure 3.A.1a shows the return behaviour of domains in the trend period. In this period we observe that 48.99% of the domains that leave Dyn return within the next 40 days. Out of these, 95.36% of domains return within 7 days of leaving. In the heat plot (Figure 3.A.1a) we can see a dark band due to domains returning within the first 7 days.

In Figure 3.A.1b we can observe the return behaviour of domains in the event period. In this period only 44.42% of the total domains that left returned within 40 days. In the heat plot (Figure 3.A.1b) we observe a huge number of domains returning a day after the attack. This shows that these domains moved to a different ANS due to service interruptions on the day of attack and returned a day after when the attack subsided. We also observe that the domains that left Dyn on the day of the attack, return continuously over the 40 day period. This shows that some domains wait for the threat of any successive attacks to pass before returning.

In the trend period of 200 days there were 23,237 occurrences when a domain stopped using Dyn. However, in the event period of 20 days there were 9,766 occurrences when a domain stopped using Dyn. 5,487 domains that stopped using Dyn services in the event period never returned with in the short term (40 days).



(a) Return Behaviour: Trend Period



(b) Return Behaviour: Event Period

Figure 3.A.1: Return Behaviour

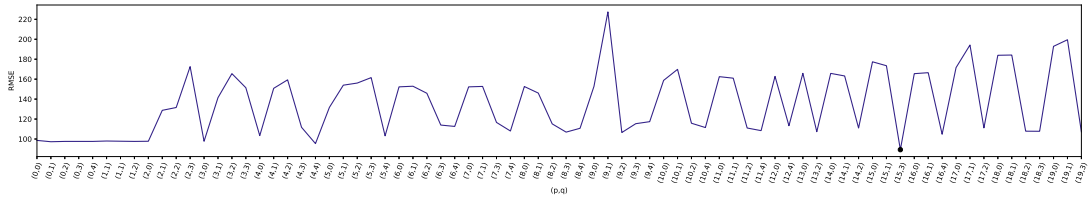


Figure 3.A.2: RMS error for each value of (p,q)

As there is constant fluctuation in the total number of domains using Dyn on each day, we feel the need for a model to determine the effective number of domains that stopped using Dyn due to the DDoS attack.

3.A.2 Estimating the effective number of domains that stopped using Dyn services

In this section, we estimate the number of domains that stopped using Dyn services due to the DDoS attack. For estimating the number of domains that stopped using Dyn due to the DDoS attack, we first need to estimate the number of domains that would have been using Dyn if there was no DDoS attack. For doing so we make use of an Auto-Regressive Integrated Moving Average (ARIMA) model [215]. An ARIMA model is a well established technique for forecasting time series (eg. demand and supply) [58, 193].

The ARIMA methodology is a generalisation of Auto-Regressive Moving Average (ARMA). ARMA is the combination of auto regression (AR) and moving average(MA). Given a time series $x(t)$ where t is an integer index and $x(t)$ are real numbers, then an ARMA(p,q) process is given by Equation 3.3.

$$\begin{aligned} x(t) = c + \phi_1 x_{t-1} + \phi_2 x_{t-2} + ...\phi_p x_{t-p} + \epsilon_t \\ - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} ... \theta_q \epsilon_{t-q} \end{aligned} \tag{3.3}$$

where c is a constant, ϕ_1, ϕ_2,...,ϕ_p are parameters for auto regression (AR) and θ_1, θ_2,...,θ_q are parameters for moving average (MA).

Statistical stationarity of a time series is a prerequisite for using ARMA as the model for forecasting. Hence we use the ARIMA(p,d,q) model in which d times difference of consecutive terms is taken in order to make the series stationary before application of the ARMA(p,q) model. Fine-tuning the ARIMA process refers to finding the best value of parameters p, d and q in order to predict the variable *Domains*.

Variable	Test-Statistic
Domains	-0.353^a
$\Delta Domains$	-3.465^b

[a] p value=0.917 [b] p value<0.05

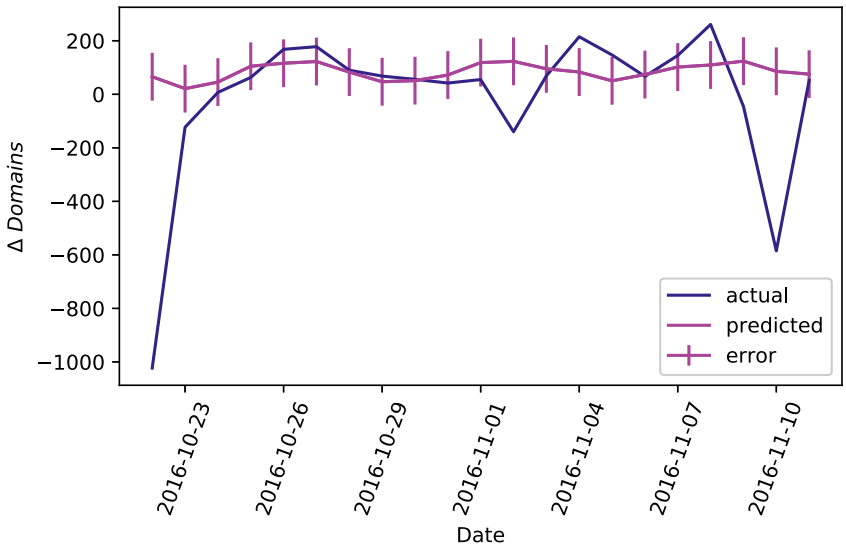
Table 3.A.1: Results of the Augmented Dickey-Fuller test

We first determine the value of parameter d (i.e. the order of difference needed to make the series stationary). To check the stationarity of behavioural variable *Domains* we apply the Augmented Dickey-Fuller test [85]. The test statistic was negative but the p value was very high as shown in Table 3.A.1. Hence, in this case we cannot reject the null hypothesis that the time series of behavioural variable *Domains* is not stationary. Now, we test the stationarity of behavioural variable Δ *Domains* (which is basically behavioural variable *Domains* with $d = 1$). In this case we found the series to be stationary at more than 95% confidence level. Hence, we can apply ARIMA with $d = 1$ in order to predict the variable *Domains* for the event period.

To determine the best values of parameters p and q we apply a number of different ARIMA models, each for parameter $p \in [0, 20]$ and $q \in [0, 4]$. Values of (p, q) in a range higher than the intervals considered in this analysis may give a marginally better estimate but will make the model computationally challenging. We fine-tune the ARIMA model by making use of the Trend Period data. We use the first 180 days (of 200 days) measurements to predict the next 20 day values and simulate the best fit ARIMA model by tuning the values of (p, q). Here, our prediction window is consistent with the Event Period defined in Section 3.3.3.2. We choose the model parameters (p, q) for which the Root Mean Square Error (RMSE) was the minimum. The RMSE for each value of (p, q) is shown in Figure 3.A.2. We find that for $p = 15$ and $q = 3$ the RMSE is the least at 89.31.

Then we use this fine-tuned model to predict the variable *Domains* for the event period. The total number of domains lost due to the attack in the event period can be estimated as the difference between actual and predicted change in variable *Domains* during the event period. Figure 3.A.3 clearly shows that in the event period the number of domains using Dyn dropped the day after the attack. Figure 3.A.3b shows the difference in the actual and predicted values of

(a) Predicted and actual *Domains*



(b) Predicted and actual change in *Domains*

Figure 3.A.3: Actual and predicted number of domains using Dyn

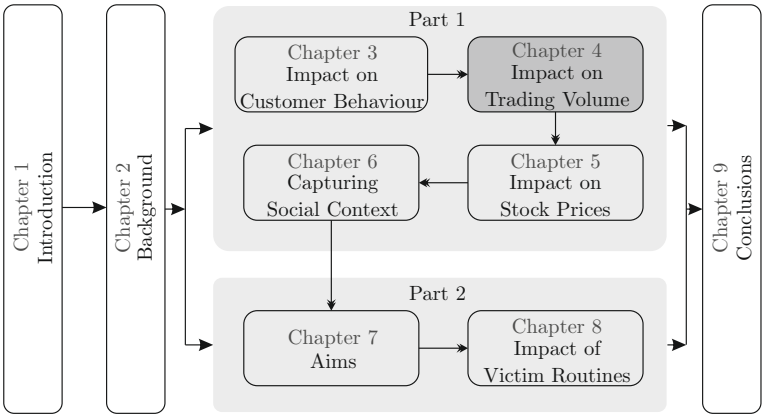
Δ *Domains*. The number of domains that stopped using Dyn services during the event period can be computed as shown in Equation 3.4.

$$domains\_stopped = \sum_{1}^{21}(\Delta\ Domains_{predicted} - \Delta\ Domains_{actual}) \tag{3.4}$$

Using the method described above we estimate total number of domains that stopped using the services of Dyn due to the DDoS attack in the event period to be 2,003 domains.

Chapter 4

Impact on Trading Volume



In this chapter, we look at another direct impact of DDoS attacks, i.e., loss of trading volume for crypto-currency exchanges. We modify the event analysis methodology to evaluate this loss. Our contributions are fourfold: Firstly, we develop an estimation model utilising ideas from behavioural finance to predict the volume of crypto-currency traded on the basis of change in prices. Secondly, we perform an event analysis to evaluate whether there is an impact of a DDoS attack on the volume traded on the exchange in 17 different cases. Thirdly, we find that in 13 cases the negative impact due to a DDoS attack is recovered within the same day by the exchange. Finally, we evaluate hourly trade data to show why in most cases the volume traded recovers within a single day.

4.1 Introduction

Another direct consequence of a DDoS attack can be on the productivity of any online platform. In this chapter, we investigate the impact of DDoS attacks on a large crypto-currency exchange. The market capitalisation of global crypto-currency markets has increased from $19 billion in the beginning of 2017 to $602 billion by the end of 2017 [46]. Crypto-currencies are digital currencies based on blockchain technology. To benefit from the increase in valuation of these digital currencies, individuals can invest with the help of online platforms known as crypto-currency exchanges. These exchanges allow their clients to buy, store (act as digital wallets) and sell crypto-currencies. The clients of these exchanges are able to trade and profit due to the fluctuation in the prices of crypto-currencies. The exchange charges them for each transaction made on its platform.

These platforms face security issues just like other online businesses. One of the biggest challenges they face is a distributed denial of service (DDoS) attack. We analyse the impact of DDoS attacks on the volume of Bitcoin traded at one such crypto-currency exchange: Bitfinex. We apply the so-called event analysis methodology to analyse this impact.

Our contributions are as follows:

- We develop an estimation model utilising ideas from behavioural finance to predict the volume of crypto-currency traded on the basis of change in prices.
- We perform an event analysis to evaluate whether there is an impact of a DDoS attack on the volume traded on the exchange in 17 different cases.
- We find that, on most occasions (13 of 17) the impact due to a DDoS attack is recovered within the same day by the exchange.
- We evaluate hourly trading data to discuss why in most cases the volume traded recovers within a single day.

Organisation of this chapter- in Section 4.2 we explain in detail the impact of a DDoS attack on the revenue stream of Bitfinex, in Section 4.3 we present the research methodology and the dataset used to test the hypothesis. In Section 4.4 we discuss the findings and elaborate on them in Section 4.5. In Section 4.7 we give the conclusions of our study.

4.2 Impact of DDoS on the revenue stream of an exchange

Just like a stock exchange, a crypto-currency exchange is a platform that facilitates trade of digital currencies. At such an exchange, a client can buy, sell and store supported digital currencies at the exchange rate. The exchange matches buyers & sellers and charges a fee for every trade made, to both parties. A DDoS attack degrades the performance of a crypto-currency exchange. In the worst case scenario, it can cause temporary unavailability of the online platform. This would mean that when the exchange is under an attack the volume of digital currency traded would decrease. As crypto-currencies can be bought from any of the hundreds of exchanges [183] that are on the web, temporary unavailability of just one of the exchanges would not have a significant impact on the price of the crypto-currency, but will have an effect on the revenues of the attacked platform. In this chapter, we analyse the impact of DDoS attacks on the volume of bitcoin traded on *Bitfinex*.

Attacks on Bitfinex: Bitfinex is a Hong Kong-based crypto-currency exchange. It was founded in December 2012 as a peer-to-peer Bitcoin exchange offering trading services all around the world. The business model of this exchange is making money from providing the matching of buyers and sellers. Bitfinex charges a fee for each trade made on the exchange.

The exchange has been a victim of DDoS attack on several occasions. In order to find the dates of attacks we make use of three different sources: 1) Bitfinex twitter feed (*@Bitfinex*), 2) Bitfinex status page [24] and 3) Google news search. To scrape all the tweets from the *@Bitfinex* twitter feed we make use of an open source python project known as Twint* [154]. We also look for mentions of DDoS attacks on Bitfinex since 2016 on Google news search and Bitfinex status page [24]. From all the sources described above we record the dates of DDoS attacks on Bitfinex. Table 4.2.1 shows the list of 17 attacks that we analyse in this chapter.

Impact on Bitfinex: A DDoS attack makes it difficult for the clients of Bitfinex to reach its online platform. This in turn affects the number of trades made on the exchange. Thus, the economic loss to the exchange will be due to the prospective trading fee that the exchange could have earned during the unavailability. Later we use this causal relationship between a DDoS attack

*It is an advanced Twitter scraping & OSINT tool written in Python that does not use Twitter's API, and allows to scrape a user's followers, following, Tweets and more while evading most API limitations.

Table 4.2.1: Table showing the list of reported attacks on Bitfinex and the damage caused.

No.	Date	Target	Damage	Source
1	20/01/2016	Bitfinex	Temporary Unavailability	Status Page
2	04/06/2016	Bitfinex/BitGo	Temporary Unavailability	Twitter and Status Page
3	07/06/2016	Bitfinex	Degraded Performance	Status Page
4	20/06/2016	Bitfinex	Temporary Unavailability	Status Page
5	26/07/2016	Bitfinex	Temporary Unavailability	Status Page
6	09/11/2016	Bitfinex	Temporary Unavailability	Status Page
7	16/11/2016	Bitfinex	Temporary Unavailability	Status Page
8	21/02/2017	Bitfinex	Degraded Performance	News,Twitter and Status Page
9	12/06/2017	Bitfinex	Degraded Performance	News, Twitter and Status Page
10	21/08/2017	Bitfinex	Degraded Performance	News and Status Page
11	26/11/2017	Bitfinex	Temporary Unavailability	Twitter and News
12	04/12/2017	Bitfinex	Degraded Performance	News, Twitter and Status Page
13	05/12/2017	Bitfinex	Degraded Performance	News, Twitter and Status Page
14	12/12/2017	Bitfinex	Temporary Unavailability	News, Twitter and Status Page
15	17/12/2017	Bitfinex	Degraded Performance	News, Twitter and Status Page
16	31/12/2017	Bitfinex	Temporary Unavailability	News, Twitter and Status Page
17	05/06/2018	Bitfinex	Temporary Unavailability	News, Twitter and Status Page

on the exchange and its impact on the commission earned by the exchange to formulate our hypothesis.

4.3 Methodology

We explain our method to evaluate the impact of DDoS attacks on a crypto-currency exchange. First we elaborate on the datasets used for conducting this study. Next, we explain the event study methodology [132] used to measure the impact of the attack. Finally, we develop our hypotheses and discuss the method of hypothesis testing.

Dataset: Two datasets were collected from *www.cryptodatadownload.com*. Both datasets provide information on the bitcoin volume traded on Bitfinex. The difference is the granularity: one provides the daily amount of volume traded on the exchange with the highest and the lowest price of the day, and the other dataset provides the same information at an hourly interval. Both longitudinal datasets start on 01-12-2015 and end on 16-06-2018. We pre-processed the datasets to remove any anomalies. For instance, the security of Bitfinex was breached and $72 million in Bitcoins were stolen on 2$^{\text{nd}}$ August 2016 [24, 18]. All trading was halted for 7 days and normal operations were resumed on the 10th of August 2016. Hence, we observe no trades on the exchange during this period.

Values for the following variables are provided by the dataset: $VolumeFrom$, $VolumeTo$, P_{High} and P_{Low}. Equation 4.1 describes the relationship between these variables and the values of variables $ActV_t$ and ΔP_t.

Event Study Analysis: To evaluate the impact of certain events on companies' stock prices a method called event analysis has been designed in finance and economics. Mackinlay [132] has discussed the method for conducting a classical event study. Abhishta, Joosten and Nieuwenhuis [9] have proposed a more robust event study method especially useful in cases when the *returns* and *abnormal returns* are not normally distributed. They have shown that the classical method of event study in the case of non-normal *abnormal returns* leads to overestimation/underestimation of losses/gains [10].

To analyse the impact of DDoS attacks on the volume of Bitcoin traded on Bitfinex, we take the following steps:

Step 1: Define estimation and event periods.

Step 2: Using the data in the estimation period to compute a model for the prediction of the volume of crypto-currency traded on Bitfinex.

Step 3: Define a null hypothesis.

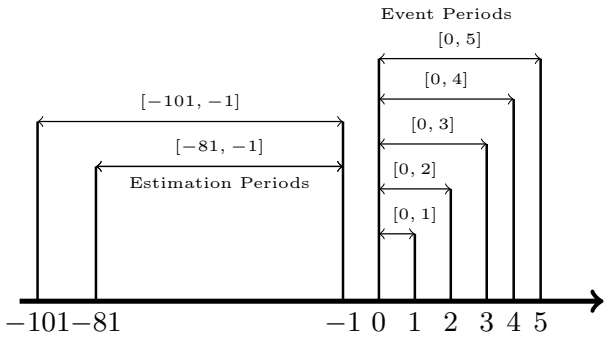


Figure 4.3.1: Estimation and event periods.

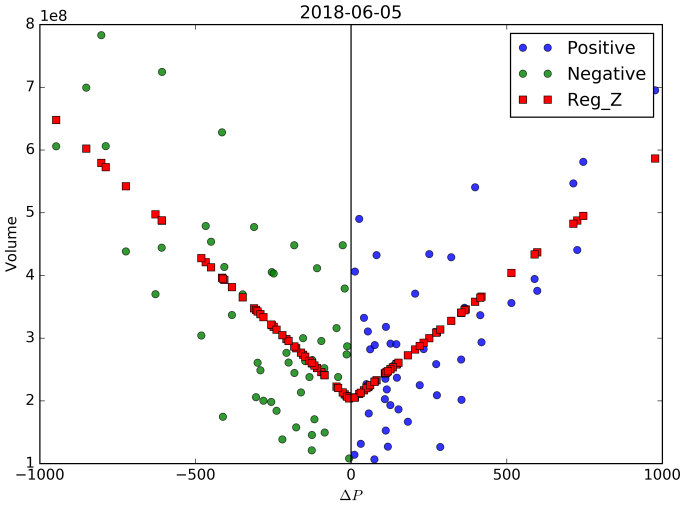
Step 4: Calculate values of *abnormal volume* and *cumulative abnormal volume* in the estimation period.

Step 5: Generate an empirical distribution by bootstrapping [60].

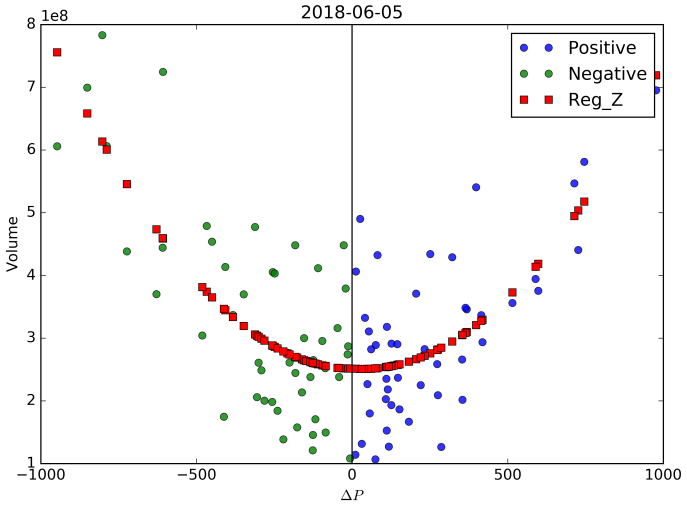
Step 6: Use the empirical distribution for hypothesis testing.

For analysing each attack event we divide our dataset in two parts as shown in Figure 4.3.1. The common practice for stock market event studies is to use 120 days for the estimation period [132]. However, as the crypto-currency market is more volatile than the stock market we consider two slightly shorter estimation periods of 100 days and 80 days. The data in the estimation period are used to estimate the parameters of the model. For further analysis we chose the estimation period that yields the highest value for the coefficient of determination (adj. R^2).

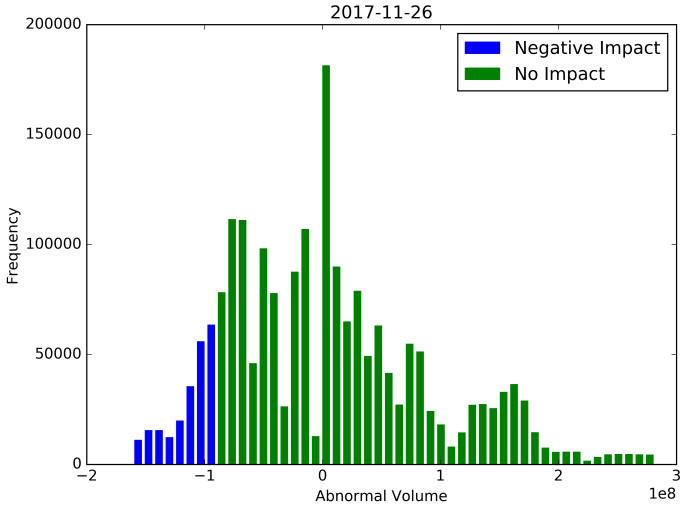
We use an additive model to predict the usual quantity of Bitcoin traded on the exchange. To determine the best estimation model we test the goodness of fit for the following two models, linear one and a quadratic one. If $|\Delta P_t|$ represents the absolute value of price change and V_t represents the volume of Bitcoin traded on day t then the linear and quadratic estimation models explaining the relationship of volume and absolute price change can be given by Equations 4.2 and 4.3 respectively. The variables $|\Delta P_t|$ and V_t can be calculated as shown in Equation 4.1 where P_t is the price of Bitcoin on day t and is calculated as the average of highest (P_{High}) and lowest (P_{Low}) price of the day. The parameters α_i, β_i and γ_i can be estimated using ordinary least square (OLS) on the basis of the data in the estimation period for an event i. OLS is chosen based on the



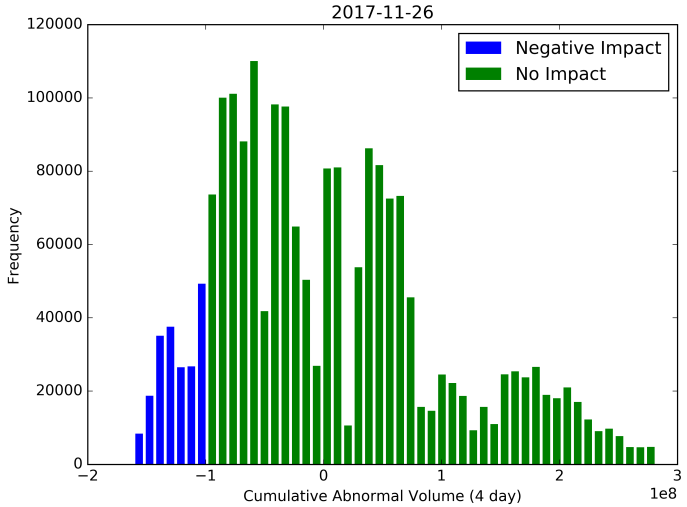
(a) Linear OLS model for the attack on 05-06-2018.



(b) Quadratic OLS model for the attack on 05-06-2018.



(c) Empirical distribution for Abnormal Volume.



(d) Empirical distribution for 4 Day Cumulative Abnormal Volume.

Figure 4.3.2: OLS models showing the dissimilar effect of negative and positive price changes and empirical distributions.

study by Karafiath *et al.* [111]. This study compared several generalized least squares and first & second order autoregressive structures and it concluded that these do not offer a material improvement over OLS in the context of event studies.

$$\begin{aligned} P_t &= \frac{P_{High} + P_{Low}}{2} \\ |\Delta P_t| &= P_t - P_{t-1} \end{aligned} \tag{4.1}$$

$$V_t = \beta_i |\Delta P_t| + \gamma_i \tag{4.2}$$

$$V_t = \alpha_i |\Delta P_t|^2 + \beta_i |\Delta P_t| + \gamma_i \tag{4.3}$$

In behavioural finance negative and positive price changes have been shown to have dissimilar effects on the volume traded [113]. This effect can also be observed in the case of crypto-currencies. In Figure 4.3.2 we show the OLS models for the attack on *05-06-2018* computed with the help of Equations 4.2 and 4.3. The plots show *Reg_Z* as the curve representing the OLS model for positive and negative price changes. In Figure 4.3.2a, we can clearly observe that in the case of linear model the slope for the resulting curve is different for positive and negative price changes. A similar effect is seen in case of the quadratic model as shown in Figure 4.3.2b.

$$\begin{aligned} {z_t}^{+ve} &= \max\{0, \Delta P_t\} \\ {z_t}^{-ve} &= \min\{0, \Delta P_t\} \end{aligned} \tag{4.4}$$

$$V_t = {\beta_i}^{+ve} {z_t}^{+ve} + {\beta_i}^{-ve} {z_t}^{-ve} + \gamma_i \tag{4.5}$$

$$\begin{aligned} V_t = {\alpha_i}^{+ve} ({z_t}^{+ve})^2 + {\beta_i}^{+ve} {z_t}^{+ve} \\ + {\alpha_i}^{-ve} ({z_t}^{-ve})^2 + {\beta_i}^{-ve} {z_t}^{-ve} + \gamma_i \end{aligned} \tag{4.6}$$

Here, we use functions as shown in Equations 4.5 and 4.6 to accommodate for the dissimilar effect of positive and negative price changes on the volume traded. Where, ${z_t}^{+ve}$ and ${z_t}^{-ve}$ represent a set of positive and negative price changes respectively and are defined as shown in Equation 4.4. We estimate the

Table 4.3.1: Table showing the adj. R^2 values for three tested models.

Attack Date	LM (100 days)	QM (100 days)	QM (80 days)
20/01/2016	0.71	0.78	0.79
04/06/2016	0.78	0.79	0.85
07/06/2016	0.75	0.76	0.80
20/06/2016	0.83	0.85	0.85
26/07/2016	0.80	0.81	0.80
09/11/2016	0.77	0.78	0.84
16/11/2016	0.71	0.78	0.82
21/02/2017	0.69	0.71	0.80
12/06/2017	0.66	0.72	0.65
21/08/2017	0.49	0.51	0.54
26/11/2017	0.36	0.42	0.40
04/12/2017	0.29	0.39	0.31
05/12/2017	0.29	0.38	0.22
12/12/2017	0.66	0.70	0.69
17/12/2017	0.64	0.65	0.65
31/12/2017	0.74	0.71	0.64
05/06/2018	0.48	0.54	0.53

LM: Linear Model and QM: Quadratic Model

coefficients ${\alpha_i}^{+ve}$, ${\alpha_i}^{-ve}$, ${\beta_i}^{+ve}$, ${\beta_i}^{-ve}$ and γ_i for event i for Equations 4.5 and 4.6 using OLS.

In Table 4.3.1 we show the values of coefficient of determination (adj. R^2) for three of the tested models. In most cases we observe that a quadratic model performs significantly better (higher values of adj. R^2) than a linear model. Also, while comparing on the basis of the estimation periods, we find that a shorter estimation period of 80 days improves the adj. R^2. Hence, we select a quadratic model with an estimation period of 80 days.

The next step in an event study analysis is to compute the deviation of the volume in the estimation period from the modelled volume. This deviation is referred to as *Abnormal Volume* and can be computed with the help of Equation 4.8. In this equation, AV_t is the abnormal volume on day t and $ActV_t$ is the actual volume of Bitcoins traded on the exchange on day t. Variable $ActV_t$ can be calculated as shown in Equation 4.7, where $VolumeFrom_t$ and $VolumeTo_t$ are the starting and the ending trading volume readings for the day t and their values can be found in the dataset.

$$ActV_t = VolumeTo_t - VolumeFrom_t \tag{4.7}$$

$$\begin{aligned} AV_t = ActV_t - ({\alpha_i}^{+ve}({z_t}^{+ve})^2 + {\beta_i}^{+ve}{z_t}^{+ve} \\ + {\alpha_i}^{-ve}({z_t}^{-ve})^2 + {\beta_i}^{-ve}{z_t}^{-ve} + \gamma_i) \end{aligned} \tag{4.8}$$

$$CAV_i^p = \sum_{t=n}^{n+p} AV_t \tag{4.9}$$

To account for more long term (more than a day) impacts of DDoS attacks we calculate a p day *Cumulative Abnormal Volume*, which can be calculated using Equation 4.9. The variable CAV_i^p represents the cumulative abnormal volume for p days after the attack event i and AV_t represents the abnormal volume on day t. As shown in Figure 4.3.1 we calculate the *Cumulative Abnormal Volumes* for the following five *event periods*:

- Day of the attack to 1 day after it $[n, n+1]$.
- Day of the attack to 2 days after it $[n, n+2]$.
- Day of the attack to 3 days after it $[n, n+3]$.
- Day of the attack to 3 days after it $[n, n+4]$.

- Day of the attack to 5 days after it $[n, n+5]$.

Finally, we formulate the null hypothesis and test it in order to evaluate the impact of DDoS attack in the *event period*.

Hypothesis: As discussed previously in Section 4.2, we expect that a DDoS attack on crypto-currency exchange would result in decreased volume of Bitcoin. Hence, our study investigates whether the daily volume of Bitcoin traded on Bitfinex on the day of the attack is significantly lower than the volume of Bitcoin traded during the *estimation* period. Thus, the null hypothesis in this case can be stated as:

H_0: There is no difference in the average volume of Bitcoin traded on Bitfinex during the estimation period and event period.

A wide spread assumption is that the abnormal returns in case of a stock market event study are distributed according to a Gaussian distribution. This assumption was challenged in the paper by Abhishta, Joosten and Nieuwenhuis [9]. In this chapter, we distance ourself from the assumption that cumulative abnormal volumes are normally distributed as well. The unknown distribution can be approximated by an empirical distribution which can be generated by bootstrapping [60]. We use a one-tailed hypothesis test to evaluate our null hypothesis (H_0). Hence, we state the alternative hypothesis (H_1) as:

H_1: The average volume of Bitcoin traded on Bitfinex during the event period is less than the average volume traded in the estimation period.

Bootstrapping and Hypothesis Testing: We make use of Monte Carlo simulation for bootstrapping the empirical distribution of *abnormal volume* and *cumulative abnormal volume*. From the set of *abnormal volume* and *cumulative abnormal volume* values that belong to the *trend period* we draw a random value two million times. To also consider the values in the vicinity of the drawn value, we introduce an error to the drawn value as shown in Equation 4.10 where x_b represents the value used in the bootstrapped distribution, x_r is the random value drawn and τ is a random number in the interval $[-0.1, 0.1]$.

$$x_b = x_r + \tau x_r \text{ where } \tau \in [-0.1, 0.1] \tag{4.10}$$

Figures 4.3.2c and 4.3.2d shows the bootstrapped distributions used to analyse the attacks on *20-01-2016* and *04-06-2016*. For testing the statistical significance of the impact we consider that if the *abnormal volume* or the *cumulative*

abnormal volume in the event periods lie in the blue portion of these distributions then the negative impact of the DDoS attack was statistically significant.

For calculating the boundaries for a significantly negative impact, we assume a confidence interval of 90%. Hence as shown in Figure 4.3.2c we consider the bottom 10 percentile of the values to be statistically significant. Hence, in terms of hypothesis testing, if the value of *abnormal volume* or *cumulative abnormal volume* lies in the bottom 10 percentile of the bootstrapped empirical distribution, then we can reject the null hypothesis.

4.4 Results

We summarise the results of our analysis in two tables: Table 4.4.1 and Table 4.4.2. In Table 4.4.1 we show the values of regression parameters ${\alpha_i}^{+ve}$, ${\alpha_i}^{-ve}$, ${\beta_i}^{+ve}$, ${\beta_i}^{-ve}$ and γ_i. The coefficient of determination (adj. R^2) for the model used to estimate the traded volume is shown in Table 4.3.1. We observe that the adj. R^2 values for estimation models in 2016 and end of 2017 are relatively high in comparison to the other values. This is due to sudden increase in Bitcoin prices in mid 2017. Looking at the adj. R^2 values in most cases we can say that more than 50% of the traded volume can be predicted on the basis of change in price of Bitcoin. In all 17 cases a low *p-value* also indicates a strong relationship between the dependent variable V_t (volume of bitcoin traded on day t) and independent variables ${z_t}^{+ve}$ and ${z_t}^{-ve}$ (positive and negative price change respectively).

We test for negative impact on the *Abnormal Volume* of Bitcoin traded on Bitfinex for five days after the DDoS attack (including the day of attack). We do this to check whether the impact seen in the *cumulative abnormal volume* is due to the DDoS attack or not[†]. The results are shown in Table 4.4.1. We observe that in case of 4 of the 17 considered events there is a significant negative abnormal volume on the day of the attack. This means that for these 4 instances the exchange was not able to recover within a day. One of the main reasons why we do not see negative abnormal returns in all cases can be due to the fact that the attack was successful for a small duration and the trading activity just after the platform recovered compensated for the volume lost due to the short unavailability. We further observe Table 4.4.2 that in case of two out of these four negative abnormal volume events, the exchange recovers within two days as the 3 day cumulative abnormal return value indicate no impact. In the other

[†]For instance, if there is a negative impact on the day of attack but no impact one day after the attack then it means that the negative impact was recovered within one day of trading if there is no impact according to 2 day *cumulative abnormal volume* value.

Table 4.4.1: Results: Model Parameters and Abnormal Volume.

	Model Parameters					Negative Impact (Abnormal Volume)				
Attack Date	$\alpha_i{}^{+ve}$	$\alpha_i{}^{-ve}$	$\beta_i{}^{+ve}$	$\beta_i{}^{-ve}$	γ_i	1st Day	2 nd Day	3rd Day	4th Day	5th Day
20/01/2016*	34529.7	-4335.9	-214450.8	880044.9	6866249.1	No	No	No	No	**Yes**
04/06/2016*	-6687.0	7872.0	1012188.0	617831.9	2070323.8	No	No	No	No	No
07/06/2016*	-6479.9	10558.9	996953.8	549459.8	2356437.1	No	No	No	**Yes**	No
20/06/2016*	3830.4	74513.7	609420.0	-80812.8	4285488.8	**Yes**	**Yes**	No	No	No
26/07/2016*	5504.5	7183.9	475625.8	543508.0	7712740.4	**Yes**	No	No	No	**Yes**
09/11/2016*	10570.1	2673.0	84391.4	171729.4	2057886.1	No	No	No	No	No
16/11/2016*	8132.7	2681.8	145909.6	169299.8	2169077.2	No	No	No	No	No
21/02/2017*	4045.8	1641.5	121104.3	167707.4	5413822.6	No	No	No	No	No
12/06/2017*	650.1	-992.0	167009.3	472802.0	18690303.7	No	No	No	No	**Yes**
21/08/2017*	178.3	-479.7	250230.1	389380.3	40050513.6	No	No	No	No	No
26/11/2017*	-645.6	550.7	494887.0	174677.4	147126862.2	**Yes**	No	No	No	No
04/12/2017*	-212.5	895.5	217365.6	-76938.6	177704654.5	No	No	No	No	**Yes**
05/12/2017*	-195.2	1036.2	199544.5	-129388.6	179678622.2	No	No	No	**Yes**	No
12/12/2017*	74.7	406.4	109273.0	116946.1	188656983.1	No	No	No	No	No
17/12/2017*	17.5	337.0	285588.6	353096.1	180600154.9	No	No	No	No	No
31/12/2017*	15.1	47.1	286913.5	708868.2	214840630.2	No	No	No	No	No
05/06/2018*	486.0	658.3	31657.9	-56779.9	231740322.9	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**

*p value$<$ 0.05

Table 4.4.2: Results: Cumulative Abnormal Volume.

Attack Date	2 Day CAV[1]	3 Day CAV[1]	4 Day CAV[1]	5 Day CAV[1]
20/01/2016	No	No	**Yes**	**Yes**
04/06/2016	No	No	No	No
07/06/2016	No	No	No	No
20/06/2016	**Yes**	**Yes**	**Yes**	**Yes**
26/07/2016	**Yes**	No	No	**Yes**
09/11/2016	No	No	No	No
16/11/2016	No	No	No	No
21/02/2017	No	No	No	No
12/06/2017	No	No	No	No
21/08/2017	No	No	No	No
26/11/2017	**Yes**	No	No	No
04/12/2017	No	No	No	No
05/12/2017	No	No	No	No
12/12/2017	No	No	No	No
17/12/2017	No	No	No	No
31/12/2017	No	No	No	No
05/06/2018	**Yes**	**Yes**	**Yes**	**Yes**

[1]CAV: Cumulative Abnormal Volume

two cases we observe that the exchange does not recover within 5 days. We investigate these points in greater detail in next.

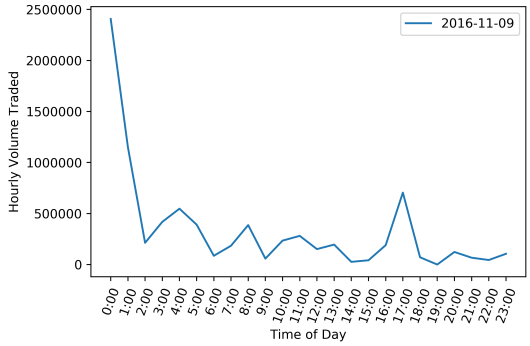
4.5 Discussion

We observe in Section 4.4 that in a majority of the cases (13 of 17) the loss of volume caused due to a successful DDoS attack is recovered by the exchange within a period of 1 day. For this reason we do not record a significant negative abnormal volume on the day of the attack. Figure 4.5.1 shows the hourly volume traded on the exchange. In Figures 4.5.1a and 4.5.1b we can observe periods when no or very little volume was being traded on the exchange. Some of these periods can be attributed to the platform issues caused due to a DDoS attack. On 9th November 2016, we observe that very little volume was traded on the exchange after the first hour of trading. However, on the basis of the total volume traded in the whole day we were unable to reject the null hypothesis. Similarly, on 21st August 2017, we observe a dip in volume traded after 13:00 hours (time of reported platform issues). The large volume of bitcoin traded in the end of the day however compensates for the loss. This quick recovery for Bitfinex can be partially attributed to the public relations (PR) strategy employed by the exchange. Bitfinex maintains and updates the status of platform availability on twitter and its own status page regularly. Hence, when the exchange resumes normal operations, all the customers are informed that they can resume trading. This can be one of the incentives for such businesses to publicly disclose DDoS attacks.

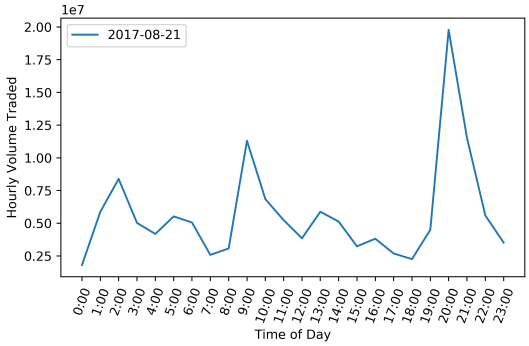
On two occasions (20th June 2016 and 5th June 2018) we observe that the negative impact lasts for more than 5 days. This is due to the fact that we have multiple days in the *event period* where the *abnormal volume* is significantly negative. In Figures 4.5.1c and 4.5.1d we can see that the exchange recovers on 21st June 2016 but the trading stops again after a few hours. This may be due to a second wave of unreported DDoS attacks or unresolved platform issues due to the first attack.

4.6 Related work

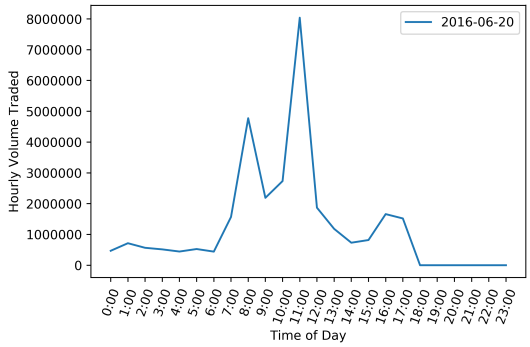
Feder *et al.* [68] also studied the impact of DDoS on crypto-currency exchanges, in particular, the Mt. Gox exchange. Mt. Gox was often targeted by DDoS attacks and was forced to close due to a serious breach that resulted in stolen funds. They measured the kurtosis and distribution of the distribution of trades that were made on the exchange when the exchange was under attack. The conclusion of the article showed a decrease in large volume trades due to a



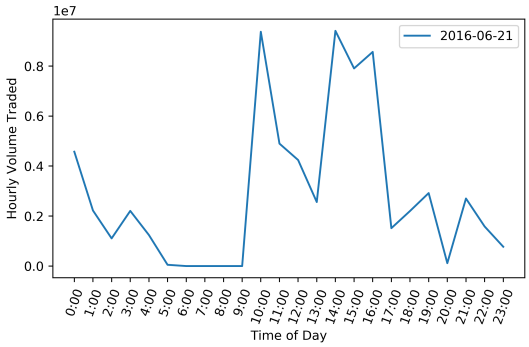
(a) Volume traded on 9$^{\text{th}}$ November 2016.



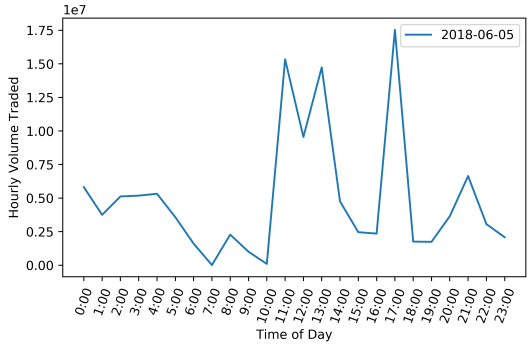
(b) Volume traded on 21$^{\text{st}}$ August 2017.



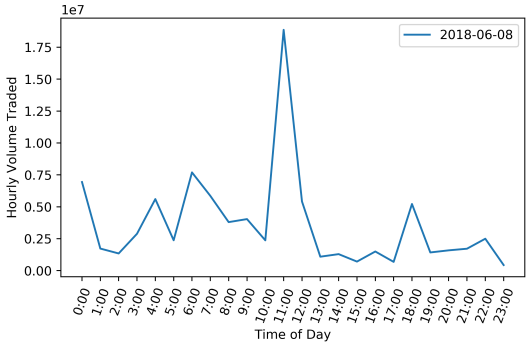
(c) Volume traded on 20$^{\text{th}}$ June 2016.



(d) Volume traded on 21$^{\text{st}}$ June 2016.



(e) Volume traded on 5$^{\text{th}}$ June 2018.



(f) Volume traded on 8$^{\text{th}}$ June 2018.

Figure 4.5.1: Hourly volume of Bitcoin traded on Bitfinex.

DDoS attack. They also suggested that other types of security breaches also had a similar kind of impact.

In another work, Johnson *et al.* [102] present a game-theoretic model for the trade-off faced by mining pools between investing in upgrades for computing infrastructure and engaging in DDoS attacks. They conclude that if attacks can be mitigated, then the size threshold for a mining pool to be safe from DDoS increases.

Similar event study methodology has also been applied to investigate the impact of DDoS attacks on stock prices and a comparison of alternatives to measure the impact of DDoS attack announcements on stock prices by Abhishta, Joosten and Nieuwenhuis [10]. This study looked at the impact of DDoS attacks on victim stock prices and concluded that most of the time the impact was not significant. This conclusion was also reached by Hovav *et al.* [90]. Only when the actual service of the company was down, it resulted in a statistically significant impact.

4.7 Conclusion

In this chapter, we present our analysis of the impact of DDoS attacks that targeted Bitfinex in the last three years. Using the data collected with the help of *www.cryptodatadownload.com* we test if there is a statistically significant negative impact on the daily volume of bitcoins that are traded on the exchange. For performing this analysis we present an event study methodology that uses the relationship between volume of bitcoin traded and change in it's price on the exchange to predict expected volume of bitcoin traded during the *event period*.

We determine the length of the *estimation period* and the degree of regression model by comparing the adj. R^2 values for multiple options. We apply our methodology to 17 different events and draw the following conclusions:

- We show that, for the investors there is a difference in the perception of positive and negative price changes. Hence, we model the impact of positive and negative price changes on the volume separately.

- We find that, on most occasions (13 of 17) the negative impact due to a DDoS attack is recovered within the same day by the exchange.

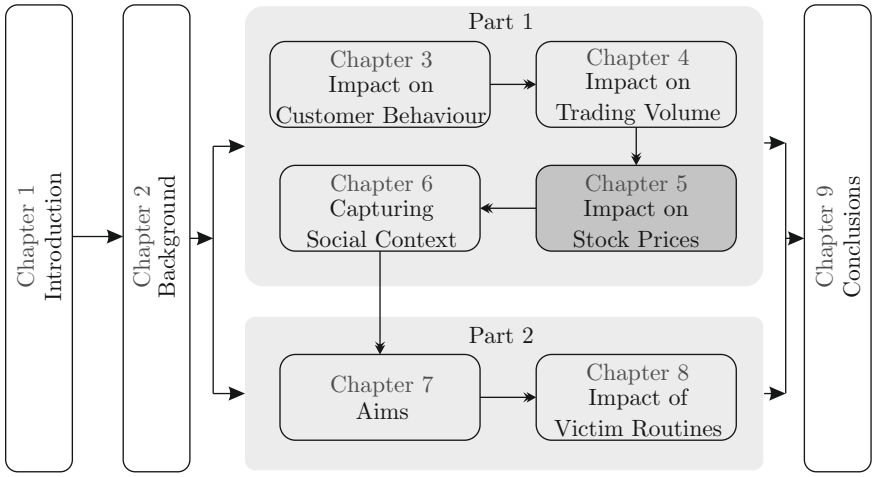
- On two instances, we find that the losses are recovered after two days of the attack.

- On two other occasions we find that the losses are not recovered within 5 days. We suppose that this is due to multiple platform un-availabilities in the *event period*.

Summarising, our study shows that in most cases this crypto-currency exchange has been able to recover from the impact of a DDoS attack within a single day. However, in the hourly data we do see the trading coming to a complete halt due to a DDoS attack. This proves that a long lasting DDoS attack can severely effect the revenues of the exchange.

Chapter 5

Impact on Stock Prices



In the previous chapters, we have dealt with measuring direct impact of DDoS attacks. Now, we look at measuring the indirect impact of DDoS attacks. One of the indirect damage due to a DDoS attack can be on the market value of the victim firm. In this chapter, we analyse the impact of 45 different DDoS attack announcements on victim's stock prices. We find that previous studies have a mixed conclusion on the impact of DDoS attack announcements on the victim's stock price. Hence, in this chapter we evaluate this impact using three different strategies and compare the results. We find that the assumption of cumulative abnormal returns being normally distributed leads to overestimation/underestimation of the impact.

5.1 Introduction

In the previous two chapters we have analysed the direct impact of DDoS attacks on the victim. In Chapter 3, we measured the impact of DDoS attacks on the behaviour of customers of large managed DNS service providers. In Chapter 4, we evaluated the impact of DDoS attacks on the volume of Bitcoins traded on a crypto currency exchange. Indirect losses include damages to company's reputation and impact at stock prices etc. In this chapter, we examine the indirect loss due to the decrease in the market value of a firm as a result of an announcement of getting hit by a DDoS attack.

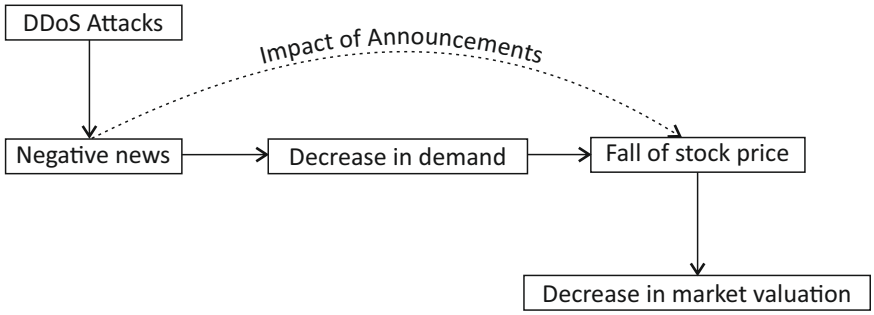


Figure 5.1.1: Impact of a DDoS attack announcement on market valuation of the firm.

The stock price of a firm is representative of its market value. In the past economists have analysed the impact of an economic event on the value of the firm [131]. A strategic business decision e.g. *merger or an acquisition* can significantly impact the future dividends. For instance, in the case of a possible negative impact on the future cash flows, it is beneficial for the investors to sell the shares and invest in a different stock.

DDoS attacks may lead to negative news articles about the firm. They come as a negative sentiment shock and can negatively influence the perception of firm's value to an investor, thus impact the demand of the victim firm's shares [197]. This in-turn leads to the fall of stock prices of the attacked company. Figure 5.1.1 shows the conceptual relationship between DDoS attack events and decrease in market valuation of the victim firm. It shows the empirical link that we evaluate in this study.

We restrict the scope of our analysis to announcements that were made after 2010. Unlike earlier studies we study the impact of DDoS attack announcements only, because these attacks do not lead to any form of information leaks and do

not pose any danger to customer data. Hence, in our sample we do not consider any of the events where DDoS has been used as a smoke screen.

We propose a less naive and more robust method for hypothesis testing in event studies. We then compare the proposed method with the traditional methods of event study and illustrate the disadvantages of using the assumptions and approximations considered by earlier methods. We then analyse the impact of DDoS attack announcements on victim stock prices using the proposed method.

5.2 Previous work

Event studies measure the impact of company related events on the market value of the firm. MacKinlay [131] discusses the procedure for conducting an event study and also the various models that can be used for estimation of normal behaviour of the market. In the past many researchers have studied the impact of information technology related events on the market value of the firm. Santos *et al.* [178] examined the impact of information technology investment announcements on the market value of the firm and suggested that there is no significant impact of these investment announcements on the market value.

Previous studies [89, 38, 40, 110] have used a one-factor market model for the estimation of stock prices as shown in Equation 5.1. Where r_{it} represents the rate of return of the stock i and r_{mt} represents the rate of return of the market index on day t. For instance, r_{it} can be calculated as $(P_{it} - P_{it-1})/P_{it-1}$, where P_{it} is the price of the stock on day t.

$$r_{it} = \alpha_i + \beta_i r_{mt} + \epsilon_{it} \tag{5.1}$$

The parameters α and β are firm dependent coefficients. $\hat{\alpha}$ and $\hat{\beta}$ are their ordinary least square (OLS) estimators. The stochastic variable ϵ_{it} is the error term with $\mathbb{E}[\epsilon_{it}] = 0$. Gordon *et al.* [76] uses a Fama-French three factor model [67] to predict the stock prices. The three factors being company size, company price-to-book ratio and market risk. The three factor model is shown in Equation 5.2.

$$r_{it} = a_i + b_i r_{mt} + s_i SMB_t + h_i HML_t + \epsilon_{it}, \tag{5.2}$$

SMB_t is the difference between the return on the portfolio of small stocks and the return on the portfolio of large stocks on day t and HML_t is the difference between the return on a portfolio of low-book-to-market stocks and the return on a portfolio of low-book-to-market stocks on day t. The parameters a_i,b_i,s_i and h_i are Fama and French three-factor model estimated firm dependent coefficients. The stochastic variable ϵ_{it} is the error term with $\mathbb{E}[\epsilon_{it}] = 0$.

	Author	Estimation Model	Sample Size	Breach Type	Conclusion	Sample Period
[89]	Hovav & D'Arcy (2003)	Market Model	23	DoS	No significant impact of DoS attacks on the capital market. Some indication of impact on firms that rely on the web for their business.	1998-2002
[38]	Campbell *et al.* (2003)	Market Model	43	Generic	Some negative stock market impact to reported information security breaches.	1995-2000
[73]	Garg *et al.* (2003)	N/A	22	Generic	Average fall in the stock price was approximately 2.9% over a 2-day and 3.6% over 3-day period.	1996-2002
[40]	Cavusoglu *et al.* (2004)	Market Model	66	Generic	Security breach announcements affect the values of the announcing firms and also the Internet security developers.	1996-2001
[110]	Kannan *et al.* (2007)	Market Model	102	Generic	Drop of 1.4% in the market valuation relative to the control group of firms.	1997-2003
[76]	Gordon *et al.* (2011)	Fama-French Model	258	Generic	Pre 9/11 information security breaches showed significant negative stock market returns but the results for the post 9/11 period were not significant.	1995-2007

Table 5.2.1: Previous works on impact on victim stock prices.

Studies such as [110, 89, 38] use abnormal returns (additive) and cumulative abnormal returns (additive) as a measure of event impact. Equations 5.3 and 5.4 show the relations used to compute abnormal returns and cumulative abnormal returns respectively. As they assume normal distribution for the CAR values hence they use Z statistic to test their hypothesis.

$$AR_{it} = r_{it} - (\hat{\alpha_i} + \hat{\beta}_i r_{mt}) \tag{5.3}$$

$$CAR_n = \sum_{t=-1}^{n} AR_{it} \tag{5.4}$$

Studies have also been conducted on evaluating the impact of information security breaches on the prices of the victim firm's shares. Table 5.2.1 lists selected works and their conclusions. In this table we also take a look on the sample size and period of the sample considered by these studies.

Researchers have had a mixed response on the impact of denial of service attacks on the stock returns of the victim firms. Garg *et al.* [73] and Hovav & D'Arcy [89] suggest that DDoS attack announcements lead to a negative abnormal returns, while Gordon *et al.* [76] deny the effect of these attacks on the market value of the firm. Spanos & Angelis [188] conducted a systematic literature review on the impact of information security events on the stock market and concluded that the events examined created a significant impact on the stock price of the firms.

5.3 Methodology

To analyse the impact of DDoS attack announcements on stock returns we follow the method as shown in Figure 5.3.1. We broadly divide the method into two sections:

- Data collection.
- Analysis.

Our contribution to the analysis is at two instances. Firstly, we use a *multiplicative model* for the estimation of return rates and secondly, we use the empirical distribution of abnormal returns by *generation of random scenarios* for the analysis. In Section 5.3.1 we explain the approach for data collection. Section 5.3.3 deals with the identification of the impact caused by these announcements.

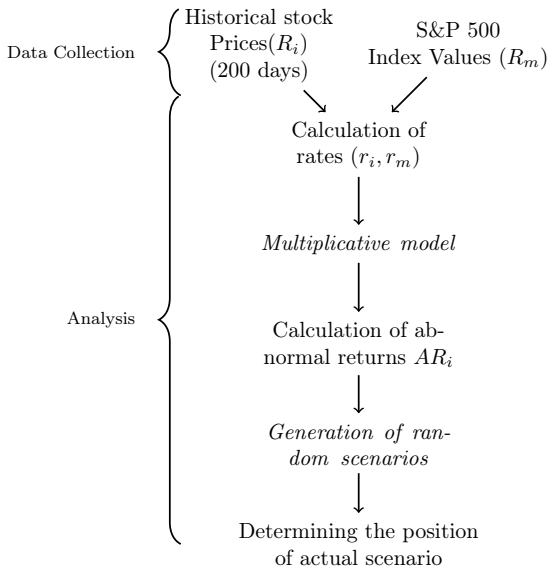


Figure 5.3.1: Method for event study. (Our contribution in *Italics*.)

5.3.1 Data collection

We consider all DDoS attack announcements that were made on the web after 'Operation Payback', launched by Anonymous in December, 2010. Table 5.3.1 shows the final list of all announcements that we analysed. For each attack we record the date of announcement, the company type and also the type of service disruption. The initial list consisted of 50 announcements.

We further filtered the list using the following criteria:

- If multiple announcements were made on consecutive days, then the earliest date was considered.
- All announcements regarding companies that were not publicly traded at the time of attack were eliminated.
- All attack announcements in which a DDoS attack was coupled with information theft were also not considered for analysis. This was done in order to be able to see the isolated effect of a DDoS attack announcements on the firm's stock price.

Table 5.3.1: Sample of DDoS attack events.

Organisation	**Announcement Date**	**Source**	**Infrastructure**	**Firm Type**
Master Card	December 8, 2010	spiegel.de	Website	Financial Services
Visa	December 8, 2010	spiegel.de	Website	Financial Services
Bank of America	December 27, 2010	infosecisland.com	Website	Financial Services
Vodafone	October 5, 2011	infosecurity-magazine.com	None	Telecommunications
Apple	May 29, 2012	att-iphone-unlock.com	Website	IT
AT&T	August 16, 2012	pcworld.com	None	Telecommunications
Wells Fargo	December 20, 2012	technologybanker.com	DNS	Financial Services
JP Morgan Chase	March 13, 2013	scmagazine.com	Website	Financial Services

Organisation	Announcement Date	Source	Infrastructure	Firm Type
TD Canada Trust	March 21, 2013	thestar.com	E Services	Financial Services
American Express Company	March 28, 2013	bankinfosecurity.com	Website	Financial Services
International Netherlands Group	April 9, 2013	nrc.nl	Payment Services	Financial Services
LinkdIn	June 21, 2013	softpedia.com	Website	Social Networking
Microsoft	November 27, 2013	scmagazine.com	DNS	IT/Gaming
Royal Bank of Scotland	December 4, 2013	theguardian.com	Banking Services	Financial Services
JP Morgan Chase	January 30, 2014	bobsguide.com	Online Banking Services	Financial Services
Bank of America	January 30, 2014	bobsguide.com	Online Banking Services	Financial Services
Facebook	February 21, 2014	nos.nl	Messageing Services	Social Networking
Activision Blizzard	March 29, 2014	ign.com	Gaming Services	Gaming

Organisation	Announcement Date	Source	Infrastructure	Firm Type
Danske Bank	July 10, 2014	ddosattacks.net	Website	Financial Services
Storebrand	July 10, 2014	ddosattacks.net	Website	Insurance Company
Gjensidige Forsikr	July 10, 2014	ddosattacks.net	Website	Insurance Company
Sony Corporation	August 24, 2014	techcrunch.com	Gaming Services	IT
Amazon	August 27, 2014	shacknews.com	Twitch Streamers	E-commerce
Activision Blizzard	November 14, 2014	eurogamer.net	Gaming Services	Gaming
Sony Corporation	November 26, 2014	wiwo.de	Gaming Services	IT
Rackspace	December 22, 2014	welivesecurity.com	DNS	Hosting
Microsoft	December 24, 2014	krebsonsecurity.com	Gaming Services	IT/Gaming
Sony Corporation	December 24, 2014	krebsonsecurity.com	Gaming Services	IT

Organisation	Announcement Date	Source	Infrastructure	Firm Type
Alibaba Group	December 25, 2014	ddosattacks.net	Cloud Services	E-commerce
Nordea Bank	January 4, 2015	ddosattacks.net	Online Banking Services	Financial Services
Facebook	January 27, 2015	gizmodo.com.au	Website	Social Networking
Amazon	March 16, 2015	scmagazineuk.com	Twitch Streamers	E-commerce
EA Sports	March 18, 2015	ibtimes.com	Gaming Services	Gaming
Ziggo	August 18, 2015	emerce.nl	DNS	Telecommunications
Overstock.com	September 3, 2015	ddosattacks.net	DNS	E-commerce
Nissan	January 12, 2016	businessinsider.com	Website	Automotive
HSBC	January 28, 2016	bbc.com	Website	Financial Services
Activision Blizzard	August 2, 2016	technobuffalo.com	Gaming Services	Gaming

Organisation	Announcement Date	Source	Infrastructure	Firm Type
Electronic Arts	August 31, 2016	pcgameshar-ware.com	Gaming Services	Gaming
StarHub	October 26, 2016	telecomasia.com	Network Infrastructure	Telecommunications
Deutsche Telekom	November 28, 2016	siliconrepublic.com	Network Infrastructure	Telecommunications

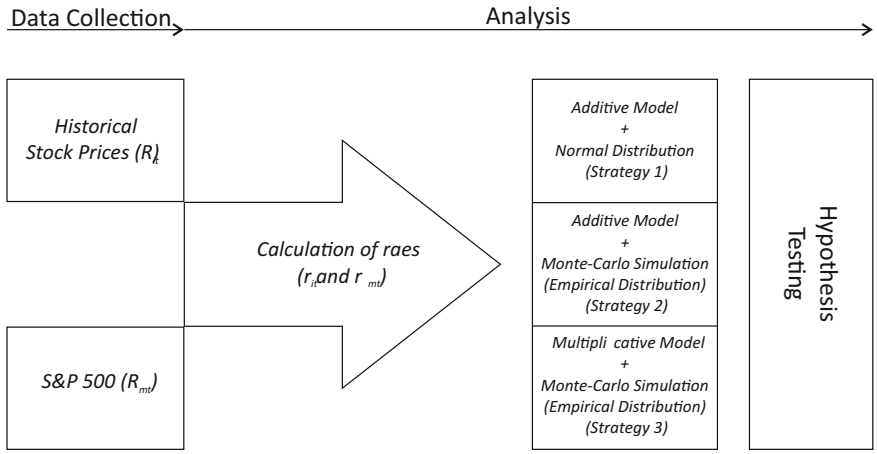


Figure 5.3.2: Methodology to compare strategy for analysis.

The stock prices for all the firms in the sample were collected by using the Yahoo! finance API. For measuring the market rate we collected the S&P 500 index values.The final sample consisted of 45 announcements.

5.3.2 Hypothesis

To evaluate the impact of DDoS attack announcements on the market value of a victim firm we establish the null hypothesis (H_0) as follows:

H_0**:** *There is no impact of DDoS attack announcements on victim stock prices.*

In order to analyse the collected data we first need to calculate the rate of return of the market index on day t (r_{mt}) and r_{it} the rate of return of the stock i on day t. The rate of return can be calculated as shown in Equation 5.5, where P_{it} and P_{mt} represent the stock price and market index for day t. The value of the market index shows the average of returns of all the firms included in the market index.

$$\begin{aligned} r_{it} &= \frac{P_{it} - P_{i(t-1)}}{P_{i(t-1)}} \\ r_{mt} &= \frac{P_{mt} - P_{m(t-1)}}{P_{m(t-1)}} \end{aligned} \tag{5.5}$$

To illustrate the benefits of the proposed model we use three different strategies to test our null hypothesis (H_0). We explain these strategies in detail in the following section.

5.3.3 Analysis

5.3.3.1 Strategy 1

In the first strategy we consider an additive model to represent the normal behavior of the market. The model can be mathematically represented as shown in Equation 5.6. This model is used to estimate the returns on a firm's stock. The variables r_{it} and r_{mt} are calculated as shown in Equation 5.5.

$$r_{it} = \alpha_i + \beta_i r_{mt} + \epsilon_{it} \tag{5.6}$$

The stochastic variable ϵ_{it} is the error term with $\mathbb{E}[\epsilon_{it}] = 0$. We use ordinary least square (OLS) in order to calculate the estimations $\hat{\alpha}_i$ and $\hat{\beta}_i$ for the firm dependent parameters α_i and β_i by considering daily returns over a period of 200 days. The estimation period starts 201 days before the date of attack announcement and ends two days before the announcement.

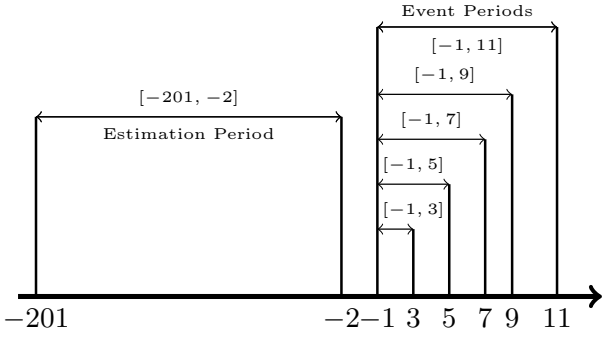


Figure 5.3.3: Estimation and event periods.

The additive abnormal return (AAR_{it}) is the measurement of the deviation of the actual returns from the ones calculated with the help of additive model equation 5.6. Hence AAR_{it} can be mathematically represented as:

$$AAR_{it} = r_{it} - (\hat{\alpha}_i + \hat{\beta}_i r_{mt}) \tag{5.7}$$

We measure the impact of DDoS attack announcements on the stock return over the following five *event periods*:

- One day prior to the announcement to 1 days after it $[t-1, t+1]$.
- One day prior to the announcement to 3 days after it $[t-1, t+3]$.
- One day prior to the announcement to 5 days after it $[t-1, t+5]$.
- One day prior to the announcement to 7 days after it $[t-1, t+7]$.
- One day prior to the announcement to 9 days after it $[t-1, t+9]$.

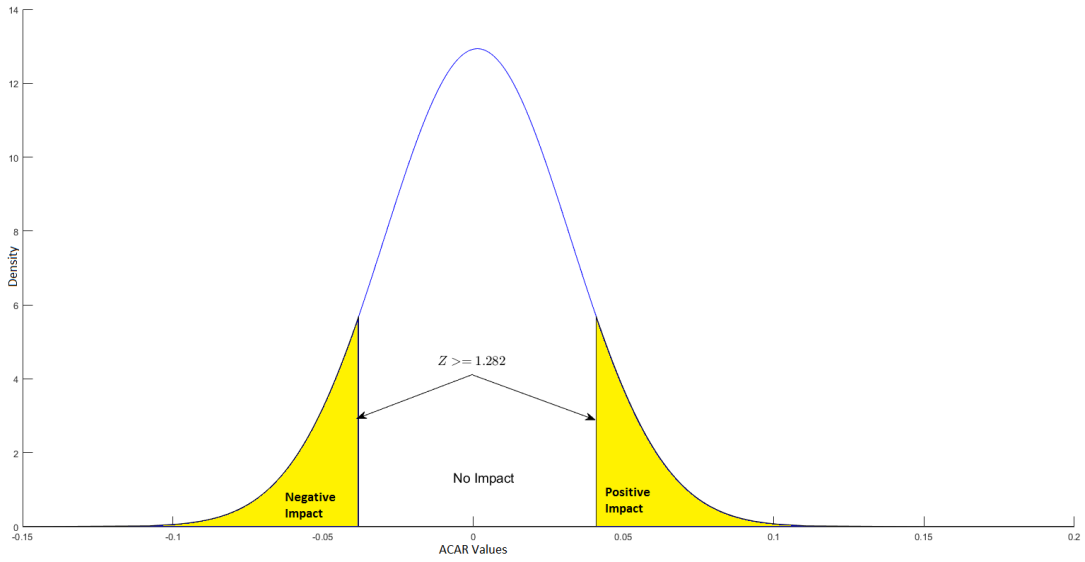


Figure 5.3.4: Normal distribution for 5 day $ACAR$ values and decision rule for impact analysis.

We keep these time periods consistent for all strategies. The *estimation period* and the *event periods* are shown in Figure 3.3. We take the event periods from one day prior to the announcements in order to compensate for any time lags. In order to calculate the combined effect over a certain number of days, we calculate the additive cumulative abnormal return ($ACAR$) as shown in Equation 5.8 for the period $[N_1, N_2]$.

$$ACAR_i = \sum_{t=N_1}^{N_2} (AAR_{it}) \tag{5.8}$$

We compute the mean $ACAR$ for 45 events in our sample as follows:

$$ACAR = \frac{1}{K}\sum_{i=1}^{K} ACAR_i \tag{5.9}$$

Where K is the number of events. We then estimate the standard deviation (σ_{ACAR}) using Equation 5.10.

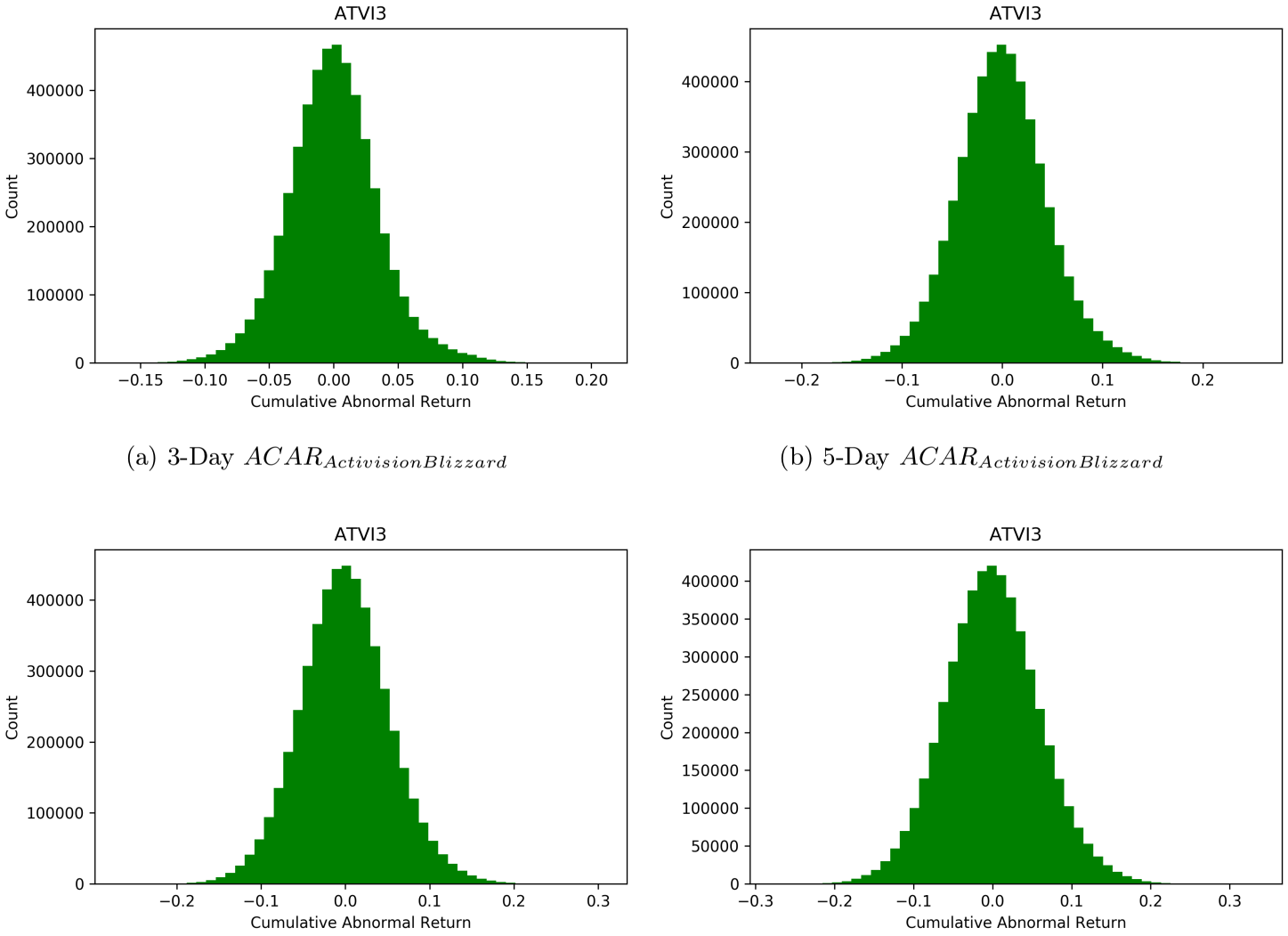
$$\sigma_{ACAR} = \sqrt{\frac{\sum_{i=1}^{K}(ACAR_i - ACAR)^2}{K-1}} \tag{5.10}$$

We now assume the $ACAR_i$ values for a given *event period* to be normally distributed and test for significance by making use of the Z-statistic at 10% confidence level. Hence we reject the null hypothesis if the $|Z| >= 1.282$ as shown in Figure 5.3.4.

5.3.3.2 Strategy 2

In this strategy we again make use of the additive estimation model as shown in Equation 5.6. We avoid the widespread assumption of short-term returns being approximately normally distributed. We also do not impose any alternative distribution to these returns. Instead, we use the technique of bootstrapping (e.g. Efron [60]). In this case we generate 5 million n-day returns by randomly drawing n one-day returns from the empirical distribution. The relative frequencies of these 5 million multi-day returns are then used as the distribution for hypothesis testing.

In order to calculate the additive abnormal returns we again employ Equation 5.7. After computing the AAR_{it}s for the estimation period and the event periods as discussed in Section 5.3.3.1 we draw 3, 5, 7, 9 and 11 one-day abnormal returns from the estimation period AARs. We then calculate the value of $ACAR_i$ for each of these scenarios with the help of Equation 5.8. Figure 5.3.5 shows the empirical distribution of $ACAR$ for Activision Blizzard. Lastly, to asses the effect of DDoS attack announcement on the stock returns we check the position of $ACAR_i$ for a certain event period in the empirical distribution of $ACAR$ for the same number of days of firm i. For example, if we are evaluating the $ACAR$ of Activision Blizzard for event period $[t-1, t+1]$ then we check the position of this $ACAR$ in the 3-day empirical distribution for Activision Blizzard. In this study we consider the 10 percentile scenarios in the left tail to be representative of negative impact and 10 percentile scenarios to the right for positive impact. Hence, if $ACAR_i$ is negative and lies in the bottom 10 percentile of the 5 million scenarios then the impact on the stock returns is considered to be negative.



(a) 3-Day $ACAR_{ActivisionBlizzard}$

(b) 5-Day $ACAR_{ActivisionBlizzard}$

(c) 7-Day $ACAR_{ActivisionBlizzard}$

(d) 9-Day $ACAR_{ActivisionBlizzard}$

Figure 5.3.5: Empirical distribution of $ACAR$ (additive) for Activision Blizzard.

5.3.3.3 Strategy 3

In this final strategy we use a multiplicative model for the estimation of stock returns. The multiplicative estimation model is:

$$(1 + r_{it}) = \alpha_i (1 + r_{mt})^{\beta_i} \tag{5.11}$$

Also, this time, we deviate from the wide spread practice of adding the corresponding single-day returns to compute the cumulative returns. Instead we calculate the exact cumulative returns*.

We linearise Equation 5.11 as Equation 5.12. The stochastic variable ϵ_{it} represents the error term with $\mathbb{E}[\epsilon_{it}] = 0$.

$$\ln(1 + r_{it}) = \widehat{\ln(\alpha_i)} + \hat{\beta}_i \ln(1 + r_{mt}) + \epsilon_{it} \tag{5.12}$$

After estimating the stock returns we use Equation 5.13 for computing the abnormal returns. As $\widehat{\ln(\alpha_i)}$ is not an unbiased estimator for α_i ($\mathbb{E}[\hat{\alpha}] \neq \mathbb{E}[e^{\widehat{\ln \alpha}}]$), we use Equation 5.14 for estimating $\hat{\alpha}$.

$$AR_{it} = \frac{(1 + r_{it})}{\hat{\alpha}_i (1 + r_{mt})^{\hat{\beta}_i}} - 1 \tag{5.13}$$

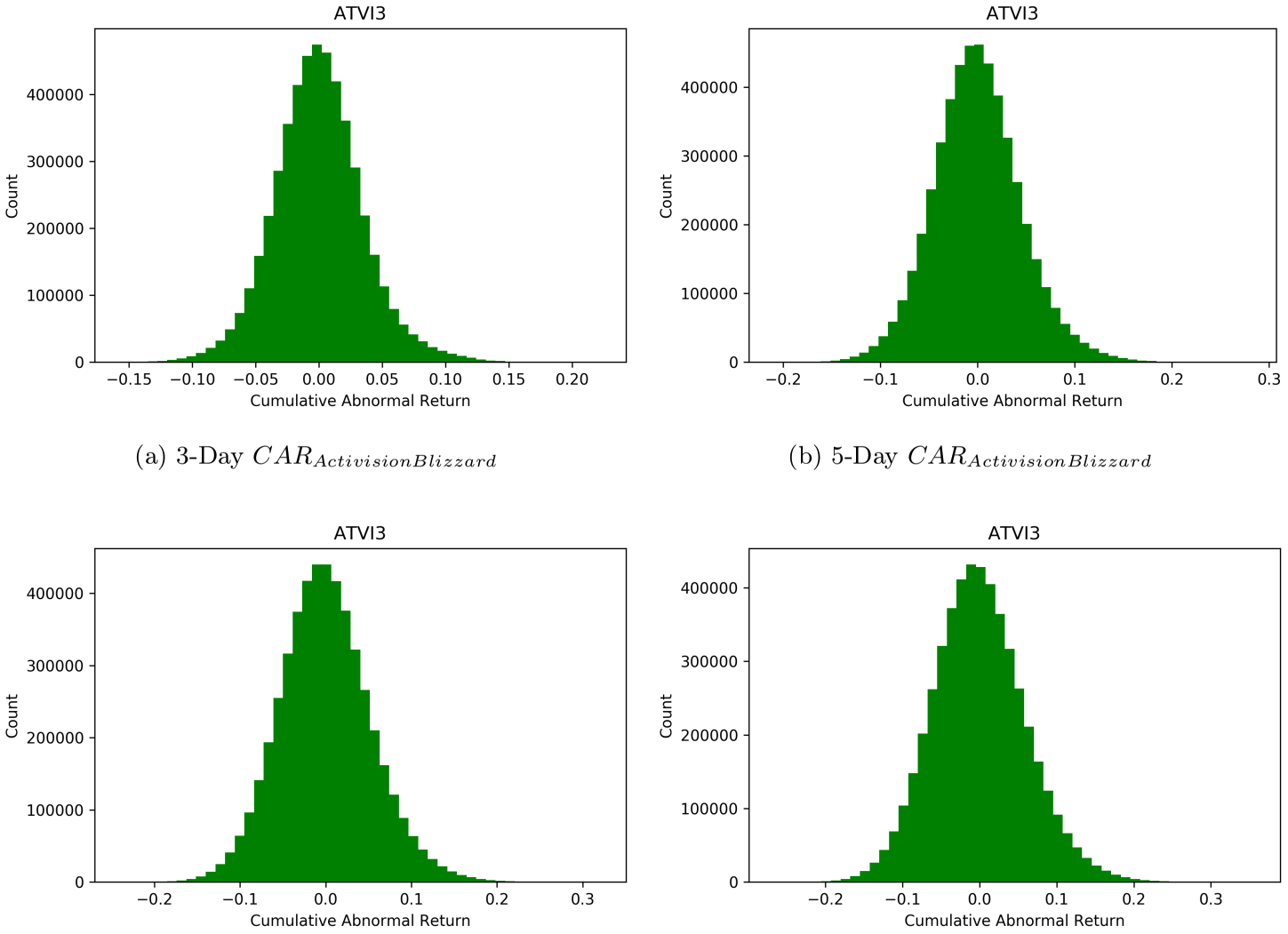
$$\hat{\alpha}_i = \frac{\sum_{t=1}^{T} (1 + r_{it})}{\sum_{t=1}^{T} (1 + r_{mt})^{\hat{\beta}_i}}, \tag{5.14}$$

After computing the AR_{it}s for the estimation period and the event periods as discussed in Section 5.3.3.1 we draw 3, 5, 7, 9 and 11 one-day abnormal returns from the estimation period ARs. As discussed earlier we then calculate the value of CAR_i for each of these scenarios with the help of Equation 5.15.

$$CAR_i = \prod_{t=N_1}^{N_2} (1 + AR_{it}) - 1 \tag{5.15}$$

Figure 5.3.6 shows the empirical distribution of CAR for Activision Blizzard. Lastly, to asses the effect of DDoS attack announcements on the stock returns we check the position of CAR_i for a certain event period in the empirical distribution of CAR for the same number of days of firm i. For example, if we are evaluating the CAR of Activision Blizzard for event period $[t-1, t+1]$ then we

*An increase of 10%, followed by a 10% decrease implies a total decrease of 1% according to the multiplicative formula $(1.1)(0.9) = 0.99$. The additive approximation yields a change of 0%, which is an overestimation of 1%.



(a) 3-Day $CAR_{ActivisionBlizzard}$

(b) 5-Day $CAR_{ActivisionBlizzard}$

(c) 7-Day $CAR_{ActivisionBlizzard}$

(d) 9-Day $CAR_{ActivisionBlizzard}$

Figure 5.3.6: Empirical distribution of CAR (multiplicative) for Activision Blizzard.

check the position of this CAR in the 3-day empirical distribution for Activision Blizzard. In this study we consider the 10 percentile scenarios in the left tail to be representative of negative impact and 10 percentile scenarios to the right for positive impact. Hence, if CAR_i is negative and lies in the bottom 10 percentile of the 5 million scenarios then the impact on the stock returns is considered to be negative.

5.4 Results

We now compare the results of our analysis. Table 5.4.1 summarizes the outcomes of using the three different strategies. The table shows the number of *event periods* in which we observed a positive or negative impact in each case. A negative event period implies that the DDoS attack announcement had an impact on investor decisions. The positive event periods on the stock price actually show that the stock was well performing and the DDoS attack announcement did not have any impact on the stock price. Later in Appendix 5.A we present the impact of each event analysed in detail.

Table 5.4.1: List of victim companies and summary of results

		Strategy								
		1			2			3		
Company Name	Date	+ve	-ve	No	+ve	-ve	No	+ve	-ve	No
Master Card	12/07/10	2	1	2	2	0	3	2	0	3
Visa	12/07/10	2	2	1	2	1	2	2	1	2
Bank of America	12/27/10	0	3	2	0	3	2	0	3	2
Vodafone	10/04/11	0	0	5	0	0	5	0	0	5
Vivendi	01/18/12	0	0	5	0	0	5	0	0	5
Bursa Malaysia	02/13/12	0	0	5	0	0	5	0	0	5
Apple	05/25/12	0	1	4	0	0	5	0	0	5
AT&T	08/15/12	0	0	5	1	0	4	1	0	4
Wells Fargo	12/19/12	0	0	5	0	0	5	0	0	5
JP Morgan Chase	03/12/13	0	0	5	3	0	2	3	0	2
TD Canada Trust	03/20/13	0	0	5	0	1	4	0	1	4
American Express	03/27/13	0	0	5	1	0	4	1	0	4

Table 5.4.1: List of victim companies and summary of results(cont...)

		Strategy								
		1			2			3		
Company Name	Date	+ve	-ve	No	+ve	-ve	No	+ve	-ve	No
ING	04/08/13	0	3	2	0	2	3	0	2	3
Linkedin	06/20/13	0	1	4	0	0	5	0	0	5
Microsoft	11/26/13	0	0	5	0	0	5	0	0	5
RBS	12/03/13	0	0	5	0	0	5	0	0	5
Electronic Arts	01/02/14	0	0	5	0	0	5	0	0	5
JP Morgan Chase	01/29/14	0	0	5	0	0	5	0	0	5
Bank of America	01/29/14	0	0	5	0	0	5	0	0	5
Facebook	02/20/14	0	0	5	0	0	5	0	0	5
Verizon	03/21/14	0	0	5	0	0	5	0	0	5
Activision Blizzard	03/28/14	1	0	4	2	0	3	2	0	3
Danske Bank	07/09/14	0	0	5	0	0	5	0	0	5
Storebrand	07/09/14	0	0	5	0	0	5	0	0	5
Gjensidige Forsikr	07/09/14	0	3	2	0	4	1	0	4	1
Sony	08/22/14	0	0	5	0	0	5	0	0	5
Amazon	08/26/14	0	0	5	0	0	5	0	0	5
Activision Blizzard	11/13/14	2	1	2	1	2	2	1	2	2
Sony	11/25/14	0	0	5	0	0	5	0	0	5
Rackspace	12/19/14	0	0	5	0	0	5	0	0	5
Microsoft	12/23/14	0	0	5	3	0	2	3	0	2
Sony	12/23/14	0	0	5	0	0	5	0	0	5
Alibaba	12/24/14	1	0	4	0	0	5	0	0	5
Nordea Bank	01/09/15	0	3	2	0	3	2	0	3	2
Facebook	01/26/15	0	0	5	0	0	5	0	0	5
Amazon	03/13/15	0	0	5	0	0	5	0	0	5
Electronic Arts	03/17/15	0	4	1	0	1	4	0	1	4

Table 5.4.1: List of victim companies and summary of results(cont...)

		Strategy								
		1			2			3		
Company Name	Date	+ve	-ve	No	+ve	-ve	No	+ve	-ve	No
Ziggo (Liberty Global)	08/17/15	2	0	3	4	0	1	4	0	1
Overstock.com	09/02/15	0	0	5	0	0	5	0	0	5
Nissan	01/12/16	1	0	4	0	0	5	0	0	5
HSBC	01/28/16	3	0	2	3	0	2	3	0	2
Activision Blizzard	08/02/16	0	1	4	0	0	5	0	0	5
Electronic Arts	08/31/16	0	1	4	0	0	5	0	0	5
StarHub	10/26/16	0	0	5	2	0	3	2	0	3
Deutsche Telekom	11/28/16	0	1	4	0	2	3	0	2	3

Strategy 2 \ Strategy 3	+ve	No	-ve
+ve	24	0	0
No	0	182	0
-ve	0	0	19

Table 5.4.2: Cross-table showing the number of differences between Strategy 2 and Strategy 3.

First we compare the differences in the results when using Strategy 2 and Strategy 3. Both strategies do not take the assumption of normal distribution for assessing cumulative abnormal returns. However, Strategy 2 uses an additive model for estimation and Strategy 3 uses a multiplicative model for the return rate estimation. We find no differences between the results of the two models in the periods analysed. Hence, we can conclude that the additive model does provide a good estimation for the computation of cumulative abnormal returns.

Then we look for differences in the results of Strategy 1 and Strategy 3. The differences between the models are as follows:

- Strategy 1 uses additive estimation model while Strategy 2 employs the multiplicative model.
- Strategy 1 computes cumulative abnormal returns by adding the successive abnormal returns where as Strategy 2 calculates them by using the multiplicative approach (Equation 5.15).
- Finally, Strategy 3 does not assume the abnormal returns or cumulative abnormal returns to be normally distributed.

Strategy 1 \ Strategy 3	+ve	No	-ve
+ve	11	3	0
No	13	169	4
-ve	0	10	15

Table 5.4.3: Cross-table showing the number of differences between Strategy 1 and Strategy 3.

Table 5.4.3 summarizes the differences between the two strategies. We believe that Strategy 3 is more accurate, or rather less inaccurate, than Strategy 1 due to the reduced number of assumptions and approximations in the model. Hence, we look at the number of times Strategy 1 overestimates or underestimates the significance of the results, i.e. gives a significant positive or negative impact when there is no impact or vice-versa. We observe that on 13 instances Strategy 1 reports a positive impact and on 4 instances a negative impact when there was no statistically significant impact. We find similar differences between the results of Strategy 1 and Strategy 2. This suggests that the assumption of normally distributed abnormal returns is the reason for inconsistencies between Strategy 1 and others.

The event wise detailed results for our study are shown in Appendix 5.A. According to the results of our analysis we observe a significant negative impact in the case of *International Netherlands Group* and *EA sports*. Whereas, a delayed negative effect is noticeable in the case of *Bank of America*, *Storebrand* and *Nordea Bank*. In most cases we do not see a negative effect on the victim stock prices.

In cases where the announcements state that the availability of the infrastructure under attack did not affect the customers, no significant impact was

noticed. For example, in the case of *Visa* and *MasterCard* the infrastructure under attack was their *website* but the customers were still able to use their cards for payment purposes. Whereas in the case of *International Netherlands Group*, customers had troubles using the payment services. Similarly, in the case of *EA Sports*, gamers were not able to log onto their on-line gaming accounts.

In the case of *Ziggo*, the customers did face troubles due to the unavailability of internet services but as the firm is a part of a bigger conglomerate *Liberty Global*, we were unable to spot any significant impact on the stock prices.

5.5 Conclusion

In this chapter, we analyse one of the indirect consequences of DDoS attacks on the victim, i.e., the impact of DDoS attack announcements on victim stock prices. We do so by using a less naive and more robust event analysis methodology. We compare the methodology proposed by us with traditional event analysis methodologies and show the benefits of using empirical distributions for hypothesis testing. Finally, we apply the proposed event analysis methodology to 45 different cases collected over a period of 5 years and discuss the results.

Our study led to two main conclusions. Firstly, by comparing the various strategies for conducting event studies we bring out the risk of overestimating or underestimating the impact of DDoS attack announcements on victim's stock prices. The choice of additive or multiplicative model does not affect the results but the assumption of normally distributed cumulative returns can lead to an incorrect estimation of the impact. Hence, we propose the use of an empirical distribution in order to check the significance of cumulative abnormal returns. Secondly, we also show that all three strategies for analysis result in a significantly negative event periods on stock price when service to the customers was hampered due to the attack. We report that the attacks on *ING* and *Nordea bank* [98, 145] resulted in significant negative returns where as *Visa* and *Mastercard* [212] resulted in no damage. We also record a delayed negative impact on the market value of gaming company *Activison Blizzard*. Similarly, in case of the attack on *Deutsche Telekom* that drove nearly 1 million of its customers offline [66], we observe a negative impact on the stock price in the 9-day and 11-day period.

As a conclusion, we can say that there is a noticeable negative impact on the stock prices of the victim firm whenever the attack causes interruptions to the services provided by the firm to its customers. However, this impact is short-lived and most of the events considered in this study recover within a 11-day event period. This drop is consistent with the results of the previous studies

[73, 89]. The marked value of the firm depends upon various other factors as well so, is not possible for us to comment on the intensity of the impact.

Appendix 5.A Impact on victim stock prices

As an appendix to this chapter, we present the impact of DDoS attacks on each of the 45 analysed cases. We present the results of using each of the three strategies over 3-day, 5-day, 7-day, 9-day and 11-day event periods.

			Strategy		
Firm	Date	Event Period	1	2	3
Bursa Malaysia	14/02/2012	3-day	No	No	No
	16/02/2012	5-day	No	No	No
	21/02/2012	7-day	No	No	No
	23/02/2012	9-day	No	No	No
	27/02/2012	11-day	No	No	No
Apple	29/05/2012	3-day	-ve	No	No
	31/05/2012	5-day	No	No	No
	04/06/2012	7-day	No	No	No
	06/06/2012	9-day	No	No	No
	08/06/2012	11-day	No	No	No
Amazon	16/03/2015	3-day	No	No	No
	18/03/2015	5-day	No	No	No
	20/03/2015	7-day	No	No	No
	24/03/2015	9-day	No	No	No
	26/03/2015	11-day	No	No	No
Amazon	27/08/2014	3-day	No	No	No
	29/08/2014	5-day	No	No	No
	03/09/2014	7-day	No	No	No
	05/09/2014	9-day	No	No	No
	09/09/2014	11-day	No	No	No
Activision Blizzard	14/11/2014	3-day	+ve	No	No
	18/11/2014	5-day	+ve	+ve	+ve
	20/11/2014	7-day	No	No	No

(to be continued on next page)

Firm	Date	Event Period	Strategy 1	Strategy 2	Strategy 3
	24/11/2014	9-day	-ve	-ve	-ve
	26/11/2014	11-day	No	-ve	-ve
	31/03/2014	3-day	No	No	No
	02/04/2014	5-day	No	No	No
Activision Blizzard	04/04/2014	7-day	+ve	+ve	+ve
	08/04/2014	9-day	No	No	No
	10/04/2014	11-day	No	+ve	+ve
	03/08/2016	3-day	-ve	No	No
	05/08/2016	5-day	No	No	No
Activision Blizzard	09/08/2016	7-day	No	No	No
	11/08/2016	9-day	No	No	No
	15/08/2016	11-day	No	No	No
	28/03/2013	3-day	No	No	No
	02/04/2013	5-day	No	No	No
American Express	04/04/2013	7-day	No	No	No
	08/04/2013	9-day	No	+ve	+ve
	10/04/2013	11-day	No	No	No
	26/12/2014	3-day	No	No	No
	30/12/2014	5-day	No	No	No
Alibaba	02/01/2015	7-day	+ve	No	No
	06/01/2015	9-day	No	No	No
	08/01/2015	11-day	No	No	No
	27/12/2010	3-day	No	No	No
	29/12/2010	5-day	No	No	No
Bank of America	31/12/2010	7-day	-ve	-ve	-ve
	04/01/2011	9-day	-ve	-ve	-ve
	06/01/2011	11-day	-ve	-ve	-ve

(to be continued on next page)

Firm	Date	Event Period	Strategy 1	Strategy 2	Strategy 3
Bank of America	30/01/2014	3-day	No	No	No
	03/02/2014	5-day	No	No	No
	05/02/2014	7-day	No	No	No
	07/02/2014	9-day	No	No	No
	11/02/2014	11-day	No	No	No
StarHub	27/10/2016	3-day	No	No	No
	31/10/2016	5-day	No	No	No
	02/11/2016	7-day	No	No	No
	04/11/2016	9-day	No	+ve	+ve
	08/11/2016	11-day	No	+ve	+ve
Danske Bank	10/07/2014	3-day	No	No	No
	14/07/2014	5-day	No	No	No
	16/07/2014	7-day	No	No	No
	18/07/2014	9-day	No	No	No
	22/07/2014	11-day	No	No	No
Deutsche Telekom	29/11/2016	3-day	No	No	No
	01/12/2016	5-day	No	No	No
	05/12/2016	7-day	No	No	No
	07/12/2016	9-day	-ve	-ve	-ve
	09/12/2016	11-day	No	-ve	-ve
Electronic Arts	18/03/2015	3-day	-ve	No	No
	20/03/2015	5-day	-ve	-ve	-ve
	24/03/2015	7-day	No	No	No
	26/03/2015	9-day	-ve	No	No
	30/03/2015	11-day	-ve	No	No
	03/01/2014	3-day	No	No	No
	07/01/2014	5-day	No	No	No

(to be continued on next page)

			Strategy		
Firm	Date	Event Period	1	2	3
Electronic Arts	09/01/2014	7-day	No	No	No
	13/01/2014	9-day	No	No	No
	15/01/2014	11-day	No	No	No
Electronic Arts	01/09/2016	3-day	-ve	No	No
	06/09/2016	5-day	No	No	No
	08/09/2016	7-day	No	No	No
	12/09/2016	9-day	No	No	No
	14/09/2016	11-day	No	No	No
Facebook	27/01/2015	3-day	No	No	No
	29/01/2015	5-day	No	No	No
	02/02/2015	7-day	No	No	No
	04/02/2015	9-day	No	No	No
	06/02/2015	11-day	No	No	No
Facebook	21/02/2014	3-day	No	No	No
	25/02/2014	5-day	No	No	No
	27/02/2014	7-day	No	No	No
	03/03/2014	9-day	No	No	No
	05/03/2014	11-day	No	No	No
Gjensidige Forsikr	10/07/2014	3-day	No	No	No
	14/07/2014	5-day	-ve	-ve	-ve
	16/07/2014	7-day	-ve	-ve	-ve
	18/07/2014	9-day	-ve	-ve	-ve
	22/07/2014	11-day	No	-ve	-ve
Activision Blizzard	29/01/2016	3-day	No	No	No
	02/02/2016	5-day	+ve	+ve	+ve
	04/02/2016	7-day	No	No	No
	08/02/2016	9-day	+ve	+ve	+ve

(to be continued on next page)

Firm	Date	Event Period	Strategy 1	Strategy 2	Strategy 3
	10/02/2016	11-day	+ve	+ve	+ve
ING	09/04/2013	3-day	-ve	-ve	-ve
	11/04/2013	5-day	-ve	No	No
	15/04/2013	7-day	-ve	-ve	-ve
	17/04/2013	9-day	No	No	No
	19/04/2013	11-day	No	No	No
JP Morgan Chase	30/01/2014	3-day	No	No	No
	03/02/2014	5-day	No	No	No
	05/02/2014	7-day	No	No	No
	07/02/2014	9-day	No	No	No
	11/02/2014	11-day	No	No	No
JP Morgan Chase	13/03/2013	3-day	No	No	No
	15/03/2013	5-day	No	No	No
	19/03/2013	7-day	No	+ve	+ve
	21/03/2013	9-day	No	+ve	+ve
	25/03/2013	11-day	No	+ve	+ve
Ziggo (Liberty Global)	18/08/2015	3-day	No	No	No
	20/08/2015	5-day	+ve	+ve	+ve
	24/08/2015	7-day	+ve	+ve	+ve
	26/08/2015	9-day	No	+ve	+ve
	28/08/2015	11-day	No	+ve	+ve
Linkedin	21/06/2013	3-day	No	No	No
	25/06/2013	5-day	No	No	No
	27/06/2013	7-day	No	No	No
	01/07/2013	9-day	No	No	No
	03/07/2013	11-day	-ve	No	No
	08/12/2010	3-day	No	No	No

(to be continued on next page)

Firm	Date	Event Period	Strategy 1	Strategy 2	Strategy 3
Master Card	10/12/2010	5-day	-ve	No	No
	14/12/2010	7-day	No	No	No
	16/12/2010	9-day	+ve	+ve	+ve
	20/12/2010	11-day	+ve	+ve	+ve
Microsoft	27/11/2013	3-day	No	No	No
	02/12/2013	5-day	No	No	No
	04/12/2013	7-day	No	No	No
	06/12/2013	9-day	No	No	No
	10/12/2013	11-day	No	No	No
Microsoft	24/12/2014	3-day	No	No	No
	29/12/2014	5-day	No	+ve	+ve
	31/12/2014	7-day	No	+ve	+ve
	05/01/2015	9-day	No	+ve	+ve
	07/01/2015	11-day	No	No	No
Nordea Bank	12/01/2015	3-day	No	No	No
	14/01/2015	5-day	No	No	No
	16/01/2015	7-day	-ve	-ve	-ve
	21/01/2015	9-day	-ve	-ve	-ve
	23/01/2015	11-day	-ve	-ve	-ve
Nissan	13/01/2016	3-day	No	No	No
	15/01/2016	5-day	No	No	No
	20/01/2016	7-day	+ve	No	No
	22/01/2016	9-day	No	No	No
	26/01/2016	11-day	No	No	No
Overstock.com	03/09/2015	3-day	No	No	No
	08/09/2015	5-day	No	No	No
	10/09/2015	7-day	No	No	No

(to be continued on next page)

Firm	Date	Event Period	Strategy		
			1	2	3
	14/09/2015	9-day	No	No	No
	16/09/2015	11-day	No	No	No
	22/12/2014	3-day	No	No	No
	24/12/2014	5-day	No	No	No
Rackspace	29/12/2014	7-day	No	No	No
	31/12/2014	9-day	No	No	No
	05/01/2015	11-day	No	No	No
	04/12/2013	3-day	No	No	No
	06/12/2013	5-day	No	No	No
RBS	10/12/2013	7-day	No	No	No
	12/12/2013	9-day	No	No	No
	16/12/2013	11-day	No	No	No
	24/12/2014	3-day	No	No	No
	29/12/2014	5-day	No	No	No
Sony	31/12/2014	7-day	No	No	No
	05/01/2015	9-day	No	No	No
	07/01/2015	11-day	No	No	No
	25/08/2014	3-day	No	No	No
	27/08/2014	5-day	No	No	No
Sony	29/08/2014	7-day	No	No	No
	03/09/2014	9-day	No	No	No
	05/09/2014	11-day	No	No	No
	26/11/2014	3-day	No	No	No
	01/12/2014	5-day	No	No	No
Sony	03/12/2014	7-day	No	No	No
	05/12/2014	9-day	No	No	No
	09/12/2014	11-day	No	No	No

(to be continued on next page)

Firm	Date	Event Period	Strategy 1	Strategy 2	Strategy 3
Storebrand	10/07/2014	3-day	No	No	No
	14/07/2014	5-day	No	No	No
	16/07/2014	7-day	No	No	No
	18/07/2014	9-day	No	No	No
	22/07/2014	11-day	No	No	No
AT&T	16/08/2012	3-day	No	No	No
	20/08/2012	5-day	No	+ve	+ve
	22/08/2012	7-day	No	No	No
	24/08/2012	9-day	No	No	No
	28/08/2012	11-day	No	No	No
TD Canada Trust	21/03/2013	3-day	No	No	No
	25/03/2013	5-day	No	No	No
	27/03/2013	7-day	No	No	No
	01/04/2013	9-day	No	-ve	-ve
	03/04/2013	11-day	No	No	No
Visa	08/12/2010	3-day	-ve	No	No
	10/12/2010	5-day	-ve	-ve	-ve
	14/12/2010	7-day	No	No	No
	16/12/2010	9-day	+ve	+ve	+ve
	20/12/2010	11-day	+ve	+ve	+ve
Vivendi	19/01/2012	3-day	No	No	No
	23/01/2012	5-day	No	No	No
	25/01/2012	7-day	No	No	No
	27/01/2012	9-day	No	No	No
	31/01/2012	11-day	No	No	No
	05/10/2011	3-day	No	No	No
	07/10/2011	5-day	No	No	No

(to be continued on next page)

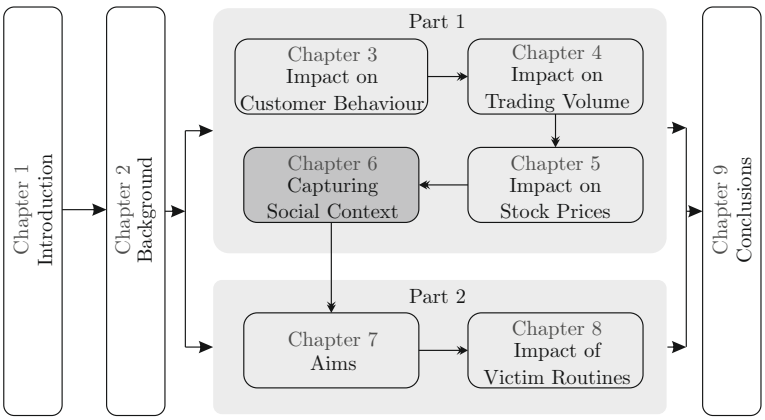
Firm	Date	Event Period	Strategy		
			1	2	3
Vodafone	11/10/2011	7-day	No	No	No
	13/10/2011	9-day	No	No	No
	17/10/2011	11-day	No	No	No
Verizon Communications	24/03/2014	3-day	No	No	No
	26/03/2014	5-day	No	No	No
	28/03/2014	7-day	No	No	No
	01/04/2014	9-day	No	No	No
	03/04/2014	11-day	No	No	No
Wells Fargo	20/12/2012	3-day	No	No	No
	24/12/2012	5-day	No	No	No
	27/12/2012	7-day	No	No	No
	31/12/2012	9-day	No	No	No
	03/01/2013	11-day	No	No	No

*The multiple events related to the same firm are sorted date wise.

This page is intentionally left blank.

Chapter 6

Capturing Social Context



Distributed Denial of Service (DDoS) attacks may lead to massive economic damages to victims. In most cases, the damage caused is dictated by the circumstances surrounding the attack (i.e. context*). One of the ways of collecting information on the context of an attack can be by using the online articles written about the attack. In this chapter, we introduce a dataset collected using* Google Alerts *that provides contextual information related DDoS attacks. We then show two case studies based on the data collected.*

6.1 Introduction

In Chapter 5 we have seen how we can measure the impact of DDoS attack announcements on the market value of a victim firm. Our results show that DDoS attacks have very short lived impact on stock prices. The stock prices tend to recover within 5-10 days of an attack. Studies have also shown that a number of network based businesses are resilient to relatively short DDoS attacks [4, 7, 140]. This is either because they are designed to be resilient or because the economic returns to the users of the service are not affected by the downtime. However, several reports by DDoS protection companies estimate the yearly costs of IT unavailability in millions of dollars [201]. These costs are usually computed using a simple linear metric or a victim survey which estimates the total damages based on subjective measures such as lost revenue, brand damage and operational cost etc. [37]. However, recent studies indicate that the average damages might not be as high as claimed by these reports. Florencio and Herley [69] find evidence that most cybercrime surveys are dominated by a minority of responses in the upper tail which leads to over estimation of losses.

The variation in the reported impact of DDoS attacks can be due to fact that these estimates do not take into account the complete *context* of an attack. *Context* is defined as *the circumstances that form the setting for an event.* As DDoS attacks only affect the availability of a network infrastructure unlike any other cyber attacks (e.g. attacks that target the confidentiality and integrity), the circumstances surrounding the attack event may dictate the consequences. In order to gather information on the *context* of an attack, we would need to interview the victims. Journalists working in the technology sector also perform such interviews. With advent of online media outlets most news is available on the web and can be used to gather contextual information on DDoS attacks. Services like *Google Alerts* can be used to collect such news articles.

In this chapter, we introduce a dataset collected with the help of *Google Alerts* that can provide *contextual* information on a publicly reported DDoS attack. We explore the different characteristics of the dataset collected in terms of number of different languages and domains (*urls*) that contribute to the dataset. We then with the help of two case studies show two possible use cases for this data. In the first case study, we show that the collected dataset is a much better source of news articles related to DDoS attacks as compared to LexisNexis. The second case study shows how the dataset can be used to track DDoS attack events. We also train and test a machine learning classifier that is able to tag attack reporting articles from the data.

6.2 Google alerts

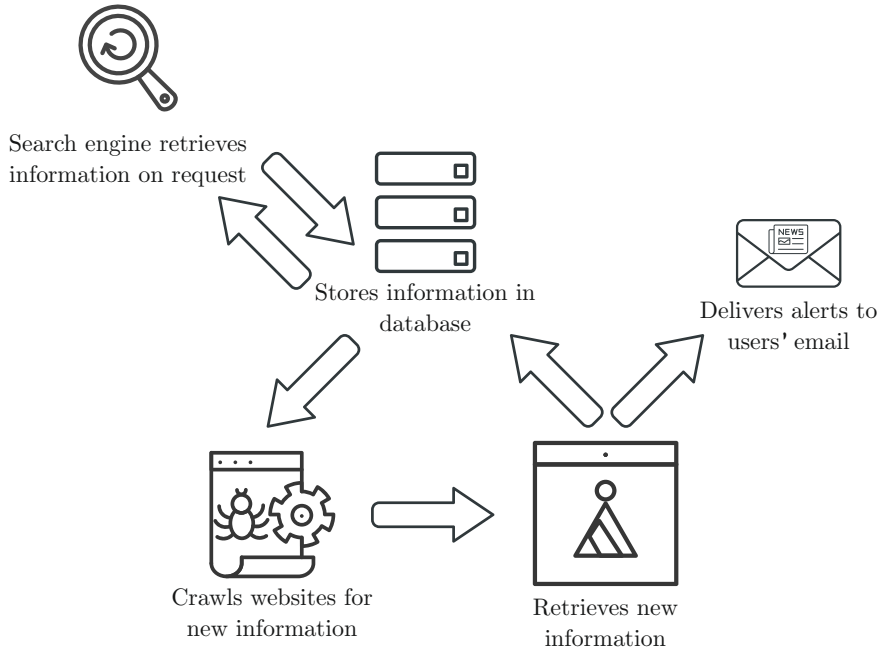


Figure 6.2.1: Generation of 'Google Alerts'.

In 2003, Google started a service named *Google Alerts* for helping users to keep themselves up to date on topics of their choice. In this chapter, we show how this service can be used to collect contextual information about current DDoS attack events. An internet bot that systematically browses the world wide web to collect new information is called a web crawler. The web crawlers of Google continuously search the world wide web for new content. If these crawlers find a new web page or change in content of an old web page, they store this information in a database for quick response. This process of making new web pages available for search engine user is known as *indexing*. The alerts are delivered via emails to the user when it finds new results, such as web pages, newspaper articles, blogs, or scientific research that match the user's search term(s). If a user has registered alerts related to a topic (trigger word or phrase), then it is possible for Google to notify a user about the new related content. Figure 6.2.1 shows the high level working of a search engine and the process of generation of 'Google Alerts'.

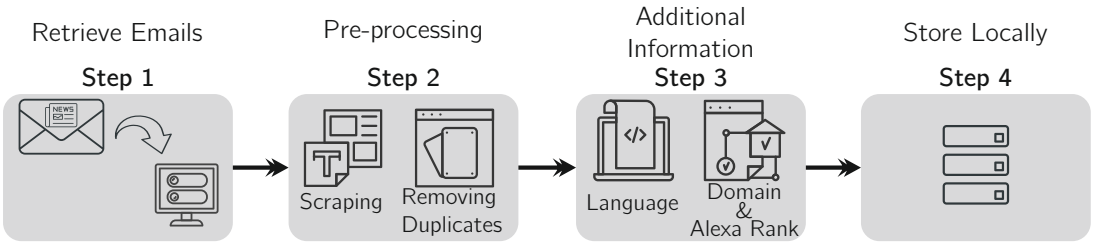


Figure 6.3.1: Data collection and processing steps.

This service can be extremely useful to someone who wants to keep themselves updated with respect to a selected topic. In this chapter, we show how we use this service to collect contextual information about DDoS attacks that are publicly reported. In Section 6.3 we explain the process of data collection and creation of dataset. Then in Sections 6.4 and 6.5 we discuss the two use cases for the collected dataset.

6.3 Google alerts dataset

Google Alerts is a content change detection and notification service. Once, a user registers the trigger word/phrase with Google, they start receiving emails with links that are relevant to the trigger(s). Here, we discuss our methodology to prepare this dataset.

6.3.1 Data collection

Using the *Google Alerts* service we collect articles on DDoS attacks. We do this by subscribing alerts on two trigger words: 1) 'denial of service' 2) 'ddos'. We select these triggers such that we are able to collect all articles related to DDoS attacks. We start collecting this data since 20th August 2015. Figure 6.3.1 shows the steps used for retrieving and processing the information from the emails before storing. These steps are as follows:

- **Step 1**: We download the alert emails from email server and store them in a local file storage for further processing. *Gmail* (i.e., email service by Google) allows a user to download selected emails as an *mbox* file.
- **Step 2**: In the pre-processing step, we scrape the text from the emails using the *mailbox* package* in Python and extract the following features for each article in an alert using *regular expressions*:

*https://docs.python.org/2/library/mailbox.html

	#Email Alerts on Trigger Word		#Articles Tagged as			
Year	'ddos'	'denial of service'	News	Web	# Domains	# Languages
2015*	2763	132	1427	3653	2467	37
2016	8084	350	4458	9387	4889	42
2017	7246	349	5805	9658	5692	44
2018	4863	313	5230	7005	5071	45

*Since 20th of August 2015

Table 6.3.1: Characteristics of the dataset.

  - Article Header
  - Associated Text
  - Type of Article (News or Web)†

  As we collect alerts using multiple trigger words and same article may be reported by both the triggers, we filter the alerts to remove any duplicate articles and proceed to step 3.

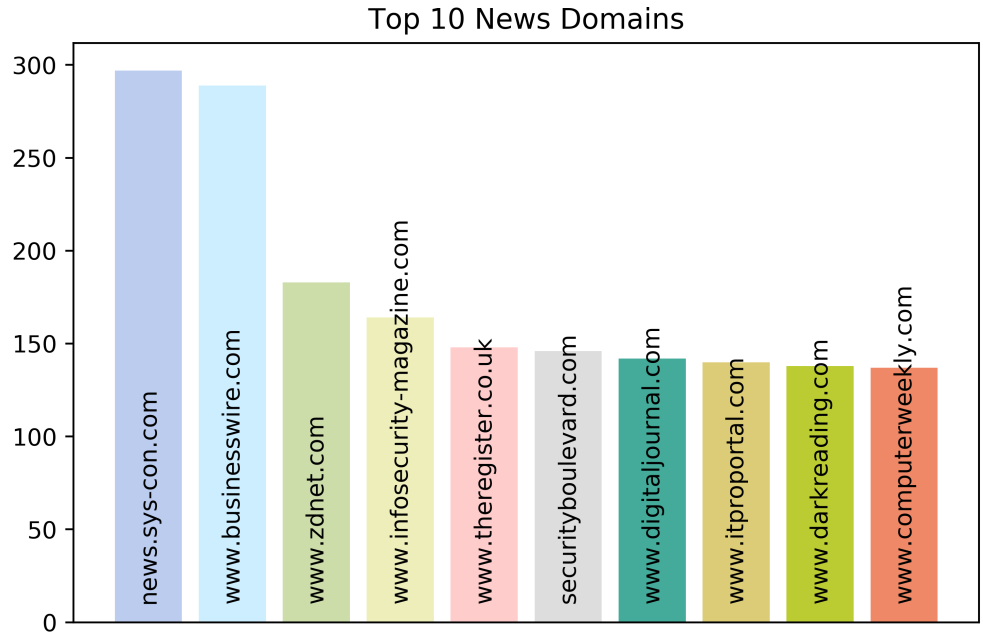
- **Step 3:** In this step, we introduce two additional features to the dataset based on the language‡ of the article and the historical alexa rank of the source (domain) of the article.

- **Step 4:** We store all data in a relational database.

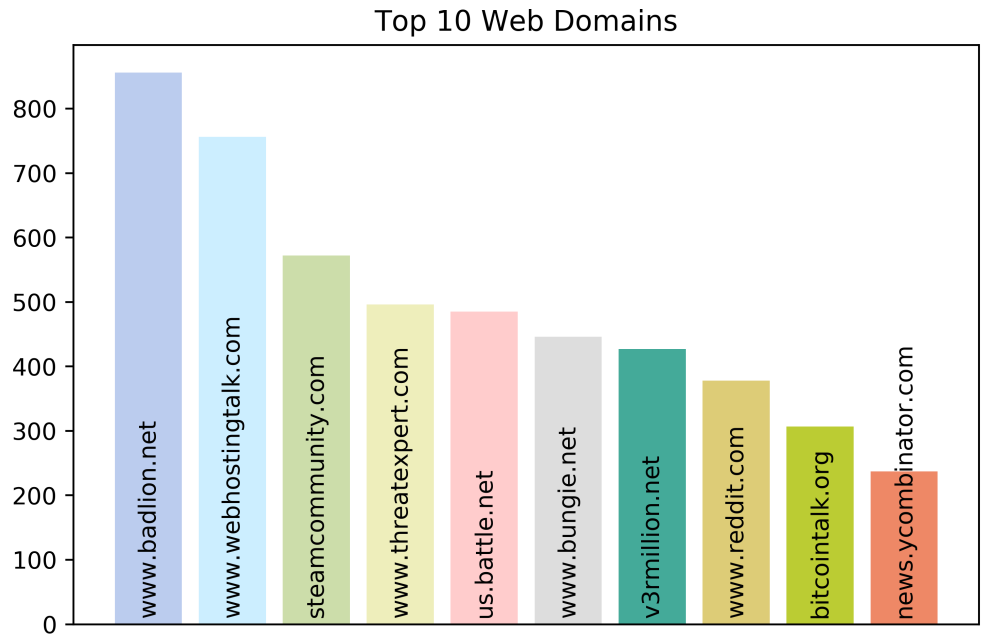
6.3.2 Characteristics of the dataset

In order to portray the breadth of data collected using the process explained above we show some of the characteristics of the data collected between 20th of August 2015 and 31st of December 2018 in Table 6.3.1. Based on the type of content posted on the websites, Google considers certain selected domains as news reporting domains. The alerts are tagged as *news* or *web* based on this classification. The collected dataset consists of a total of 41,543 alerts with 15,493 news alerts and 25,050 web alerts. These alerts are in a total of 47 different languages. Figure 6.3.3a shows the 10 most frequent languages in the dataset. It comes as no surprise that just as most of the content on the web, 87.7% of the alerts that we collect are in *English*. Other major languages in the dataset are: *Chinese, Russian, Japanese, French and Spanish*. The collected

†Google tags the articles depending on the source of information.
‡https://pypi.org/project/googletrans/

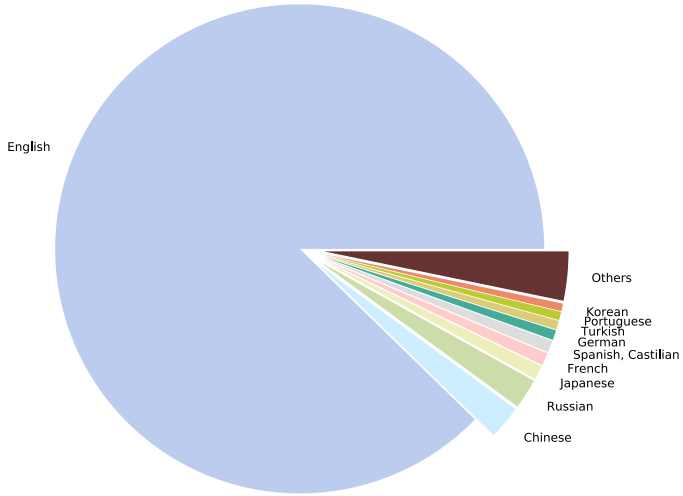


(a) Top 10 domains for News alerts.

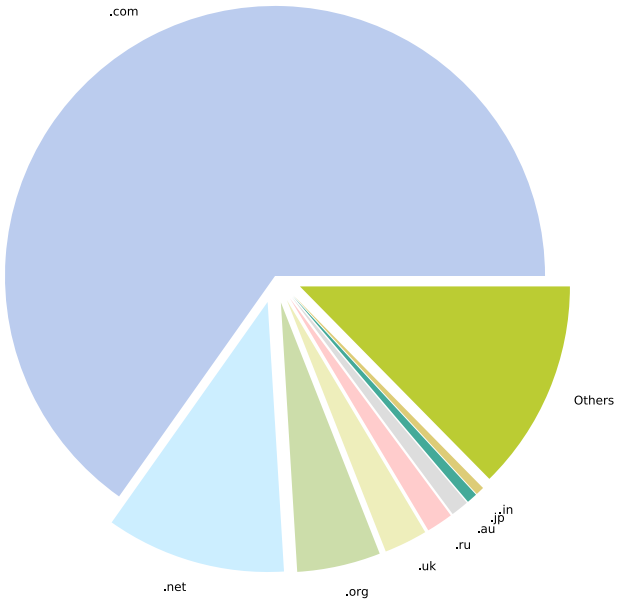


(b) Top 10 domains for Web alerts.

Figure 6.3.2: Most frequent domains in data collected between 20th of August 2015 and 31st of December 2018.



(a) 10 most frequent language of alerts.



(b) 8 most frequent top-level domains of source *urls*.

Figure 6.3.3: Most frequent languages and top-level domains in data collected between 20$^{\text{th}}$ of August 2015 and 31$^{\text{st}}$ of December 2018.

data belonged to 14,423 different sources (domains); 11,128 web sources and rest news sources. One of the benefits of using a service provided by a large search engine such as Google, is that we are able to gather data from a large number of different domains. Figures 6.3.2a and 6.3.2b show the top 10 news and web sources respectively. IT news websites such as *news.sys-con.com, businesswire.com and zdnet.com* form the top three news websites reporting about DDoS attacks and gaming and technology blogs such as *badlion.net, webhostingtalk.com and steamcommunity.com* are the top three web sources for DDoS attack related alerts. These websites belong to 288 different top level domains (TLDs). Figure 6.3.3b shows the distribution of the websites among top level domains. Most of the domains in the dataset belong to *'.com', '.net' and '.org'* TLDs. This characteristic of the data is also consistent with the most popular top level domains [55].

6.4 Case Study 1: Comparison with LexisNexis

Table 6.4.1: Characterstics of dataset used in case study 1.

Dataset	**Start Date**	**End Date**	**#Articles**	**#Articles (*en*)**
LexisNexis	20^{th} of August 2015	10^{th} of August 2018	441	405
Google Alerts			42,861	37,748

Researchers make use of services such as LexisNexis for a wide variety of studies that involve analysing news articles or their impact [157, 191]. LexisNexis is a media monitoring, risk management and research service. In this first of three case studies, we compare the dataset collected using Google Alerts with the data available on LexisNexis related to DDoS attacks. Such a comparison will help us in benchmarking our dataset.

We collect the data from LexisNexis using the following two keywords: 1) DDoS and 2) Denial of service. We then filter all articles dated between 20^{th} of August 2015 and 10^{th} of August 2018. Table 6.4.1 shows the characteristics of the two datasets. As the dominant language in both the datasets is English, for simplicity in comparison process we only consider the entries in English. We compare the two datasets by finding entries in the Google Alerts dataset that are similar to the entries in LexisNexis dataset. For doing so we compute the Levenshtein similarity ratio [127].

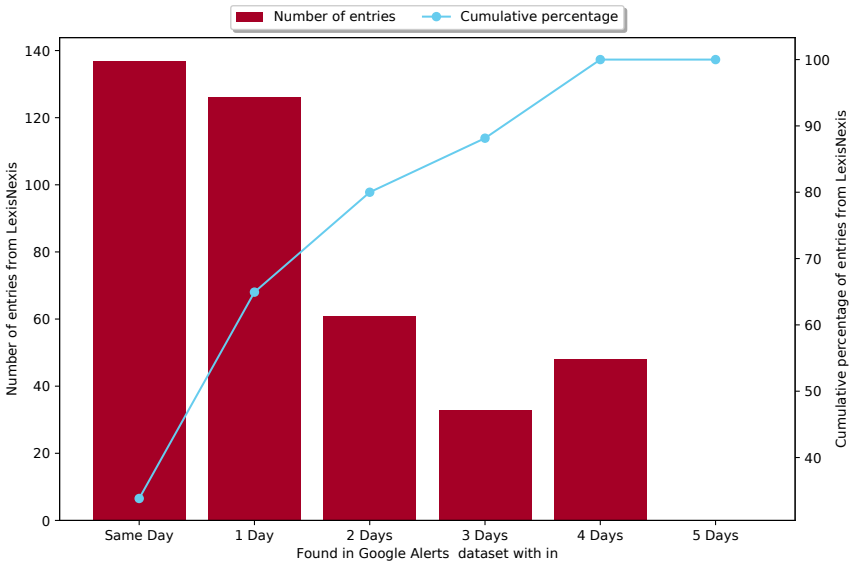


Figure 6.4.1: Number of entries found with in 5 days of the date in LexisNexis.

$$lev_{a,b}(i,j) = \begin{cases} max(i,j) & \text{if } min(i,j) = 0 \\ min \begin{cases} lev_{a,b}(i-1,j) + 1 \\ lev_{a,b}(i,j-1) + 1 \\ lev_{a,b}(i-1,j-1) + 1_{(a_i \neq b_j)} \end{cases} & \text{otherwise} \end{cases} \tag{6.1}$$

Levenshtein distance is a metric to measure the similarity/dissimilarity between two strings. It measures the minimum number of edits that one has to do in order to change one word sequence into another. The Levenshtein distance between first i characters of string a and first j characters of string b can be computed as shown in Equation 6.1. Where $1_{(a_i \neq b_j)}$ is a indicator function that is equal to 0 when $a_i = b_i$ and is otherwise 1. The Levenshtein similarity ratio can be calculated as shown in Equation 6.2. Where $|a|$ and $|b|$ are lengths of strings a and b respectively. For example, the Levenshtein distance between *"request"* and *"reject"* is 3, as minimum 3 edits are needed to change one into the other. An "edit" is defined by either insertion, deletion or replacement of a character.

$$\text{Levenshtein Similarity Ratio} = \frac{(|a| + |b|) - lev_{a,b}(i,j)}{|a| + |b|} \tag{6.2}$$

As LexisNexis data is based on news papers and most modern ones are also publicly available online (hence, crawled by Google), we assume that similar entries in the two datasets will have the same order of words. Hence, we use the Levenshtein similarity ratio to compute the similarity of each LexisNexis entry with the all Google Alerts entries that were recorded five days before or after the date mentioned in LexisNexis entry. We then manually cross check the most similar entry from Google Alerts dataset to verify if the article reported on the same event.

We observe that on all 405 instances, we found a similar entry in the Google Alerts database as that of LexisNexis. On nearly 15% instances, the Levenshtein similarity ratio was 1, and in case of 56.3% entries it was greater than 0.6. This result also justifies the use of Levenshtein ratio as the comparison metric. Figure 6.4.1 shows the number of similar entries found with in 5 days of the date in LexisNexis. Our results show that with the proposed data collection method we were able to collect 65% of the entries on LexisNexis related to DDoS attacks with in 1 days of the date mentioned on LexisNexis. Of course, with the help of the proposed method we are able to collect data on more number of events, than the selected ones recorded by LexisNexis. This limited coverage of LexisNexis has also be pointed out by some other studies [207].

6.5 Case Study 2: Tracking articles on DDoS attack events

Data collected using Google Alerts can be used to track DDoS attack events. Many times DDoS attacks are publicly reported by a number of websites. With the help of this dataset it is possible to track these articles. As a case study, we analyse the metadata of the articles related to four major DDoS attack events within the first 20 days of an attack. Using a regular expression based word search we calculate the number of articles on attacks on *Pokemon Go*, *OVH*, *Dyn* and *Github*. Fig. 6.5.1 shows the number of articles related to each of the attacks within 20 days of the attack. We observe a few hits in the dataset corresponding to the attack before the actual date of attack. These are definitely not the articles reporting a DDoS attack. Hence, we need to develop a filter to select only attack reporting articles. In order to track these articles we use off-the-shelf supervised machine learning algorithms that can be used for multi-class text categorisation.

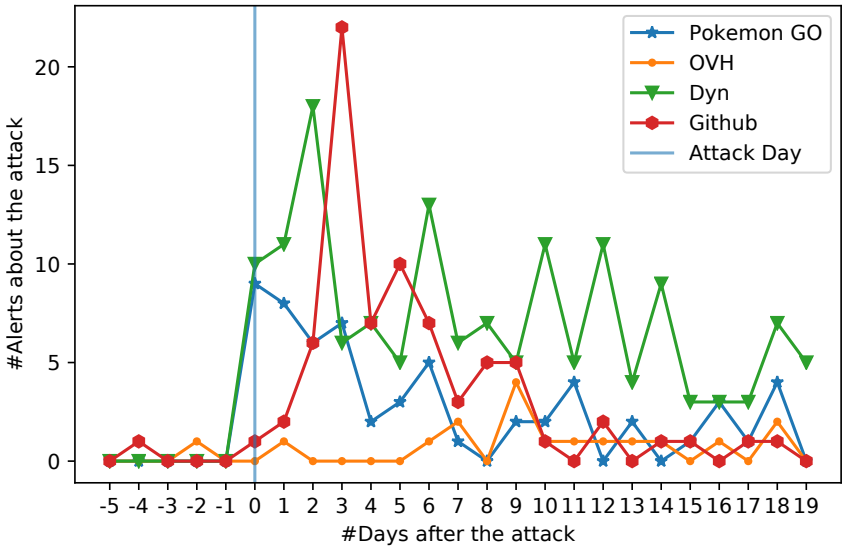


Figure 6.5.1: Tracking DDoS attack events using a simple word search without the *machine learning* filter.

6.5.1 Methodology

6.5.1.1 Machine Learning Classification

We use supervised machine learning classification algorithms, as opposed to unsupervised learning, these algorithms are able to learn rules based on a training dataset. It requires known labels (the corresponding correct outputs), so it can learn how to classify based on the targeted output class. We use the annotated classes as output data. We use the following two class definitions to prepare the training dataset:

1. **Attack**: describes an occurrence of a DDoS attack. For example: *"Dutch tax office, banks hit by DDoS cyber attacks (...)"*

2. **Other**: all other results not fitting in the attack category.

We create a balanced training dataset with 500 attack reporting articles and 500 other articles. We then test the effectiveness of various machine learning algorithms.

6.5.1.2 Used Algorithms

Kotsiantis [118] described the most-used supervised machine learning classifiers in detail. Kotsiantis categorized the algorithms into 6 categories: Decision Trees, Neural Networks, Naive Bayes, k-Nearest Neighbors, Support Vector Machines and Rule Learners. We compare algorithms from all categories, except k-Nearest Neighbors and Rule Learners because of their intolerance of noise and large computational time needed for classification [118]. We use the following 8 supervised machine learning algorithms:

- Logistic Regression *(LR)*: one of the most widely used algorithms for classification in the industry, performing well on linearly separable classes [162]. By default LR is a binary model, that is, only capable of separating two classes. To enable multi-class classification we make use of the One-versus-Rest (OvR) technique.

- Support Vector Machine *(SVM)*: introduced by Cortes and Vapnik, it is designed to maximize the so-called *margin* [47]. The margin is defined as the distance between the separating hyperplane (the decision boundary) and the training samples that are closest to this hyperplane, which are the so-called support vectors [162]. Models with a maximum margin tend to have a lower generalization error, whereas smaller margins are more prone to over-fitting. It occurs when the generalization a model creates to classify unseen data corresponds to closely to the training data. This results in poor performance, as the model takes the specific characteristics of the training data into account too much. Like LR, we use the OvR technique to enable multi-class classification, as SVM is binary by default.

- Decision Tree *(DT)*: introduced by Breiman, it is an easily interpretable model, designed to maximize the information gain [162, 29]. The DT makes decisions based on a series of questions the algorithm learns. This forms a tree of questions, in which each question is called a node. Each node leads to a decided category or another question. The question with the highest information gain, that is, the question that has most influence on the decided class, is on top. Pruning is performed to limit the size of the tree, thus avoiding over-fitting. The implementation we use is binary - meaning that every parent node is split up into two child nodes.

- Random Forest *(RF)*: can be intuitively seen as an ensemble of decision trees, in which a voting system for the most popular class is present. By combining multiple decision trees with the same distribution, each

suffering from high variance, the generalization error is decreased and the RF is less susceptible to over-fitting [30, 162].

- Extremely Randomized Trees *(ET)*: in short: Extra Tree. It was introduced by Geurts, Ernst and Wehenkel as a variation on the Decision Tree [75]. It randomizes both attribute and cut-point choices when splitting a tree node. The strong point of the ET is, besides a high predictive accuracy, the computational efficiency.

- Two Naive Bayes algorithms: Multinomial Naive Bayes *(M-NB)* and Bernoulli Naive Bayes *(B-NB)* Naive Bayes algorithms make the strong assumption that each prediction variable is independent from the others. In text classification this means that each word is seen as independent from all other words. In practice, Naive Bayes systems can work surprisingly well, even when the conditional independence assumption is not true [173].

  The Multinomial and Bernoulli prepositions refer to the distribution of the probability a feature belongs to a certain class. The Multinomial Naive Bayes is used for discrete data, so in our case for the total number of times a certain word appears. The Bernoulli Naive Bayes assumes the feature distribution is binary i.e., a word either exists (1) or not (0).

- Multi-layer Perception *(MLP)*: a network consisting of multiple layers containing single neurons. Neurons are referred to as ADAptive LInear NEurons (ADALINE), first published in 1960 by Widrow and Hoff. An interesting fact about ADALINE algorithm is that it focuses on minimizing continuous cost functions, and updates *weights* based on a linear activation function. Gradient descent optimization is used to learn the weight coefficients. Part of this gradient descent optimization is the learning technique called back propagation, introduced by [172]. We make use of Adam, a simple and computationally efficient algorithm for gradient-based optimization of stochastic objective functions [116].

6.5.1.3 Pre-Processing

Next, before applying the algorithms on the dataset, we performed preprocessing. First, we tokenized and lowercased the alerts. Next, we used *term frequency-inverse document frequency* (TF-IDF) weighting, as described by Salton and Buckley, for feature extraction [174]. This calculates the importance of terms in a document, and in this case, a category. Chi-Square

was used select the most outstanding correlated terms as identified by TF-IDF in each of the categories [182]. We ignored words that appeared in over 50% (max_df=0.5) of all alerts (maximum document frequency = 50%). We assume that these words don't significantly contribute to correct determination of a certain class. English stopwords have been ignored as well, using the English stopword dictionary[§] by the Glasgow Information Retrieval Group.

6.5.1.4 Analysis

To avoid over-fitting, we use the stratified k-fold cross-validation method as described by [117]. This method splits up the training dataset into k equal size subsets (folds). One fold is kept as validation data, while the other k-1 folds are used for training. The validation process will repeat k times so that every fold will be used as validation data once. The effectiveness is then the average of the k times the process has been repeated. We will stratify the folds "so that they contain approximately the same proportions of labels as the original dataset" [117]. This yields better bias and variance estimates, especially in cases of unequal class proportions [41]. In our case, we will vary the number of folds, starting with 4 folds. This means that we will create 4 folds of 250 Google Alerts each, and validate the algorithms with those folds accordingly.

To prevent information leakage, we used pipelines and a holdout dataset. Information leakage appears when information from the test data is used for training the model. This makes the model biased towards the test data, resulting in a too optimistic performance estimation. Pipelines were used to perform pre-processing solely on training data. Additionally, a part of the dataset was held out, only to be used for final testing, after the model has been trained.

Lastly, to find the best-fitting hyper-parameters for each of the algorithms and pre-processing methods, we have performed hyperparameter optimization. A hyperparameter is a parameter from a prior distribution; it captures the prior belief, before data is observed [164]. For example, we have varied the number of Chi-Square features to be used, and the maximum document frequencies of words. Additionally, algorithm-specific parameters have been tuned. Final testing as described in the previous paragraph has been performed using the found optimal set of hyperparameters.

6.5.1.5 Performance Measurement

The most simple way for measuring the performance of a machine learning algorithm is by looking at the number of articles that are correctly classified.

[§]http://ir.dcs.gla.ac.uk/resources/linguistic_utils/stop_words

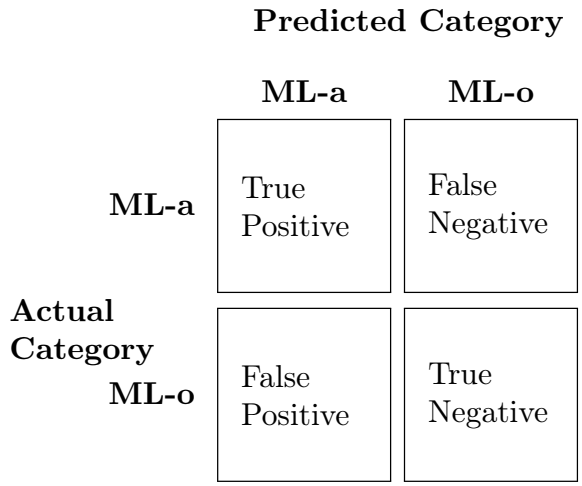


Figure 6.5.2: Confusion matrix.

One way of doing that is with the help of a confusion matrix [200]. In this case the confusion matrix is comprised of the following parameters:

True Positive (T_P) The percentage of articles in the testing dataset that were correctly classified as *attack* reporting.

True Negative (T_N) The percentage of articles in the testing dataset that were correctly classified as *other*.

False Positive (F_P) The percentage of articles in the testing dataset that were incorrectly classified as *attack* reporting.

False Negative (F_N) The percentage of articles in the testing dataset that were incorrectly classified as *other*.

A confusion matrix for a machine learning algorithm *ML* for classification in categories attack ('a') and other ('o') is shown in Figure 6.5.2. The confusion matrix reporting the parameter values for all 8 algorithms is shown in Appendix 6.A.

We also use other metrics for measuring the performance and robustness of machine learning algorithms. Each of these metrics test the The performance of the algorithms is measured using the F-score, which is the harmonic mean of precision and recall. *Precision* is the number of positive predictions divided by the total number of positive class values predicted. *Recall* is the number of

Table 6.5.1: Performance indicators for tested algorithms.

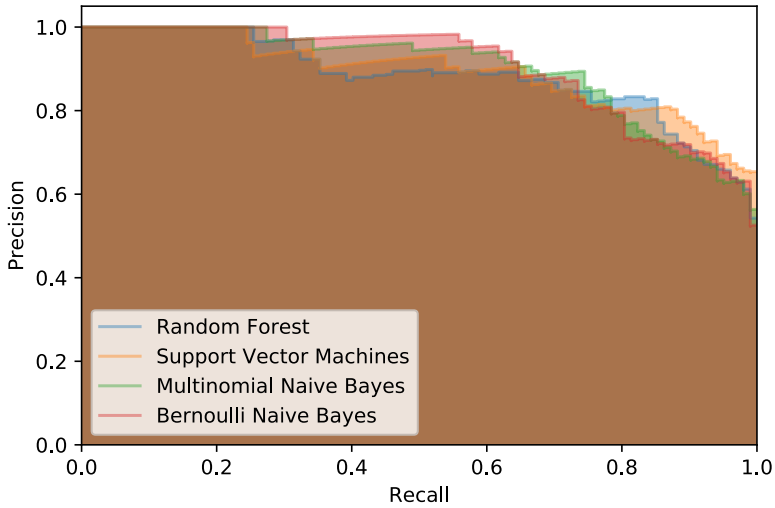
Algorithm	Score K1	Score K2	Score K3	F-Score	Recall	Precision	Accuracy	Skill
Random Forest	0.768	0.766	0.843	0.813	0.873	0.761	0.799	0.886
Support Vector Machine	0.778	0.831	0.855	0.809	0.853	0.770	0.799	0.899
Multinomial Naive Bayes	0.783	0.834	0.813	0.831	0.892	0.778	0.819	0.888
Bernoulli Naive Bayes	0.780	0.846	0.819	0.825	0.902	0.760	0.809	0.891
Logistic Regression	0.778	0.831	0.858	0.819	0.863	0.779	0.809	0.904
Decision Tree	0.721	0.716	0.757	0.709	0.706	0.713	0.711	0.704
Extra Tree	0.389	0.438	0.551	0.671	0.971	0.513	0.525	0.546
Multi-layer Perception	0.763	0.808	0.843	0.802	0.833	0.773	0.794	0.892

positive predictions divided by the number of positive class values in the test data [31]. Because the F-score takes both false positives and false negatives into account, it is a relatively better measure. The algorithm with the highest F-score can be considered to be most effective for classifying Google Alerts into the two categories. We further measure classification performance by computing True and False Positive Rates (TPR/FPR) for each algorithm. Based on the values of True and False Positive Rates we plot Receiver Operating Characteristic (ROC) curve. We further report the area under this curve that represents the *skill* of an algorithm in classifying the data.

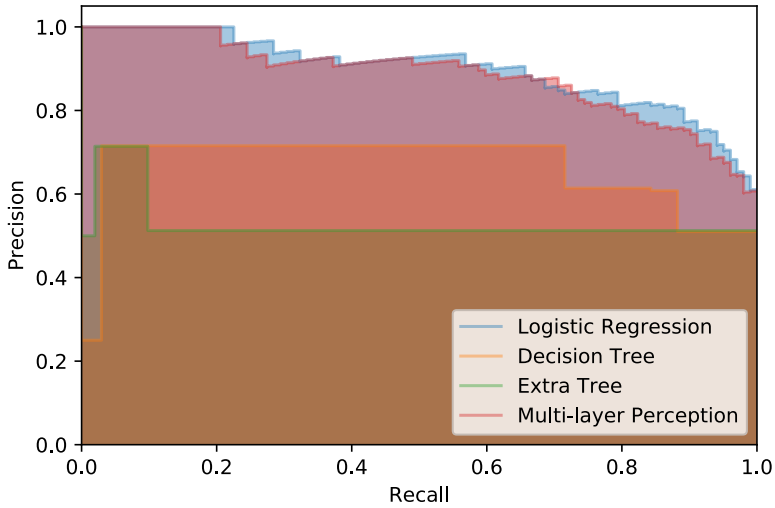
In Appendix 6.B, we show the words that were most definitive in classification of alerts in the two categories. We do so by computing the χ^2 statistic for each of the words in the articles.

6.5.2 Results

We test the performance of 8 different machine learning classifiers. In Table 6.5.1, we report the values of all the performance indicators discussed in the previous section. In order to avoid over fitting we use a three-fold cross validation. Column *Score K1* in Table 6.5.1 shows the cross validation score when the first fold is used as the testing dataset. We observe that except for *Extra Tree* algorithm, all other algorithm perform relatively well for each of the three folds. Also, there is not much difference in the values of *Score K1*, *Score K2* and *Score K3*. Hence, the models do not perform extremely better or worse for predicting any one of the three folds. While examining the *Precision* values, we notice that all the algorithms except *Decision Tree* and *Extra Tree* have a value higher than 0.75. This shows that 6 of the 8 algorithms that we test are able to classify *attack* reporting alerts atleast 75% of the times. The same can also be seen with the help of ROC curves as shown in Figure 6.5.4. In terms of

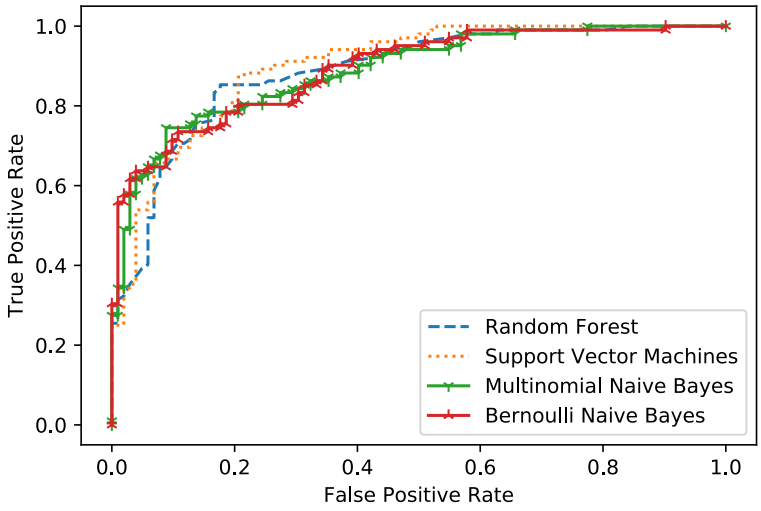


(a) Precision-Recall Curves (Part 1)

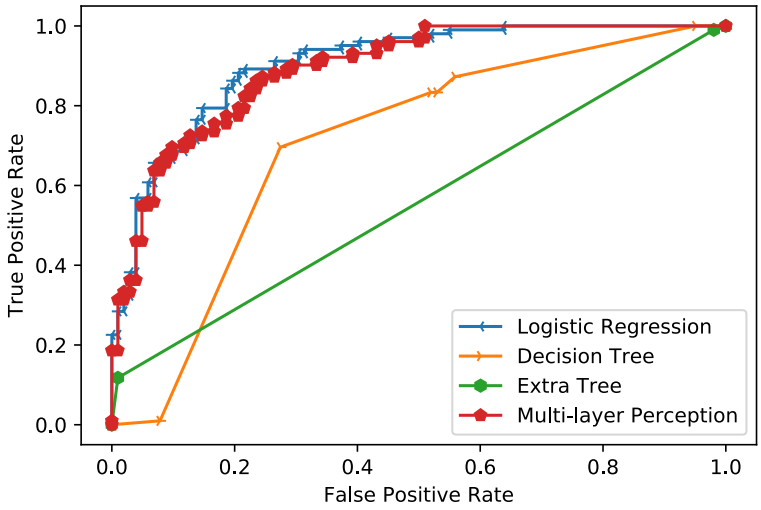


(b) Precision-Recall Curves (Part 2)

Figure 6.5.3: Precision-Recall curves.



(a) ROC Curves (Part 1)



(b) ROC Curves (Part 2)

Figure 6.5.4: Receiver operating curves (ROC).

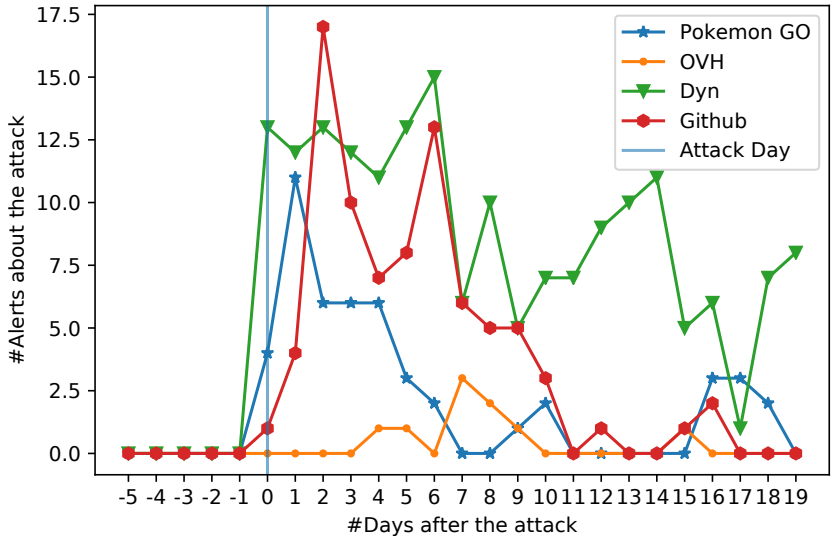


Figure 6.5.5: Tracking DDoS attack events using a simple word search with the *machine learning* filter.

Precision, *Logistic Regression* performs the best and is marginally better than *Multinomial Naive Bayes*. *Recall* wise *Extra Tree* performs the best, this is because it classifies most of the alerts as *attack* reporting. We can also observe this with the help of the confusion matrix as shown in Table 6.A.1. According to *F-Score* and overall *Accuracy*, Naive Bayes classifiers perform the best, they perform a slightly better than *Logistic Regression* and *Random Forrest* algorithms.

In order to chose, the most robust algorithm from four best performing ones (i.e., *Random Forrest, Naive Bayes and Logistic Regression*), we consider the *Skill* scores. They are defined as the area under the Precision-Recall curves (Figure 6.5.3). In terms of robustness, *Logistic Regression* classifier performs the best.

We use the *Logistic Regression* classifier to filter the entire dataset for *attack* reporting alerts. With the help of Fig. 6.5.5 we can clearly see that we are able to remove all noise from our dataset as there are no attack reporting articles before the attack day. We observe that we record a relatively large number of articles just after the attack day. This shows that we are able to successfully track articles reporting DDoS attack using our data collection strategy. Also, the fact that more articles discussed the attack on *Pokemon Go* than attack on

OVH shows that the popularity of an attack on web forums is not proportional to the intensity of an attack.

6.6 Concluding remarks and future works

In this chapter, we present a dataset that provides *contextual* information related to DDoS attacks. We explore the various characteristics of the dataset such as, number of languages, news sources,top level domains etc. to show the breadth of the data collection. With the help of two simple case studies we show two probable use cases for the presented dataset. In the first case study we compare our dataset with LexisNexis. We use Levenshtein similarity ratio to compare the articles from the two datasets and then manually validate the most similar articles. We show that our dataset covers all the articles that are indexed on LexisNexis. Using the proposed methodology, we collect all the articles within 4 days of LexisNexis. We also show that on nearly 56% of the cases, the similarity ratio was greater than 0.6 (i.e., the wordings of the articles compared was nearly the same). Hence, our dataset can be used as a trusted source of articles on DDoS attacks.

In the second case study, we test the accuracy and robustness of 8 different supervised machine learning algorithms for classifying the collected alerts as *attack* or *other*. We use 3-fold cross validation to counter over-fitting of classifiers. We find that *Logistic Regression* classifier performs the best. We are able to classify the alerts with an overall accuracy of 80.9%. Finally, we show the effectiveness of filtered (for *attack* reporting alerts) dataset for tracking DDoS attack events.

The goal of this chapter is not only to showcase our data collection methodology and the breadth of the dataset collected but also to invite other researchers for collaboration and to inform them about the dataset. Furthermore, we have already started using this data in several of our projects where contextual information about a DDoS attack event could improve our analysis. We use the dataset in Chapter 7 to identify publicly reported DDoS attacks in 2016. Currently, we update our dataset on a weekly basis and plan to openly publish the data on a website soon.

Appendix 6.A Confusion matrices

Table 6.A.1: Confusion Matrices for all 8 algorithms

Real Category \ Predicted Category	RF-a	RF-o	SVM-a	SVM-o	MNB-a	MNB-o	BNB-a	BNB-o	LR-a	LR-o	DT-a	DT-o	ET-a	ET-o	MLP-a	MLP-o
RF-a	87.3%	12.7%														
RF-o	27.5%	72.5%														
SVM-a			85.3%	14.7%												
SVM-o			25.5%	74.5%												
MNB-a					89.2%	10.8%										
MNB-o					25.5%	74.5%										
BNB-a							90.2%	9.8%								
BNB-o							28.4%	71.6%								
LR-a									86.3%	13.7%						
LR-o									24.5%	75.5%						
DT-a											73.5%	26.5%				
DT-o											28.4%	71.6%				
ET-a													99.0%	1.0%		
ET-o													92.2%	7.8%		
MLP-a															84.3%	15.7%
MLP-o															25.5%	74.5%

As a first appendix to this chapter, we present the confusion matrix showing the percentage of correctly and wrongly classified alerts for each algorithm in Table 6.A.1.

Appendix 6.B χ^2 statistic

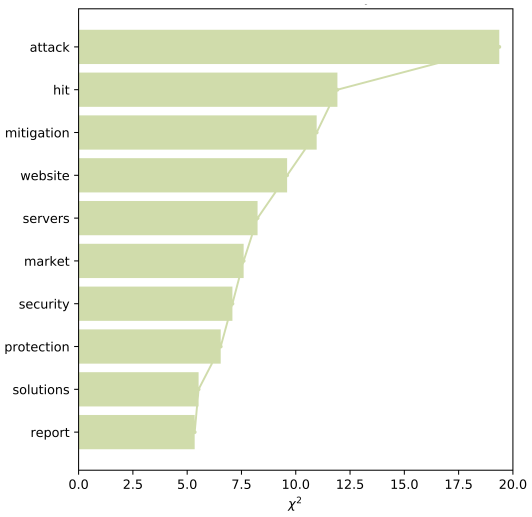


Figure 6.B.1: Top 10 χ^2 words.

The χ^2 statistic is a way to compute the lack of independence between the word w and document category i [129]. If n is the number of documents in the collection, $p_i(w)$ is the conditional probability of the category i for documents that contain w, P_i is the global fraction of documents that belong to category i, and $F(w)$ is the global fraction of the documents that contain the word w, the χ^2 statistic between word w and class i can be defined as:

$$\chi_i^2(w) = \frac{n.F(w)^2.(p_i(w) - P_i)^2}{F(w).(1 - F(w)).P_i.(1 - P_i)} \tag{6.3}$$

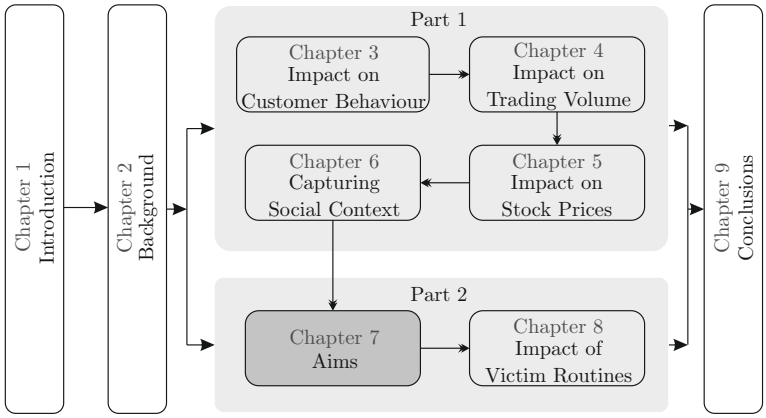
In Figure 6.B.1, we show the top 10 words by their χ^2 §[2]We use the python package provided by Pedregosa et al.[151] to make these computations. statistic. χ^2 values can be used to select dominant features in order to improve classification.

Part II

Attacker Aims

Chapter 7

Aims



Aims of an attacker for using DDoS attacks can determine whether an attacker is looking to inflict severe economic damage or not. In this chapter, we analyse the attacker aims for the use of DDoS attacks. We propose a model that can be used to evaluate news articles for determining probable aims of attackers. Thereafter, we apply this model to evaluate 27 distinct attack events from 2016. We make use of a DDoS specific longitudinal news database to select these attack events. We find the proposed model useful in analysing attack aims. We also find that in some cases attackers might target a web infrastructure just because it is virtually invincible.

7.1 Introduction and background

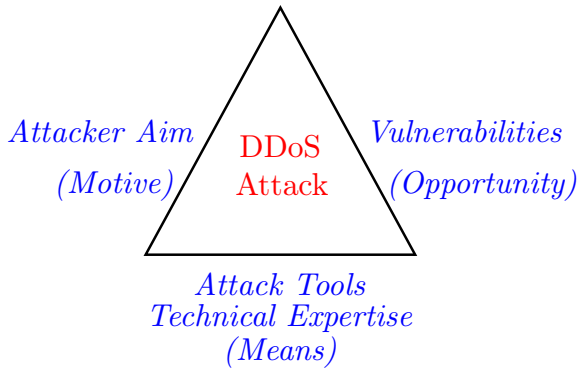


Figure 7.1.1: Aspects of a DDoS Attack

In order to protect itself a firm needs to evaluate its vulnerabilities and threats so as to plan its defence strategy [77]. These threats can be realised by acknowledging the various reasons for which the firm's IT infrastructure might become a target. Hence, it is important to investigate the aims of attackers for the use of DDoS attacks. Cyber criminals use DDoS attacks for a variety of reasons that may or may not be aimed at direct monetary gains [219].

Several theories have evaluated conventional crimes such as burglary, theft etc. to identify the drivers for criminals. Some of these theories have linked the characteristics of offenders with the type of crime they commit [42, 163]. Routine activity theory (RAT) emphasises on the circumstances in which a crime is committed to analyse criminal aims [43]. A conventional crime has three aspects that need to be proven before a wrongdoing is determined: Means, Motive and Opportunity. Just like conventional crimes, DDoS attacks require a means to execute, a motive to select the target and an opportunity to attack. In this case, *means* refers to the attack tools or the necessary technical expertise needed to execute the attack, the aim of the attacker points towards the reason for the attacker to act and vulnerabilities in the network provide the opportunity for the attack. Figure 7.1.1 shows the three aspects of a DDoS attack.

In this chapter, we focus on analysing attacker aims for the use of DDoS attack. The obvious way to investigate the aims of attackers is to interview them. However, it is also possible to model the probable aims based on the information reported by journalists in news articles related to the attack event.

Taking into account the socio-cultural, political and economic dimensions of DDoS attacks and the postulates of routine activity theory (RAT), we propose a model to analyse the content of news articles related to an attack. We then use this model to analyse probable attacker aims in 27 different cases from 2016.

7.2 Previous works

A few studies have tried to evaluate the attacker aims behind DDoS attacks. Hutchings & Clayton [94] discuss the incentives for booter owners. Paulson & Webber [149] evaluate the use of DDoS attacks for extortion against online gaming companies. Nazario [142] discuss politically motivated DDoS attacks. Later, Sauter [179] highlights the use of DDoS attacks for hacktivism purposes. Finally Zargar *et al.* [219] listed the probable incentives for attackers to use DDoS attacks as follows:

- Financial/economical gain: This is the motive when an attacker gets paid for the assault.
- Revenge: The motive of an attacker in this category is to DDoS for retribution.
- Ideological belief: The attackers in this category attack usually as a portrayal of disagreement.
- Intellectual Challenge: The attackers in this category experiment and learn from their activities. They are usually hackers who wish to show off their capabilities.
- Cyber warfare: The attackers in this category belong usually to a military or terrorist group.

However, Zargar *et al.* [219] do not provide any evidence for most of the listed motives. Some other studies also evaluated the non-technical characteristics of cyber attacks as a whole. Liu & Cheng [128] discuss the reasons for cyber attacks to happen. They also explain who these attackers are and how they conduct these attacks. Gandhi *et al.* [72] discuss the socio-cultural, political and economic (SPEC) dimensions of cyber attacks. They analyse selected security events between 1996 and 2010 on the basis of SPEC criteria. Sharma *et al.* [185] proposed a social dimensional threat model by using historical cyber attack events. On the basis of their model they evaluate 14 different news articles concerning cyber attacks. Geers *et al.* [74] analyse the nation-state motives

Table 7.3.1: Characteristics of the dataset.

Dates		#Articles		#Articles/day		Standard Deviation	
Start	End	Web	News	Web	News	Web	News
01-01-2016	31-12-2016	9387	4458	25.6	12.18	7.55	8.67

behind cyber attacks. Kumar & Carley [124] used network analysis on the data from Arbor network's digital attack map and World Bank to study the aims behind DDoS attacks. They conclude that there is an increase in the probability of attacks on the country if there are negative sentiments towards the country on social media.

All of the above mentioned studies show that not all attacks are carried out for economic gains. As booters have made DDoS attacks an easy weapon for nearly everyone, a number of aims can trigger attackers to launch an attack. These studies either evaluate the aims of attackers with respect to the SPEC criteria, or assume an aim and provide evidence to show the relevance of the aim in certain attacks. We believe that in case of DDoS attacks, attackers have to make two choices; 1) The victim (company or the individual they wish to attack). 2) Network infrastructure of the victim they wish to target. We propose a hybrid strategy for evaluating attacker aims by analysing the victim with respect to SPEC criteria and analysing the choice of infrastructure by considering the postulates of routine activity theory.

7.3 Methodology

Here, we discuss the characteristics of the dataset and the sampling strategy used by us to extract DDoS attack events. We then explain the proposed model for content analysis of news articles.

7.3.1 Dataset and sampling

The dataset is a collection of *Google Alerts* on DDoS attacks *. The collection process and some possible uses of the dataset are mentioned by Abhishta *et al.* [5]. Table 7.3.1 shows the characteristics of the dataset used in this research.

*Google Alerts is a content change detection and notification service. A user of this service can keep themselves updated about the topic of their choice. The service notifies with two types of alerts: 1) News 2) Web. News alerts report about content posted on news websites, all others are categorised as web alerts.

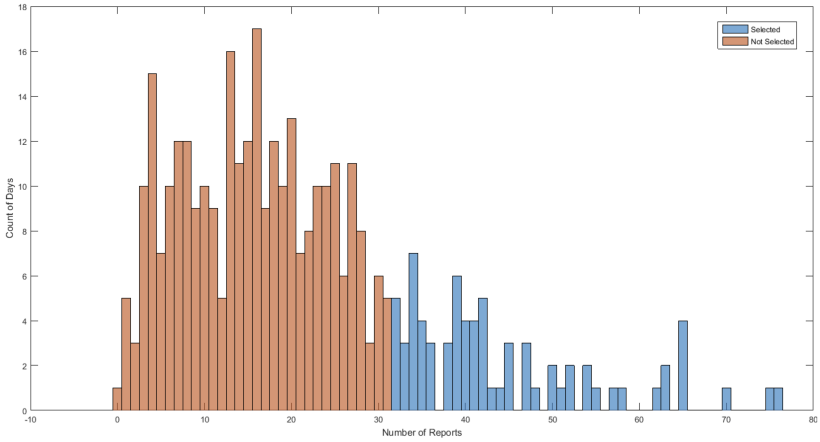


Figure 7.3.1: Histogram depicting selection criterion for *eventful* days.

In this chapter, we are looking for a sample of DDoS attack events that were discussed at length in the media. Hence, the goal of sampling is to extract the most reported DDoS attacks of 2016. We divide event sampling process into two parts: (1) We identify *eventful* days (2) We evaluate the 'News' alerts of an eventful day to extract attack events.

The statistical criteria for identification of 'eventful days' is based on the methodology also used by Kallus [108]. We consider the days on which the number of alerts were greater than θ as 'eventful'. In order to calculate the threshold θ we make use of the empirical distribution of number of alerts generated each day. Figure 7.3.1 shows the empirical distribution of number of 'News' alerts that are generated daily over the year. In this chapter, we consider the threshold to be at 20 percentile. If we consider top 10 percentile of the alerts then most of the eventful days lie in the second half of 2016 this is due to an enormous increase in reporting of DDoS attacks in the later half of the year. In this case, θ is calculated to be at 31.92 alerts. *Thus, if in a single day greater than or equal to 32 'News' alerts are reported then we consider that as an eventful day.* With this method, we are able to select 43 *eventful* days. We consider the alerts generated on *eventful* days for our study.

In order to identify the events responsible for the generation of abnormally high number of alerts on *eventful* days, we evaluate the text of all alerts on an *eventful* day and record the reported events as DDoS related events (non-attack) and DDoS attack events. We find that these news alerts report either an attack or an activity associated to an attack e.g. a research report by a DDoS

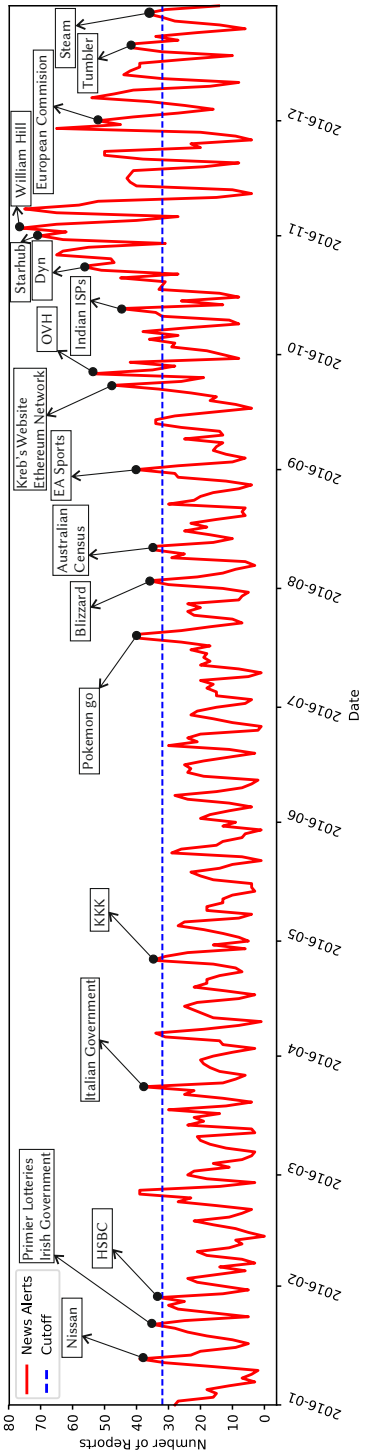


Figure 7.3.2: Attack time-line showing the extracted attack events for $\theta = 32$.

protection company, or steps taken by law enforcement agencies. We manually tag the content of the alerts on selected days to identify attack reporting alerts. The extracted attack events are shown in Figure 7.3.2. For this research we only consider the articles reporting a DDoS attack. We identify 27 separate attack events being discussed in these news articles.

7.3.2 Content analysis

The decision of the attacker to choose a target for a DDoS attack can be broken down in the following two components: 1) Choice of victim organistion to target. 2) Choice of network infrastructure to target. Figure 7.3.3 shows the model followed by us to analyse attacker aims. In Gandhi *et al.* [72] have shown that social, political, economic and cultural circumstances influence the choice of victim for an attacker. Hence, we evaluate the attacker's choice of victim using the SPEC criteria suggested by Gandhi *et al.*. For the choice of network infrastructure, we assume that the attackers are rational i.e. the attacker choose to launch an attack [49]. This assumption enables us to make use of the postulates of RAT. According to Cohen and Felson's (1979) [43] routine activities theory, direct contact predatory victimization occurs with the convergence in both space and time of three components: a motivated offender, the absence of a capable guardian, and a suitable target. According to routine activity theory, the suitability of a infrastructure for predation can be estimated using its fourfold constituent properties: value, inertia, visibility and accessibility, usually rendered in the acronym VIVA [217]. VIVA dimensions can be described as follows:

Value The importance of the infrastructure to the victim. For example, depending on the online sales of a company, a website can be more or less valuable to the company.

Inertia The degree of resistance posed by the infrastructure to an effective predation. So, a high inertia infrastructure will be the ones employing better protection strategies against DDoS attacks or the ones that can sustain high intensity network traffic (e.g. distributed servers, websites hosted in the cloud etc.).

Visibility The visibility of the objects an offender wishes to steal [126]. High visibility web infrastructures are mostly public facing such as, a public website.

Accessibility The ability of an offender to get to the target and get away from the scene of crime. An example of a high accessibility infrastructure can

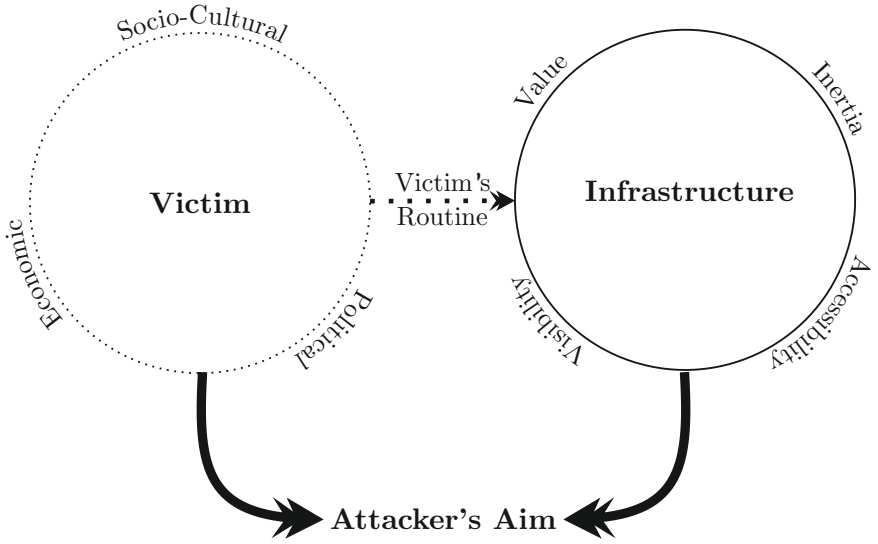


Figure 7.3.3: Model for analysing attacker aims using news articles.

be servers whose `ip address` can be easily accessed and are setup without intrusion detection systems or network monitoring applications.

With the help of the concepts discussed above, we develop a model for analysing the probable aims behind attack events. We analyse news articles related to 27 distinct attack events using this model to understand the attacker aims.

7.4 Results and discussion

Figure 7.3.2 shows the DDoS attack events reported on *eventful* days. As a result of filtering, a total of 43 dates were selected as *eventful* days. We evaluate all the alerts on these days and select DDoS attack events on the basis of the criteria mentioned in Section 7.3.1. The number of alerts collected on *eventful* days is 1929. Hence, these 11.75% of the days of the calender year account for nearly 43% (((*Number of news alerts on eventful days*)/(*Number of news alerts in the whole year*))*100) of the total 'news' alerts. This result supports the findings of Johnson [103] with respect to the concentration of traditional crimes, as traditional crime is also very much concentrated in time and space.

Table 7.4.1 summarises the components of each of the selected attack event i.e., victim, attacked infrastructure, SPEC variables and VIVA characteristics of the infrastructure. In the following paragraphs we discuss these attack reports

Table 7.4.1: Analysis of each of the selected attack event.

Date	*Reference*	*Victim*	**Socio-Cultural**	**Political**	**Economic**	*Infrastructure*	**Value**	**Inertia**	**Visibility**	**Accessibility**
13/01/2016	[15]	Nissan Motors	✓			Website	Low	Low	High	High
22/01/2016	[101]	Primier lotteries			✓	Ticket machines and Website	High	Low	High	High
22/01/2016	[80]	Irish government	✓	✓		Website	Low	Low	High	High
29/01/2016	[91]	HSBC			✓	Online Banking Server	High	High	Low	Low
26/02/2016	[14]	Italian government	✓	✓		Website	Low	Low	High	High
26/04/2016	[83]	Ku Klux Klan	✓			Website	Low	Low	High	High
20/07/2016	[153]	Pokemon go			✓	Gaming Server	High	Low	Low	High
03/08/2016	[25]	Blizzard			✓	Gaming Server	High	Low	Low	High
11/08/2016	[17]	Australian Census	✓			Website	High	Low	High	High
01/09/2016	[21]	EA sports			✓	Gaming Server	High	Low	Low	High
23/09/2016	[120]	Brian Krebs		✓		Website	Low	High	High	High
23/09/2016	[63]	Ethereum network			✓	Servers	High	Low	Low	Low
29/09/2016	[208]	OVH			✓	Hosting Server	High	High	Low	High
18/10/2016	[100]	ISPs in India			✓	Network Devices	High	High	Low	High
21/10/2016	[86]	Dyn			✓	Servers	High	High	Low	High
27/10/2016	[52]	StarHub			✓	Network Devices	High	High	Low	High
02/11/2016	[213]	William Hill			✓	Website	High	Low	High	High
08/11/2016	[56]	Canadian migration		✓		Website	Low	Low	High	High
08/11/2016	[211]	Wikileaks		✓		Website	High	Low	High	High
08/11/2016	[159]	Trump and Clinton		✓		Website	Low	Low	High	High
29/11/2016	[62]	Eir			✓	Email Server	High	Low	Low	Low
25/11/2016	[66]	Deutsche Telekom			✓	Network Devices	High	High	Low	High
30/11/2016	[65]	European Commission	✓	✓		Website	Low	Low	High	High
15/12/2016	[198]	Black lives matter	✓			Website	Low	Low	High	High
15/12/2016	[23]	BTC exchange			✓	Servers	High	Low	Low	Low
21/12/2016	[202]	Tumblr			✓	Website	High	Low	High	High
23/12/2016	[190]	Steam			✓	Gaming Servers	High	Low	Low	High

in detail and report our findings in accordance with the criterion discussed in Section 7.3.2.

In our analysis we see that the selected attack events can be broadly classified in 6 categories: 1) Attacks on large manufacturing companies 2) Attacks targeting public figures and ideological groups 3) Attacks targeting governments 4) Attacks on gaming and gambling platforms 5) Attacks on internet service providers and hosting service providers and 6) Attacks on financial institutions.

The first category includes the attack on Nissan Motors, all the global websites of the automotive company Nissan [15] were reported to suffer downtime. As Nissan does not sell cars online, the website is of relatively low value to the company. However, it was reported that the attack was carried out during Detroit auto show. During auto shows, car manufacturers expect attendants to visit their website to know more about the vehicle. Hence, even though Nissan doesn't sell cars online, the website has a high visibility during this period. Later reports suggested that Anonymous (hacker group) targeted the website to protest against whale hunting in Japan (justifies choice of the Nissan as a victim). Hence, high visibility of the website was the key input for the choice of target.

The second category include attacks on Ku Klux Klan, website of Brian Krebs, Black Lives Matter, Wikileaks, Donald Trump and Hillary Clinton [83, 120, 198, 159, 211]. The websites of this category of victims are easy targets and have high visibility. As a result of a protest against racism 'Anonymous' attacked the website of Ku Klux Klan [83]. According to the reports, websites on Wikileaks, Donald Trump and Hillary Clinton were targeted on the day of election result, showing socio-cultural reasons for the attacks.

The next category comprises of attacks on websites of Irish, Italian and Australian government [14, 80, 17]. These attacks could have been launched for both socio-cultural and political reasons as government websites usually do not cater online services. Italian government websites [14] were targeted by hacker group 'Anonymous'. The motivation behind the attack was to protest against the participation of local bodies in the Trans Adriatic Pipeline (TAP) project. However, the attack on Australian government website was clearly targeted to interrupt census data collection.

The fourth category includes online gaming platforms and gambling websites. The news sources reported an attack on the Irish lottery website [101] and vending machines that lead to the disruption of the sale of tickets. According to the reports this time the lottery jackpot was the highest in 18 months (*high value*). Hence, more people were expected to buy the tickets (*high visibility*). In July 2016, when the game 'Nintendo Pokemon Go' [153] was very popular (*high visibility*), another hacker group 'PoodleCorp' attacked the servers of the game. They took responsibility of the attack thus gaining a lot of publicity. Just after this online assault an attack on the servers of Blizzard was reported that made the Warcraft servers inaccessible for the gamers.

The fifth category comprises of attacks on ISPs and web hosting providers. In September and October 2016 attacks on ISPs in India [100], OVH (web hosting provider) [208] and Dyn (DNS service provider) [86] were reported. Usually ISPs form a high inertia targets for DDoS attacks. A new internet of things (IoT) based botnet, 'Mirai', whose code was released online was used for these attacks. Each of these attacks were bigger than the other in intensity.

The final category includes the attack on HSBC online banking services. As the attack was launched on last Friday of the month (salary day), the reasons for the attack was clearly economic. This is another example in our sample when the routine period affected the value of the infrastructure.

7.5 Conclusions and future work

In this chapter, we propose a model for analysing the attacker aims for using DDoS attacks. This model uses SPEC criteria for evaluating the reasons for choosing the victim and then studies the VIVA characteristics of the choice of infrastructure. We use this model to evaluate news articles related to 27 attack events that were reported in 2016. Our main conclusions are as follows:

- News articles are able to put DDoS attacks in context. Using the proposed model it is possible to evaluate the decisions made by the attacker to chose the victim and infrastructure.
- Companies need to monitor their socio-cultural and political environment at all times, not all attackers look for personal economic gains.
- All infrastructure connected to the internet is vulnerable to DDoS attacks. Companies must be aware of the degree of visibility and accessibility of the infrastructure. They should also consider their routine periods while analysing the VIVA characteristics of the infrastructure.
- Attacks on high inertia targets such as Dyn [86] show that, sometimes attackers target infrastructures because they are virtually invincible.

In this study, we only use data from 2016, hence we cannot derive conclusions on how often attackers are motivated by a particular aim. In the future, we would like to analyse a larger and more representative sample of all reported attacks. We hope to use the proposed model as a base for automatically detecting attacker aims from news articles reporting DDoS attacks.

Appendix 7.A Complete list of identified events.

Date	News Item	Event Type	Query Keywords
13-1-2016	Europol arrests key suspects of DD4BC extortion group.	Related News	
13-1-2016	Attack on Nissan website.	Attack Event	nissan
22-1-2016	Attack on Irish lottery Site and ticket machines.	Attack Event	irish, lottery
22-1-2016	Attack on irish government websites.	Attack Event	irish, govt
29-1-2016	Kasperisky lab released a report on DDoS attacks.	Related News	
29-1-2016	Attack on HSBC online banking.	Attack Event	hsbc
25-2-2016	Google's Project Shield starts protecting news websites.	Related News	
26-2-2016	Attack on Italian government websites.	Attack Event	italian, government
24-3-2016	US to charge Iran for attacks against banks.	Related News	
7-4-2016	Github suffers major outage.	Related News	
26-4-2016	Attack on KKK website.	Attack Event	kkk

(to be continued on next page)

Date	News Item	Event Type	Query Keywords
20-7-2016	Attack on pokemon go.	Attack Event	pokemon
3-8-2016	Attack on Blizzard's servers.	Attack Event	blizzard
11-8-2106	DDoScoin is introduced.	Related News	
11-8-2016	Attack on Australian Census Website.	Attack Event	Australian, census
12-8-2016	Attack on Australian Census Website.	Related News	
1-9-2016	EA sports servers suffer DDoS attack.	Attack Event	ea, sports, battlefield
13-9-2016	Two teens from Israel arrested for running a booter website. Vdos gets taken down.	Related News	
14-9-2016	Two teens from Israel arrested for running a booter website. Vdos gets taken down.	Related News	
23-9-2016	Attack on Brian Kreb's website.	Attack Event	brian, kreb, website
23-9-2016	IBM held responsible for failing the attack on Australian Census Website.	Related News	

(to be continued on next page)

Date	News Item	Event Type	Query Keywords
23-9-2016	Ethereum network under computational DDoS attack.	Attack Event	ethereum
26-9-2016	Hijacked IOT devices used for the attacks.	Related News	
26-9-2016	Google saves Brian Kreb's website.	Related News	
29-9-2016	Attack on hosting provider OVH.	Attack Event	ovh, hosting
29-9-2016	Hijacked IOT devices used for the attacks.	Related News	
5-10-2016	Mirai IOT malware responsible for attack on Brian Kreb's website.	Related News	
5-10-2016	Feds accuse two 19-year olds for lizard stresser and poodlecorp.	Related News	
7-10-2016	Feds accuse two 19-year olds for lizard stresser and poodlecorp.	Related News	
7-10-2016	Reports on Mirai botnet.	Related News	
13-10-2016	Reports on Mirai botnet.	Related News	

(to be continued on next page)

Date	News Item	Event Type	Query Keywords
13-10-2016	Singtel and Akamai announce strategic partnership to fight DDoS attacks.	Related News	
18-10-2016	Attacks on ISPs in India.	Attack Event	mumbai, pune
21-10-2016	Attack on Dyn.	Attack Event	dyn
24-10-2016	New World Hackers take responsibility for Dyn attack.	Related News	
24-10-2016	Reports on Dyn attack.	Related News	
25-10-2016	Xiongmai recalls 10000 webcams.	Related News	
25-10-2016	Reports on Dyn attack.	Related News	
27-10-2016	Reports on Dyn attack.	Related News	
27-10-2016	Attack on StarHub broadband.	Attack Event	starhub
1-11-2016	Reports on StarHub attack.	Related News	
1-11-2016	British Teen charged for Spamhaus attack. 2013.	Related News	
1-11-2016	Reports on Dyn attack.	Related News	
2-11-2016	William Hill website under attack.	Attack Event	william, hill
3-11-2016	Reports on Mirai botnet.	Related News	

(to be continued on next page)

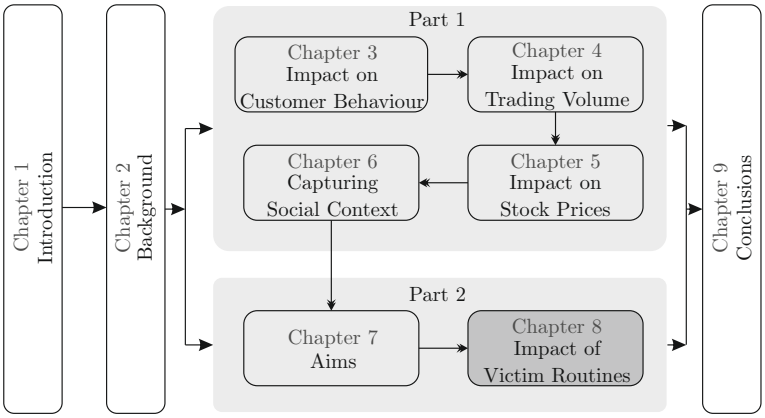
Date	News Item	Event Type	Query Keywords
8-11-2016	Canadian migration website attacked.	Attack Event	canadian, migration
8-11-2016	Attack against wikileaks.	Attack Event	wikileaks
8-11-2016	Attempted DDoS against Trump and Clinton's website.	Attack Event	trump, clinton
16-11-2016	Reports on IOT security.	Related News	
22-11-2016	Oracle buys Dyn.	Related News	
23-11-2016	Reports on oracle acquiring Dyn.	Related News	
29-11-2016	Eir's email system under attack.	Attack Event	eir
29-11-2016	Attack on Deutsche Telekom.	Attack Event	deutsche, telekom
30-11-2016	Attack against European Commission.	Attack Event	european, commission
1-12-2016	AWS launches shield against DDoS attacks.	Related News	
7-12-2016	Hackers gamify DDoS attacks.	Related News	
7-12-2016	New Mirai variant infecting home routers.	Related News	
13-12-2016	UK police crack down on people paying for DDoS attacks.	Related News	

(to be continued on next page)

Date	News Item	Event Type	Query Keywords
15-12-2016	FBI bust indian student for conducting DDoS attacks.	Related News	
15-12-2016	Attack on black lives matter website.	Attack Event	black, lives, matter
15-12-2016	BTC exchange taken down by an attack.	Attack Event	btc, exchange
16-12-2016	Reports on the attack on BTC.	Related News	
21-12-2016	Attack on Tumblr.	Attack Event	tumblr
23-12-2016	Attack on steam servers.	Attack Event	steam, servers
29-12-2016	Student charged for conducting DDoS attacks.	Related News	

Chapter 8

Impact on Victim Routines



In this chapter, we study the influence of daily routines of Dutch academic institutions on the number of DDoS attacks targeting their infrastructure. We hypothesise that the attacks are motivated and harness the postulates of Routine Activity Theory (RAT) from criminology to analyse the data. We define routine periods *in order to group days with similar activities and use 2.5 years of NetFlow alerts data measured by SURFnet to compare the number of alerts generated during each of these periods. Our analysis we shows clear correlation between academic schedules and attack patterns on academic institutions. This leads us to believe that most of these attacks are not random and are initiated by someone who might benefit by disrupting scheduled educational activities.*

8.1 Introduction

In the previous chapter we used RAT to develop a model that explains attacker aims for the use of DDoS attacks. Here, we use the postulates of routine activity theory to analyse the attacks targeting academic institutions. Over the years academic institutions have become more and more dependent on information and communications technology (ICT) to impart education. Today majority of the assignments submitted by students are via the web and a number of examinations are conducted online. Several e-learning strategies [168] are used by teachers to develop interactive content for students. Hence, ICT has become an indispensable resource for modern day educational institutions.

Network resources form the backbone of communication technologies and are under a constant scare of cyber attacks. Distributed denial of service (DDoS) attacks constantly threaten the availability of network resources. Even attackers with no prior knowledge of cyber attacks can order a DDoS attack using Booters [177]. In the recent years several academic institutions have become a victim of such attacks [203]. This raises the question: *why academic institutions are being targeted by DDoS attacks? Are these just random attacks on their network infrastructure or do attackers target them in a planned manner?*

In this chapter, we answer this question by analysing the timing of attacks that in the past have targeted the network infrastructure of SURFnet*. We hypothesise that the attacks are motivated and harness the postulates of Routine Activity Theory (RAT) from criminology to analyse the data.

Many studies in the field of criminology have shown the impact of attacker routines on crime rates [45]. Routine activity theory (RAT) suggests that changes in crime rates should be associated with days that affect the daily routines [44]. Holidays not only have an impact on attacker routines but also the routines of the victim. For instance, in the case of academic institutions all teaching related activities (classes and examinations) are on a halt during holidays. If the attacker's aim is to disrupt teaching related activities by means of a cyber attack then there is no incentive in launching such an attack during holidays. During vacations and weekends no lectures or examinations are scheduled. We leverage this feature of academic institutions to analyse if statistically significant number of attacks are driven by academic routines. We hypothesise that as greater disruption can be caused to academic activities during working days, we would observe more number of attempted DDoS attacks during this period.

*SURFnet is the primary supplier of advanced networking to Colleges, universities and research institutions.

Maimon, Kamerdze, Cukier and Sobesto [133] tested a similar hypothesis using the Intrusion Prevention System (IPS) data of a single university and showed that attacks on university are more likely to happen during business hours. In this we look to generalise the findings by Maimon, Kamerdze, Cukier and Sobesto by using data collected by SURFnet. As SURFnet provides network services to all academic institutions in The Netherlands, they are able to record all the attacks on Dutch academic institutions. On the basis of our analysis we show:

- how routine activity theory can be used to evaluate the influence of victim routines on attack patterns.
- that most of the attacks on academic institutions are not random. Daily routines of academic institutions heavily influence the rate of attack alerts.
- that the number of denial of service attacks targeting academic institutions in the Netherlands are significantly (statistically) higher during the working hour of working days as compared to holidays.
- that attack patterns do not change significantly (statistically) with type of holidays.

8.2 Method

The data in this research consists of alerts based on 1/100 sampled netflow using two different software: 1) NfSen [82] 2) Arbor Peakflow [150]. Both software were used to measure different alerts to avoid double counting. The alerts were based on packet rate triggers from both the software and are indicative of an attempted denial of service attack. The data were measured by SURFnet between 12:00:00 a.m. on 1st January 2015 and 12:00:00 a.m. on 30th June 2017. Thus, we make use of 2.5 years of attack alerts to test our hypothesis.

To analyse the impact of daily routines of academic institutions on the number of denial of service attacks we follow these steps:

Step 1: Define routine periods on the basis of academic calendar.

Step 2: Clean the data by filtering anomalies and exceptions.

Step 3: Group alerts in one hour periods and use dummy variables (1,0) to prepare the dataset for hypothesis testing.

Step 4: Formulate hypotheses using the postulates of RAT.

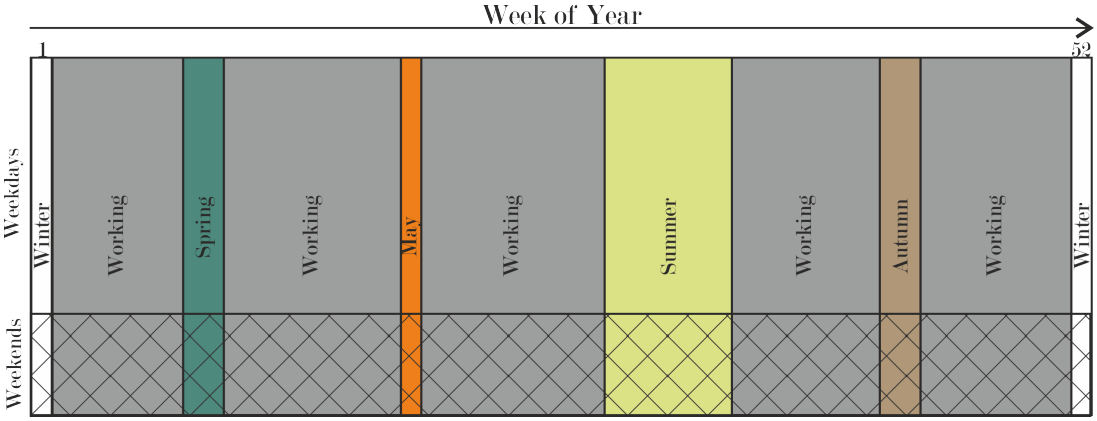


Figure 8.2.1: Routine periods in a calender year.

Step 5: Use an apt statistical method to test the null hypotheses.

In order to prepare this data for hypothesis testing, it is important for us to define the *routine periods*. The *routine periods* used in this study are shown in Figure 8.2.1. According to the academic calendar [143] in the years 2015, 2016 and 2017 we divide the calendar year in the following routine periods:

Winter Vacation:	1 and 52
Spring Vacation:	8 and 9
May Vacation:	18
Summer Vacation:	28,29,30,31,32,33 and 34
Autumn Vacation:	42 and 43
Working Weeks:	All other weeks

On a few occasions there was a week's difference between the start of vacations in the north and other regions of the Netherlands, in such a case we have considered the union of the vacation weeks from all regions as a *routine period*. In total 691 days of data belonged to working weeks and 221 days of data belonged to vacation weeks. We divide each week into *weekdays* and *weekends* (Saturday and Sunday) as the routines of academic institutions will be dissimilar in these periods. We group all the days belonging to the *vacation periods* and weekends as *holiday period* and others as *working period*. Based on this more broad division we get 417 days in the *holiday period* and rest of the 495 days in the *working period*.

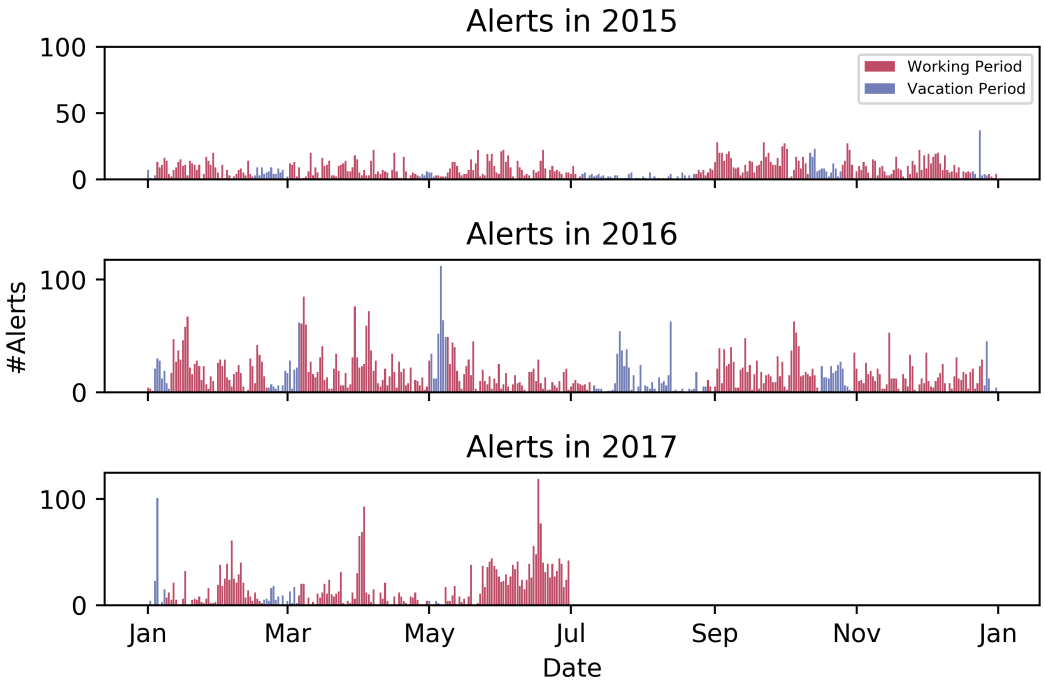


Figure 8.2.2: #Alerts collected in Working and Vacation periods.

In Section 8.2.1 we explain the steps taken by us to prepare the dataset, then in Section 8.2.2 we formulate our hypothesis and finally explain the statistical test used to test the hypothesis in Section 8.2.3.

8.2.1 Dataset

Here we discuss the dataset, the assumptions used by us to filter the anomalies in the data and method used to prepare the dataset for hypothesis testing. The 2.5 year long dataset consisted of 13,337 alerts before filtering and 11,777 alerts after filtering for multi-vector attacks and other anomalies. The year-wise distribution of the alerts are shown in Table 8.2.1. Figure 8.2.2 shows the number of alerts collected on each day.The number of attacks substantially increased since 2016. We can see that on an average greater number of alerts were collected in 2016 and 2017 as compared to 2015. However, no changes in measurement systems were carried out by SURFnet.

In order to count multi-vector attacks as a single attack, we merge alerts having the same time-stamp as a single alert. To account for larger attacks

Table 8.2.1: Dataset.

Year	#Alerts	#Alerts/day
2015	2780	7.62
2016	6022	16.45
2017	2975	16.44

that might generate multiple alerts, we merge all alerts where the difference in time-stamps is less than 5 minutes. If we encounter a *Large SYN* alert (alert generated due an oversized SYN packet) followed by *TCP SYN* alerts, then we filter these alerts as it is indicative of an active botnet on SURFnet's network that might have been used to attack some other network infrastructure. In this chapter, we assume that this filtered dataset provides the number of alerts are representative of the number of attempted denial of service attacks on SURFnet's infrastructure. To prepare the data were then grouped into one-hour periods by calculating the total number of alerts generated each hour.

Dummy variable (1,0) coding was used to assess the differences between each of the *routine periods*. A similar coding was done to distinguish between the larger groups; *holiday period* and *working period*. We also code the dataset to show working and non-working hours of the day. As most educational activities are planned between 8:00 a.m. and 6:00 p.m., we consider these hours as *working hours* and others as *non-working hours*.

8.2.2 Hypotheses

Hypothesis testing is required to test the statistical significance of the differences that we might observe with the help of descriptive statistics. In this section, we develop the hypotheses and formulate the corresponding null hypotheses that we will test using the dataset described in the previous section. We base all our hypotheses on the following postulate of RAT: *change in victim routines will impact rate of attacks on the victim*. Hence, in case of academic institutions we hypothesise that *routine periods* will impact the number of denial of service attacks targeting their network infrastructure. In total we develop 9 different hypothesis to compare the number of attempted attacks in each of the routine periods. Table 8.2.2 shows the hypotheses and the corresponding null hypotheses. The null hypothesis assumes that there is no significant difference between the routine periods subjected to a statistical test. Hence, in cases where

Table 8.2.2: Hypotheses and corresponding null hypotheses*.

Hypothesis	**Null Hypothesis**
$H1$: The average number of attack alerts generated during the *working period* is higher than in the *holiday period*.	$H1_0$: There is no significant difference in the average number of alerts generated during the *working period* and the *holiday period*.
$H2$: The average number of attack alerts generated during the weekdays of *working weeks* period is higher than in the weekends.	$H2_0$: There is no significant difference in the average number of alerts generated during the weekdays and weekends of *working weeks* period.
$H3$: There is no significant difference in the average number of alerts generated during the weekdays and weekends of vacation period.	$H3_0$: There is no significant difference in the average number of alerts generated during the weekdays and weekends of vacation period.
$H4$: The average number of attack alerts generated during the *working weeks* period is higher than in the Vacation *routine periods*.	$H4_0$: There is no significant difference in the average number of alerts generated during *working weeks* period and the Vacation *routine periods*.
$H5$: There is no significant difference in the average number of alerts generated during the weekends of vacation and *working weeks* period.	$H5_0$: There is no significant difference in the average number of alerts generated during the weekends of vacation and *working weeks* period.
$H6$: There is no significant difference in the average number of alerts generated during any of the *vacation periods*.	$H6_0$: There is no significant difference in the average number of alerts generated during any of the *vacation periods*.
$H7$: The average number of alerts generated in the working hours of *working weeks* period are higher than in the non-working hours.	$H7_0$: There is no significant difference in the average number of alerts generated during the working and non-working hours of *working weeks* period.
$H8$: There is no significant difference in the average number of alerts generated during the working and non-working hours of vacation periods	$H8_0$: There is no significant difference in the average number of alerts generated during the working and non-working hours of vacation periods.
$H9$: There is no significant difference in the average number of alerts generated during the working and non-working hours on the weekends.	$H9_0$: There is no significant difference in the average number of alerts generated during the working and non-working hours on the weekends.

*The hypothesis and the null-hypothesis is the same when in accordance with the theory we do not expect a difference in group averages.

on the basis of RAT we expect no difference in attack patterns, the hypothesis and the null hypothesis are the same.

For the 1^{st} hypothesis we consider the two large groups: *holiday period* and *working period*. As greater damage can be done to an academic institution when it is a working day, for the second hypothesis we consider the weekdays and weekends of the *working weeks* period. Alternatively, one could also argue that in the vacation weeks statistically there is going to be no difference in the rate of attacks on academic institutions during weekdays and weekends. In case of $H3$ we test this aspect of the dataset. As vacation *routine periods* and *working weeks* period have contrasting routines for the academic institutions, we formulate 4^{th} hypothesis on this basis. For the 5^{th} hypothesis we compare the number of alerts generated during the weekends of vacation and *working weeks* period. In the 6^{th} hypothesis, we test if type of vacation period (summer, spring, etc.) has an impact on the number of alerts.

In the next three hypothesis, we analyse the impact of hour of the day on attack pattern. As mentioned in the previous section, we group the hours of a day in working hour and non-working hour category. The routines of academic institution vary during working and non-working hours, there are several other businesses where this might not be the case (e.g. e-commerce). In the 7^{th} hypothesis we analyse the difference in attack patterns during the working and non-working hours on a weekday in the *working weeks* period. Through $H8$ and $H9$ we analyse if there is an impact of working and non-working hour categories on the weekend and vacation periods.

8.2.3 Testing

In order to test the null hypotheses we make use of Analysis of Variance (ANOVA), a collection of statistical models and their associated estimation procedures (such as the "variation" among and between groups) used to analyse the differences among group means in a sample. Studies have used ANOVA to analyse NetFlow samples to detect anomalies [114]. A *student's t-test* may also be used to analyse the differences among means of two samples but it cannot be used for more than two samples as required in the case of hypothesis $H6$. We use a one-way ANOVA in order to test the statistical significance of differences between *routine periods* [104].

8.3 Results

In this section we discuss the results of the statistical tests. First, we look at the descriptive statistics, then we discuss the results of ANOVA to establish

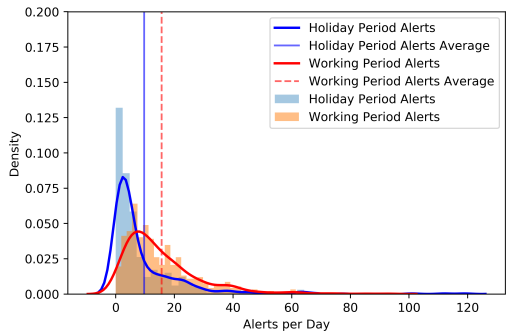
Table 8.3.1: Results of ANOVA.

Hypothesis	F-Statistic	Significance (p-value)	Null Hypothesis Status
$H1$	32.911	0.000	Rejected
$H2$	22.570	0.000	Rejected
$H3$	0.000	0.989	Not Rejected
$H4$	12.470	0.000	Rejected
$H5$	0.000	0.985	Not Rejected
$H6$	1.774	0.151	Not Rejected
$H7$	33.475	0.000	Rejected
$H8$	0.506	0.477	Not Rejected
$H9$	8.739	0.003	Rejected

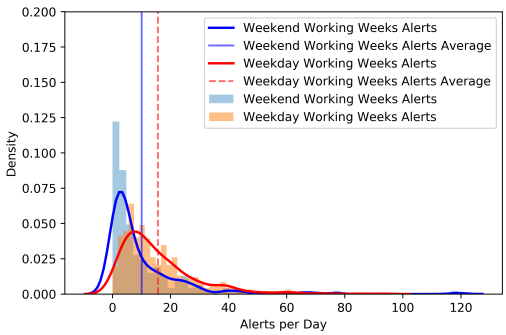
statistical significance for each hypothesis. Figures 8.3.1 and 8.3.2 shows the descriptive statistics for each of the pair of *routine periods* for which a hypothesis is tested. In each sub-figure we plot number of alerts on the *x-axis* and density (proportion of days on which corresponding number of alerts were generated). We also show the average number of alerts in each *routine period* in the plot. Table 8.3.1 shows the test statistic (F-statistic) and the significance of the test statistic. It also shows if on the basis of the results we are able to reject the null hypothesis or not.

Figure 8.3.1a shows the descriptive statistics for the pair of periods considered in $H1$. With the help of this figure we can clearly observe that the average number of alerts generated in the *working period* is considerably higher than the average number of alerts generated in the *holiday period*. Based on the density plot we can see that there are more number of days with a higher number of alerts in the working period, most of days in the holiday period has very few number of alerts recorded. ANOVA analysis of $H1_0$ resulted in a F-statistic of 32.911 and a p-value of 0.000. This shows that the difference between the average number of alerts generated during the *working period* and *holiday period* is very high. A low p-value shows high confidence of the statistical test. Hence, we can reject null hypothesis $H1_0$.

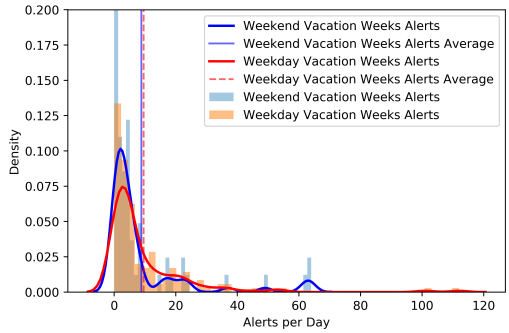
The descriptive statistics related to $H2_0$ are shown in Figure 8.3.1b. We observe that the average number of alerts generated in the weekdays of *working*



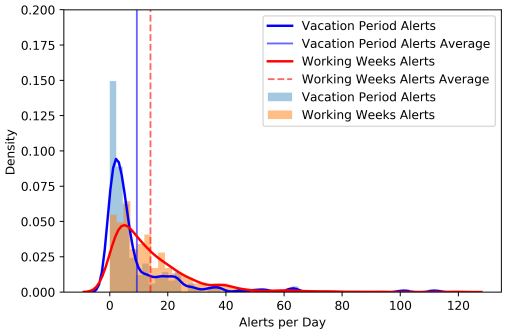
(a) Difference between *working* period and *holiday period*



(b) Difference between Weekdays and Weekends (Working Weeks)



(c) Difference between Weekdays and Weekends (Vacation Weeks)



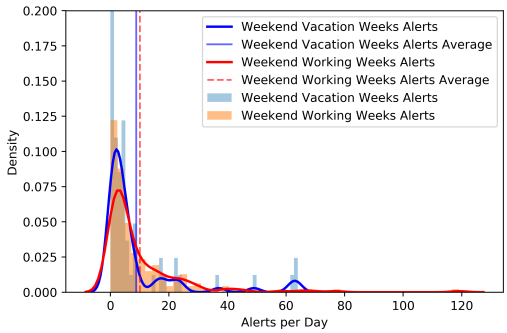
(d) Difference between Vacation Weeks and Working Weeks

Figure 8.3.1: Empirical Distributions showing difference in the number of alerts generated per day during various *routine periods*.

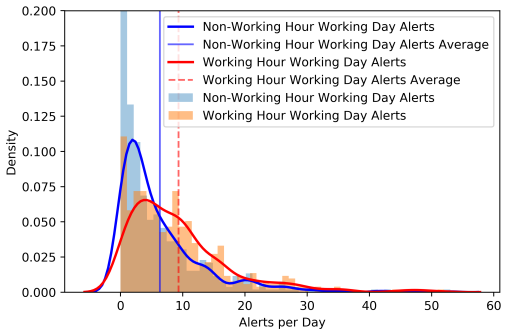
weeks period are greater than the average number of alerts generated in the weekends of the same period. The density plots again show that high number of alerts are generated on more occasions on weekdays. ANOVA analysis of $H2_0$ resulted in a F-statistic of 22.570 and a p-value of 0.000. Hence, in this case as well we reject the null hypothesis with high degree of confidence.

Figure 8.3.1c shows the descriptive statistics of alerts generated during the weekdays and weekends of vacation periods. According to this figure we observe that the average number of alerts generated during the weekday in the vacation period is nearly equal to the average number of alerts generated on a weekend (Saturday or Sunday) in the vacation period. Both density plots in this case

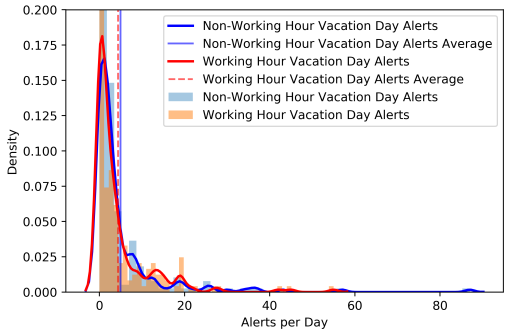
also show a considerable overlap. ANOVA analysis of $H3_0$ also resulted in a F-statistic of 0.000 and a high p-value 0.989. Thus, we cannot reject the null hypothesis.



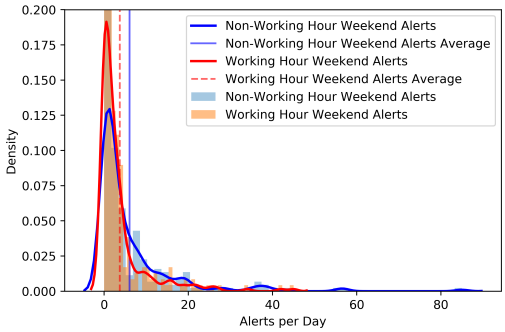
(a) Difference between the weekends of Vacation periods and Working Weeks Period



(b) Difference between Working hour and non-working hour Alerts (Working Weeks)



(c) Difference between Working hour and non-working hour Alerts (Vacation Weeks)



(d) Difference between Working hour and non-working hour Alerts (Weekends)

Figure 8.3.2: Empirical Distributions showing difference in the number of alerts generated per day during various *routine periods*.

For analysing null hypothesis $H4_0$ we take help of Figure 8.3.1d. Here we compare the number of alerts generated during the *working weeks* period and the vacation periods. The descriptive statistics in this care are similar to the ones in the case of $H1_0$ and $H2_0$. The average number of alerts generated in the *working weeks* period is higher that the average number of alerts generated in

the vacation periods. As the ANOVA analysis also resulted in a high F-statistic of 12.470 and a low p-value of 0.000, we reject the null hypothesis $H4_0$.

Figure 8.3.2a compares the number of alerts generated during the weekends of vacation periods with the number of alerts generated during the weekends of working weeks periods. We hardly observe any difference in the average number of alerts generated in the two periods. We also see a significant overlap in the density plots. In this case, the ANOVA analysis resulted in a low F-statistic of 0.000 and a high p-value of 0.985. Hence, we do not reject the null hypothesis $H5_0$.

With the help of null hypothesis $H6_0$ we test if there is a difference between the average number of alerts generated during the five vacation periods. The ANOVA analysis in this case resulted in a low F-statistic of 1.774 and a relatively high p-value of 0.151. As the p-value is greater than 0.05, it is not possible to reject the null hypothesis.

Figure 8.3.2b differentiates between number of alerts generated during the working and non-working hours of the *working weeks* period. We observe that a significantly higher number of attack alerts are generated during the working hour of a working day as compared to the non-working hour of a working day. The ANOVA analysis of $H7_0$ resulted in a F-statistic of 33.475 and p-value of 0.000. Hence, in this case with high confidence we reject the null hypothesis.

The difference between the number of alerts generated during the working and non-working hours of vacation weeks is shown in Figure 8.3.2c. We observe negligible difference in the average number of alerts generated in the two periods. Considering the F-statistic of 0.506 and a high p-value of 0.477, we are unable to reject the null hypothesis $H8_0$.

We show the descriptive statistics for comparing the number of alerts generated in the working hour and non working hour of weekends in Figure 8.3.2d. The average number of alerts generated in the non-working hours is slightly greater than the average number of alerts generated in the working hours. ANOVA analysis resulted in a F-statistic of 8.739 and a p-value of 0.003. Hence, we reject the null hypothesis $H9_0$.

8.4 Discussion

Based on RAT it was hypothesised that change in daily routines of the victim will have an impact on the attack pattern. We base our hypotheses on the principle that a motivated attacker looking to disrupt the educational activities of an academic institution would target the network infrastructure during the working weeks period.

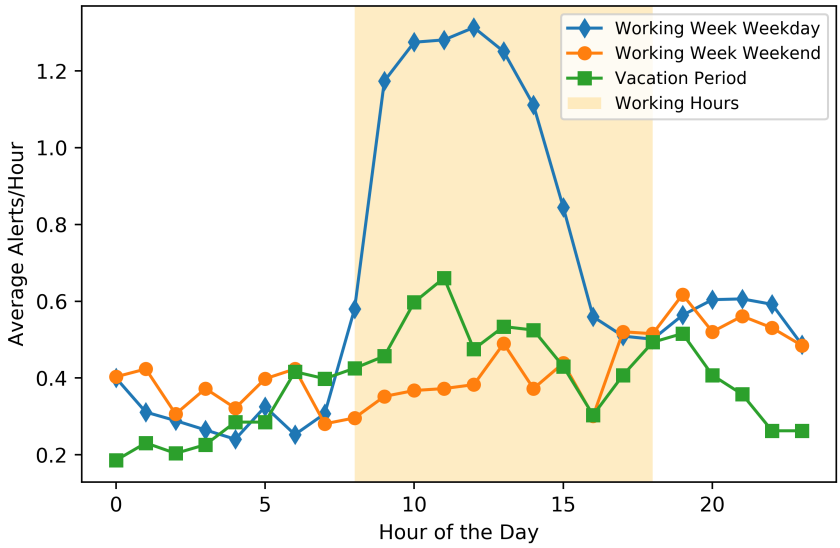


Figure 8.4.1: Number of alerts per hour in various routine periods.

With the help of hypotheses $H1$, $H2$, $H4$ and $H5$, we test if there are more attacks during the days when there are planned educational activities. As discussed in the previous section, using the dataset we were able to reject the null hypothesis in each case. We were also able to show that statistically more attacks happen on the working day of a work week. Hence, we can say that majority of the attacks on the dutch educational network target working days.

Next with the help of hypotheses $H3$ and $H6$, we compare the average number of attacks that happen during different types of holidays. In this case, we were not able to reject any of the null hypothesis. This is an indication that attack patterns are not influenced by type of vacation. This outcome also supports the central theme: *attack patterns change with daily routines.*

Finally, in hypotheses $H7$, $H8$ and $H9$, we compare the number of attacks on SURFnet's infrastructure during different hours of a day. Figure 8.4.1 shows the average number of alerts generated during each hour of a day in three different routine periods. The figure clearly shows that educational institutions in the Netherlands get targeted more often during working hour of a working day. With the help of statistical test we also find that more attacks target the network during the non-working hour of weekends. However, the difference in the average number of alerts generated in the two periods is much smaller as compared to the case of hypotheses $H1$, $H2$, $H4$ and $H5$.

8.5 Related work

We divide the papers in this section in two categories: 1) papers that discuss the motives behind DDoS attacks. 2) papers from criminology that have studied the impact of daily routines on crime patterns.

Past studies [219] showed the various incentives that can be there for a hacker to launch DDoS attacks. Nazario [142] in his study analysed the major events in case of political DDoS attacks. Segura & Lahuerta [184] tried to model the economic incentives that can behind DDoS attacks. Sauter [179] in her paper analysed the motivation of activists to use DDoS attacks as a tool to portray civil disobedience. Paulson & Weber [149] discussed the use of DDoS attacks as an effective cyber extortion weapon aginst online gaming companies. In this chapter, we show how targeted attacks can be driven by daily routines of the victim.

A few studies in criminology have studied the impact of type of holiday on type of crime. Templer, Brooner and Corgiat [196] have shown that calls for police service were more frequent on national and local holidays in Fresno. Similarly, Cohn and Rotton[45] concluded that crimes of expressive violence were significantly more prevalent on major holidays, whereas property crimes were less frequent on those days. Maimon, Kamerdze, Cukier and Sobesto [133] have shown that more attacks are likely to target academic institutions based on the data collected at a single university. In this chapter, we further generalise these findings by using data from all academic institutions in The Netherlands.

8.6 Conclusion

In this chapter, we evaluate NetFlow based attack alerts measured by SURFnet on its infrastructure. We analyse these alerts to study the impact of daily routines of academic institutions on the rate of denial of service attacks. On the basis of RAT we formulate nine hypotheses considering similar and dissimilar daily routines that we test using one-way analysis of variance (ANOVA) method. On the basis of this analysis we show how routine activity theory can be used to evaluate the influence of victim routines on attack patterns and prove that most of the attacks on academic institutions are not random. Daily routines of academic institutions heavily influence the rate of attack alerts. We also show that attack patterns do not change significantly (statistically) with type of holidays. In view of these results we can draw the following conclusions:

- We should not look at DDoS attacks in isolation, but also consider the societal aspects.

- There is a clear correlation between academic schedules and attack trends.
- This can inform decisions for selecting the type mitigation services.

Our results provides proof for the fact that most attacks on academic institutions in the Netherlands are initiated to disrupt educational activities (e.g. lectures, evaluations, etc.). If we speculate on who might benefit from these disruptions, one of the clear contenders are students.

8.7 Limitations and future work

This study also comes with some limitations. Netherlands is ranked 7^{th} on the ICT development index list [96]. This means that institutions in the Netherlands highly depend upon ICT infrastructure for day to day activities. Hence, availability of ICT services is of critical importance for academic institutions. If such a study is repeated in countries with low levels of ICT integration, we might not see similar results.

Due to unavailability of institution specific data, we could not narrow down upon the educational activities that can lead to greater number of attacks (e.g. exams or open days). Modelling the daily routine of academic institutions is more straight forward than modelling the routines for many other business models (eg. e-commerce websites). In the future it would be interesting to study if daily routines of other businesses also influence the rate of attacks targeting them.

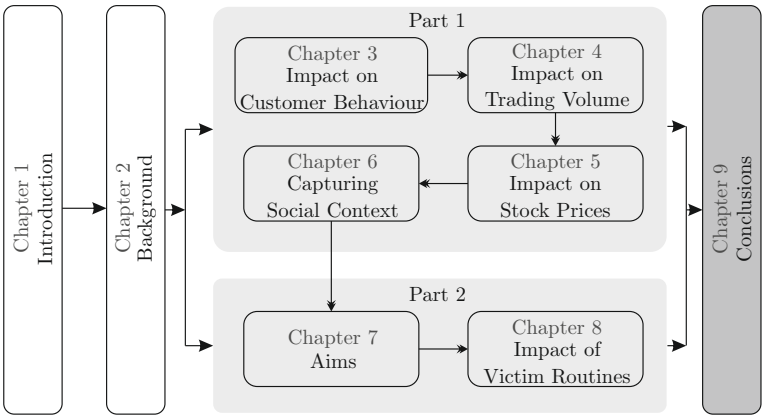
8.8 Acknowledgements

We would like to thank Dr. Simone Ferlin, Dr. Anna Sperotto, Dr. Roland van Rijswijk-Deij and the anonymous reviewers for their suggestions. This work would not have been possible without the support of Xander Jansen and Bart Bosma from SURFnet.

This page is intentionally left blank.

Chapter 9

Conclusions



In this chapter, we discuss the conclusions of this thesis. We first present the main conclusions by providing answers to the main research question posed in Chapter 1. We then revisit each of the sub-questions and provide answers based on previous chapters. Finally, we discuss our view on the directions of future research.

9.1 Main conclusions

Both public and private firms are heavily dependent on IT services for guaranteeing and improving the efficiency of operations. Malicious actors can target the network infrastructure of these organisations with DDoS attacks and cause IT unavailability. To invest wisely in DDoS mitigation, organisations need to measure the impact of these attacks. In Chapter 1 we pose the main research question analysed in this thesis as follows:

> *What are the economic impacts of DDoS attacks on public/private organisations?*

To begin our investigation, we needed to know the major parties involved in an attack. DDoS attacks have implications not only for operations of the victim firm but also for other stakeholders in an attack. In Chapter 2 we identify the stakeholders involved in a DDoS attack and discuss the sources of revenue for an attacker and damages to a victim.

Based on the framework proposed by Anderson et al. [13], we divide these damages done to public or private organisations in the following two categories:

- Direct impacts.
- Indirect impacts.

In Chapters 3 and 4 we evaluate the direct impacts of DDoS attacks. One of the direct consequences of DDoS attacks to a victim firm can be measured in terms of the change in customer behaviour. Clients of IT service providers expect the companies to protect them from cyber attacks. Especially, when the service providers advertise DDoS protection. In Chapter 3 we evaluate two such cases in which managed DNS service (MDNS) provider's services were interrupted due to DDoS attacks. We develop a framework for measuring the behaviour of domains that used the services of attacked MDNS providers. Our results show that although it leads to higher costs, using a second DNS service provider is a good strategy to guarantee availability at all times. Another direct implication of DDoS attacks is the loss of productivity due to downtime of online platforms. In Chapter 4 we measure the impacts of DDoS attacks on the volume of bitcoins traded on a large cryptocurrency exchange. Our results show that in nearly 75% of the considered cases, a negative impact of a DDoS attack on trading volume is recovered within a day. In both studies, we observed that recovery of the losses begin as soon as the services are back online. However, depending on the duration of unavailability, parts of the losses can be permanent.

It is possible that the perception of the investors about a company's future change due to public reports of DDoS attacks. In Chapter 5 we evaluate the impact of DDoS attack announcements on the market value of a victim firm. We apply a more robust and less naive event study methodology to measure this impact in 45 different cases. We observe that the negative impact of these announcements are short lived and in 42 of these events, the stock price recovers within 10 days after the first report on the attack is published. We observe negative impact (sometimes a delayed impact) on the market value of victim firms, if the services to the customers of the victim firm were interrupted. This shows that the indirect economic impact of DDoS attacks is highly context dependent. Before attributing a monetary value to the indirect consequences, we need to know more about the value of the attacked infrastructure to the company and its clients.

The indirect impact of DDoS attacks can also be measured in terms of online popularity of an attack. In order to evaluate the latter, it is important to track the publicly reported DDoS attack events systematically. The online reports on these attacks will not only help us in analysing the popularity but also collecting the circumstances (social, economic, political etc.) of an attack. In Chapter 6 we describe the methodology for collecting such a dataset with the help of Google Alerts*. We compare 3.5 years of data collected using the proposed methodology with the data available on LexisNexis and show that the proposed technique helps in collecting a more complete dataset. We also show how machine learning algorithms can be used to filter attack reporting news articles and track publicly reported attack events.

DDoS attacks have been used as a tool by hacktivists, cyber criminals, gamers and many other groups for variety of different aims [180, 175, 84]. In Chapter 7 we propose a model that considers the postulates of routine activity theory (RAT) [43] and socio-cultural, economic and political (SPEC) [72] dimensions of a DDoS attack event and they can be used to evaluate the decision making of attackers. We then apply this model to analyse 27 different attack events tracked with the help of the dataset presented in Chapter 6. We show that the model helps in understanding the decision making of an attacker. In Chapter 8 we validate one of the hypotheses based on the postulates of RAT with data on the attacks on academic institutions in the Netherlands. We show a clear correlation between academic schedules and DDoS attack trends. The studies presented in Chapters 7 and 8 again show that we should not look at DDoS attacks in isolation and consider the *context* of an attack to calculate the

*Google Alerts is a content change detection and notification service.

true economic impact. Basing the *investment in mitigation strategy* decision on actual measurements of past attacks will help organisations in making educated choices and save them from over/under investments.

9.2 Revisiting sub-questions

In Chapter 1 we divided the main research question in 5 sub-questions to help us in systematically answering the main question. The first three sub-questions are answered in Part I and the rest are answered in Part II of this thesis. To provide a more detailed view of all the results, we revisit each of these sub-questions in this section.
The first sub-question was:

> ***SQ 1:*** *Who are the major stakeholders in a DDoS attack? How are they affected by a DDoS attack?*

We answered this question in Chapter 2. In Section 2.2 we study the historical evolution of DDoS attacks. We discuss how the phenomenon of denial of service attacks have been defined in literature. We explain the various techniques used and vulnerabilities exploited by attackers to produce high intensity attacks. Then we study the defence strategies that are available to organisations to protect themselves against DDoS attacks. In Section 2.3 we also discuss the models most used by organisations to support security investment decisions. Based on the knowledge gathered by studying the techniques used by attackers to launch DDoS attacks, and strategies used by organisations to defend themselves, we identify four main stakeholders in a DDoS attack. These are:

- The attacker.
- The victim.
- Customers of the victim.
- DDoS protection companies.

In Figure 2.3 we also show how these stakeholders interact. We answer the second part of our first sub-question in Section 2.5. We identify the revenue streams of an attacker by studying the business model of a botnet. On the basis of the technical capabilities of the practitioners involved in development of a botnet, they can be divided in 4 tiers [79]. We use this tier distribution to develop a *botnet ecosystem* (Section 2.5.1.1), which gives a snapshot of how the botnet economy functions at a macro level. We explain the *botnet assembly chain*

[28] and *botnet life cycle* [167] to understand the partners, skills and investments required to develop a botnet business from scratch. Based on four different case studies, we also summarise the profits that a botnet owner can make. A botnet owner can use the botnet for a number of illicit activities to make money e.g., click fraud, sale of spam services and sale of booter services etc. According to one of these case studies, a booter owner can make as much as $ 26,000 monthly, for a median of 24 months [32]. Finally, we use the *business model canvas* framework to depict the business of developing, using and maintaining a botnet (Figure 2.6).

To study the damages caused by DDoS attacks, we follow the framework provided by Anderson et al. [12]. Based on the framework, we discuss the implications of a DDoS attacks for a victim organisation in Section 2.7. To analyse the direct consequences of DDoS attacks on a victim firm, we state the second sub-question as:

SQ 2: *How can we measure the direct consequences of a DDoS attack?*

We answer this sub-question in Chapters 3 and 4. We show two examples of how we can empirically measure the direct consequences of DDoS attacks on victim organisations. In Chapter 3 we analyse the impact of DDoS attacks on customer behaviour of two large managed domain name service (MDNS) providers. We study the added value provided by an MDNS provider over an ordinary authoritative name server (ANS) and show with the help of a value flow diagram (Figure 3.1)), how a MDNS promises greater availability to its customers. We use the OpenINTEL dataset [166] to measure the impact of DDoS attacks on the customers of the two MDNS providers i.e., *NS1* and *Dyn*. OpenINTEL project collects unique long-term datasets with daily DNS measurements for all domains that belong to the major top level domains (TLDs). It covers nearly 60% of the global DNS name space every 24 hours. In Section 3.3.3 we develop a framework that can be used to measure the customer behaviour of DNS service providers with the help of OpenINTEL dataset. We then use this framework to measure the behaviour of *domains* that used the services of Dyn and NS1 before and after the attack. We show that a significant number of MDNS customers that were using the attacked MDNS's services exclusively started using the services of a secondary DNS service provider (Section 3.4). We observe that most of the newly *non-exclusive* customers after the attack on Dyn and NS1 use another MDNS provider as a secondary DNS. However, we do not record any change in the behaviour of *domains* that were already *non-exclusive*. This suggests that in terms of risk management using multiple providers is a good

strategy. In Appendix 3.A we zoom into the analysis of the attack on Dyn. We first analyse the return behaviour of domains that stopped using the services of Dyn just after the attack and compared it to the return behaviour of domains that stopped before the attack. We observe that the domains that stop using the services of Dyn were slightly less likely to return as compared to domains that stopped using the services before the attack. In Section 3.A.2 we use an ARIMA model to predict the number of Domains that might have stopped using the services of Dyn permanently due to the attack. We estimate that Dyn lost nearly 2,000 domains due to the attack.

In Chaper 4, we measure another direct consequence of DDoS attacks i.e., impact on the volume of Bitcoins traded on a large crypto-currency exchange. We modify the traditional *event analysis* methodology to measure the impact of 17 different DDoS attack events on the daily volume of Bitcoins traded on Bitfinex. In Section 4.3 we discuss the modified methodology and choose the most apt parameters for the estimation model. We also observe that positive and negative price changes are perceived differently by investors and hence, in our model the impact of positive and negative price changes is captured by separate variables (Equation 4.6). We find that only on 4 of 17 instances there was a statistically significant impact of DDoS attacks on volume of bitcoin traded on the exchange that lasted more than a single day. In Section 4.5 we discuss the hourly trading data for these 4 attacks. On two occasions (20^{th} June 2016 and 5^{th} June 2018) we observe that the impact lasts for more than 5 days. This was because of multiple days in the *event period* where the traded volume was lower than expected which can be due to successive attacks or maintenance. Bitfinex uses a *Twitter* account and a *service status page* to regularly update its customers about platform downtime, this reduces the information asymmetry on the side of customers and helps them to resume trading as soon as the platform is back online. This may also be a key reason why Bitfinex recovers the lost volume within a short period of time.

These two chapters shed some light on how empirical measurements can be used to calculate the direct impact of DDoS attacks on a victim. For measuring the true impact, it is key to understand the value of the infrastructure under attack to the victim firm. It is also essential to take into account the resilience of the victim organisation/attacked infrastructure towards short term unavailability. Next, in order to analyse the indirect impact of DDoS attacks on a victim organisation, we posed the third sub-question as:

SQ 3: *How can we measure the indirect consequences of a DDoS attack?*

We provide answer to ***SQ3*** in Chapters 5 and 6. We discussed in Section 2.5.2 that one of the indirect losses to a firm victimised by a DDoS attack could be due to change in perception of the investors of firm's market value. In Chapter 5, we test if public announcements of DDoS attack events lead to a negative change in stock prices of a victim firm. We use event analysis methodology (Section 5.3) to evaluate the impact of DDoS attacks in 45 separate attacks over a period of 5 years. In order to make the methodology more robust, we make two main changes to the traditional method. Firstly, we use a multiplicative model in place of an additive model to estimate the value of stock price. Secondly, we also avoid the widespread assumption of short-term returns being approximately normally distributed. Instead, we use the technique of bootstrapping to generate an empirical distribution, that we use for hypothesis testing. By comparing the results with the traditional method of event analysis, we show that the proposed method leads to less number of incorrect conclusions (i.e., it more accurately calculates the statistical significance). We find that in most cases DDoS attack announcements do not lead to negative impact on the market value. Only in some cases where attacks led to disruption of services provided to victim's clients we found a short-lived (recovered within five days) impact on the stock prices. These findings are consistent with the ones presented by Hovav & D'Arcy [89].

Economic impact of DDoS attacks is also related to the circumstances in which an organisation is attacked. One way of gathering information on these circumstances or *context* is by interviewing victims. However, as journalists working in the technology sector also perform such interviews, a dataset based on online media sources reporting DDoS attacks may also provide us with contextual information. In Chapter 6, we present one such dataset that can be used to learn the context of a DDoS attack. We use the content change detection and notification service known as *Google Alerts* to collect this dataset. In Section 6.3.1 we explain the method used to collect data. We then explore the characteristics of the data collected and show that using the proposed method within 3.5 years we are able to collect data from nearly 14,000 domains in 47 different languages. We show probable uses of the dataset using two simple case studies. In the first case study we compare our dataset with the data reported on LexisNexis. We find that all the news article reported on LexisNexis are also available in our dataset. Hence, our dataset can be used as a trusted longitudinal dataset for articles on DDoS attacks. In the second case study we show how the dataset can be used to track DDoS attack events. In order to filter attack reporting news article, we used a supervised machine learning classifier (Section 6.5) and show that it performs with an accuracy of 80.9%. By track-

ing 4 different attack events with the help of the dataset, we demonstrate the effectiveness of the filtered data in tracking DDoS attack events.

In Section 1.4 we discussed the importance of analysing attacker aims in understanding the damages an attacker is hoping to inflict on the victim. The next two sub-questions deal with analysing the aims of attackers for the use of DDoS attacks. The 4^{th} sub-question is phrased as:

> ***SQ 4:*** *What are the various aims of attackers to use DDoS attacks? How can classical theories in criminology be used to explain the aims of attackers?*

We answer the first part of this question in Section 1.4. We discuss the various aims for which attackers use DDoS attacks as mentioned in literature and classify them as monetary and non-monetary as shown in Figure 1.3. We have seen that apart for the obvious economic gains, DDoS attacks are often used for demonstration/hacktivism purposes. In Section 2.6 we have seen how routine activity theory (RAT) can be instrumental in analysing the aims of traditional criminals. In Chapter 7, we show how postulates of RAT can also be used to analyse attacker aims behind DDoS attacks. We break the decision of an attacker to choose a target for a DDoS attack down into two components: 1) Choice of victim organisation and 2) Choice of network infrastructure to target. We model the choice of victim organisation based on the socio-cultural, economic and political (SPEC) dimensions i.e., the attacker chooses a victim considering its SPEC variables. We model the choice of infrastructure based on value, inertia, visibility and accessibility (VIVA) characteristics of RAT i.e., an attacher chooses that target infrastructure which ranks high on one or more of these characteristics. The proposed model also suggests that the choice of infrastructure depends on the daily routines of the victim. We use this model to analyse the choices of attackers in case of 27 distinct attack events. We conclude that companies should analyse their socio-cultural, economic and political environment continuously to be prepared for forthcoming DDoS attacks. We empirically test the impact of victim routines on DDoS attack trends in the next sub-question. Our final sub-question is:

> ***SQ 5:*** *How can we use the postulates of classical theories in criminology to explain DDoS attack trends?*

We answer the final sub-question in Chapter 8. RAT suggests that changes in crime rates should be associated with days that affect the daily routines. Based on this postulate we hypothesise that as greater disruption can be caused to academic activities during working days, we would observe greater number of DDoS attacks on the network infrastructure of schools, colleges and universities

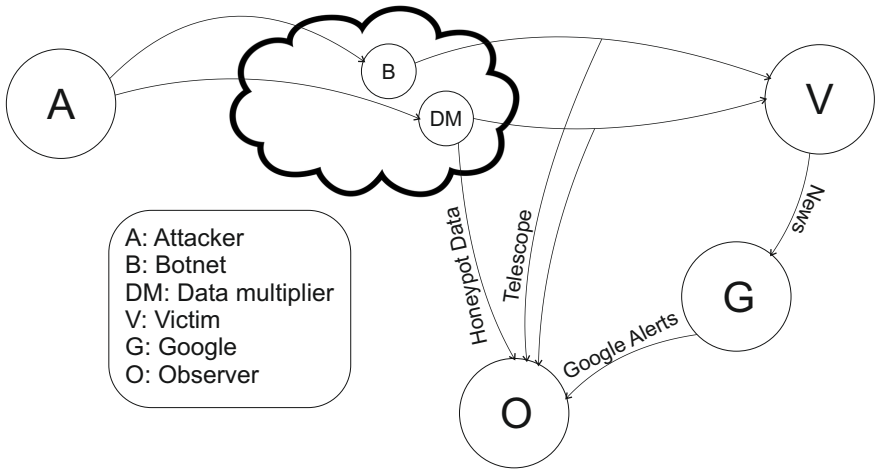


Figure 9.3.1: Datasets to measure DDoS attacks and their impacts empirically.

during working days. To test our hypothesis we use 2.5 years of data consisting of alerts based on 1/100 sampled netflow using two different software: 1) NfSen 2) Arbor Peakflow. The data were collected by SURFnet and hence, represents all attempted DDoS attacks targeting academic institutions in the Netherlands in 2.5 years period. We test several hypotheses to establish the relationship between the daily routines of academic institutions in the Netherlands and change in number of attacks. We show that most of the attacks on academic institutions are not random and daily routines of the victim heavily influence the rate of attack alerts. In Figure 8.4.1 we can clearly see that the number of attempted attacks were considerably higher in the working hours. We also observe that unlike the results in traditional criminology [45], attack patterns do not change with the type of holidays. If such patterns are visible in attack trends, then companies can save money on DDoS mitigation and keep most of their daily traffic private by opting for DDoS mitigation service providers that need you to reroute your traffic only during attacks.

9.3 Directions for future research

In the final section of this thesis we discuss the directions for future research. We divide them in two main categories. In the first category we discuss using empirical measurements to facilitate decision making for the users of DDoS mit-

igation services. The second set of studies are aimed at analysing the resilience of organisations towards IT downtime.

- *Using multiple datasets to facilitate informed decision making among the users of DDoS mitigation services:* in this thesis we use empirical measurements to analyse the impact of DDoS attacks on organisations. In the past few years, several measurement studies have been carried out to measure DDoS attacks and their impact. At the same time more datasets that can help us in studying the consequences of IT downtime empirically have been created. Figure 9.3.1 shows some of the datasets based on these measurements. Jonker *et al.* [106] has leveraged data from four independent data sources to characterise DDoS attacks. There is a need to further analyse the industry wise distribution of these attacks which can provide a guideline for companies looking to invest in DDoS mitigation. Frameworks such as Dmap [216] can be used to categorise domain names according to industry classes. Combining this information with attack data can help us in understanding the threat of these attacks on each industry sector. Results in Chapter 7 and 8 motivate us to believe that these attacks are not random and are related other societal factors. Empirically analysing the industry wise attack trends will reduce the information asymmetry between the users and providers of DDoS protection services and will provide a better view of the threat posed by DDoS attacks on the users of these services.

- *Studies to analyse the resilience of organisations towards IT downtime:* In the early 20^{th} century motion and time study was used to measure the efficiency of processes in the manufacturing industry [19]. This helped researchers and practitioners to standardise assembly lines and design effective plans to deal with mechanical downtime on the shop floor. In a 21^{st} century IT dependent workplace, organising a traditional motion and time study is difficult. However, by using data on how individuals use IT infrastructure and applications (e.g., system logs etc.), we can measure the usage patterns. These data can be used to measure the impact of unscheduled events such as DDoS attacks on an organisation and can help us to understand how resilient organisations are to such events.

Bibliography

[1] *18 U.S. Code § 1030 - Fraud and Related Activity in Connection with Computers | US Law | LII / Legal Information Institute*. URL: https://www.law.cornell.edu/uscode/text/18/1030 (visited on 13/06/2017).

[2] *2017 Cost of Cybercrime Study*. Accessed on 05/14/2019. 2017. URL: https://www.accenture.com/nl-en/insight-cost-of-cybercrime-2017?src=SOMS.

[3] *2019 Cost of Cybercrime Study*. Accessed on 05/14/2019. 2019. URL: https://www.accenture.com/us-en/insights/security/cost-cybercrime-study?src=SOMS.

[4] A. Abhishta, R. Joosten, S. Dragomiretskiy and L. Nieuwenhuis. 'Impact of Successful DDoS Attacks on a Major Crypto-currency Exchange'. *2019 27th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*. United States: IEEE, 2019, pp. 379–384.

[5] A. Abhishta, R. Joosten, M. Jonker, W. Kamerman and L. Nieuwenhuis. 'Poster: Collecting Contextual Information About a DDoS Attack Event Using Google Alerts'. 2019. Poster presented at 40th IEEE Symposium on Security and Privacy, San Francisco, CA.

[6] A. Abhishta, M. Junger, R. Joosten and L. Nieuwenhuis. 'Victim Routine Influences the Number of DDoS Attacks: Evidence from Dutch Educational Network'. *2019 IEEE Security and Privacy Workshops (SPW)*. 2019, pp. 242–247.

[7] A. Abhishta, R. van Rijswijk-Deij and L. Nieuwenhuis. 'Measuring the Impact of a Successful DDoS Attack on the Customer Behaviour of Managed DNS Service Providers'. *Computer Communication Review* 48.5, 2018, pp. 70–76.

[8] A. Abhishta, M. Junger, R. Joosten and L. J. Nieuwenhuis. 'A Note on Analysing the Attacker Aims Behind DDoS Attacks'. *International Symposium on Intelligent and Distributed Computing*. Springer. 2019, pp. 255–265.

[9] Abhishta, R. Joosten and L. J. M. Nieuwenhuis. 'Analysing the Impact of a DDoS Attack Announcement on Victim Stock Prices'. *Proc. of 25th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP'17), St. Petersburg,Russia*. United States: IEEE, 2017, pp. 354–362.

[10] Abhishta, R. Joosten and L. J. Nieuwenhuis. 'Comparing Alternatives to Measure the Impact of DDoS Attack Announcements on Target Stock Prices'. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)* 8.4, 2017, pp. 1–18.

[11] Abhishta, R. van Rijswijk-Deij and L. Nieuwenhuis. 'Measuring the Impact of a Successful DDoS Attack on the Customer Behaviour of Managed DNS Service Providers'. *WTMC '18*. ACM Press, 2018, pp. 1–7.

[12] R. Anderson, C. Barton, R. Böhme, R. Clayton, M. J. G. van Eeten, M. Levi, T. Moore and S. Savage. 'Measuring the Cost of Cybercrime'. *The Economics of Information Security and Privacy*. Ed. by R. Böhme. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 265–300.

[13] R. Anderson, C. Barton, R. Böhme, R. Clayton, M. J. Van Eeten, M. Levi, T. Moore and S. Savage. 'Measuring the Cost of Cybercrime'. *The Economics of Information Security and Privacy*. Springer, 2013, pp. 265–300.

[14] *Anonymous Attacks Italian Government Portals Because of Gas Pipeline Project*. 2016. URL: `http://news.softpedia.com/news/anonymous-attacks-italian-government-site-because-of-gas-pipeline-project-500977.shtml`.

[15] *Anonymous takes down Nissan website in protest of Japanese whale killings*. 2016. URL: `http://www.businessinsider.com/anonymous-attacks-nissan-website-to-protest-japanese-whale-killings-2016-1?international=true&r=US&IR=T`.

[16] K. Arora, K. Kumar, M. Sachdeva et al. 'Impact Analysis of Recent DDoS Attacks'. *International Journal on Computer Science and Engineering* 3.2, 2011, pp. 877–883.

[17] *Australian 2016 census sabotage puts a question mark on private cloud.* 2016. URL: http://www.computerweekly.com/news/450302728/Australian-2016-census-sabotage-puts-a-question-mark-on-private-cloud.

[18] C. Baldwin. *Bitcoin Worth $72 Million Stolen from Bitfinex Exchange in Hong Kong.* URL: https://www.reuters.com/article/us-bitfinex-hacked-hongkong/bitcoin-worth-72-million-stolen-from-bitfinex-exchange-in-hong-kong-idUSKCN10E0KP (visited on 20/08/2018).

[19] R. M. Barnes. 'Motion and Time Study', 1949.

[20] J. Barney, M. Wright and D. J. Ketchen Jr. 'The resource-based view of the firm: Ten years after 1991'. *Journal of management* 27.6, 2001, pp. 625–641.

[21] *Battlefield 1 Beta: You Have Lost Connection to EA Servers.* 2016. URL: http://www.pcgameshardware.de/Battlefield-1-2016-Spiel-54981/News/Beta-Server-down-Verbindungsabbrueche-DDOS-1206368/.

[22] K. Beevers. *A note from NS1's CEO: How we responded to last weeks's major, multi-faceted DDoS Attacks.* Blog. 2016. URL: http://ns1.com/blog/how-we-responded-to-last-weeks-major-multi-faceted-ddos-attacks.

[23] *Bitcoin Exchange BTC-e Resumes Services after Latest DDoS Attack.* 2016. URL: https://www.cryptocoinsnews.com/bitcoin-exchange-btc-e-resumes-services-latest-ddos-attack/.

[24] *Bitfinex Status Page.* URL: https://bitfinex.statuspage.io (visited on 20/08/2018).

[25] *Blizzard hit with another DDoS attack, Overwatch, WoW, Hearthstone and more down.* 2016. URL: https://www.technobuffalo.com/2016/08/23/blizzard-ddos-battlenet-down-august-23-sombra-theory/.

[26] L. D. Bodin, L. A. Gordon and M. P. Loeb. 'Evaluating Information Security Investments Using the Analytic Hierarchy Process'. *Communications of the ACM* 48.2, 2005, pp. 78–83.

[27] R. Bolstridge. *Dyn DDoS Attack: Wide-Spread Impact Across the Financial Services Industry (Part 1).* Blog. 2016. URL: https://blogs.akamai.com/2016/10/dyn-ddos-attack-wide-spread-impact-across-the-financial-services-industry-part-1.html.

[28] G. Bottazzi and G. Me. 'The Botnet Revenue Model'. *Proceedings of the 7th International Conference on Security of Information and Networks.* ACM. 2014, p. 459.

[29] L. Breiman. *Classification and Regression Trees.* Wadsworth statistics/-probability series. Wadsworth International Group, 1984.

[30] L. Breiman. 'Random Forests'. *Machine Learning* 45.1, 2001, pp. 5–32.

[31] J. Brownlee Ph.D. *Classification Accuracy is Not Enough: More Performance Measures You Can Use.*

[32] R. Brunt, P. Pandey and D. McCoy. 'Booted: An Analysis of a Payment Intervention on a DDoS-for-Hire Service'. *Workshop on the Economics of Information Security.* 2017.

[33] P. H. Bucy, E. P. Formby, M. S. Raspanti and K. E. Rooney. 'Why Do They Do It: The Motives, Mores, and Character of White Collar Criminals'. *St. John's Law Review* 82, 2008, pp. 401–571.

[34] A. Bushatz. *Report: Hack of Adultery Site Ashley Madison Exposed Military Emails.* (Accessed on 05/13/2019). URL: https://www.military.com/daily-news/2015/08/19/report-hack-adultery-site-ashleymadison-exposed-military-emails.html.

[35] S. A. Butler. 'Security Attribute Evaluation Method: A Cost-Benefit Approach'. *Proceedings of the 24th international conference on Software engineering.* ACM. 2002, pp. 232–240.

[36] J. Caballero, C. Grier, C. Kreibich and V. Paxson. 'Measuring Pay-Per-Install: The Commoditization of Malware Distribution.' *Usenix security symposium.* 2011, pp. 13–13.

[37] *Calculate the Cost of DDoS Attacks.* URL: https://www.akamai.com/uk/en/products/security/calculate-the-cost-of-ddos-attacks.jsp.

[38] K. Campbell, L. A. Gordon, M. P. Loeb and L. Zhou. 'The Economic Cost of Publicly Announced Information Security Breaches : Empirical Evidence from the Stock Market'. *Journal of Computer Security* 11, 2003, pp. 431–448.

[39] B. Cashell, W. D. Jackson, M. Jicklin and B. Webel. *The Economic Impact of Cyber-Attacks*. 2004. URL: `http://www.au.af.mil/au/awc/awcgate/crs/rl32331.pdf`.

[40] H. Cavusoglu, B. Mishra and S. Raghunathan. 'The Effect of Internet Security Breach Announcements on Market Value: Capital Market Reactions for Breached Firms and Internet Security Developers'. *Int. J. Electron. Commerce* 9.1, 2004, pp. 70–104.

[41] G. C. Cawley and N. L. Talbot. 'On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation'. *J. Mach. Learn. Res.* 11, 2010, pp. 2079–2107. URL: `http://dl.acm.org/citation.cfm?id=1756006.1859921`.

[42] M. R. Chaiken and B. D. Johnson. *Characteristics of different types of drug-involved offenders*. US Department of Justice, 1988.

[43] L. E. Cohen and M. Felson. 'Social Change and Crime Rate Trends: A Routine Activity Approach'. *American Sociological Review* 44.4, 1979, pp. 588–608.

[44] L. E. Cohen and M. Felson. 'Social Change and Crime Rate Trends: A Routine Activity Approach (1979)'. *Classics in Environmental Criminology*. CRC Press, 2016, pp. 203–232.

[45] E. G. Cohn and J. Rotton. 'Even criminals take a holiday: Instrumental and expressive crimes on major and minor holidays'. *Journal of Criminal Justice* 31.4, 2003, pp. 351–360.

[46] CoinMarketCap. *Cryptocurrency market capitalizations*. URL: `https://coinmarketcap.com/currencies/bitcoin/`.

[47] C. Cortes and V. Vapnik. 'Support-vector networks'. *Machine Learning* 20.3, 1995, pp. 273–297.

[48] M. Cremonini and P. Martini. 'Evaluating Information Security Investments from Attackers Perspective: The Return-on-Attack (ROA)'. *Workshop on Economics of Information Security*. 2005.

[49] P. Cromwell and J. N. Olson. 'The reasoning burglar: Motives and decision-making strategies'. *their own words: Criminals on crime (an anthology)*, 2005, pp. 42–56.

[50] J. J. Cronin, M. K. Brady and G. T. M. Hult. 'Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments'. *Journal of retailing* 76.2, 2000, pp. 193–218.

[51] *Cyber Security on the Offense : A Study of IT Security Experts*. 2012. URL: https://security.radware.com/uploadedFiles/Resources%5C_and%5C_Content/Attack%5C_Tools/CyberSecurityontheOffense.pdf.

[52] *DDoS attacks caused StarHub broadband outages*. 2016. URL: http://www.telecomasia.net/content/ddos-attacks-caused-starhub-broadband-outages.

[53] L. Demetz and D. Bachlechner. 'To Invest or Not to Invest? Assessing the Economic Viability of a Policy and Security Configuration Management Tool'. *The Economics of Information Security and Privacy*. Springer, 2013, pp. 25–47.

[54] C. Dietzel, A. Feldmann and T. King. 'Blackholing at ixps: On the effectiveness of ddos mitigation in the wild'. *International Conference on Passive and Active Network Measurement*. Springer. 2016, pp. 319–332.

[55] *Domain Name Statistics*. 2019. URL: https://domainnamestat.com/statistics/tldtype/all.

[56] *Donald Trump sweeping American Polls, Canadian migration website down*. 2016. URL: http://www.techworm.net/2016/11/donald-trump-sweeping-american-polls-canadian-migration-website.html.

[57] T. Dubendorfer, A. Wagner and B. Plattner. 'An Economic Damage Model for Large-Scale Internet Attacks'. *13th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*. IEEE. 2004, pp. 223–228.

[58] V. Ş. Ediger and S. Akar. 'ARIMA forecasting of primary energy demand by fuel in Turkey'. *Energy Policy* 35.3, 2007, pp. 1701–1708.

[59] M. van Eeten. *The Value of Cyber Risk Quantification*. (Accessed on 05/13/2019). 2016. URL: https://securitytalent.nl/events/the-value-of-cyber-risk-quantification-event-2016-10-13.

[60] B. Efron. 'Bootstrap methods: another look at the jackknife'. *Breakthroughs in Statistics*, 1992, pp. 569–593.

[61] M. Ehrenhard, B. Kijl and L. Nieuwenhuis. 'Market adoption barriers of multi-stakeholder technology: Smart homes for the aging population'. *Technological Forecasting and Social Change* 89.Supplement C, 2014, pp. 306–315.

[62] *Eir's webmail affected by DDoS attack.* 2016. URL: https://www.rte.ie/news/business/2016/1125/834480-eirs-webmail-affected-by-ddos-attack/.

[63] *Ethereum's network is currently suffering from a computational DDoS attack.* 2016. URL: http://www.ibtimes.co.uk/ethereum-network-hit-by-computational-ddos-attack-1582935.

[64] M. Ettredge and V. J. Richardson. 'Assessing the risk in e-commerce'. *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, 2002.

[65] *European Commission Hit By DDoS Attack.* 2016. URL: https://www.infosecurity-magazine.com/news/european-commission-hit-by-ddos/.

[66] *Failed Mirai botnet attack downed 900000 Germans' internet access.* 2016. URL: https://www.siliconrepublic.com/enterprise/mirai-botnet-deutsche-telekom.

[67] E. Fama and K. French. 'Common risk factors in the returns of stocks and bonds.' *Journal of Financial Economics* 33.1, 1993, pp. 3–56.

[68] A. Feder, N. Gandal, J. Hamrick and T. Moore. 'The Impact of DDoS and Other Security Shocks on Bitcoin Currency Exchanges: Evidence from Mt. Gox'. *Journal of Cybersecurity* 3.2, 2018, pp. 137–144.

[69] D. Florencio and C. Herley. *Sex, Lies and Cyber-crime Surveys.* 2011.

[70] D. Florêncio and C. Herley. 'Sex, Lies and Cyber-Crime Surveys'. *The Economics of Information Security and Privacy III.* Springer, 2013, pp. 35–53.

[71] B. A. Forouzan. *Cryptography & Network Security.* McGraw-Hill, Inc., 2007.

[72] R. A. Gandhi, A. C. Sharma, W. Mahoney, W. Sousan and Q. Zhu. 'Dimensions of Cyber-Attacks: Cultural, Social, Economic, and Political'. *IEEE Technology and Society Magazine* 30.1, 2011, pp. 28–38.

[73] A. Garg, J. Curtis and H. Halper. 'Quantifying the financial impact of IT security breaches'. *Information Management & Computer Security* 11.2, 2003, pp. 74–83.

[74] K. Geers, D. Kindlund, N. Moran and R. Rachwald. *World War C: Understanding Nation-State Motives Behind Today's Advanced Cyber Attacks.* 2013.

[75] P. Geurts, D. Ernst and L. Wehenkel. 'Extremely randomized trees'. *Machine Learning* 63.1, 2006, pp. 3–42.

[76] L. A. Gordon, M. P. Loeb and L. Zhou. 'The impact of information security breaches : Has there been a downward shift in costs?' *Journal of Computer Security* 19, 2011, pp. 33–56.

[77] L. A. Gordon and M. P. Loeb. 'The Economics of Information Security Investment'. *ACM Trans. Inf. Syst. Secur.* 5.4, 2002, pp. 438–457.

[78] L. A. Gordon, M. P. Loeb and W. Lucyshyn. 'Information Security Expenditures and Real Options: A Wait-and-See Approach'. *Computer Security Journal* 19.2, 2003, pp. 1–7.

[79] J. Gosler and L. Von Thaer. 'Resilient Military Systems and the Advanced Cyber Threat'. January, 2013, pp. 1–146. URL: `http://www.acq.osd.mil/dsb/reports/ResilientMilitarySystems.CyberThreat.pdf`.

[80] *Govt websites forced offline in DDoS attack*. 2016. URL: `http://www.rte.ie/news/2016/0122/762161-cyber-attack`.

[81] A. D. de Groot. *Methodologie*. Vol. 6. Hague: Mouton, 1969.

[82] P. Haag. 'Watch your Flows with NfSen and NFDUMP'. *50th RIPE Meeting*. 2005.

[83] *Hacker group Anonymous shuts down KKK website*. 2016. URL: `http://www.telegraph.co.uk/technology/2016/04/25/hacker-group-anonymous-shuts-down-kkk-website/`.

[84] *Hackerangriff auf PlayStation und Xbox*. 2014. URL: `http://www.wiwo.de/technologie/digitale-welt/sony-und-microsoft-betroffen-hackerangriff-auf-playstation-und-xbox/11161500.html`.

[85] R. I. Harris. 'Testing for unit roots using the augmented Dickey-Fuller test: Some issues relating to the size, power and the lag structure of the test'. *Economics letters* 38.4, 1992, pp. 381–386.

[86] S. Hilton. *Dyn Analysis Summary of Friday October 21 Attack*. Blog. 2016. URL: `https://dyn.com/blog/dyn-analysis-summary-of-friday-october-21-attack/`.

[87] S. Hilton. *Dyn Analysis Summary Of Friday October 21 Attack | Dyn Blog*. 2016. URL: `http://dyn.com/blog/dyn-analysis-summary-of-friday-october-21-attack/` (visited on 13/06/2017).

[88] N. Hoque, D. K. Bhattacharyya and J. K. Kalita. 'Botnet in DDoS Attacks: Trends and Challenges'. *IEEE Communications Surveys Tutorials* 17.4, 2015, pp. 2242–2270.

[89] A. Hovav and J. D'Arcy. 'Impact of Denial-of-Service attack announcements on the market value of firms'. *Risk Management And Insurance Review* 6.2, 2003, pp. 97–121.

[90] A. Hovav, J. Han and J. Kim. 'Market Reaction to Security Breach Announcements: Evidence from South Korea'. *SIGMIS Database* 48.1, 2017, pp. 11–52.

[91] *HSBC online banking is 'attacked'*. 2016. URL: `http://www.bbc.com/news/business-35438159`.

[92] C. D. Huang, Q. Hu and R. S. Behara. 'An Economic Analysis of the Optimal Information Security Investment in the Case of a Risk-Averse Firm'. *International Journal of Production Economics* 114.2, 2008, pp. 793–804.

[93] M. Al-Humaigani and D. B. Dunn. 'A Model of Return on Investment for Information Systems Security'. *2003 46th Midwest Symposium on Circuits and Systems*. Vol. 1. IEEE. 2003, pp. 483–485.

[94] A. Hutchings and R. Clayton. 'Exploring the Provision of Online Booter Services'. *Deviant Behavior* 37.10, 2016, pp. 1163–1178.

[95] T. Ibragimov, O. Kupreev, E. Badovskaya and A. Gutnikov. *DDoS Attacks in Q2 2018*. (Accessed on 05/04/2019). 2018. URL: `https://securelist.com/ddos-report-in-q2-2018/86537/`.

[96] *ICT Development Index 2017*. URL: `http://www.itu.int/net4/itu-d/idi/2017/index.html`.

[97] IHS. *IoT: Number of Connected Devices Worldwide 2012-2025*. (Accessed on 05/15/2019). 2019. URL: `https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/`.

[98] *ING ondanks maatregelen getroffen door nieuwe DDos-aanval*. 2013. URL: `http://www.nrc.nl/nieuws/2013/04/10/ing-nieuwe-cyberaanval-sneller-afgeslagen-door-maatregelen`.

[99] P. Institute. 'Trends in the Cost of Web Application & Denial of Service Attacks', 2017.

[100] *Internet providers claim cyber attack, to meet senior cop*. 2016. URL: `http://www.nyoooz.com/mumbai/635360/internet-providers-claim-cyber-attack-to-meet-senior-cop`.

[101] *Irish lottery site and ticket machines hit by DDoS attack*. 2016. URL: http://www.bbc.com/news/technology-35373890.

[102] B. Johnson, A. Laszka, J. Grossklags, M. Vasek and T. Moore. 'Game-Theoretic Analysis of DDoS Attacks Against Bitcoin Mining Pools'. *Financial Cryptography and Data Security*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 72–86.

[103] S. D. Johnson. 'A brief history of the analysis of crime concentration.' *European Journal of Applied Mathematics* 21(4-5), 2010, pp. 349–370.

[104] E. Jones, T. Oliphant and P. Peterson. {*SciPy*}*: open source scientific tools for* {*Python*}. 2014. URL: https://www.%20scipy.%20org.

[105] M. Jonker, A. King, J. Krupp, C. Rossow, A. Sperotto and A. Dainotti. 'Millions of Targets Under Attack: a Macroscopic Characterization of the DoS Ecosystem'. *Internet Measurement Conference (IMC)*. ACM, 2017.

[106] M. Jonker, A. King, J. Krupp, C. Rossow, A. Sperotto and A. Dainotti. 'Millions of Targets Under Attack: A Macroscopic Characterization of the DoS Ecosystem'. *Proceedings of the 2017 Internet Measurement Conference*. ACM. 2017, pp. 100–113.

[107] M. Jonker, A. Sperotto, R. van Rijswijk-Deij, R. Sadre and A. Pras. 'Measuring the Adoption of DDoS Protection Services'. *Proceedings of ACM SIGCOMM Internet Measurement Conference 2016*. Santa Monica, CA, USA: ACM Press, 2016.

[108] N. Kallus. 'Predicting Crowd Behavior with Big Public Data'. *Proceedings of the 23rd International Conference on World Wide Web*. WWW '14 Companion. Seoul, Korea: ACM, 2014, pp. 625–630.

[109] C. Kanich, N. Weaver, D. McCoy, T. Halvorson, C. Kreibich, K. Levchenko, V. Paxson, G. M. Voelker and S. Savage. 'Show Me the Money: Characterizing Spam-Advertised Revenue.' *USENIX Security Symposium*. 2011, pp. 15–15.

[110] K. Kannan, J. Rees and S. Sridhar. 'Market Reactions to Information Security Breach Announcements: An Empirical Analysis'. *International Journal of Electronic Commerce* 12.1, 2007, pp. 69–91.

[111] I. Karafiath. 'Detecting cumulative abnormal volume: a comparison of event study methods'. *Applied Economics Letters* 16.8, 2009, pp. 797–802.

[112] S. Karnouskos. 'Stuxnet Worm Impact on Industrial Cyber-Physical System Security'. *IECON 2011-37th Annual Conference of the IEEE Industrial Electronics Society*. IEEE. 2011, pp. 4490–4494.

[113] J. M. Karpoff. 'The Relation between Price Changes and Trading Volume: A Survey'. *Journal of Financial and Quantitative Analysis* 22.1, 1987, pp. 109–126.

[114] C. Kemp, C. Calvert and T. Khoshgoftaar. 'Utilizing Netflow Data to Detect Slow Read Attacks'. *2018 IEEE International Conference on Information Reuse and Integration (IRI)*. IEEE. 2018, pp. 108–116.

[115] Z. King, D. Henshel, L. Flora, M. Cains, B. Hoffman and C. Sample. 'Characterizing and Measuring Maliciousness for Cybersecurity Risk Assessment'. *Frontiers in Psychology*, 2018, p. 39.

[116] D. P. Kingma and J. Ba. 'Adam: A Method for Stochastic Optimization'. *CoRR* abs/1412.6980, 2014. arXiv: 1412.6980. URL: http://arxiv.org/abs/1412.6980.

[117] R. Kohavi. 'A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection'. *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*. IJCAI'95. Montreal, Quebec, Canada: Morgan Kaufmann Publishers Inc., 1995, pp. 1137–1143. URL: http://dl.acm.org/citation.cfm?id=1643031.1643047.

[118] S. Kotsiantis. 'Supervised Machine Learning: A Review of Classification Techniques'. 31, 2007, pp. 3–24.

[119] L. Krämer, J. Krupp, D. Makita, T. Nishizoe, T. Koide, K. Yoshioka and C. Rossow. 'Amppot: Monitoring and defending against amplification ddos attacks'. *International Workshop on Recent Advances in Intrusion Detection*. Springer. 2015, pp. 615–636.

[120] B. Krebs. *KrebsOnSecurity Hit With Record DDoS*. 2016. URL: https://krebsonsecurity.com/2016/09/krebsonsecurity-hit-with-record-ddos/.

[121] B. Krebs. *Most Malware Tied to 'Pay-Per-Install' Market - MIT Technology Review*. 2011. URL: https://www.technologyreview.com/s/424241/most-malware-tied-to-pay-per-install-market/ (visited on 13/06/2017).

[122] B. Krebs. *Ragebooter: 'Legit' DDoS Service, or Fed Backdoor? — Krebs on Security*. 2013. URL: https://krebsonsecurity.com/2013/05/ragebooter-legit-ddos-service-or-fed-backdoor/ (visited on 13/06/2017).

[123] T. Kreing and H. Modderkolk. *Ongrijpbaar aan de Zwarte Zee*. 2017.

[124] S. Kumar and K. M. Carley. 'Understanding DDoS cyber-attacks using social media analytics'. *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*. IEEE. 2016, pp. 231–236.

[125] O. Kupreev, E. Badovskaya and A. Gutnikov. *DDoS Attacks in Q1 2019*. 2019. URL: https://securelist.com/ddos-report-q1-2019/90792/ (visited on 13/06/2017).

[126] E. R. Leukfeldt and M. Yar. 'Applying Routine Activity Theory to Cybercrime: A Theoretical and Empirical Analysis'. *Deviant Behavior* 37.3, 2016, pp. 263–280.

[127] V. I. Levenshtein. 'Binary Codes Capable of Correcting Deletions, Insertions, and Reversals'. *Soviet physics doklady*. Vol. 10. 8. 1966, pp. 707–710.

[128] S. Liu and B. Cheng. 'Cyberattacks: Why, What, Who, and How'. *IT Professional* 11.3, 2009, pp. 14–21.

[129] T. Liu and J. Guo. 'Text Similarity Computing Based on Standard Deviation'. *Advances in Intelligent Computing*. Ed. by D.-S. Huang, X.-P. Zhang and G.-B. Huang. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 456–464.

[130] R. F. Lusch, S. L. Vargo and M. Tanniru. 'Service, value networks and learning'. *Journal of the Academy of Marketing Science* 38.1, 2010, pp. 19–31.

[131] A. C. Mackinlay. 'Event Studies in Economics and Finance.' *American Economic Association* XXXV.March, 1997, pp. 13–39.

[132] A. C. MacKinlay. 'Event Studies in Economics and Finance'. *Journal of Economic Literature* 35.1, 1997, pp. 13–39.

[133] D. Maimon, A. Kamerdze, M. Cukier and B. Sobesto. 'Daily Trends and Origin of Computer-Focused Crimes Against a Large University Computer Network: An Application of the Routine-Activities and Lifestyle Perspective'. *British Journal of Criminology* 53, 2013, pp. 319–343.

[134] MalwareTech. *Mapping Mirai: A Botnet Case Study | MalwareTech*. 2016. URL: https://www.malwaretech.com/2016/10/mapping-mirai-a-botnet-case-study.html (visited on 13/06/2017).

[135] C. Miller. *Kim Jong-il and Me: How to Build a Cyber Army to Attack the U.S.* DEF CON 18. 2010.

[136] J. Mirkovic and P. Reiher. 'A taxonomy of DDoS attack and DDoS defense mechanisms'. *ACM SIGCOMM Computer Communication Review*, 2004.

[137] D. Moore, C. Shannon, D. J. Brown, G. M. Voelker and S. Savage. 'Inferring internet denial-of-service activity'. *ACM Transactions on Computer Systems (TOCS)* 24.2, 2006, pp. 115–139.

[138] C. Morales. *1 Terabit DDoS Attacks Become a Reality; Reflecting on Five Years of Reflections*. (Accessed on 05/16/2019). 2018. URL: https://www.netscout.com/blog/asert/1-terabit-ddos-attacks-become-reality-reflecting-five-years.

[139] G. C. Moura, R. d. O. Schmidt, J. Heidemann, W. B. de Vries, M. Muller, L. Wei and C. Hesselman. 'Anycast vs. DDoS: Evaluating the November 2015 Root DNS Event'. *Proceedings of the 2016 Internet Measurement Conference*. IMC '16. Santa Monica, California, USA: ACM, 2016, pp. 255–270.

[140] G. Moura, J. Heidemann, M. Müller, R. de O Schmidt and M. Davids. 'When the Dike Breaks: Dissecting DNS Defenses During DDoS'. *Proceedings of the Internet Measurement Conference 2018*. ACM. 2018, pp. 8–21.

[141] M. Müller, G. C. M. Moura, R. de O. Schmidt and J. Heidemann. *Recursives in the Wild: Engineering Authoritative DNS Servers*. Tech. rep. ISI-TR-720. Available: https://www.isi.edu/~johnh/PAPERS/Mueller17a.pdf. USC/Information Sciences Institute, 2017.

[142] J. Nazario. 'Politically motivated denial of service attacks'. *Cryptology and Information Security Series*, 2009.

[143] *Netherlands School Holidays*. URL: https://www.schoolholidayseurope.eu/school-holidays-holland/.

[144] L. H. Newman. *A 1.3-Tbs DDoS Hit GitHub, the Largest Yet Recorded*. (Accessed on 05/14/2019). 2016. URL: https://www.wired.com/story/github-ddos-memcached/.

[145] *Norway banks hit in largest-ever DDoS attack, Anonymous takes credit.* 2014. URL: http://ddosattacks.net/norway-banks-hit-in-largest-ever-ddos-attack-anonymous-takes-credit/.

[146] A. Osterwalder and Y. Pigneur. *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers.* John Wiley & Sons, 2010.

[147] P. Paganini. *Botnets, How do they Work? Architectures and Case Studies – Part 2.* 2013. URL: http://resources.infosecinstitute.com/botnets-how-do-they-work-architectures-and-case-studies-part-2/%7B%5C#%7Dgref (visited on 13/06/2017).

[148] P. Paganini. *ProtonMail paid a $6000 Ransom to stop DDoS AttacksSecurity Affairs.* URL: http://securityaffairs.co/wordpress/41775/cyber-crime/protonmail-paid-ransom-ddos.html (visited on 12/11/2015).

[149] R. A. Paulson and J. E. Weber. 'Cyberextortion: an overview of distributed denial of service attacks against online gaming companies'. *Issues in Information Systems* 7.2, 2006, pp. 52–56.

[150] A. Peakflow. *IP Traffic Flow Monitoring System.* URL: http://www.%20arbornetworks.%20com/index.%20php.

[151] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay. 'Scikit-learn: Machine Learning in Python'. *Journal of Machine Learning Research* 12, 2011, pp. 2825–2830.

[152] J. Pieters. *Ziggo: More Cyber Attacks Expected.* Blog. 2015. URL: https://nltimes.nl/2015/08/20/ziggo-cyber-attacks-expected.

[153] *Pokemon Go down: Hacking group claims credit for taking down servers 'with DDOS attack'.* 2016. URL: http://www.independent.co.uk/life-style/gadgets-and-tech/gaming/pokemon-go-down-servers-ddos-attack-hackers-poodlecorp-game-unavailable-a7140811.html.

[154] F. Poldi. *TWINT - Twitter Intelligence Tool.* URL: https://github.com/twintproject/twint (visited on 20/08/2018).

[155] Ponemon Institute. '2016 Cost of Cyber Crime Study & the Risk of Business Innovation'. October, 2016.

[156] L. Ponemon. *Calculating the Cost of a Data Breach in 2018, the Age of AI and the IoT*. (Accessed on 05/14/2019). 2018. URL: https://securityintelligence.com/ponemon-cost-of-a-data-breach-2018/.

[157] R. Poynder. 'LEXIS-NEXIS: Past and future'. *Online and CD-Rom Review* 22.2, 1998, pp. 73–80.

[158] T. C. Pratt, K. Holtfreter and M. D. Reisig. 'Routine Online Activity and Internet Fraud Targeting: Extending the Generality of Routine Activity Theory'. *Journal of Research in Crime and Delinquency* 47.3, 2010, pp. 267–296.

[159] *Presidential candidate websites targeted*. 2016. URL: http://techaeris.com/2016/11/08/presidential-candidate-websites-targeted-unsophisticated-ddos-attacks/.

[160] C. Putman, Abhishta and L. J. Nieuwenhuis. 'Business Model of a Botnet'. *2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*. IEEE. 2018, pp. 441–445.

[161] Radware. *Mirai Botnet: The Rapid Evolution of DDoS Attacks | Radware Security*. 2016. URL: https://security.radware.com/ddos-threats-attacks/threat-advisories-attack-reports/mirai-rapid-evolution/ (visited on 13/06/2017).

[162] S. Raschka and V. Mirjalili. *Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-learn, and TensorFlow*. Expert insight. Packt Publishing, 2017.

[163] R. Ressler and A. Burgess. 'Crime scene and profile characteristics of organized and disorganized murders'. *FBI Law Enforcement Bulletin* 54.8, 1985, pp. 18–25.

[164] C. Riggelsen. 'Approximation Methods for Efficient Learning of Bayesian Networks'. *Proceedings of the 2008 Conference on Approximation Methods for Efficient Learning of Bayesian Networks*. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2008, pp. 1–137. URL: http://dl.acm.org/citation.cfm?id=1563844.1563846.

[165] R. van Rijswijk-Deij. 'Improving DNS Security: A Measurement-Based Approach'. PhD thesis. University of Twente, 2017.

[166] R. van Rijswijk-Deij, M. Jonker, A. Sperotto and A. Pras. 'A High-Performance, Scalable Infrastructure for Large-Scale Active DNS Measurements'. *IEEE Journal on Selected Areas in Communications* 34.6, 2016, pp. 1877–1888.

[167] R. Rodriguez-Gomez, G. Maciá-Fernández and P. Garcia-Teodoro. 'Analysis of Botnets through Life-Cycle'. *Proceedings of the International Conference on Security and Cryptography*. IEEE. 2011, pp. 257–262.

[168] M. J. Rosenberg and R. Foshay. 'E-learning: Strategies for delivering knowledge in the digital age'. *Performance Improvement* 41.5, 2002, pp. 50–51.

[169] T. J. Rothenberg. 'Approximating the distributions of econometric estimators and test statistics'. *Handbook of econometrics* 2, 1984, pp. 881–935.

[170] M. Rouse and M. Haughn. *What is Command-and-Control Servers (C&C Center)?* 2017. URL: `http://whatis.techtarget.com/definition/command-and-control-server-CC-server` (visited on 13/06/2017).

[171] L. Ruddock. 'ICT in the construction sector: Computing the economic benefits'. *International Journal of Strategic Property Management* 10.1, 2006, pp. 39–50.

[172] D. E. Rumelhart, G. E. Hinton and R. J. Williams. 'Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1'. Ed. by D. E. Rumelhart, J. L. McClelland and C. PDP Research Group. Cambridge, MA, USA: MIT Press, 1986. Chap. Learning Internal Representations by Error Propagation, pp. 318–362. URL: `http://dl.acm.org/citation.cfm?id=104279.104293`.

[173] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. 3rd. Upper Saddle River, NJ, USA: Prentice Hall Press, 2009.

[174] G. Salton and C. Buckley. 'Term-weighting Approaches in Automatic Text Retrieval'. *Inf. Process. Manage.* 24.5, 1988, pp. 513–523.

[175] J. J. Santanna, R. van Rijswijk-Deij, R. Hofstede, A. Sperotto, M. Wierbosch, L. Z. Granville and A. Pras. 'Booters - An Analysis of DDoS-as-a-Service Attacks'. *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)*. 2015, pp. 243–251.

[176] J. J. Santanna. 'DDoS-as-a-Service: Investigating Booter Websites'. PhD thesis. University of Twente, 2017. URL: `http://jairsantanna.com/papers/jjsantanna_thesis.pdf`. published.

[177] J. J. Santanna, R. van Rijswijk-Deij, R. Hofstede, A. Sperotto, M. Wierbosch, L. Z. Granville and A. Pras. 'Booters—An analysis of DDoS-as-a-service attacks'. *2015 IFIP/IEEE International Symposium on Integrated Network Management*. IEEE. 2015, pp. 243–251.

[178] D. Santos. 'Technology Investment Announcements on the Market Value of the Firm', 1993.

[179] M. Sauter. '"LOIC Will Tear Us Apart": The Impact of Tool Design and Media Portrayals in the Success of Activist DDOS Attacks'. *American Behavioral Scientist* 57.7, 2013, pp. 983–1007.

[180] M. Sauter. *The Coming Swarm*. Bloomsbury, 2014.

[181] A. Savitzky and M. J. Golay. 'Smoothing and differentiation of data by simplified least squares procedures.' *Analytical chemistry* 36.8, 1964, pp. 1627–1639.

[182] F. Sebastiani. 'Machine Learning in Automated Text Categorization'. *ACM Comput. Surv.* 34.1, 2002, pp. 1–47.

[183] K. Sedgwick. *The Number of Cryptocurrency Exchanges Has Exploded*. URL: `https://news.bitcoin.com/the-number-of-cryptocurrency-exchanges-has-exploded/` (visited on 03/09/2018).

[184] V. Segura and J. Lahuerta. 'Modeling the Economic Incentives of DDoS Attacks: Femtocell Case Study'. *Economics of information security and privacy*, 2010.

[185] A. C. Sharma, R. A. Gandhi, W. Mahoney, W. Sousan and Q. Zhu. 'Building a social dimensional threat model from current and historic events of cyber attacks'. *2010 IEEE Second International Conference on Social Computing*. IEEE. 2010, pp. 981–986.

[186] L. I. Shelley and J. T. Picarelli. 'Methods and Motives: Exploring Links Between Transnational Organized Crime and International Terrorism'. *Trends in Organized Crime* 9.2, 2005, pp. 52–67.

[187] W. Sonnenreich, J. Albanese, B. Stout et al. 'Return on Security Investment (ROSI) - A Practical Quantitative Model'. *Journal of Research and Practice in Information Technology* 38.1, 2006, pp. 45–56.

[188] G. Spanos and L. Angelis. 'The impact of information security events to the stock market : A systematic literature review'. *Computers & Security* 58, 2016, pp. 216–229.

[189] S. M. Specht and R. B. Lee. 'Distributed Denial of Service: Taxonomies of Attacks, Tools, and Countermeasures.', 2004, pp. 543–550.

[190] *Steam connection servers down in probable DDOS attack.* 2016. URL: http://www.pcinvasion.com/steam-connection-servers-probable-ddos-attack.

[191] G. M. Steede, C. Meyers, N. Li, E. Irlbeck and S. Gearhart. 'A Content Analysis of Antibiotic use in Livestock in National US Newspapers'. *Journal of Applied Communications* 103.1, 2019, p. 6.

[192] J. Stewart and L. Stein. *WWW Security FAQ: Securing Against Denial of Service Attacks.* URL: http://www.w3.org/Security/Faq/wwwsf6.html (visited on 25/09/2015).

[193] L. Suganthi and A. A. Samuel. 'Energy models for demand forecasting—A review'. *Renewable and sustainable energy reviews* 16.2, 2012, pp. 1223–1240.

[194] L. J. Tallau, M. Gupta and R. Sharman. 'Information Security Investment Decisions: Evaluating the Balanced Scorecard Method'. *International Journal of Business Information Systems* 5.1, 2010, pp. 34–57.

[195] P. P. Tallon, K. L. Kraemer and V. Gurbaxani. 'Executives' Perceptions of the Business Value of Information Technology: A Process-Oriented Approach'. *Journal of Management Information Systems* 16.4, 2000, pp. 145–173.

[196] D. I. Templer, R. K. Brooner and M. D. Corgiat. 'Geophysical variables and behavior: XIV. Lunar phase and crime: Fact or artifact'. *Perceptual and Motor Skills* 57.3, 1983, pp. 993–994.

[197] P. C. Tetlock. 'Giving Content to Investor Sentiment: The Role of Media in the Stock Market'. *The Journal of Finance* 62.3, 2007, pp. 1139–1168.

[198] *The DDoS vigilantes trying to silence Black Lives Matter.* 2016. URL: https://arstechnica.com/security/2016/12/hack_attacks_on_black_lives_matter/.

[199] D. R. Thomas, R. Clayton and A. R. Beresford. '1000 days of UDP amplification DDoS attacks', 2017.

[200] K. M. Ting. 'Confusion matrix'. *Encyclopedia of Machine Learning and Data Mining*, 2017, pp. 260–260.

[201] *Trends in the Cost of Web Application & Denial of Service Attacks.* URL: https://content.akamai.com/us-en-pg10029-ponemon-cost-of-ddos-web-app-report.html.

[202] *Tumblr Goes Down After Hacker Attack.* 2016. URL: http://news.softpedia.com/news/tumblr-goes-down-after-hacker-attack-511251.shtml.

[203] *University DDoS attack leads to $8.6 million fine, house arrest for New Jersey man.*

[204] Verisign. *Verisign: Managed DNS Services.* 2017. URL: https://www.verisign.com/assets/pdf/resource-center/datasheet-mdns-overview.pdf.

[205] J. Wang, A. Chaudhury and H. R. Rao. 'Research Note—A Value-at-Risk Approach to Information Security Investment'. *Information Systems Research* 19.1, 2008, pp. 106–120.

[206] S. Weagle. *Financial Impact of Mirai DDoS Attack on Dyn Revealed in New Data.* Blog. 2017. URL: https://www.corero.com/blog/797-financial-impact-of-mirai-ddos-attack-on-dyn-revealed-in-new-data.html.

[207] D. A. Weaver and B. Bimber. 'Finding news stories: a comparison of searches using LexisNexis and Google News'. *Journalism & Mass Communication Quarterly* 85.3, 2008, pp. 515–530.

[208] *Web Host Hit by DDoS of Over 1Tbps.* 2016. URL: http://www.infosecurity-magazine.com/news/web-host-hit-by-ddos-of-over-1tbps/.

[209] A. Welzel, C. Rossow and H. Bos. 'On measuring the impact of DDoS botnets'. *Proceedings of the Seventh European Workshop on System Security.* ACM. 2014, p. 3.

[210] B. Widrow and M. E. Hoff. 'Adaptive Switching Circuits'. *1960 IRE WESCON Convention Record, Part 4.* Institute of Radio Engineers. New York: Institute of Radio Engineers, 1960, pp. 96–104. URL: http://www-isl.stanford.edu/~widrow/papers/c1960adaptiveswitching.pdf.

[211] *WikiLeaks comes under 'unrelenting' cyber attack that briefly prevents it from releasing more emails linked to Hillary Clinton on Election Day.* 2016. URL: http://www.dailymail.co.uk/news/article-3917996/WikiLeaks-comes-unrelenting-cyber-attacks-briefly-prevented-releasing-emails-linked-Hillary-Clinton-Americans-polls-Election-Day.html.

[212] *WikiLeaks supporters disrupt Visa and MasterCard sites in 'Operation Payback'.* 2010. URL: https://www.theguardian.com/world/2010/dec/08/wikileaks-visa-mastercard-operation-payback.

[213] *William Hill website under siege from DDoS attacks.* 2016. URL: http://www.theregister.co.uk/2016/11/02/william_hill_ddos/.

[214] *Worldwide Infrastructure Security Report, Arbor Networks.* 2015.

[215] J. Wu and S. Wei. *Time series analysis.* Hunan Science and Technology Press, ChangSha, 1989.

[216] M. Wullink, G. C. Moura and C. Hesselman. 'Dmap: Automating Domain Name Ecosystem Measurements and Applications'. *2018 Network Traffic Measurement and Analysis Conference (TMA).* IEEE. 2018, pp. 1–8.

[217] M. Yar. 'The Novelty of Cybercrime An Assessment in Light of Routine Activity Theory'. *European Journal of Criminology* 2.4, 2005, pp. 407–427.

[218] S. T. Zargar, J. Joshi and D. Tipper. 'A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks'. *IEEE Communications Surveys Tutorials* 15.4, 2013, pp. 2046–2069.

[219] S. T. Zargar, J. Joshi and D. Tipper. 'A survey of defense mechanisms against distributed denial of service (DDOS) flooding attacks'. *IEEE Communications Surveys and Tutorials*, 2013.

[220] V. A. Zeithaml, L. L. Berry and A. Parasuraman. 'The behavioral consequences of service quality'. *the Journal of Marketing*, 1996, pp. 31–46.

[221] C. Zott and R. Amit. 'Business Model Design: An Activity System Perspective'. *Long Range Planning* 43.2-3, 2010, pp. 216–226.

About the Author

Abhishta was born in Meerut, a city in the province of Uttar Pradesh in India on 4th of December 1991. During that period both his parents were working as Lecturers of Economics at a government funded institution in Pasighat which is a city in the north-eastern region of India. He received his primary education in Pasighat. When Abhishta was 8 years old he moved with his parents to Bomdila, another city in the same province as Pasighat. He got his secondary education in Bomdila. Four years later he moved with his parents to the centre of India to Meerut, where he received rest of his school education.

He chose to follow a five year integrated bachelor in industrial engineering and master of business administration at Thapar Institute of Engineering and Technology in Patiala, India. During his undergraduate he worked as an intern at the production plant of Behr GmbH, a manufacturer of heat exchangers where he was involved in the design of ergonomic assembly lines. For his summer internship during his MBA education, he visited University of Twente for a period of 3 months.

Other than carrying out financial measurement studies, Abhishta likes to experiment with modern and traditional ways of cooking. He also enjoys reviewing food and restaurants. He has been playing table tennis at a competitive level since 2015 and has a current NTTB (De Nederlandse Tafeltennisbond) rating of 1080.

List of Publications

The list of peer-reviewed publications (in chronological order) co-authored by Abhishta during his doctoral research are as follows:

- A. Abhishta, M. Junger, R. Joosten and L. J. Nieuwenhuis. 'A Note on Analysing the Attacker Aims Behind DDoS Attacks'. *International Symposium on Intelligent and Distributed Computing*. Springer. 2019, pp. 255–265

- A. Abhishta, R. Joosten, M. Jonker, W. Kamerman and L. Nieuwenhuis. 'Poster: Collecting Contextual Information About a DDoS Attack Event Using Google Alerts'. 2019. Poster presented at 40th IEEE Symposium on Security and Privacy, San Francisco, CA

- A. Abhishta, M. Junger, R. Joosten and L. Nieuwenhuis. 'Victim Routine Influences the Number of DDoS Attacks: Evidence from Dutch Educational Network'. *2019 IEEE Security and Privacy Workshops (SPW)*. 2019, pp. 242–247

- A. Abhishta, R. Joosten, S. Dragomiretskiy and L. Nieuwenhuis. 'Impact of Successful DDoS Attacks on a Major Crypto-currency Exchange'. *2019 27th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*. United States: IEEE, 2019, pp. 379–384

- A. Abhishta, R. van Rijswijk-Deij and L. Nieuwenhuis. 'Measuring the Impact of a Successful DDoS Attack on the Customer Behaviour of Managed DNS Service Providers'. *Computer Communication Review* 48.5, 2018, pp. 70–76

- Abhishta, R. van Rijswijk-Deij and L. Nieuwenhuis. 'Measuring the Impact of a Successful DDoS Attack on the Customer Behaviour of Managed DNS Service Providers'. *WTMC '18*. ACM Press, 2018, pp. 1–7
- C. Putman, Abhishta and L. J. Nieuwenhuis. 'Business Model of a Botnet'. *2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*. IEEE. 2018, pp. 441–445
- Abhishta, R. Joosten and L. J. Nieuwenhuis. 'Comparing Alternatives to Measure the Impact of DDoS Attack Announcements on Target Stock Prices'. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)* 8.4, 2017, pp. 1–18
- Abhishta, R. Joosten and L. J. M. Nieuwenhuis. 'Analysing the Impact of a DDoS Attack Announcement on Victim Stock Prices'. *Proc. of 25th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP'17), St. Petersburg,Russia*. United States: IEEE, 2017, pp. 354–362